

Knowledge from experience: essays

Emmanuel Ordóñez Angulo

Balliol College, Oxford

A thesis submitted for the degree of DPhil in Philosophy

Michaelmas Term 2021

Para Darío, Bárbara y Manuel.

En memoria de Gudelia.

Acknowledgements

For my first intellectual mentorship, thanks to Jolanta Klyszcz. For my first intellectual friendship, thanks to José Antonio Pérez Robleda. For encouragement to keep studying philosophy after a first taste of it, then much support throughout the journey, thanks to Michael Martin. For invaluable teaching, advice, and patience, thanks to Timothy Williamson. For their love and wisdom, thanks to my family, Darío, Bárbara and Manuel, and especially Bárbara, to whom this thesis owes its existence in many ways.

Contents

Introduction	6
Is the concept <i>human</i> a priori?	14
Transformative torture	45
Seeing the infinite	98
What do we see when we see absences?	187

Abstract

In a series of four essays, I explore some ways in which sense experience may deliver knowledge about aspects of reality that have been thought to lie beyond its remit. The opening essay, 'Is the concept human a priori?', deals with moral knowledge, specifically with the question of whether some key moral concepts – namely, those championed by neo-Aristotelian meta-ethics – are, as that theory holds, a priori. The second essay, 'Transformative torture', deals with various kinds of knowledge presumed to be delivered by cinema, most notably, knowledge of the suffering of others. The third essay, 'Seeing the infinite', deals with mathematical knowledge, specifically, knowledge of infinity. Finally, 'What do we see when we see absences?' deals with the question of whether absence experience is veridical.

Introduction

There is a story in Jorge Luis Borges's *Fictions* about a man who perceived the world in full detail, and who remembered everything:

A circumference on a blackboard, a rectangular triangle, a rhomb, are forms which we can fully intuit; the same held true with Ireneo for the tempestuous mane of a stallion... He could perceive I do not know how many stars in the sky... [He remembered] every crevice and every moulding of the various houses which surrounded him.

Secluded in his home's courtyard in a small Uruguayan town, Ireneo Funes spent his days watching clouds and reconstructing entire days – over the course of another entire day – with the precise instruments that were his sense organs and his memory. Unfortunately for him, these odd skills were accompanied by a deficient intellectual department. While his mnemonic and perceptual capacities were overdeveloped, Funes was unable to rise above the individual details of every individual object he perceived:

It was not only difficult for him to understand that the generic term dog embraced so many unlike specimens of differing sizes and different forms; he was disturbed by the fact that a dog at three-fourteen (seen in profile) should have the same name as the dog at three-fifteen (seen from the front).

Our narrator diagnoses Funes's problem as an unawareness of 'platonic ideas'. This unawareness, he tells us, rendered Funes incapable of thinking at all: 'to think is to forget a difference, to generalize, to abstract. In the overly replete world of Funes

there were nothing but details, almost contiguous details'. This identification cannot be right. To say that to think *is just* to generalise, to abstract, seems like an oversimplification of both abstraction and thinking. But the story's premise is the more interesting idea that awareness of such things as 'platonic ideas' would be impossible for someone who only had the faculty of sense perception. This would mean that all we do in perceiving is to concretise, to zoom in, and so that the knowledge delivered by perception is limited to concrete, particular things. Now is *this* contention right?

Borges's setup betrays the belief that the realm of the senses is exclusively populated by concrete particulars, and that their opposite, generalities, belong instead to the realm of thought. It is an honourably ancient belief. The idea that the abstract and the concrete are fundamentally distinct goes back to Plato, for whom the world was divided into empirical reality – the realm of space and time, discoverable through the senses – and the realm of Forms, which the intellect alone can reveal to us. In addition to individual red roses, beautiful people and pairs of things, there exist on this view such entities as the colour red, the property of beauty and the number two. But do we ever see this number? Or bump into redness, or trip over beauty? If the senses only tell us about concrete, individual objects, it follows that our grasp of 'Forms' is not perceptual. So it must be intellectual.

Thus formulated, Plato's theory has not survived. Not many people these days believe reality is divided between spatiotemporal and non-spatiotemporal sectors. But with or without Plato, our knowledge of what isn't concrete remains puzzling. We still draw a distinction between what we can know by sensing and what we cannot. Indeed, one thing *we have* inherited from Plato is the sense that there isn't just a difference in kind but also a normative difference between perceptual and non-perceptual knowledge. Beliefs of the kind we associate with the 'Platonic realm' seem

to have a distinctive, privileged profile with regards to our justification in believing them. When we in fact are so justified, this justification doesn't seem to depend on any course of experience, nor does it seem vulnerable to being defeated by any further such course – the discovery that our perceptual apparatus wasn't functioning as we thought, say, or that we were the victims of a deceitful demon, etc.¹ Knowledge about numbers, moral values and the like – knowledge that, if had, is had regardless of such circumstances – thus receives the label 'a priori'. It is achievable through pure reason. This robustness, and the contrasting frailty of knowledge that *is* empirical, is present in Plato's metaphor that sensible particulars are mere shadows of Forms, as it is in the subsequent thought that perception only gives us a veiled grasp of reality – the same we would have of rich objects from seeing their mere shadows on a wall.

*

A man like Funes would have access to the concrete layer of reality only. Or so Borges, following Plato, tells us. The essays in this thesis explore ways in which Funes may have enjoyed a richer epistemic life than Borges would allow. In these essays I probe the line between what we can know by sense experience and what we cannot.

It is helpful to start by acknowledging what Tim Williamson (2000, p. 165) calls a standard distinction between the two epistemic roles of sense experience. When a piece of knowledge has been acquired without the use of the senses except for acquiring the constituent concepts, the role of experience is 'enabling'. 'All vixens are foxes', for example, is a proposition for which experience only contributed the concepts *vixens* and *fox*. Experience plays a more substantial, 'evidential' role in cases where the knowledge in question could not have been acquired by reflection alone. 'That vixen is mauve', for example, is a proposition we cannot know except by looking; 'my feet are cold', a proposition we cannot know except by feeling. ('My

¹ I'm adapting a characterisation of a priori justification in terms of defeat, as found in e.g. Field (2005).

neighbour's pet vixen is mauve' can be known by testimony in addition to observation, maybe, but our example involves a demonstrative, 'that vixen', which typically fixes its reference by, well, demonstration, that is, by the subject's pointing to the target object, which typically involves her ability to see it.)

Williamson goes on to argue that the line between the two roles is fuzzy, such that for some pieces of knowledge it is not clear whether sense experience played one role, the other, or both. Still, there are epistemic domains we paradigmatically associate with one or the other of the roles. Funes would have been, in these terms, banned from knowing anything beyond that for which experience can provide evidence. What cases would those be? Well, experience might be unable to play an evidential role for a body of knowledge because the entities such knowledge is about are not perceptible (which might in turn be because it isn't clear those entities exist, or because, if they exist, they might not be apt to feature in experience), or because, as a matter of fact, there is no experience that would allow one to arrive at the relevant kind of knowledge.

Take, for instance, knowledge about absences. Absences are a kind of entity with a controversial ontological status: it isn't clear we should accept them among existents. If existence is necessary for perceptibility, the latter is even more fraught a question for absences. Funes is predicted to be unable to know about them, since he only knows about what he can see. Would he have mourned his pets, then, or his mother, once they were gone? How would he react to total darkness?

Take, on the other hand, mathematical and moral knowledge. Mathematical objects and moral values are perhaps less suspect than absences, but even a realist about them would typically concede that they are not the kind of thing with which humans can enter into perceptual relations, at least if we take a view of perception along the lines of Grice (1961) on which a causal link between perceiver and

perceived entity is necessary for perception. Would Funes have known right from wrong, then? Or grasped the relation between a single apple, a pair of apples, and a triple of them?

Interestingly, Borges's story is explicit about the mathematical case. Funes was indeed acquainted with some numbers: he knew of the thirty-three patriots who lost their lives for the independence of his country. But for every number Funes 'knew' he assigned a proper name, and proper names function precisely to single out particular things in a way that leaves out their relations to others:

Instead of seven thousand thirteen (7013), he would say, for instance, "Máximo Pérez"; instead of seven thousand fourteen (7014), "the railroad"; other numbers were "Luis Melián Lanifur", "Olimar", "sulfur"... instead of five hundred (500), he said "nine".

On what the most popular contemporary view of mathematics, though, the number 1 is not a 'thing' to be christened like dogs and people but just a place in the structure of the natural numbers, namely the second place (the first is zero). If that is so, Funes was not only unaware of the way all natural numbers are related but also unaware of the kind of entity any particular number is, and so unaware of each particular number he had a name for and *thought* he knew. 'Maximo Pérez' was not Funes's idiosyncratic name for what he knew to be the 7014th place in the natural number structure. With that expression, he wasn't naming anything.

Finally, take knowledge about the experiences of others. Knowledge of our own minds – in a comprehensive sense of the term to include not just propositional attitudes but sensory and affective experiences – can indeed be achieved, at least in part, through sense experience. We know that we are cold, hot or angry when we

experience numbness in our limbs, our bodies soaking in sweat, our heart racing. But knowledge of the thoughts and feelings of others is mainly delivered by another method, testimony, and perhaps distal perception. Yet it is unclear Funes would have been able to understand that ‘I am bored’ in someone else’s mouth, or even someone else’s sighs, which he could see, had some relation to what he himself felt listening to the radio all afternoon, every afternoon, alone in his courtyard.

So knowledge of absences, moral and mathematical knowledge, and knowledge of the experiences of others seem cases where sense experience plays non-evidential roles. These domains are confined, if not to the armchair, at least to the intellectual realm. The essays in this thesis question this verdict. In ‘Is the concept human a priori?’, my subject are certain moral concepts. I ask whether those key concepts championed by neo-Aristotelian meta-ethics – namely, *life-form* and *human* – are really a priori, as that theory requires in order keep worries about its letting ‘empirical concepts into ethical theory’ at bay. In ‘Transformative torture’, I deal with various kinds of knowledge presumed to be delivered by cinema, most notably knowledge of the suffering of others. This knowledge is largely thought to be inaccessible from a third-personal perspective. My question here is whether novel technologies manage to bridge the boundary between viewers and characters that is the screen, and if so, whether this means that, as recent film critics have suggested, those technologies finally give viewers first-personal knowledge of the experiences of others. In ‘Seeing the infinite’, I ask whether perceptual knowledge of infinity is possible. This knowledge is cashed out in terms of the ability to conceive of a model for the Dedekind-Peano axioms of arithmetic, such that the question becomes whether perceptual knowledge about those axioms is possible. Finally, in ‘What do we see when we see absences?’, my topic is the perceptual experience of absence. A growing body of literature about the phenomenological and epistemic role of the

structural features of experience, as well as on the topic of absence experience itself, point toward the verdict that absences are not entities we perceive but, much like Platonic Forms, something we non-perceptually know. I insist this contradicts ordinary discourse according to which we take ourselves to speak truthfully when we attribute observable properties to absences just in the way we do to present entities, and ask whether ordinary discourse can be vindicated in the face of such pessimism.

*

Borges's story ends on a sad note. Having conversed with him all night, the narrator only sees Funes's face at daybreak. He is struck by it:

Ireneo was nineteen years old; [yet] he seemed to me as monumental as bronze, more ancient than Egypt, older than the prophecies and the pyramids.

Unable to turn his mind from the empirical world, Funes had grown old too quickly. The inability 'to generalise, to abstract' had resulted in something like sensory fatigue. If the spirit of these essays is right, the story needs a different ending. One way to see the lesson would be to think that our experience of the world does give us access, after all, to something like the realm of Forms. Or perhaps, on the contrary, we should think empirical reality is richer than Borges – and Plato, and much philosophy – has supposed. There may be other lessons, too. Perhaps the significance of sense experience is best appreciated not in isolation but, contrary to Borges's set-up, by seeing it as a cognitive faculty, that is, as part of the knowledge-yielding capacity that thrives from the interaction between perception, reasoning and the imagination. Another lesson is that there may be links between mathematics, morality, first-person thought, and practical knowledge (knowledge about our

intentions and action) that can be exploited in new ways. Some of these links come up repeatedly throughout the essays, as the reader will soon see. In trusting the epistemic gateway of his senses, at any rate, Funes would *not* have been ‘the solitary and lucid spectator of a multiform, [...] almost intolerably precise world’, or not only this; even from ‘his poor South American suburb’, he would have been a powerful epistemic agent, one with the ability to entertain general and abstract thoughts about himself and about the larger world just like the rest of us, in being perceptually open to that world, are and have the ability to do.

References

- Field, H. 2005. Recent Debates about the A Priori. In T.S. Gendler & J. Hawthorne (eds.), *Oxford Studies in Epistemology* 1. Oxford: OUP
- Grice, H. P. 1961. The causal theory of perception. *Aristotelian Society Supplementary Volume* 35, doi: 10.1093/aristoteliansupp/35.1.121
- Williamson, T. 2000. *Knowledge and its limits*. Oxford: OUP

Is the concept *human* a priori?

Abstract

According to Michael Thompson's rendition of neo-Aristotelian naturalism in meta-ethics, the scope of our moral propositions (e.g. 'it is impermissible to do A' or 'one has reason to do B') is the same as the scope of our 'natural-historical judgements': judgements, in our case, about the specifically human life-form (2004, p. 62). The worry arises that this amounts to 'introducing something empirical or even biological into ethical theory' (p. 69). In response, Thompson suggests that, contrary to appearances, (i) '[t]he concept *life-form* is a pure or a priori, perhaps a logical, concept', and that (ii) '[t]he concept *human*, as we human beings have it, is an a priori concept' (p. 57). Here I contend that Thompson's argument for (ii) is unsound, and hope to thereby offer some light on the a priority of concepts more generally.

“¡Ah de la vida!” ... ¿Nadie me responde?

– Francisco de Quevedo²

Introduction

There is probably no more succinct characterisation of the neo-Aristotelian naturalist view in meta-ethics than Philippa Foot’s slogan, early in *Natural Goodness*, that ‘the grounding of a moral argument is ultimately in facts about human life’ (2001, p. 24). Eyebrows were bound to be raised. Wasn’t moral knowledge, like logical and mathematical knowledge, supposed to be a priori? What is an empirical concept, *human life*,³ doing in its grounding?

This is the worry Michael Thompson (2004, 2012) sets out to assuage in his project of lending epistemological support to the naturalist’s position. According to Thompson, the scope of our moral propositions (e.g. ‘it is impermissible to do A’ or ‘one has reason to do B’) is the same as the scope of our ‘natural-historical judgements’: judgements, in our case, about the specifically human life-form (2004, p. 62). Thompson puts the corresponding worry as the thought that this amounts to ‘introducing something empirical or even biological into ethical theory’ (p. 69). In response, he suggests that, contrary to appearances, (i) ‘[t]he concept *life-form* is a pure or a priori, perhaps a logical, concept’, and that (ii) ‘[t]he concept *human*, as we human beings have it, is an a priori concept’ (p. 57).

² ‘*¡Ah de la vida!*’ is Quevedo’s repurposing of an old Spanish call used by newcomers to enquire about the presence of someone where they have arrived: at a house (*¡Ah de la casa!*), at a castle (*¡Ah del castillo!*). The comedic effect in the sonnet is achieved by the phrase after the pause (*¿nadie me responde?*), which highlights the pointlessness of making this kind of enquiry about life itself. Isn’t the enquirer his own answer?

³ I use italics to denote concepts.

This article seeks to advance three claims: that (1) the concept *human* is not a priori, even if (2) the concept life form is; and that (3) Thompson's argument for the a priority of *human* is unsound. But my hope is that the discussion sheds some light on the nature of a priori concepts more generally.

1

What is, to begin with, an a priori concept, and why are a priori concepts important to neo-Aristotelian naturalism? We should begin by noting that the 'a priori' qualification is most usually applied to propositional knowledge. The most natural way to read the claim that moral truths are knowable a priori is as the claim that the epistemic justification of moral truths is independent of experience, to the notion of non-empirical knowledge. To say that moral knowledge is a priori comes down, then, to saying that it is non-empirical. Thus no course of experience could settle for one the question of whether it is impermissible to do A, for instance, in the way it *can* settle the question of whether to do A is quick or pleasurable. Moral naturalism claims that moral facts are knowable a priori even though they are grounded in facts about something that, one would think, is part of empirical reality. the human life-form. This is prima facie problematic because the justification of a truth, one would think, tends to accord with the justification of the truth that grounds it. So how can a seemingly empirical truth like 'humans are F' ground a non-empirical one like 'it is impermissible to do A'?

Thompson's answer is that the method by which we acquire knowledge of truths about the human life-form is not, in fact, empirical. 'Human beings are characteristically in possession of some general substantive knowledge of the human life form which is not founded on observation of members of their kind, and thus

not “biological”, he writes (2004, pp. 57-8). To say this is to say something about the epistemology of the human life-form, not about its subject matter. It is to say, in other words, that such knowledge is not like knowledge of other portions of empirical reality – which *is* founded on observation – but instead like moral knowledge. Both moral truths and truths about the human life-form are about observable, concrete entities, Thompson is telling us, namely moral agents and their actions. The difference is just that where the content of truths about the human life-form may be entirely empirical (‘humans have two feet’, ‘humans care for each other’), moral truths predicate of such concrete entities non-observable properties (‘John is blameworthy’, ‘killing is impermissible’).

So much seems incontestable. What worries Thompson is that the connection between these two kinds of claims – the grounding relation between them – that naturalism posits is an object of suspicion. ‘There is in the larger literature a kind of fear or dread of any appeal to this sort of concept in ethical theory’, Thompson writes, speaking of a concept, like *human*, that ‘attaches to a definite product of nature, one which has arisen on this planet, quite contingently, in the course of evolutionary history’. ‘[B]ut if the naturalist is right, the concept in question is everywhere nipping at his heels’ (op cit., p. 59). The solution Thomson envisages to show that the naturalist is indeed right is to demonstrate that the concept naturalists ‘propose to make central to ethical theory’ is acquired independently of observing what it attaches to. ‘The concept *human*, as we humans have it, is an a priori concept’ (p. 58).

Why should this solution be attractive to the naturalist? Naturalism, as Thompson sees it,⁴ is the view that the central concept of ‘ethical theory is not the

⁴ Thompson’s sources are, besides Foot’s *Natural Goodness*, ‘the concluding paragraphs of Elizabeth Anscombe’s “Modern Moral Philosophy” and [...] the last part of Rosalind Hursthouse’s book *On Virtue Ethics*’ (Thompson 2004, p. 58).

abstract idea of a rational being or a person’, so not ‘what Kant meant in speaking of “humanity”’ (p. 58) but rather the concept of something non-abstract, ‘the idea of the *human* – that is, the concepts of a *human being* and of the specifically human life form and of so-called human nature’ (ibid.). This sounds like a revisionary view on the subject matter of ethics, but perhaps the revision is on a different level. Views that make the central concept of ethics that of *person* may agree with naturalism on the extension to which the relevant concepts apply: human beings, us. But where to theorise in terms of the concept *person* is akin to theorising about *object* or *concept* (an ‘abstract idea’), to theorise in terms of *human* is akin to theorising about *water* (‘a product of nature’). Just as no course of experience can tell us whether something is a Fregean *object* or a Fregean *concept*, so no course of experience can tell us whether *personhood* is borne by human adults only or by foetuses – primates, aliens, other animals – too. Those are matters of a priori reasoning. And *human* is, we would think, prima facie not like this. To theorise in terms of *human* is to theorise in terms of a concept that applies or not to a concrete individual as a matter of empirical fact. The issue is that moral claims, while in fact speaking about those empirical beings, are not meant to be justified by observing them because the properties with which they are concerned are indeed like *person* or *object* or *concept* in that no course of observation will tell us whether they apply to them – whether an action is *impermissible*, an agent *blameworthy*. Rendering *human* like *object* or *concept* in this respect would smooth things out for moral naturalists. It would protect moral naturalism from the ‘fear or dread’ others might see in its use of *human* as the central moral concept; it would legitimise its view of ethics as a theory that does not, after all, tread on empirical – ‘biological’ – evidence.

So this is why the a priority of *human* is important for Thompson. Now, Thompson himself does not say much about how he understands the notion of an a

priori concept. Compared to the notion of a priori knowledge, the notion of a priori concepts has received little attention,⁵ perhaps because it's routinely taken for granted. One natural first thought would be to fix the notion of a priori concepts analogously to that of a priori knowledge: '[a]n a priori concept is one that can be acquired independently of experience, which may – but need not – involve its being innate' (Baehr 2003). What exactly does this amount to? It is certainly intuitive that there should be a property of a priori knowledge that has an equivalent in a priori concepts. However, what is independent of experience in a priori knowledge is not learning or acquisition but justification. A piece of knowledge might be acquired through experiential means without its justification depending on experience. Consider, for example, learning, through Cantor's argument, which involves visual engagement with a grid, that the set of natural numbers has the same cardinality as the set of rational numbers. The learning method was experiential, but the learnt truth is a priori.

Similarly, a concept might be acquired by experiential means without the concept itself depending, for being available in our conceptual repertoire, on experience. Take the concept *triangle*. Understanding triangularity does not depend on experience because the contents of geometrical concepts, like those of all mathematical concepts, can be specified without mention of anything empirical. *Triangle* is just the concept of a closed figure which has three angles.⁶ Yet *triangle* can be acquired not only by geometrical reasoning but also by encountering instances of the concept in the world – halves of sandwiches, obelisk tops. Thus, *triangle* can be

⁵ This is true of the notion as I'll understand it. Understood as the notion of innate concepts, it can be traced back to Locke and Leibniz's debate on innate 'ideas', i.e. on whether we are born already with something of 'whatever it is that the mind can be employed about in thinking' (Locke 1689/1975, Ch. 1, §8).

⁶ This is not meant to be an argument for the a priority of *triangle*. Someone sceptical of the priority of *triangle* would presumably also be sceptical of the a priority of *closed figure*, *angle*, and *three*.

said to be an observable concept even if it is not ‘observational’, the latter being Steve Yablo’s label for concepts the possession of which *does* require the ‘ability to work out [their] extension in *perceptually* (as opposed to intellectually) presented scenarios’ (2002, p. 461, emphasis original). Yablo’s example of an observational concept is interesting because it may seem mathematical, thus a priori, but it is not: the concept *oval*. ‘[W]hat marks a figure as oval is not its satisfaction of some objective geometric condition, but the fact that when you look at it, it looks egg-shaped’ (ibid., p. 465).⁷

Now, learning an a priori concept by experience is sometimes not possible because it is neither observable nor observational. The concept *continuous curve*, for instance, which is essential in geometrical reasoning, is impossible to derive from visual experience, perceptual or imaginary, because, as Peter Epstein (2018a, b) shows, there is no visual difference between the properties of continuity and mere denseness. The gaps in a dense curve are by definition invisible, such that no feat of the imagination, let alone of perception, could distinguish between the two. Mention of anything that can be experienced, then, isn’t merely unnecessary but impossible in a specification of the content of the concept *continuous curve*.

I propose to take a view of a priori concepts as those which may or may not pick out things we can encounter in experience (as *triangle* and *continuous curve*, respectively, show) but which in any case do not require experience for being available to be deployed in thought. Because specification of neither of our geometrical examples – *triangle* and *continuous curve* – mentions empirical terms but only ‘objective geometrical conditions’, as Yablo puts it, and these conditions seem

⁷ Chalmers (2002, p. 190) voices a reasonable doubt about Yablo’s example. Is *oval* not just like *triangle* in that it’s a shape concept? But Yablo implicitly draws something like a distinction between shape and what I’ll call *geometrical-shape* concepts. Only the latter could be involved in geometrical reasoning: ‘Why are the oval things picked out experientially? There is no in-principle reason, but only a practical one: we have no other way [no theoretical way] of roping in the intended shapes’ (Yablo op cit., p. 466).

akin the to introduction and elimination rules of logical concepts, we can say that a priori concepts can become part of our conceptual repertoire just thanks to our mastery of some such rules or conditions, of which the mastery may – but need not – be innate.⁸

Another feature of a priori concepts, that their content does not vary across contexts or worlds, will arise in the next section. With the above characterisation in place for now, though, we are ready to note Thompson wants *human* to be like this. Even though *human* is instantiated in empirical reality, that is, even though it picks out ‘products of nature’, the concept’s availability for use in thought doesn’t require observation of those natural entities any more than the availability of the concept *person* – or *object*, or *concept* – requires observation of persons, objects, concepts. Or so Thompson wants to say. ‘The concept *human* is a pure concept of the understanding devoid of even the least empirical accretion’, he writes (2004, p. 69). If this is right, it would put the naturalist’s view of ethics on the same a priori footing as *person*-centred views, and indeed as other a priori fields, such as logic and mathematics.

A brief look at the literature on concept possession will help to frame our assessment of Thompson’s defence of such an ambitious epistemological point.

2

If we take concepts to be sub-propositional mental representations, then to possess a concept is to have certain representational capacities. In virtue of allowing us to think about things by representing them in certain ways, concepts allow us to undertake

⁸ This doesn’t require the ability to articulate such rules, just to conform to them. More on this shortly.

certain epistemic tasks. One such task is putting things into categories, which the literature on concept possession calls ‘sorting’. The idea is that when a subject can sort things into red and non-red, say, she meets one condition for possessing the concept *red*.⁹ Categorising things is perhaps more obviously a function of empirical concepts like *red*, but it is only less obviously so of geometrical concepts like *triangle* and *continuous curve*, and even logical concepts, like *object* and *concept*. Frege (1892/1984) himself famously categorised the referent of ‘the concept *horse*’ as under *object*, for example.

Now, one may successfully categorise together triangular things while regarding them not as triangular but as falling under the distinct, co-extensional concept *three-sided*. So, additionally to the capacity to get the concept’s extension right, in order to count as possessing *triangle* the subject has to be disposed to infer that triangular things have angles rather than being disposed to infer that they have closed sides, which disposition comes, instead, with possession of the concept *three-sided*. Possessing concepts endows one with the ability to make the inferences that are allowed by the concept’s ‘constitutive structure’: *angle* is structurally constitutive of *triangle*, and *three* and *side* of *three-sided*.¹⁰ This second feature of concepts – call it inference licensing – is of particular importance with regards to our interest here, a priori concepts. Take the logical concept *and*. Jerry Fodor writes: ‘a sufficient condition for a speaker’s meaning *and* by ‘and’ [is] that, ceteris paribus, he takes ‘P and Q’ to be true iff he takes ‘P’ to be true and ‘Q’ to be true’ (1990, p. 111). One understands the concept *and* if one makes inferences that conform to its rules of introduction and elimination.

⁹ See e.g. Fodor (1990, 1994), Peacocke (1992), Weiskopf and Bechtel (2004), Ginsborg (2006).

¹⁰ Weiskopf and Bechtel (op cit), p. 52.

Finally, a third condition for concept possession, or perhaps a condition underlying the other two, is that the subject must be aware that her way of sorting things and reasoning about them are appropriate to what she is sorting and reasoning about. Hannah Ginsborg (2006) calls this an awareness of normativity. This condition is a little tricky. Concepts themselves come with norms, as we have just seen, so Ginsborg's condition may seem to presuppose what it is meant to help to analyse, but she distinguishes between the 'primitive normativity' this condition involves and the normativity that does come with possessing the concept (call the latter 'conceptual normativity'). The priority of primitive normativity is theoretical, but it can also be chronological. Take an infant who is learning the concept *triangle* by being exposed to triangular and square objects. Before actually mastering the concept *triangle*, she might start sorting the triangular objects together and apart from the square objects. At this point the infant would meet Ginsborg's condition that 'she take it that she ought to perceive [the object] *this* way, [...] where her taking it that she ought to perceive the object in this way does not depend on any prior appreciation – implicit or explicit – of how it ought to be perceived' (ibid).¹¹ The condition does not involve, then, a judgement about the objects themselves – a judgement about things themselves being Fs – but a judgement about what one is doing, about one's way of representing things. It is a species of self-awareness, something like the judgement '*this* way of representing is appropriate'.

In short, concepts allow us to correctly reason about and categorise things in virtue of allowing us to represent things in a way we know to be appropriate. An a priori concept allows us to display this behaviour independently of experience. Differently put, a subject's ability to hold knowledgeable beliefs featuring these kinds

¹¹ Ginsborg puts this in terms of perceiving because this particular discussion of hers is concerned with empirical concepts, but the point generalises if we replace 'perceive' with 'represent'.

of concepts – featuring *object*, *concept*, *triangle* – would not be predicted by any course of experience in the way experiences of oval and red things might indeed predict such abilities with respect to *oval* and *red*. Our question is whether *human* is in the same class of concepts as *object* and *concept* rather than *oval* and *red*.

To delve into Thompson's views now: in order to argue that *human* is a priori, Thompson proceeds by showing first that the concept *life-form* – 'the notion, that is, of a living kind' (Thompson 2004, p. 47) – is a priori too. This is because *life-form* will figure largely in the argument about *human*. So Thompson starts by suggesting that the empirical criteria that have been suggested by scientists to distinguish living things from non-living ones (he pays special attention to reproduction and homeostasis) assume that living things are a special *kind of kind* of thing. Hence, those criteria don't say what marks this kind of kind as special. The biological notions of reproduction, understood as *self-reproduction* (or 'form-transmission') and homeostasis, understood as *keeping oneself the same* (or 'form-maintenance') assume that the living thing's *form* is unlike other forms that can also be transmitted or maintained but of which the processes of transmission or maintenance is not a *vital* process, a phenomenon of life. Take crystals, which have been observed to be involved in form-transmission processes not unlike the process amoeba exhibit when *they* reproduce themselves. Yet, crystals are not regarded as alive. Corpses can be made, at least conceivably (Thompson's is a thought experiment) to tend to equilibrium more or less in the way living things do when we say they engage in form-maintenance; yet, we regard corpses as no longer alive. This last example of Thompson's is reminiscent of Donald Davidson's (1987, pp. 443-4) point that if a molecule-for-molecule replica of him were accidentally created by a lightning's strike in a swamp, the Swampman could not be regarded as having thoughts, with which Thompson evidently agrees (Thompson 2012, p. 60). But

Thompson agrees because he thinks Swampman should not to be regarded as a living thing at all. Thinking, we will witness Thompson argue later, is one among many processes characteristic of living – and indeed, human – beings only. So what is it that Swampman lacks, if it has all the molecule-for-molecule properties the original Davidson does? This all is meant to show that in trying to specify what empirical properties a thing must exhibit to count as living, scientists have already assumed what special kind of *form* it is that the process of reproduction *form*-transmits and of homeostasis *form*-maintains, or in other words, they have assumed the concept *life-form*.

That is the negative part of Thompson's argument. The upshot is that assuming the concept *life-form* in propositions about living kinds of things – including the propositions that were meant to mark off living kinds in the first place – is no different from assuming the Fregean concepts *object* and *concept* in any proposition. Thompson's argument for the a priority of *life-form* is thus something of an argument by analogy. There is no such thing as giving an empirical account of what makes something a Fregean object or concept; no such thing, that is, as giving an account of objecthood or concepthood in terms of empirically knowable properties. Rather, *object* and *concept* characterise what falls under them as suitable to play certain linguistic roles. Specifying what an object or a concept is makes no mention of anything empirical: something is an *object* if it can be designated by the subject term of a proposition of the form $\ulcorner \alpha \text{ is } \Psi \urcorner$, say,¹² and so if predicates can be true or false of it. And a *concept* is whatever is designated by the predicate in such a proposition. Similarly, the concept *life-form* characterises whatever falls under it as suitable to play a certain role in a certain form of proposition. Like all kind-concepts, life-form-concepts are suitable to occur in the subject position of 'general

¹² Where α is a singular term and Ψ a predicate.

judgements’, but when a general judgement is about this special kind of kind, the judgement’s ‘grammar’ will differ from the grammar of other general judgements, e.g. from universally quantified sentences (‘all triangles have three sides’) or other generic sentences (‘gold is a precious metal’).¹³ Thompson’s name for propositions about this special kind of kind is ‘natural-historical judgements’. Examples are ‘amoeba are/have/do F’ and ‘humans are/have/do G’.

Thompson’s suggestion that *life-form* characterises whatever falls under it as suitable for a certain role in propositions with certain rules of inference echoes what we’ve called the ‘inference licensing’ condition of concept possession. To say that natural-historical judgements have a different grammar from other forms of propositions is to say, in the terms of the previous section, that they involve different rules of inference. A subject who deploys a life-form-concept in a judgement of the form ‘Fs are/have/do G’ is not disposed to judge that if some x doesn’t do G, it ipso facto counts as a non-F. This is because natural-historical judgements are not universal quantifications. Instead, the subject is disposed to judge that a non-G-ing F is a defective F. Where the grammar of propositions like ‘all triangles have three sides’ *does* dispose one to judge that a non-three-angled shape is simply not a triangle, the grammar of propositions like ‘cats meow’ disposes us to judge that a mute cat is still a cat, albeit a defective one. These two are simply the kinds of inferences that universally quantified and natural-historical judgements, respectively, license.^{14,15}

On the other hand, Ginsborg’s condition of awareness of normativity sits well too with Thompson’s view on *life-form*. It may be what explains scientists’

¹³ Thompson (2012), pp. 25-7, 76-82.

¹⁴ This rule of inference is extracted from Thompson’s (2012, Ch. 3, §2) characterisation of the ‘grammar’ or formal properties of natural-historical judgements.

¹⁵ Many a reader will recall that something like this is prefigured by Kripke when he writes that ‘[t]he phrase “a three-legged tiger” is not a *contradictio in adjecto*’ (Kripke 1972/2001, p. 119). This is just one of several points at which the discussion is, not surprisingly, reminiscent of Kripke’s.

disposition to make such confident natural-historical judgements – judgements about living kinds – in the very process of (hopelessly) attempting to pin down the empirical features that mark living kinds off from non-living kinds. Here’s a smaller-scale version of the phenomenon: consider a toddler sorting crabs and stranded jellyfish together, and apart from sea debris, while walking on the beach. Later, when the child learns that crabs and jellyfish typically reproduce themselves and tend to homeostasis, she will become disposed to make certain judgements and not others about them in virtue, now, of not just thinking that her way of representing crabs and jellyfish was correct but also in virtue of thinking that they are – hence, ought to be conceived as instantiating – certain forms of life.

The upshot is that Thompson seems to be right that *life-form* is, like *object* and *concept*, part of what he calls, in Kantian phrase, our ‘pure concept[s] of the understanding’ (2004, p. 69).

3

In light of the above considerations, I will have nothing to say against Thompson’s contention that *life-form* is a priori. My target will be Thompson’s claim that one of the concepts we deploy in natural-historical judgements, the concept *human*, is a priori too. Now this is more contentious. On the face of it, *human* is a natural kind concept, and Thompson acknowledges that this is a reasonable first thought (2012, p. 47): natural kind concepts are typically thought to be acquired upon exposure to the natural world. But this is, of course, the empiricist position he sets out to oppose.

Thompson’s strategy to argue for the a priority of *human* is altogether different from his strategy for *life-form*. To put things in context: the a priority of *human* is Thompson’s response against the general empiricist view that ‘[c]oncepts of

particular life-forms [...] are invariably empirical, or observation-dependent, concepts' (2004, p. 57), which is, in turn, the consequence of another empiricist view, the view that '*singular* representations of individual organisms are invariably empirical representations' (ibid., my italics). Thompson's argument against these two views works in the singular-to-general direction. It is true, Thompson says, that the singular representations we deploy of living organisms are usually based on observation. But there is one exception: 'the peculiar singular representation mature humans are able to effect, a representation, namely, of themselves through the first person, or through the *I*-concept [henceforth *I*]' (p. 66). When I use *I*, Thompson wants to say, I am representing a spatiotemporal individual, the one located in this chair, in a way that is nevertheless not based on observation of it. That this representation of an individual organism, achieved through *I*, is not empirical, provides a counterexample to the second empiricist view above, which in turn demonstrates that the *I*-concept is a priori.¹⁶

So Thompson takes himself to have secured two truths about a priority before arguing for the a priority of *human*. First, (i) the a priority of *life-form*, and second, (ii) the a priority of *I*. A third point needs to be made before tackling *human*:

¹⁶ The above can seem too swift a move on Thompson's part, but much of the work for the a priority of *I* has been done for him by a history of complaints against Descartes, on the one hand, and a Wittgensteinian scepticism about the first-person, on the other. Consider Kant's view that, as Peacocke puts it, Descartes 'is mistaking this "merely logical subject" [i.e. the first person deployed in self-ascriptions] as something more substantial' (2014, p. 158). Russell, on his part, complained that the word 'I' is not a proper name. About the argument 'I think, therefore I exist', he wrote: "[h]ere the word "I" is really illegitimate; he ought to state his ultimate premiss [i.e. 'I exist'] in the form "there are thoughts." The word "I" is grammatically convenient, but does not describe a datum' (1945, p. 567). Later, Wittgenstein argued that 'I' or 'my' don't only fail to refer; there is nothing to be found in one's experience of the world for those words to aim, and fail or not, to refer to (1922/1971, 5.631). Anscombe (1981) followed Wittgenstein with the suggestion that the first person refers to neither a physical nor a non-physical substance, that is, to nothing at all. So Thompson is not alone in thinking that, either for referential or extra-referential reasons, *I* is void of empirical content. A difference between a priori concepts and *I* that respects their epistemological similarity will be discussed shortly, however.

(iii) that in representing an individual organism or its activities, one is implicitly representing its life-form. Thompson argues for this point thus:

Even the most pedestrian case of life-description, say, that the cat is drinking the milk, must make an implicit claim about the relevant ‘form’ or ‘context’—that *for it*, or *in it*, the events before us add up to drinking; or that what the creature is doing is drinking, *for such as it is* (Thompson 2012, p. 57).

The ‘context’ is the organism’s life-form. What counts as drinking for a cat might count as vomiting for a reverse-Earth-cat. So one cannot assert of an individual that it is A-ing without assuming that it is an F. This is what it means for the representation of a life-form to be implicit in the representation of any phenomenon of life.¹⁷ Two such phenomena, Thompson tells us, are self-conscious thought and speech. Thus, when holding a self-conscious thought, I am implicitly tokening the concept *human*, not in virtue of referring to myself but just in virtue of representing what I am doing, thinking, as *thinking*,¹⁸ which very same activity might count (who knows?) as vomiting for a weird-Earth-human but counts as thinking for me, *such as I am*.

With all three points in place, Thompson is in a position to make his key claim in defence of the a priority of *human*: that *human* is effectively tokened whenever any *I*-thought is formed. And because an *I*-thought needn’t involve empirical content, nor does *human* (pp. 66-7). The argument proceeds in two steps. First, Thompson submits the claim as a consequence of the points (i) and (ii). In an *I*-thought, I represent myself, and when I represent myself, I represent my life-form,

¹⁷ Thomson (2012), pp. 56-62.

¹⁸ I use *thought* and *thinking* interchangeably.

which is to say that I deploy the concept *human*. The point extends to any indexical concept of the sort, naturally: '[e]ach of us can lay hold of this item [the concept *human*] under the title 'my life-form' or 'the life-form I bear', descriptions which contain no empirical content at all' (ibid.).

Thompson considers here the objection that a deployment of *I* or of *my life-form* can be empty. In this scenario, the *I*-thought in question would not, as he says, 'carry [the concept *human*] onto the scene' in the way it does when it succeeds in referring. His response to this will be to say that *I* needn't appear in a self-conscious thought for this thought to nevertheless deploy the concept *human*. Thompson doesn't elaborate on the details of such a scenario, but here is a similar one that could help us to appreciate the point. In a scenario where I am seeing a full-length picture of a man who looks just like me – a picture of my lost twin – but I think I'm seeing a mirror, I might judge 'I am wearing a red tie', and be mistaken about that. If I feel pain, by contrast, and think 'I am in pain', I cannot be similarly mistaken because being in pain is the kind of property that is constituted by the subject's very experience of it.¹⁹

What if I have always been in a situation like that of the man who mistakenly thinks he's wearing a red tie while being, unbeknownst to me, a non-living thing, a ghost? This would be a case, as Thompson says, in which 'the definite description "the life form I bear" could fail to answer to anything, as the description "the present King of France" does' (p. 67). So would this possible failure of reference threaten the a priori nexus Thompson sees between *thought* and *human*? His response is, of course, no. To think so would be, Thompson says, 'to overlook that deployments of the idea of my life-form [or of *I*] will be in acts of self-conscious thought or speech. [*And t]hought and speech are phenomena of life if anything is'* (ibid., my italics). So

¹⁹ Paraphrasing Wittgenstein (1958), pp. 66-70.

Thompson is suggesting that the concept *human* is implicitly deployed in any self-conscious thought not in virtue of its deploying *I* or *my-life-form* but simply in virtue of self-conscious thought representing itself as a phenomenon of life. And that's all it needs. To represent myself as thinking is, in other words, already to represent an activity – thinking – that members of my life-form – *humans* – do.

The appeal to a condition met by all self-conscious thought renders Thompson's line of reasoning something of a transcendental argument. Here is an attempt at regimenting its steps:

1. To hold a self-conscious thought is to represent oneself as *thinking*.
2. Representing any organism's activities – any phenomenon of life – entails representing the organism's life-form.
3. (1) counts as representing an organism's – one's – activities. So,
4. (1) entails representing one's life-form.
5. My life-form is human.
6. So, representing what I do as thinking entails deploying the concept *human*.
7. (1) needn't involve empirical content. So,
8. Deploying the concept *human* doesn't involve empirical content.

The conclusion amounts to saying that the concept *human* is a priori:

9. If deploying a concept one does not involve representing empirical content, the concept is a priori.
10. So, *human* is a priori.

But Thompson seems to be equivocating here between what David Kaplan (1989) would call the character of a term, the way it gets its content when deployed, and the content itself.²⁰ My objection will thus amount to blocking premise (6) above. I will object to the view that self-conscious thought necessarily deploys the concept *human* just in virtue of deploying *thought*, which would be to say that in no possible world *thought* is deployed without *human*. For this to be the case, the concept *thought* should have the concept *human* in its ‘constitutive structure’, much like, as we saw, *triangle* has *three* and *angle* in its own. But could *thought* really be analysable as some complex concept like *human mental activity*?

It is bears noting that Thompson is tempted – and is aware of being tempted – by a similar move on the topic of *I*:

Everyone grants that it is in the nature of the first person concept to refer to the thinker of the thought in which it is deployed; [...] that is its character, in Kaplan’s language. It would be inept, though, to analyse *I* as the complex concept *the thinker of this very thought*. But only slightly inept: it is an equivalent concept, at least in ordinary extensional contexts; it can replace *I* in any such thought (Thompson 2004, p. 68-9).

Why might an analysis of *I* as the complex concept *the thinker of this very thought* is inept? One first hypothesis might be that *this* covertly carries a reference to the first person, if *this* is analysable as *what I am demonstrating*, say, which would render the analysis circular. But let us grant a distinction between reference to the

²⁰ I follow here the view that, if concepts are sub-propositional mental representations that can be shared between subjects and participate in inference, then concepts can be attributed certain linguistic properties, either because concepts form a mental correlate of sentences (as proposed by Fodor 1975) or because they form actual, albeit internalised utterances (as proposed by Dennett 1991).

first person, as occurs in *de se* thought, and reference to the place of demonstration, as occurs in what we might call *de hinc* thought.²¹ With this distinction in place, the proposed would not be circular, but it would mean any occurrence of *I* refers to something we could refer to in terms involving *this*. Which it may not. Recall the twin picture scenario introduced earlier on. In that scenario, when I think ‘I am wearing a red tie’, the mistake is not to represent a blue-wearing individual as red-wearing; rather, the mistake is to represent a red-wearing individual, correctly, as red-wearing, but identify it with the individual doing the representation.²² Hence, in that scenario, the content of my token of *I* is not the content of *the thinker of this very thought* (which is, instead, a blue-tie-wearing individual, myself, who have forgotten what colour he chose that morning).²³

But Thompson’s hunch that he can permit himself this ‘only slightly inept’ analysis of *I* is instructive. It prompts the question: why would the case of *I* occur to one in positing an a priori nexus between *thought* and *human*? My guess is that it is because there is a feature of the *I*-concept that is similar to the one Thompson wants to attribute to *thought*: what Santiago Echeverri (2020) has helpfully called ‘guarantee’. Echeverri characterises this property of *I* by contrasting *I* – or utterances of ‘I’ – with utterances of, say, proper names. Where Echeverri’s use of the proper name ‘Santiago’ is reflexive – that is, refers to him – only *de facto*, he says borrowing Peacocke’s (2014) terminology, his use of the *I*-concept is reflexive *de jure*, where the

²¹ I borrow the phrase from Schellenberg (2016), who employs it to distinguish between *de se* and *de hinc* contents of experience.

²² Imagine the thinker in this scene is a shopper trying a new jacket on. As he walks towards what he thinks is a mirror, he videocalls a friend to ask her opinion of the jacket. ‘But does it go well with your tie?’, she asks, remembering he was wearing a blue tie that morning. He looks ahead and says, ‘well, yes, look! I am wearing a red tie’, mentally referring to what he sees ahead.

²³ One might put this paragraph’s point in Kripke’s (op cit.) linguistic terms and say that indexical expressions such as ‘I’ are not synonymous with the descriptions one might associate with them, such as ‘the thinker of this very thought’. The former is a rigid designator; the latter isn’t. See Kripke (op cit., pp. 116-143) for discussion.

Latin conveys that in getting its reference, the *I*-concept follows a *rule*, specifically, an indexical rule, something like: ‘a token of *I* in an act of thinking stands for the subject of that thinking’ (Echeverri op cit., op. 480). This is precisely what Thompson tells us Kaplan would call the character of the *I*-concept: the rule it follows to fix its reference when deployed, or in Kaplan’s own words, ‘what determines the content in varying contexts’ (Kaplan op cit., p. 505).

My suggestion is that the relation Echeverri’s tokens of the *I*-concept bear to him is akin to the relation any human’s tokens of *thought* bear to *human*. Just as the *I*-concept is guaranteed to refer to its thinker, *thought* is guaranteed to make covert reference to *human*. We might put this by saying that just as the character of *I* is ‘to refer to the thinker of the thought in which it is deployed’, as Thompson says, so is part of the character of phenomena of life concepts (such as *thought*) to implicitly deploy a concept for the relevant organism’s life-form (such as *human*). But this is insufficient to regard *thought* as holding an a priori nexus to *human*, just as my tokens’ of *I* being guaranteed to refer to myself is insufficient to regard *I* as equivalent to *the thinker of this very thought*.

Let me elaborate. The nexus between *I* and myself is one that holds in the actual world and in normal conditions. In a non-actual world, where my proprioceptive system is wired to a chimpanzee’s body, say,²⁴ I might think ‘I am cross-legged’ but refer to the chimpanzee’s body by that *I*-token, not my own, which means I would be mistaken not just about the reference of ‘I’ but also about the property I am mentally representing, which might, for a chimpanzee, count as being cross-hind-armed (who knows?). In abnormal conditions, such as the red-tie-wearing twin scenario, a similar mistake is possible, as we have seen. The phenomenon of *I* bringing the thinker ‘onto the scene’ is thus like the phenomenon of *thought* bringing

²⁴ A type of scenario famously thought up by Evans (1984), p. 235.

human onto the scene in that both hold in the actual world and in normal conditions. Yet Thompson wants to see in the *thought-human* link something stronger than Echeverri's guarantee. He wants to see something of the strength of immunity to error through misidentification (IEM), a kind of property of judgements that holds *some similarity* to, but goes beyond, guarantee. As Echeverri writes,

It is easy to conflate guarantee and IEM because both of them exclude identity mistakes. Nevertheless, guarantee and IEM highlight two different ways in which identity mistakes can be excluded. Guarantee excludes the possibility that a thinker can (coherently) take the referent of her own token of *I* to be different from herself. IEM excludes a different possibility: that a thinker can (coherently) take the bearer of a property that seems to be instantiated to be different from the referent of her own token of *I* (her) (Echeverri op cit., p. 480).

Another way of putting the epistemic upshot of guarantee is to say that 'if an *I*-thinker can follow [the above reflexive rule] by performing a psychologically basic, self-referring act, she does not need to *find out* a suitable subject to stand in the thinking relation to a token of *I*' (Echeverri op cit., p. 487, emphasis original). The redundancy of 'finding out' a suitable referent for '*I*' renders it a concept whose extension can be known without empirical investigation. Similarly, the epistemic upshot of Thompson's view of *thought* can be put by saying that if a thinker follows the rule according to which *thought* – or *thinking* – makes covert reference to the thinker's life-form, one does not need to 'find out' which life-form the organism that is being representing as *thinking* bears. One has already assumed it.

The epistemic upshot of IEM, by contrast, is stronger, and guarantee can hold of an *I*-thought that does not enjoy IEM. When I see a picture of my twin that I think is a mirror and I think ‘I am wearing a red tie’, I am falling in the trap IEM excludes, the trap of misidentifying the bearer of the relevant property. For reasons mentioned earlier on, such a mistake is excluded in judgements that enjoy IEM, such as ‘I am in pain’. Thompson wants something like the latter to hold between *thought* and *human*. He wants it to be the case that when the concept *thought* is deployed, the possibility of a different life-form than *human* being implicitly deployed is excluded. But this possibility is not in fact excluded. Echeverri’s recourse of putting the difference between guarantee and IEM in terms of what doubts can coherently arise in each case is helpful here. Just as I can coherently form the doubt, when I notice an unfamiliar mole on the face I see, ‘wait, am *I* wearing a red tie?’, so can Pinocchio coherently form the doubt, when he notices he doesn’t bleed, ‘wait, am *I* human? Is that my life-form?’ If this counts as a self-conscious thought, Pinocchio is deploying *thought* without assuming – now – that the relevant organism’s life-form is *human*, just as in the former scenario, I am deploying *red-tie-wearing* without assuming – now – that the bearer is myself.

But why would Thompson want the strength of IEM to hold of the link between *thought* and *human*? Because there seems to be a connection between a priority of concepts and this strength of mistake exclusion. Just as in no possible world could my proprioception-based judgement ‘I am in pain’ be mistaken about who’s in pain, so in no possible world could a representation of something as *object* fail to pick out something that can be referred to by a singular term, or a representation of something as *person* fail to attribute, say, moral agency or worth to it (if that is the character of the concept *person*). The level at which the link between concept and extension is secured in the case of a priori concepts is at the level of

character, not just of content. Because character determines content, this means that in the case of a priori concepts, content does not vary across contexts or worlds in the way it does vary in the case of the *I*-concept.²⁵ Hence the room, in the latter case, for (coherent) doubts such as ‘wait, am I wearing a red tie?’ but not, in the former case, for (coherent) doubts such as ‘wait, does this person have moral significance?’ or ‘wait, can I refer to this object via a singular term?’. To represent something as an *object* is already to represent it as a potential referent of a singular term, and to represent something as a *person* is already to represent it as a moral agent/subject. If it were part of the character of the concept *thought* to deploy *human*, there would be similarly no room for a (coherent) doubt such as ‘am I *human*?’ in any context where self-conscious thought occurs, just as there is no room, in any context, for a (coherent) doubt of the form, ‘wait, am I thinking right now?’ because this is already to think.

To sum up. Thompson’s argument for the a priority of *human* is that because *thought* is deployed in any self-conscious thought, and any concept for a phenomenon of life implicitly deploys a concept for the relevant organism’s life form, *human* is deployed in any self-conscious thought. Now, this argument seems to assume actual-world, normal conditions: conditions, that is, in which self-conscious thought is something only humans are deemed to be capable of.²⁶ Given such a

²⁵ That content may be invariant in the case of other non-indexical concepts is of course true as well. Noting that it is specifically true in the case of a priori concepts is relevant because part of my task is to show how an a priori concept would differ from the *I*-concept even though in both cases observation can be unnecessary for knowledge of their extensions.

²⁶ One wonders if Kripke’s (op cit.) suggestion regarding the role of descriptions in fixing the reference of names and natural kind terms didn’t inspire Thompson’s notion of natural-historical judgements, and perhaps, unfortunately, bleed into Thompson’s mistaken seeing of an a priori nexus between *thought* and *human*. Even though the notion of self-conscious thought cannot be referred to as ‘that activity only humans do when they think’ any more than the notion of tiger can be as ‘that large carnivorous quadrupedal feline’, *those* descriptions may have historically helped to fix our knowledge of those life-forms, as

background, the link between *thought* and *human* holds in the same way in which the link between my tokens of *I* and myself holds; both are, in Echeverri's terms, 'guaranteed', and so both are cases in which no empirical investigation is needed to have access to the relevant concept's – *I*, *human* – extension. But guarantee concerns epistemic access to which object *I* refers to, in the first case, and which life-form *thought* is assuming, in the other. To show *human* to be a priori, Thompson would need to show that it isn't just epistemic access to the extension of *human* what is independent of empirical investigation but indeed possession of the concept itself. He would need to show, that is, that which life-form a token of *thought* is implicitly deploying cannot coherently be doubted (i.e. that it would be logically impossible for *thought* to implicitly deploy any other concept than *human*), just as it cannot coherently be doubted whether an individual represented as *person* has moral agency or worth, or an individual represented as *triangle* has three angles.

Importantly, Thompson's qualification that *human* is a priori 'as we humans have it' does not help. What Thompson wants to say is that our knowledge of the human life is fundamentally first-personal, not observational. This is a claim, perhaps a true one, about our knowledge of our own form of life. But to say that *human* is a priori is to make a claim about the nature of the concept itself, and to say that there is an a priori nexus between *thought* and *human* is to say that this nature is such that in no possible world is *thought* deployed without *human*. It is easy to conceive of such a world: I have done so already with the Pinocchio example. Now, perhaps one feature of a priori concepts is that they can be acquired by reflection alone. Thus, by geometrical reflection, we might arrive at *continuous curve* and *triangle*; by reflection on the structure of thought, we might arrive at *object* and *concept*. There is a link of IEM-level strength between those concepts and their extensions; that link is secured,

recorded in, precisely, natural-historical judgements: 'tigers eat meat and have four feet'; 'humans deploy self-conscious thoughts when they think'.

as we saw, at the level of character. So perhaps it was part of Thompson's intention, in positing an a priori nexus between *thought* and *human*, to allow for the possibility of humans' arriving at *human* by reflection on the phenomenon of thought alone. We humans are non-observationally aware of our own thinking, after all. But the link between *thought* and *human* does not enjoy IEM-level strength. *Human*'s being deployed along with *thought* is merely 'guaranteed'. For all a self-conscious thinking subject (non-observationally) knows, reflection on her own thinking may lead not to *human* but to *living toy*. And consider this further point, if I may stretch the limits of the paper's scope a little. The question might occur to us now: how about the possibility of arriving at *human* by reflection on other acts of ours than thinking? The knowledge that we have of our behaviour may be 'non-observational' in the sense of needing no investigation of external reality, as Thompson insists, following Anscombe. Still, here again, this doesn't mean it is a priori. To 'non-observationally' know facts like 'humans have two feet', we need to look inwards, if not to our thoughts, exactly, to the sensations our limbs feel. And to know 'humans care for each other', do we not need to engage – or at least form the intention of engaging – in caring action, at least with respect to one 'product of nature', ourselves?

Conclusion

One advertised end of this paper was to show Thompson's argument for the a priori of the concept *human* to be unsound. In order to do so, I have proposed to understand a priori concepts as those which can be available for use in thought independently of experience, even if they do, sometimes, feature in beliefs about the experienced world, as geometrical concepts do. In the latter part of the discussion, I noted a similarity between our deployment of a priori concepts and the *I*-concept:

that epistemic access to their extension is independent of experience. However, in the case of a priori concepts the link between concept and extension is secured at the level of character, which means that their content does not vary across contexts or worlds in the way the content of the *I*-concept does vary. This explains why, despite the epistemological similarity, the excluded mistakes are different. The mistake excluded in the deployment of an a priori concept is that of taking what one represents with that concept to fail to fall under its extension, which a Kantian might put by saying that a priori concepts are *pre-conditions* for the very possibility of the mental representations in question. The mistake excluded in the deployment of the *I*-concept is less radical. It is the mistake of taking the referent of one's own token of *I* to be distinct from oneself. I call this a less radical mistake because associating oneself to one's token of *I* is not a pre-condition for deploying *I*. One can deploy *I* while not associating oneself to that token of it, as when I think 'I am cross-legged' and bring the chimpanzee, not myself – not the organism that I am – 'onto the scene' due to the cross-wiring.

Now, the other thesis of Thompson's I've discussed, that *life-form* is a priori, does seem to be vindicated by this latter part of the discussion. When we represent something as alive, as instantiating a life-form, we are already exercising the sort of representational capacity the concept *life-form* allows: we will be inclined to count it as a member of its life-form even if it fails to meet some of the relevant criteria (having x number of legs, making x sounds), and so on. Thus there seems to be no room for a (coherent) doubt such as 'do these Fs reproduce themselves?' when F is a life-form, because to represent some object as a living organism is already to think of it as the kind of thing that reproduces itself (among other features of that special kind of kind).

Let me conclude by asking where my conclusions leave Thompson's defence of neo-Aristotelian moral naturalism. The answer is that it is not clear it is much affected. While Thompson has failed to provide an a priori nexus between *thought* and *human*, a nexus that holds across possible worlds, self-conscious thought does seem, in the actual world, represent the human life-form in virtue of representing itself as thought. And it isn't clear, I want to suggest, Thompson needs more than this. Thompson's stated goal was not to *reduce* moral knowledge to knowledge about human life, or to render the former derivable from the latter, in the way Frege's goal was indeed to reduce mathematics to logic, or to render the former derivable from the latter. Rather, Thompson's goal was to fix the scope of moral propositions – typically presumed to be a priori – in accordance with the scope of propositions about human life – typically presumed to be empirical. The view that our knowledge of the latter is not observational but first-personal doesn't need *human* to be a priori; not any more, at any rate, than my knowing who my tokens of *I* refer to needs the *I*-concept to have a stronger epistemic property than guarantee. This means, perhaps, that Thompson's goal of putting ethics on the same a priori footing as *person*-centred views of ethics, or as geometry, has failed. But was it a desirable goal in the first place? Ethics is, as naturalists see it, a theory about how we, humans, ought to live. It is concerned with human practical knowledge. And Thompson concedes practical knowledge is first-personal. So shouldn't he be satisfied with the result that knowledge about the human life-form is first-personal too – *rather than* a priori, that is? Doesn't this, in fact, suggest that moral naturalism calls for a sort of empiricism about ethics? Acquaintance with ourselves is a form of experience, too. Perhaps we needn't be anti-empiricists about *human* to agree with Thompson that 'we have [...] ways of knowing things about our own life-form "from within"' (2004, p. 72).

References

- Anscombe, E. 1981. The First Person. In *Metaphysics and the Philosophy of Mind*.
Oxford: Blackwell
- Baehr, J. 2003. A priori / a posteriori. *Internet Encyclopedia of Philosophy*
- BonJour, L. 1985. *The Structure of Empirical Knowledge*. Cambridge, MA: Harvard
University Press
- Chalmers, D. 2002. Does conceivability entail possibility? In T. Gendler and J.
Hawthorne (eds.). *Conceivability and Possibility*. Oxford: Oxford University
Press
- Davidson, D. 1987. Knowing One's Own Mind. *Proceedings and Addresses of the
American Philosophical Association* 61, pp. 441–58
- Dummett, M. 1973. *Frege: Philosophy of Language*. Cambridge, MA: Harvard
University Press
- Echeverri, S. 2020. Guarantee and Reflexivity. *Journal of Philosophy* CXVII (9)
- Epstein, P. F. 2018a. A Priori Concepts in Euclidean Proof. *Proceedings of the
Aristotelian Society* 118 (3)
- Epstein, P. F. 2018b. Shape Perception in a Relativistic Universe. *Mind* 127 (506)
- Fodor, J. 1990. A theory of content II. In Jerry Fodor, *A Theory of Content*.
Cambridge: MIT Press.
- Fodor, J. 1994. Concepts: A potboiler. *Cognition*, 50, 95–113.
- Foot, P. 2001. *Natural Goodness*. Oxford: OUP
- Frege, G. 1892/1984. On Concept and Object. In *Collected Papers on Mathematics,
Logic, and Philosophy*, ed. B. F. McGuinness. Oxford: Basil Blackwell
- Giaquinto, M. 2001. Knowing Numbers. *Journal of Philosophy* 98, doi:
10.2307/2678417

- Ginsborg, H. 2006. Empirical concepts and the content of experience. *European Journal of Philosophy*, 14(3), doi: 10.1111/j.1468-0378.2006.00230.x
- Kaplan, D. 1989. Demonstratives. In J. Almog, J. Perry and H. Wettstein (eds.), *Themes From Kaplan*. New York: Oxford University Press
- Kripke, S. 1972/2001. *Naming and Necessity*. Cambridge, MA: Harvard University Press
- Locke, J. 1689/1975. *An Essay Concerning Human Understanding*. Oxford: Clarendon Press
- MacBride, F. 2008. Can Ante Rem Structuralism Solve the Access Problem? *The Philosophical Quarterly* 58, pp.155-64
- Marshall, O. 2017. Giaquinto on Acquaintance with Numbers. *Journal of Philosophy* 114 (1), doi: 10.5840/jphil201711413
- Peacocke, C. 1992. *A Study of Concepts*. Cambridge: MIT Press.
- Peacocke, C. 2014. *The Mirror of the World*. Oxford: OUP
- Russell, B. 1945. *A History of Western Philosophy And Its Connection with Political and Social Circumstances from the Earliest Times to the Present Day*. New York: Simon and Schuster
- Shapiro, S. 2011. Epistemology of Mathematics: What are the Questions? What count as Answers? *The Philosophical Quarterly* 61, doi: 10.1111/j.1467-9213.2010.658.x
- Shoemaker, S. 1968. Self-Reference and Self-Awareness. *Journal of Philosophy* 65: 555–67.
- Shoemaker, S. 1970. Persons and their Pasts. *American Philosophical Quarterly*, 7(4): 269–285.
- Strawson, P.F. *Individuals*. London: Methuen & Co. Ltd.

- Thompson, M. 2004. Apprehending Human Form. In A. O’Hear (ed.), *Modern Moral Philosophy*. Cambridge: Cambridge University Press, 47-74.
- Thompson, M. 2012. The Representation of Life. In *Life and Action*. Cambridge, MA: Harvard University Press, 25-82.
- Weiskopf, D. and Bechtel, W. 2004. Remarks on Fodor on Having Concepts. *Mind and Language* 19, 48-56
- Wikforss, Å. 2017. Incomplete Understanding of Concepts. *Oxford Handbooks Online*, doi: 10.1093/oxfordhb/9780199935314.013.49
- Williamson, T. 2007. *The philosophy of philosophy*. Oxford: Blackwell.
- Williamson, T. 2013. How deep is the distinction between a priori and a posteriori knowledge? In A. Casullo & J. Thurow (Eds.), *The a priori in philosophy*. Oxford: Oxford University Press.
- Wittgenstein, L. 1922/1971. *Tractatus Logico-Philosophicus*, trans. D. Pears and B. McGuinness. London: Routledge.
- Wittgenstein, L. 1958. *The Blue and Brown Books*. Oxford: Blackwell
- Yablo, S. 2002. Coulda, Woulda, Shoulda. In T. Gendler and J. Hawthorne (eds.). *Conceivability and Possibility*. Oxford: Oxford University Press

Transformative torture

Abstract

A long-standing aim of cinema – in particular of ‘extreme’, ‘unwatchable’ or ‘feel-bad’ cinema – has been to acquaint viewers with extreme suffering. In this article I offer an explication of this aim in terms of recent work in the philosophy of mind, then employ the resulting framework to examine recent claims to the effect that a recent technological development, Virtual Reality, provides cinema’s best shot at achieving that aim.

No 'we' should be taken for granted when the subject is looking at other people's pain.

– Susan Sontag

Introduction

'Movies are the most powerful empathy machine of all the arts', the popular film critic Roger Ebert once declared. 'When I go to a great movie [...] I can walk in somebody else's shoes. I can see what it feels like to be a member of a different gender, a different race, a different economic class, to live in a different time, to have a different belief'. It is a strong but widespread sentiment. The value it attributes to cinema's being an 'empathy machine' seems to be epistemic as well as aesthetic. It resides in both the knowledge cinema gives the viewer – knowledge of the experiences of others – and what it *feels like* to gain such knowledge, what it feels like to 'walk in somebody else's shoes'. But the value is also ethical. 'The great movies enlarge us, they civilize us, they make us more decent people', Ebert added next.²⁷

The drive to see in art a vehicle for empathy is not new to movies, to be sure,²⁸ but an 'empathy machine' is presumably most useful when the target experience is radically unfamiliar to the empathizer, and the power of film to steer viewers' phenomenology seems particularly well poised to put viewers through just such an unfamiliar experience – experiences they wouldn't otherwise be able to undergo. This logic seems to underlie a long-standing aim of cinema, in particular of 'unwatchable' (Grønstad 2012), 'feel-bad' (Lübecker 2015) or 'extreme' (Frey 2016)

²⁷ From Ebert's (2005) speech upon accepting a Hollywood Walk of Fame star.

²⁸ See Hammond (2014) on the view that the interest in empathy from early psychology was intimately linked to the very rise of modernism in narrative art.

cinema: the aim of acquainting viewers with extreme suffering, the kind of suffering with which viewers of the arthouse-cinema-watching kind are likely to be unfamiliar. Here I offer an explication of this aim in terms of recent work in the philosophy of mind, then employ that framework to explore recent claims to the effect that a recent technological development, Virtual Reality, provides cinema's best shot at achieving that aim.

The spirit of the paper is exploratory, however, and ultimately critical of the purported power of VR. I start by building a case in favour of such power in the first section. To this end, I draw a connection between the notion of cinema's potential to offer a transformative experience to the viewer, as suggested in recent philosophy of film, and the notion of empathetic experience, as explored in recent philosophy of mind and fiction. The result is a criterion that a film – or an audiovisual work, more generally – purporting to be an 'empathy machine' in the above sense should want to meet: the ability to elicit a thought in the viewer that takes her own mental state to be of the same type as the target subject's. As a case study, I introduce towards the end of this section Alejandro González Iñárritu's *Carne y Arena* (2017), a piece that has been met with great enthusiasm by film critics who seem to share Ebert's intuition. In the second section, I employ the theoretical framework discussed in the first section to try and explain the empathetic power critics have attributed to *Carne y Arena*. Eventually, this will lead to some issues that might compel us to curb our enthusiasm about such power, as seemingly made possible (at long last!) by VR.

1. 'He is as I am, like this'

1.1 Abusing the viewer

A brief background on extreme cinema will be useful in introducing the subject. To this end it will be useful, in turn, to focus on the films of Michael Haneke, by taking him as the stereotype of cinema that punishes its audience²⁹ not only because Haneke is often thus taken by scholars and critics (e.g. Pérez 2012, Weigel 2013) but also because he has made his motivations explicitly clear. A guide to those motivations, helpfully exposed by the critic Moira Weigel in her (2013), might be Haneke's own slogan that his work aims to 'rape the spectator into autonomy'.³⁰ The 'raping' metaphor³¹ stands for Haneke's *means*, and the 'autonomy' part stands for his *end*, which is twofold, epistemic and ethical.

The core of Haneke's means is, as noted, extreme violence, which I'll understand as an act of punishment or torture the result of which is extreme suffering. Extreme cinema's violent methods range from narrative techniques (the abuse remaining unexplained, for example) to the explicit display of extreme violence on-screen: whatever contributes, as Haneke puts it, to breaking 'all the rules that usually make the viewer go home happy and contented'.³² In *Funny Games* (1997/2007), we see at length and in excruciating detail the violence that two young men physically inflict on a wealthy family; in *Amour* (2012), we see it being inflicted on an old couple's minds. In both physical and psychological varieties, the display only ends when suffering has reached its peak and there's no more of it left to squeeze. Being shown extreme violence of this sort does not leave the viewer

²⁹ In the way stereotypes are understood as helping us to fix a concept by embodying the most typical, though not any or all, of its features (Putnam 1975, Fodor 1998).

³⁰ As quoted in Weigel (op cit).

³¹ I reproduce Haneke's metaphor for accuracy, but see Fraser (2018) on its problemat�city.

³² What about these films' violence makes it *aesthetic* exactly is a thorny question. I will assume 'aesthetic' can qualify any media that is used as part of the (higher-order) artistic medium of cinema, 'the medium of the moving image', as Gaut (2012) characterises it. I subscribe to Gaut's (2012, Ch. 1) view on the diverse range of lower-order (or 'nesting') media that can constitute the medium of cinema, not all of which are exclusive of it or artistic by themselves.

untouched. The fact that Haneke's victims are middle class, as is his typical audience are,³³ and that they have placeholder names – Georges and Anne – for the audience to substitute their own completes the familiar view that Haneke's goal is not only to display extreme suffering but to torture the viewer herself.

The first consequence of this experience is meant to be epistemic. The viewer comes out the other end knowing something she didn't know before: what extreme suffering is like. This epistemic end is, in turn, ethically motivated. To recall, Haneke's goal is the viewer's autonomy. Media violence, Haneke thinks, makes real violence indiscernible, so his mission is 'to set right a dangerous "commingling and indistinguishability" of *real* and *media* death' (Weigel op cit .p. 136, my italics). Viewers' newly gained acquaintance with real suffering is supposed to put them in a position to engage in conversation about the phenomenon they now have knowledge of, which eventually should 'lead to social and political change' (ibid., p. 142).

Both the affective means and the twofold – epistemic and ethical – end are necessary to make extreme cinema what it is. They are what distinguishes it from films that are similarly violent but do not share the same epistemic and ethical ends. Contrasting extreme cinema with films like *Hostel* (2005, Eli Roth) or *Saw* (2004, James Wan), Asbjørn Grønstad writes that extreme films

insult our sensibilities [... and make] us want to avert our eyes [...in order] to force us to reconsider our investments [...]. The aim is not emotional and cognitive discomfort for its own sake; the production of unpleasure is rather a means to an epistemological-ethical end (Grønstad op cit, p. 15).

³³ The 'bourgeois' epithet for Haneke's target audience is employed by Haneke himself and reproduced widely (cf. Weigel and Pérez). I avoid it here because this paper will generally refrain from engaging in the political discussions associated with extreme cinema.

That the end is both epistemic and ethical is reflected in Haneke's intentions. First, viewers must acquire knowledge of – *understand* – the extreme suffering presented to them. Only then, in possession of this understanding, will viewers be able to *take action* that will 'lead to social and political change'. Let us count as 'extreme' any film that follows Haneke's violent-*means-for-an-epistemic-ethical-end* slogan. This characterisation may include films like Virginie Despentes' *Baise-moi* (2000), Gaspar Noé's *Irreversible* (2002), and various of Lars von Trier's films, among others.³⁴

To be sure, the urge to give audiences first-personal knowledge of the suffering of others by audiovisual means for ethical concerns precedes Haneke and company. Perhaps most significantly, the question whether suffering must be experienced to be known took centre stage in the debate on representations of suffering that arose in the wake of the Holocaust. Against the filmmaker Claude Lanzmann's conviction that no visual representation of those sufferings could do them justice, such that they should only be conveyed discursively (which was his choice in his nine-hour-long documentary *Shoah* (1985)), various authors deemed discursive representation to be inappropriate. Georges Didi-Huberman, for example, wrote in praise of an exhibition of photographs of the camps – which he called 'silent witnesses' – and against Lanzmann's interdiction of images that 'to know something, one must be able to picture it' (2001, p. 219). Along similar lines, Giorgio Agamben wrote that faithful discursive representation was impossible, since a 'true witness' of

³⁴ Haneke's complaint against media violence focuses on its artifice, its eagerness to please, the experience's failure to be traumatic for the audience. His films thus often bring out the fact that media images of violence block the bourgeois viewers' acquaintance with real suffering. In this he differs from other extreme filmmakers, for whom the notion of images of violence is not a recurring theme, as it is for him (see e.g. Benny's *Video* (1992) and *Caché* (2005)). I will gloss over this admittedly major feature of Haneke's work because I assume that Haneke's concern about the absence of a trauma in the bourgeois' experience of media violence can be cashed out as a concern about absence of first-personal knowledge of it.

the sufferings would be the kind who, precisely, the experience had rendered speechless (1999, p. 150). And years before, Hannah Arendt had written: '[t]he more authentic [verbal reports by survivors] are, the less they attempt to communicate things that evade human understanding and human experience – sufferings, that is, that transform men into “uncomplaining animals”’ (1973, p. 439).

Relating these authors' concerns to extreme cinema isn't exactly news,³⁵ but my point is that Lanzmann's critics aren't simply saying that showing is better than telling because it is epistemically more efficacious but because that would fare better in promoting their ethical and political cause. This connects straightforwardly to the view that cinema as a medium has a unique potential not only for the transmission of knowledge that is unfit for discursive representation – call it phenomenal knowledge³⁶ – but also for the transmission of moral knowledge. While words are paradigmatically fit for knowledge-*that*, cinema may have fitter means for knowledge-*what-it-is-like* and knowledge-*what-to-do*. Hence Robert Sinnerbrink, who has also written on Haneke (2011), tells us in *Cinematic Ethics* (2015) that film's unique power to 'engage our emotions, exercise our moral imaginations and question our beliefs' (p. 17) renders it a medium with the 'transformative potential [... to] *sharpen* our moral perception, challenge our beliefs through experiential means, and thus *enhance* our understanding of moral-social complexity' (p. 16, my italics). Sharpen, enhance: improve our knowledge-what-to-do.

The key words here are 'transformative' and 'experiential'. Sinnerbrink puts a fair amount of emphasis on the fact that cinema is a phenomenally rich sensory experience, which renders it a medium for 'ethical understanding' (p. 6) that is best appreciated in contrast it with other media, for example, discursive means of representation, the approach of, e.g., moral philosophy. Tackling the subject of

³⁵ See e.g. Saxton (2008c) and Gustafsson (2014).

³⁶ After Loar's (1990) phenomenal/theoretical distinction.

cinema's ethical upshot by focusing on its experiential nature serves Sinnerbrink to bring out cinema's capacity to 'express [...] the synergy between affective attunement, emotional engagement, and moral evaluation' (p. 95), where *affect* and *emotions* are, again, phenomenal or experiential. For Sinnerbrink, the jump from cinematic experience to novel and enhanced moral understanding basically happens via a combination of empathy and sympathy. He calls it 'cinempathy'.

Is it a hyperbole for Sinnerbrink to call the moral effects of cinema 'transformative'? It is a common thought that one's moral beliefs contribute to who one is, i.e. to constitute one's self. If this is so, then having one's moral compass impacted may reasonably change one's self. Further, if an ordinary way of having one's moral compass transformed is *thought*, as when one reads or listens to others, then having it challenged by the power of *experience* – emotion, affect – may in fact so impact one's moral compass that one winds up being transformed.

Thus, for all we've said, extreme filmmakers seem to be on the right track. They don't only follow Agamben and Arendt's advice to show us suffering rather than tell us about it but also intuit that suffering is the kind of thing you can't really understand unless you yourself go through it. Hence their efforts to punish us.

1.2 Cinematic transformative experience

Sinnerbrink's conviction that cinema has a 'transformative potential' is recognisable in the extreme filmmaker's hope that acquainting the viewer with extreme suffering will give her morally significant knowledge – knowledge that, somehow, will lead to moral action. It can be helpful to flesh this out in terms of what Laurie Paul (2014, 2015, 2018) has also called 'transformative experience'. Both authors coincide in that

the crux of the notion is the action-enabling³⁷ upshot of the phenomenal knowledge novel experiences provide us with, cinematic or otherwise. For Paul, however, this involves a distinction Sinnerbrink doesn't draw and which may prove crucial. According to Paul, an experience can be epistemically and/or personally transformative:

An *epistemically* transformative experience teaches you something you could not have learned without having that kind of experience. [...] A *personally* transformative experience changes you in some deep and personally fundamental way, for example, [...] by changing the way you understand your desires and the kind of person you take yourself to be (Paul 2015, p. 761).

There is an echo here of Laurence Nemirow's view that 'understanding an experience [amounts to the] ability to place oneself, at will, in a state representative of the experience'.³⁸ Paul also puts her Nemirow-like view by saying that the *de se* or 'subjective' – as opposed to 'descriptive' – mode of presentation of a fact, or knowing it 'from the inside', is necessary for either kind of transformation.³⁹ However, while

³⁷ Paul doesn't go into moral matters; the normative standard in her discussion is rationality. Presumably, Sinnerbrink would require moral action to be rational as well.

³⁸ As approvingly quoted by Lewis (1990).

³⁹ Paul's view is controversial on various fronts. On the one hand, metaphysicians might ask how exactly what is transformed is individuated (is it the self? Is it the person?), and what exactly the transformation consists in (is gaining new phenomenal knowledge, and having one's moral compass affected in consequence, like the ship of Theseus getting a new mast?). On the other hand, even if these questions are answered, philosophers of mind might ask if the term 'transformation' doesn't overstate the epistemic value of first-hand experience and understate our capacity to achieve knowledge by testimony and other media. I want to remain neutral on these debates by simply using the notions of personal and epistemic transformation technically, that is, in Paul's sense. This will be useful to eventually characterize the sort of 'transformation' VR viewers may undergo which film viewers may fail to. One mark of personally but not of epistemically transformative experiences, in Paul's sense, again, is the correlation of two kinds of knowledge: phenomenal and *de se*. More on this in the final section.

no walk in the park, watching Haneke's films is not comparable to being hit to death by deranged trespassers in one's own home, as in *Funny Games*, or to watching one's partner's mind rot and ending her life, as in *Amour*. Because cinematic experience – even *extreme* cinematic experience – falls short of the personal transformation involved in real extreme suffering, perhaps extreme films are, or aim to be, epistemically transformative, experiences that will teach the viewer as much as possible about extreme violence by submitting her to a similar, though of course not identical type of suffering to the one endured by the characters on-screen (call it, following Sinnerbrink's term, suffering by cinempathy).

There is reason to think this falls short of empathy, however. Consider Paul's related (2018) distinction between two levels of empathetic experience. Beyond the kind of empathy one might engage in by acquiring knowledge of how the empathizee feels when, for example, one sees her suffer on-screen (call that 'affective empathy'), there is a further, richer empathy of a cognitive sort that involves 'knowing what it [would] be like to be [that subject by means of] understanding enough about the nature and content of her *first-personal conscious perspective*' (ibid., my italics). Call this latter type 'cognitive empathy'. It is, as Paul puts it, like 'stepping into the shoes' of the subject. It requires one to 'occupy [the subject's] psychologically centered conscious perspective' to access her 'way of understanding the world as a located, *immersed*, conscious experiencer' (ibid.).⁴⁰

⁴⁰ As I (and Paul) understand it, the *de se* mode of access corresponds to what Blanke and Metzinger (2009) call 'strong first-person perspective'. Their notion of 'weak first-person perspective', by contrast, is simply a 'geometrical feature' of a system's representation of its environment: its 'point of projection functioning as its origin'. It 'is not linked with theoretically more charged notions such as 'subject of experience' (the conscious self) or 'epistemic subject (knowing self)' (p. 8). The strong first-person perspective thus encompasses not only a perceptual point of view but also associated perceptual beliefs and other attitudes. The distinction corresponds roughly to Schellenberg's (2016) distinction between *de hinc* and *de se* contents of experience, and to Alsmith's (2018) distinction

Cognitive empathy is radically different from affective empathy in that occupying an individual's conscious perspective is a way to acquire knowledge of facts about her in a mode of presentation exclusive to her: the *de se* mode of presentation. Where simply witnessing an event generates a *de re* thought, a thought indexed to the individual one sees, the *de se* mode of presentation generates a *de se* thought, which is indexed to the subject's *self*. We can illustrate this difference schematically. One first way for subject A to learn the fact that subject B is in extreme pain is discursive testimony. A second alternative is to learn it by visual perception, in which case A might infer that B is in pain. In such a case, A perhaps also enters a state of pain herself, but her pain will differ from B's pain not only in degree (if A is sitting in a comfy cinema seat, for starters) but also in kind: it will not be 'pain by being beaten to death' but 'pain by cinempathy'. And then a third theoretical alternative is that A learns the fact that B is in extreme pain in the same way B herself learns it: by occupying B's conscious experience. In that case, A will learn the fact expressed by the proposition 'B is in extreme pain' in a way that can be expressed by the proposition 'I am in extreme pain'. Assuming that an individual's being in pain constitutes the existence of *that* instance of pain, this means there will be now two tokens of the same type of suffering: B's and A's, which means there will be now two distinct facts: A's being in extreme pain and B's. The additional fact, however, does not preclude but rather allows A to learn the target fact – B's being in extreme pain – in the *de se* mode of presentation, whereby she will ascribe the extreme pain not just to some individual distinct from her, as in the two previous epistemic methods, but also to herself.

The failure of the first and second methods and success of the third to meet Paul's standards for cognitive empathy can be mapped onto Kendall Walton's (2015)

between perceptual experience being simply 'perspectivally structured' and its being also 'self-locating'. More on this in the second section.

view that to genuinely (as opposed to ‘sort of’) empathise is neither to judge that the subject one is empathising with is feeling thus – or judging that ‘the empathizee feels like *that*’ – nor to judge that one is feeling thus and so is the subject – ‘I am panicked and so is he’ – but rather to judge that the other’s mental state is of the same type as one’s current state – ‘he is *as I am*, like *this*’. Only this third judgement is genuinely empathetic, and it features three indexicals referring to (i) the target person, ‘he’, (ii) one’s own self, ‘I’, and (iii) the type of state one is in, ‘like *this*’.

At first glance, the Paul-Walton bar for empathy is very high. But it is not as high as a view on which to empathise with a subject means to imagine being *them* and not simply to imagine oneself in their situation, to ‘enter as it were into his body’, as Adam Smith (1759/1976) puts it, to ‘become in some measure the same person with him’. This latter experience, representable as a judgement in which there are not two subjects, as in Walton’s empathetic thought, but one, the target subject (only referred to by the first-person indexical), is what Rae Langton (2018) calls ‘imaginary identification’, and it is *not*, I want to suggest, what extreme cinema aims at. The focus of the extreme cinema aesthetic is not the suffering individual, her token state; rather, it is the type of experience an individual’s state exemplifies. By this I mean that extreme cinema does not aim at what film scholars call ‘character identification’, about which worries have been raised in the philosophical literature (Carroll 1990, Wollheim 1999)⁴¹ that one might further associate with what Langton calls the problem of auto-alienation (see again Langton *op cit.*, pp. 97-101). Now, there are ways to resist or circumvent this worry (Gaut 2012, Ch. 6.; Smith 1995, 2017). Here, however, I want to steer clear of the debate by suggesting that identification is simply not extreme cinema’s target. This is because extreme cinema

⁴¹ The specific view that extreme cinema aims to elicit a disgust-like reaction (Blümlinger 2010) dovetails with Carroll’s general worries against identification. For a response, see Smith (2017, esp. pp. 203-4).

does not aim to put the viewer in any *particular* character's shoes. As Pérez (op cit.) has observed, Haneke's concern is not what *individuals* experience but what whole classes of them do, where a class of individuals is more or less defined by their socio-economic circumstances. Pérez cites several features of Haneke's work to back this claim; of note are the facts that, as mentioned, Haneke's characters have placeholder names and that the detachment with which Haneke shows violence is intended to prevent individual identification with any particular sufferer: 'our identification with them as individuals', Pérez writes, would 'blur the focus on social class' (ibid.). Indeed, Haneke himself talks of fighting the "indistinguishability" of real and media death' as a general phenomenon.

But the question then arises: under what conditions exactly might two subjects count as being in the same kind of state? Walton's framework is fairly permissive in this respect. The crucial condition for holding the genuinely empathetic judgement is that one be able to 'us[e] one's current mental state as a sample' of the target kind of experience (Walton op cit., p. 14). The question for Walton arises, in turn, of how the empathizer picks out 'one aspect or property of her current mental state, as opposed to others, which she attributes to [the target]' (ibid., p. 8). Walton's own answer is that this is likely a matter of context sensitivity, and that, in any case, it is a question that he (Walton) is not alone in facing: it is 'also [a question] for those who use the notion of phenomenal concepts for other purposes, of course' (ibid.). For the purposes of our discussion, the question can be put like this. Suppose the extreme filmmaker somehow does achieve her goal. Suppose, that is, that she succeeds in making the viewer mentally point at her own state, as well as the character's, who she sees undergo the same experience, and utter or think, 'he is as I am, like this'. What phenomenal property of her current mental state does 'like this' refer to? Certainly not to the property of sitting comfortably on a cinema seat, or

of laying on the ground, bleeding, because neither of those is shared. We can add to Walton's suggestion about context sensitivity that the phenomenal concept in question has to be general enough. For example, in Walton's illustration, the empathizer learns her target is trapped in a cave and she (the empathizer) is thereby prompted to imagine being trapped in a cave too, by drawing, perhaps, on memories from her speleology days. She feels her pulse accelerating, starts to sweat. She reaches a state that she could express by saying, 'He is as I am, like this'; or else, '[s]he can appropriately say, "I know how it is with him" or "I know how he feels," where "know" carries a connotation of intimacy, *acquaintance*' (ibid, p. 9, my italics). Her experience of being trapped is imaginative rather than real, of course, and in this respect she is not like her target, nor is she like her target in that her experience is occurring in broad day light rather than in the dark and so on. Because she knows what being trapped does to one, though, she can confidently pick out the relevant phenomenal properties 'like this' will refer to: the ones elicited by that situation. She can thus use Walton's demonstrative in a way that the extension of the concept it expresses covers both her own state and the target's.

It bears noting that one consequence of this liberal view of the conditions under which an empathizer can deploy the triply indexed judgement is that because her mental state will, avowedly, not be *identical* to the target subject's, there is bound to be a difference between the emotions each feels, too. If the target subject's emotion in Walton's trapped-in-a-cave case is panic, the empathizer's emotion will be, in Walton's terminology (1978), 'quasi-panic'. The reason the empathizer can nevertheless use her mental state as a sample of the target's is that the phenomenology of 'quasi-emotions' matches that of their actual counterparts on Walton's view; what differs is just their cognitive profile (one involves belief; the other *make-believe*). In other words, the fact that the empathizer's emotion arose

from what she knows to be an *imaginary* exercise while her target's arose from what *he* knows to be a real situation does not preclude the empathizer's ability to use her state – specifically, the phenomenal properties relevant to the situation: a sense of alert, a feeling of suffocation, and so on – as a sample of what the target subject is experiencing.

Now, part of what Walton wants to show is that the imagination needn't play a role in empathetic experience. This, he thinks, can occur when the empathizer and the subject are in sufficiently similar situations. Walton writes:

That being in a situation similar to the target's helps one to empathize with him is certainly not news. But the usual idea is that being in a similar situation makes it easier than it would be otherwise to “put oneself in his shoes” or to “imaginatively identify” with him.” My proposal is that this last step is unnecessary. Actually being in a parallel situation is enough (Walton op cit., p. 11).

If I am grieving for a friend of mine and you are grieving for yours, I can truthfully hold the empathetic proposition – ‘I am as you are, like this’ – without the need to imagine grieving for *your* friend rather than mine. The specifics, in other words, may not matter. To return to our discussion, this is significant for the extreme filmmaker because their view is precisely that the imagination is powerless when it comes to emulating an experience so alien to the viewer as extreme suffering. Reading the news, Haneke might say, or watching the Holocaust survivors' interviews in Lanzmann's documentary, will not stimulate the imagination in such a way that the

subject can really use her mental state as a sample of those sufferers'.⁴² This is why the extreme filmmaker aims to inflict violence to the viewer, to put her in a situation that is not identical but 'parallel' to the character's such that her 'quasi-suffering', if that's what it is, can be employed in her deployment of the genuinely empathetic judgement.

Paul's discussion, on its part, focuses on empathy with one's future selves rather than with other subjects. In such a case, the imagination can indeed play a part, but only because it may be aided by the memory of the subject's own past experiences. Generally, we're warranted to be in a similar situation to our future selves at least in that that our future selves' psychology is similar to our current self's. When the target type of experience is extreme suffering, however, the empathizer – the filmgoer – will have no mnemonic resources to draw on in an attempt to make the imaginative jump. Extreme suffering is, by assumption, unlike anything the viewer has experienced; her ignorance of extreme violence is precisely what motivated extreme filmmakers to make their films in the first place. Cinema hinges on visual perception, the second epistemic method mentioned above, which yields, in turn, Walton's second insufficient thought – 'the target feels like *that*' – rather than his 'truly empathetic' thought, 'he is *as I am*, like *this*'.

If only there were a means to put the viewer in the extreme sufferer's shoes...

1.3 Virtual transformative experience

Upon first entering the installation, all you hear is the wind and your own steps on the actual sand under your feet. Then you hear murmurs and panting. The bushes move. One person appears, then another. And children.

⁴² Readers might recognise here the same intuition that drives Paul to write that what testimony gives you is not what is required for decision making in radically unfamiliar scenarios (Paul 2014 p. 77 ff; 2015, p. 481 ff).

Immigrants drag themselves under the sun; they are led by their smuggler. They walk so closely past you that you might feel compelled to touch them, upon which you realise they can't feel or see you—but you've tried. Within minutes, night falls. The sound of a helicopter quickly goes from faint to unbearable, and all of a sudden it has landed, patrol agents sprawling from it like spiders. Stifled by the helicopter and the cocking of guns, you only hear the screams of those who are near you. Because you were invisible before, you resist the agents' orders to "get down!" Little do you know that the system *can* in fact track your position and that you will not be spared. One of them points his gun at you. You obey. You know the hangar's floor is covered in sand but at this point you don't think about your Cannes-ready outfit because you feel, as film critic Peter Bradshaw describes it, 'lowered, lessened, [...] subhuman, without even a criminal's civilian rights'.⁴³

Alejandro González Iñárritu's *Carne y Arena* (*Flesh and Sand*) premiered in 2017 as Cannes' first Virtual Reality inclusion. Reviewers claimed that the six-minute, photo-realistic⁴⁴ piece pushed the limits of cinema (Bradshaw 2017, Zeitchik 2017), that it transcended them (Thompson 2017), and that it was 'the VR revolution we've been waiting for' (Gleiberman 2017). *Carne y Arena* puts the viewer in a group of Latin American immigrants attempting to cross the US border. It was first hosted at Cannes in a spacious hangar, where the audience wore Oculus Rift headsets and walked around the virtual Arizona desert, guarded by staff, one at a time. It has since been hosted in similarly large but more accessible venues in Milan

⁴³ From Ordóñez Angulo (2017).

⁴⁴ Rarely for a first-person VR environment, *Carne y Arena*'s imagery is not computer-generated but cinematography. A taster of the process can be seen in the trailer: <https://youtu.be/zF-focK30WE>

(the Prada Foundation), Los Angeles (the Los Angeles County Museum of Art), and Mexico City (the Tlatelolco Cultural Centre).

The first thing to say is that I'm not interested in the art-type ontology of *Carne y Arena*. Although, reportedly, González Iñárritu has preferred to call it not cinema but an 'art installation' (Robey 2017), several critics have wondered whether the piece pushes out of *cinema's* limits rather than into a different art-form's (cf. above), while others have suggested it constitutes the first instance of a *new* type (Utichi 2017). The Cannes website itself called it 'an experience some are already comparing to that of the first ever spectators in the history of film' (Khaldi 2017). So consensus seems difficult to reach. Let us then call *Carne y Arena* extreme 'audiovisual fiction'.⁴⁵

Why might *Carne y Arena* count as belonging to the extreme cinema tradition? Consider, first, the echo of Haneke's intuitions in González Iñárritu's. 'Until you feel it', González Iñárritu has said, 'until you feel what it's like to be 20 years old, not left wing or right wing or any wing, going through something like this, you can't really talk about it'.⁴⁶ And consider the echo of Haneke's concern in González Iñárritu's talk of the experiences of (social) classes rather of individuals: 'the urgency is about people understanding that 60% or 65% of these people are kids or women running for their lives', he has said;⁴⁷ and elsewhere: 'it's like a Holocaust. That's the reality'.⁴⁸

Carne y Arena contrasts with forerunning VR pieces that have similar ethical motivations but do not share the extreme cinema aesthetic because, being

⁴⁵ That said, I am sympathetic to Gaut's (2012) view on why novel technologies (including interactivity) do not disallow calling an artwork cinema but in fact illustrate the plasticity of the medium.

⁴⁶ As quoted in Zeitchik (2017).

⁴⁷ As quoted in Lang (2017).

⁴⁸ As quoted in Thompson (2017).

documentary, they do not aim to punish the audience. Take, most notably, Nonny de la Peña's work, which she calls 'immersive journalism' and which uses VR to report on humanitarian issues, such as Guantanamo Bay. Where de la Peña's goal is to have viewers 'feel an extraordinary emotional connection as *witnesses*' (de la Peña 2017, my italics), González Iñárritu's goal is to put viewers in the position of the sufferer themselves. To see the effect, just look at critics' reports. Arendt's characterisation of extreme suffering is strikingly echoed in Bishop's (2017) remark that he was made to feel like a 'piece of cattle', as well as in Bradshaw's description cited above: '[y]ou become lowered, [...] subhuman'; Bradshaw also had, '[f]or the first time, [...] an inkling of what it must be like [to have a gun pointed at you]'. Unsurprisingly, the adjective 'harrowing' abounds (Robey op cit., Bishop op cit., Farago 2017). Most critics also highlight the normative upshot of the experience, either in personal terms (e.g. Gleiberman 2017 'felt a primal empathetic connection to the immigrants, and [he] also never felt more like a privileged white person who had never known an experience like this'), or more directly in political terms (Lang 2017, Farago op cit).

Long story short, *Carne y Arena* aims to make viewers suffer not by watching others suffer but by occupying the sufferer's first-person perspective. This means the viewer is poised to learn the fact that immigrants suffer as a *de se* truth, because the viewer's own self suffers, which means, in turn, that their experience gets closer to Paul's cognitive and Walton's genuine empathy than any prior work has ever made possible. The feat is meant to be achieved thanks to the system's unprecedented ability to track the viewers' eye level, which allows it to centre the experience around her (including, importantly, the agents' violent confrontation). If *Carne y Arena* is only the first VR work put to this use, and if the technology advances, then the extreme cinema aesthetic might indeed become *really* extreme and *really*

unwatchable, which is to say that, far from threatening to dead-end, as some critics have suggested (Weigel *op cit.*), it might finally really thrive.

2. A matter of perspective

Let me briefly take stock. I have suggested that extreme films are a manifestation of the faith in cinema's power to be an 'empathy machine'. By showing us extreme violence being inflicted upon characters, extreme filmmakers aim to provide us with epistemic access to – in particular, an understanding of, in Nemirow's sense – a type of experience that would otherwise remain unknown to us. To count as understanding this experience, however, it has been suggested that one should be able to index the suffering not only to the subject one sees but also to oneself. The fact that traditional cinema most readily allows *de re* beliefs about the suffering – '*he* feels like *that*' – while VR allows *de se* beliefs gives us reason to think VR could increase filmmakers' chances to realise extreme cinema's ends. Film critics' reports on their experience of *Carne y Arena* lend some credit to the view that González Iñárritu has, in fact, succeeded.

In the remainder of this article, I explore some ways of explaining this apparent success. My working hypothesis is that there is a link between a distinctive feature of VR, its giving a sense of agency to the viewer, and its allowing her to index the suffering to herself, as per the Walton-Paul criterion (Section 2.1). But issues will arise along two dimensions concerning this link. On the one hand, while the link between a sense of agency and *de se* thought may explain the VR viewer's distinctive experience, the fact that – as we will see – VR still engages the viewer's imagination to yield the target phenomenal knowledge makes room for the thought that the difference between VR and traditional cinema is one of degree rather than kind

(Section 2.2). On the other hand, even if VR does provide knowledge of the sufferings of others in virtue of its ability to put the viewer in the target subject's perspective, the question will arise about what the moral significance of this exactly is (Section 2.3).

2.1 Agency and the *de se*

By 'sense of agency' I mean something fairly intuitive: the subject's awareness that what she experiences is *up to her* to some extent. In the case under discussion, this is the result of head-tracking, which allows the visual field of the viewer whose head is being tracked to shift as she moves. But a lower extent of this phenomenon is already present in an older technology: first-person shooter games. Even though one's head is not being tracked in the latter case, first-person shooter games as a medium have a different way of allowing the viewer to feel she is in charge of her visual experience: through the joystick, the computer's keyboard, the console's controller. There being a connection between the viewer's sense of agency even at this minimal level and her ability to hold *de se* thoughts, explicit or implicit, should be evident to anyone who recalls playing first-person shooter games as a kid and physically responding to the perceptual stimuli (ducking, recoiling, leaning forward to throw a punch).

One way of explaining this phenomenon might be to see how *de se* thought arises in ordinary experience generally. As Adrian Alsmith's (2018) has observed, there are two main hypotheses as to how that is. According to the 'intellectual account', *de se* thought arises in virtue of self-consciousness; according to the 'agentive account' (his choice), it does in virtue of the subject's ability to act upon objects from her location.⁴⁹ Now, to have the ability to act upon objects is a stronger

⁴⁹ See Alsmith (2018, p. 263) for a survey of theorists who endorse each.

condition than the ability to shift one's visual field at will, as video game players do. But it is plausible that the latter is at the end of a continuum of which the former is higher up. In *Carne y Arena*, there is a higher sense of agency than the one enjoyed by gamers even though the viewer in *Carne y Arena* cannot actually act upon the objects around her. This is, to be sure, for several reasons: her movements are not constrained to handling a joystick, the shifts in visual experience are more substantial, etc; significantly, though, she can experience something like the converse of the sense that she can act upon her environment: the potential for objects in the environment to act upon *her*.

To elaborate on this suggestion, it will be helpful to contrast *Carne y Arena* with a virtual environment from a laboratory setting: Mavi Sanchez-Vives and Mel Slater's⁵⁰ virtual experience of domestic violence, that is, of suffering it from a first-person perspective, after which experience experimental subjects who used to abuse their partners acquire or improve recognitional capacities⁵¹ the lack of which is thought to lie behind their prior violent behaviour. In this setting, the subject's sense of agency is reinforced by her undergoing an exercise that consists in manipulating virtual objects – plain, floating cubes, spheres – via a tracked glove before the virtual experience of abuse begins (Slater and Sanchez-Vives 2016, p. 4).⁵² Why would the *Carne y Arena* subject enjoy a similar sense of agency if she cannot move things around her? Only her head, not her hands, is being tracked! Head-tracking suffices for what Sanchez-Vives, Slater et al deem the key factor for immersion, 'the ability to perceive through natural sensorimotor contingencies', by which they mean 'whether you can turn around 360°, all the while seeing a very low-latency continuous update

⁵⁰ The main study I refer to is published as Sanchez-Vives, Slater et al (2018), but I'll refer to related outputs by the same research team throughout.

⁵¹ The authors suggest e.g. the capacity to recognise fear.

⁵² I am citing the authors' description of this exercise, but I underwent it myself.

of your visual field in correspondence with your gaze direction' (2016, p. 4). And this already may help, as I suggested just above, to simulate the converse of the sense of being able to manipulate objects in one's environment. In both the experience of domestic violence and *Carne y Arena*, in contrast with, say, Nonny de la Peña's work, the subject experiences a virtual agent – the domestic abuser in one case, the patrol agent in the other, both of whose gazes follow the subject's own – as being able to act upon *her*: to physically hit or shoot her, respectively. Indeed, forming attitudes about our ability to act upon our environment *and* the environment's ability to act upon one may not be a matter of experiencing those actual actings – of experiencing actually being hit or shot – but of experiencing dispositions. Alsmith's characterisation of the condition under which self-locating thought arises in visual experience is phrased in modal terms: it is the condition of an experience 'represent[ing] an object as the focus of its subject's *possible* action' (2018, p. 264, my italics). Thus it remains open, I want to suggest, whether the *Carne y Arena* viewer's experiencing the environment as *able* to harm her goes some way towards achieving what is achieved by the offender's experiencing, in Sanchez-Vives and Slater's experimental setting, the virtual abuser's ability to harm her plus experiencing a similar disposition of her own, *exercised*, when she in fact manipulates the virtual shapes via the tracked glove prior to the virtual abuse.

Perhaps what I'm suggesting depends on whether visual perception allows us to form attitudes about the relevant unexercised capacities (either ours or our environment's relative to us) or, in other words, whether perception-based *de se* beliefs can involve modal contents, 'I can x; I can be x-ed'; and if so, whether the experienced environment being *virtual* makes a difference. While Alsmith of course thinks the answer to the former question is affirmative (the agentic account rests on it), he himself thinks the answer to the latter question is negative:

In some cases, the subject may represent an object as the focus of her possible action even if the object is [...] merely virtual or imaginary. [...]

Nevertheless, to the extent that her experience interfaces with her practical capacities, and in a manner that is equivalent to those cases in which the genuine article is present, the object is represented as the focus of possible action (ibid., p. 275).

Now, Alsmith's answer takes for granted that in visual experience we 'represent' possibilities, in which view he is not alone: that experience itself represents modal contents has been explored elsewhere (Gibson 1986, Nanay 2011, 2012). But it has recently been suggested that one needn't rely on representational views of experience – and so on the view that modal contents are perceptually represented – to get the result that perception yields non-inferential attitudes about our abilities to interact with our environment (Strohlinger 2015). This is of course a matter of further investigation, but the subjects' behaviour in both *Carne y Arena* and Sanchez-Vives and Slater's experiment seem to support Alsmith's view. Primed by the seeing of exercised violence (the patrol agents abusing fellow immigrants; the domestic abuser smashing a telephone), then being verbally abused themselves, the *Carne y Arena* viewer drops to the floor; the experimental subjects 'felt very upset [...] they felt that their peripersonal space was invaded even when of course they knew there was nobody there.'⁵³ By allowing for the illusion of interactivity with the

⁵³ Quote from Sanchez-Vives' (2018). While Sanchez-Vives, Slater et al's original report (2018) does not discuss the subject's affective experience but only its cognitive upshot, in similar studies they have recorded, in addition to self-reports, increased stress levels during virtual experiences of 'negative social interaction' (Slater et al 2020a) and of a 'stressful VR situation' (Slater et al 2019). On a later report regarding the domestic violence experiment,

environment, then, understood in this broad sense, and by allowing viewer-initiated changes in visual point of view, head-tracking places VR on the higher end of the immersiveness spectrum, where the degree of immersiveness may be understood roughly as the extent to which the relevant *de se* thoughts ('I am in front of the monster', in video games; 'I am being abused', in the cases under consideration) are given credence by the subject.⁵⁴

A discussion of how different media afford different degrees of immersiveness merits a separate occasion.⁵⁵ But one important consequence is that the transformativeness of a virtual experience may be a matter of degree, too. To this we must add that the transformativeness of *any* novel experience comes in degrees, and other factors such as the viewer's individual psychology and even her social background can impinge on this in surprising ways (Barnes 2015). The details of this contingency are beyond the scope of this paper, but compatibly with it, the result seems to be that if Nemirow's notion of what it is to understand an experience is right, then there is a mode of presentation that understanding suffering requires, the subjective mode, and which VR, in virtue of its affording a sense of agency to the viewer, seems well-placed to meet.

however, they do report 'that actions in the space surrounding the artificial body are perceived as if they were close to the real body' (Sanchez-Vives, Slater et al 2020).

⁵⁴ For further details on what a VR system's level of immersion depends on, see Ramirez (2018), pp. 5-8; Slater (2018), pp. 431-2; Barbot and Kaufman (2020).

⁵⁵ One way to think about the spectrum is this: *if* the media on it are viewed as aiming to represent the strong first-person perspective, then the experiential responses they elicit might depend partly on the degree of representational accuracy. There are differences within each medium. Even within traditional film, for instance, which can only represent a character's weak first-person perspective, the sight-link shot seems to be less accurate than the point-of-view shot (cf. Cumming, Greenberg et al forthcoming). Further along the line, 3D is phenomenologically directed towards the viewer in a way traditional cinema is not, although here, too, authors distinguish between 'non-immersive' and 'immersive' 3D (Kozhevnikov and Dhond 2012). And then there are, as we've seen, differences between non-first-person-perspective VR, such as de la Peña's work, and first-person-perspective VR; and within the latter, differences between tracking the viewer's hands in addition to their head and not tracking them.

2.2 On the epistemology of suffering

The aim of allowing the viewer to index the suffering to herself stems from an epistemological intuition, that suffering is epistemically subjective. Let me very briefly flesh this intuition out. All mental states are ontologically subjective because they depend on their subject to exist, but propositional mental states such as beliefs and desires are *epistemically objective* whereas affective states are *epistemically subjective*.⁵⁶ A subject's beliefs and desires can be known by a third party: that is what discursive testimony helps to do. Meanwhile, the intuition goes, suffering cannot be accessed by any parties other than by the sufferer herself. Imagine a large amount of desires: there it is, expressed in a pile of letters to Santa. Now imagine a large amount of suffering: there *you* are, feeling awful.

This is no doubt to a very strong epistemological view. It seems to rest on something like a scepticism about other minds, the implausibility of which is shown by a *reductio ad absurdum* of the unimaginability of pain except as experienced by oneself: 'I will understand "pain" to mean "my pain" and the supposition that pain can be ascribed to anything other than myself will be unintelligible to me' (Thornton 2018). What this *reductio* brings out is that on such an epistemological view, an experience can only appear in one's mental state if one indexes it to oneself. But it isn't clear that this is so. It isn't clear, for example, that for some experience to feature in the contents of a subject's imaginative exercise, the subject herself must be the subject of *that* – the imagined – experience. This is easier to see in the case of visual experiences (cf. Williams 1973), but experiences of suffering have also been regarded by aestheticians to be made epistemically accessible by art (specifically by tragic

⁵⁶ The labels are Searle's (1992).

artworks) without the viewer's imagination delivering a representation of herself in the sufferer's circumstances (cf. Schier 1983, Friend 2007). Grant that tragic films can do this. The consensus seems to be that it has a moral consequence: 'confronting us with painful realities in a way that prompts ethical reflection' (Friend op cit., 190). This, in turn, will manifest in 'moral feeling', or feelings of *sympathy* for the sufferers, either real or fictional (ibid., pp. 189-90). That sympathy is the result of experiencing tragic artworks seems right. Schier contrasts it with an experience of 'mere perception' of suffering (1983, p. 84), which may elicit non-moral reactions. Think of witnessing an actual assault in the street and feeling the urge to flee or look away. If this is right, the question arises of why the extreme filmmaker would insist in producing a *de se* experience of suffering for her ethical ends. Why wouldn't Walton's 'sort of' empathy or Sinnerbrink's 'cinempathy' be a morally satisfying result? Why must, in other words, the target be to elicit Walton's 'true' empathetic thought, 'he is as I am, like this'?

One possible response from the extreme filmmaker – or González Iñárritu, at this point – might be to acknowledge that even if the epistemic access to suffering does not require the *de se* mode of presentation – even if tragic artworks can provide such access – there is a moral import to empathy but not sympathy for the suffering of others, a moral import, in other words, to feeling *with* rather than *for* the sufferers, to adopting *their* perspective. What such moral import is will be a question for the next section. To move on for now, let me just illustrate, without evaluating, the consequences that the above intuition on the epistemology of suffering may have on Haneke's and González Iñárritu's viewers. We can, once again, do this schematically. If Haneke's viewer has never seen someone extremely suffer before, after seeing his films she will have learned not what extreme suffering is like – not the phenomenal truth 'this is what extreme suffering is like' – but what *seeing someone else suffer* is like.

In contrast, if González Iñárritu's viewer has never been reduced to cattle, pointed a gun at, felt subhuman and so on before, after experiencing the piece she will have discovered not only a phenomenal truth ('this is what extreme suffering is like') but also a *de se* truth ('I am suffering'). If suffering is epistemically subjective, in fact, one cannot entertain the relevant phenomenal truth without entertaining the *de se* truth, so here we'd have a relation of implication between the phenomenal and the *de se* truths discovered; and one mark of personally but not of epistemically transformative experiences, according to Paul, is the correlation of both.⁵⁷

I need to delegate the defence of this point to Paul. Even if she and extreme filmmakers are right, however, the epistemological question may remain: why is it assumed that it is not possible to reach the *de se* truth 'I am suffering' by means of the imagination in some way? The response is that the imagination is indeed at work in *Carne y Arena*. No matter how charitably we read film critics' accounts about how *real* things felt to them, their experience remains of the imaginative kind. The viewer is not, after all, an immigrant fleeing across the US-Mexico border; she is just imagining being one. However strong the sensory and affective dimensions of the experience are, it remains in the domain of make-believe. But this does not preclude recognising the significance of the viewer's perceptual phenomenology thanks to which the target *de se* knowledge is achieved.

As Tim Williamson (2016) has pointed out, the imagination is not opposed to but rather works along perceptual input in the acquisition of knowledge. This is indeed, he speculates, how it thrives. To appreciate this point, consider how the imagination can work in both voluntary and involuntary modes. A group of early

⁵⁷ That correlation, unlike in our case, is typically a *leading-to* relation between the phenomenal and *de se* discoveries (in that direction), as in Paul's example of a congenitally blind subject who becomes sighted. Discovering 'what it is like to see' *leads to* discovering 'what it is like for *one* to be sighted' (Paul 2018).

humans who are distracted might be instantly pushed to imagine the dangers of the forest they're entering when they suddenly perceive the sounds of wolves (Williamson p. 115). This is a scenario the imagination could conjure up on its own but is *pushed* to by an occurrent perception. Another case is perception *enabling* the imagination to conjure up a scenario it wouldn't be able to conjure up otherwise. A hiker might want to imagine jumping a stream that is on a path he wants to take in order to strategize whether to jump it or find an alternative path instead. This imaginative exercise, however, requires actually going to the stream to gather some necessary perceptual knowledge, knowledge about 'the width of the stream, [...] the awkwardness of the place from which he would have to launch himself, the slipperiness of the rocks on which he would have to land, how tired he is, and so on' (Williamson p. 116).⁵⁸

The gloss offered earlier on of the extreme filmmakers' intuitions on the epistemology of suffering needs refinement in this light. The view that one does not understand extreme suffering without suffering oneself – that suffering is epistemically subjective – does not mean that the *de se* truth 'I am suffering' cannot be imagined. It means that the imagination cannot deliver this truth just from watching someone else suffer because those are not the appropriate perceptual cues,⁵⁹ just as the hiker cannot imagine jumping over the river without the necessary perceptual knowledge about its features and its surroundings.⁶⁰ By contrast, the

⁵⁸ This is a modified version of Williamson's hunter example.

⁵⁹ The characteristics of the required perceptual cues have to do, as the reader may suspect, with subject-centeredness. More on this in the next section.

⁶⁰ This point may be seen as an objection to Paul's view, according to whom *de se* knowledge of certain facts can only be provided by actually experiencing them. One first way to respond is that the point that such knowledge can indeed be arrived at via the imagination is, rather, just a point about what 'actually experiencing' something can amount to. In the case under consideration, the *de se* knowledge of suffering is arrived at via *perception*-guided imagination, after all, without which perceptual experience such an achievement wouldn't have been possible. Thus, experience ends up being crucial anyway, in accordance with Paul's view.

stimuli provided by *Carne y Arena* don't only allow but *push* the imagination to conjure up the target *de se* truth much in the way the hearing of wolves pushes the early humans to imagine the truth 'I am surrounded by wolves'.

That the imagination interacts in these ways with perception is not surprising.⁶¹ Williamson's examples are geared towards showing the evolutionary advantage of the two faculties' interaction. Both the hiker's imagining of jumping the stream, which will elicit vertigo (or quasi-vertigo), and the early humans' imagining of being surrounded by wolves, which will elicit fear (ditto), result in their taking appropriate actions, which will keep them safe. But the advantage cuts both ways. It would also be disadvantageous for the imagination *not* to have such a cognitive gateway; for there to be no limit to what one can imagine unaided by perception, or with less aid from it:

An imagination that clutters up the mind with a bewildering plethora of wildly unlikely scenarios is almost as bad as no imagination at all. It is better to have an imagination that concentrates on fewer and more likely scenarios. One's imagination should not be completely independent of one's knowledge of what the world is like (Williamson op cit., p. 115).

Frank Jackson (1986) illustrated most famously, perhaps, the importance of perceptual experience for one's imaginative capacities. 'All the imagining [Mary]

Then again, perhaps an upshot of this paper is that the line between what counts as providing a 'transformative experience' in Paul's sense and what does not is blurred when art – and technology! – is concerned.

⁶¹ Indeed, the present point about perceptual input's pushing and allowing relations (in that direction) to the imagination's operation locks into the more general view that the imagination's operation follows perception's in that they employ the same cognitive capacities, albeit offline and online, respectively. See e.g. Kosslyn (1994), Williamson (2007, Ch. 5), Yablo (2002), Jackson (2018).

could do would not be enough for her to know what it was like to see color', Paul herself writes early on in her book (2014, p. 8). This perceptual constraint is usually viewed in a negative light, but the case under discussion shows it can be a good thing, from an evolutionary perspective, that humans are not so easily capable of representing to themselves the *de se* truth 'I am extremely suffering'. We wouldn't want the imagination to elicit such strong emotional responses as those elicited by *Carne y Arena* every time we see an extreme film (or see someone crying, say, or so much as read the news), not just for the cognitive load this would carry but also because of the actions such experiences might continually motivate us to do or avoid doing. The emotions elicited by imaginative experiences can be motivationally powerful, and for good reason: the hiker's vertigo and the early humans' fear pushed them to act in the way that would keep them safe. This bolsters, perhaps, the suspicion that extreme cinema can't reach their goal via their traditional medium. Extreme filmmakers are not only up against the industry that produces 'media death' and blinds audiences to 'real' violence, as Haneke puts it; they're also up against human nature itself.

2.3 On morality and perspective

The acknowledgement that *Carne y Arena* owes its success, such as it is, not only to the perceptual experience it offers but also to the imagination's work prompts the thought that perceptual cues other than the ones distinctive of VR environments may also be able to stimulate *de se* imaginings of suffering, albeit in their own, different ways.⁶² From this, it would follow that VR is less distinctive a medium than what critics' reports on *Carne y Arena* would have one believe. Let me elaborate. On the

⁶² One might think, for example, that one feat of the advertising industry has been to prompt viewers to imagine *themselves* experiencing the situations ads show them.

one hand, as discussed, what seems to explain VR's ability to ground a *de se* imagining of suffering is something VR shares with another medium, first-shooter person games. It is the ability to give the viewer a sense of agency (albeit to a different extent than videogames). On the other hand, the fact that the target *de se* knowledge is reached via an interaction of both sensory faculties, perception and the imagination, is something VR presumably shares with those games as well as another medium, traditional film.⁶³ This means that the power of *Carne y Arena* should not be cashed out as that of *uniquely* providing *de se* knowledge of suffering to the audience but providing such knowledge, if it does, in its own distinctive way – by offering a perceptual experience that *perspectivally* parallels the target sufferer's.

Acknowledging this gives rise to a further question. Even if the *Carne y Arena* viewer shares the immigrants' perspective, the contents of the two experiences are not qualitatively identical. Although the imagery is not computer-generated but shot with actual, if specialised cinematography techniques, once you get close to the characters, the image's quality decreases. In Sanchez-Vives and Slater's experimental virtual experience of domestic violence, on the other hand, the imagery *is* visibly computer-generated: the quality resembles an old videogame's. Importantly, too, non-visual aspects of the overall experience are largely dissimilar from the represented experience in both cases. The *Carne y Arena* viewer is not being kicked in the gut, nor starved for days, nor painfully thirsty. Sanchez-Vives, Slater et al's experimental subjects are not traumatised by prior abuse; they are not physically sore from the previous episode

⁶³ Indeed, it is important to note that the target type of suffering may be relevant here. Consider the suffering of the protagonist in Julian Schnabel's *The Diving Bell and the Butterfly* (2008), which stems from the character's incapacity to move as a product of a car accident. It isn't clear that the sense of agency afforded by VR would help to stimulate empathy with this type of suffering.

of beating, and so on. In short, the sensory information from other channels than the visual one is not altered in these virtual settings.⁶⁴

Yet the fact remains that both can have powerful effects, not only in affective but also, it seems, moral terms. Where the evidence in the former case consists in the presence of ethical vocabulary in film critics' reports, in the latter case it consists in the documented shift in the subjects' moral behaviour. This shift, which Sanchez-Vives and Slater phrase in epistemic terms – the 'knowledge that their behavior is wrong' (p. 7) – was not possible through watching educational films: 'perspective-taking techniques based on watching movies [...] can be problematic [...] Although Offenders might obtain abstract knowledge that their behavior is wrong, this may not actually result in behavioral changes' (ibid.).⁶⁵ In both cases, the sought moral effects seem to have been achieved even though the phenomenology of the virtual experience is pretty different from the real one. How to explain this?

One first thing to say is that this may be an effect of visual dominance in the overall phenomenology of cross-modal perception. Visual dominance is a phenomenon well-known to most of us (how many times have we heard we're an overly visual culture?), and which Christopher W. Robinson, Marvin Chandra & Scott Sinnett (2016) put in slogan form thus: 'humans are inclined to focus on visual information when presented with compounded [...] stimuli' (p. 1104). If the phenomenology of conscious experience depends largely on what conscious attention is focused on, it is unsurprising that the phenomenology of the *Carne y Arena* and the experimental setting's viewers would be largely informed by what pulls their

⁶⁴ There are, as mentioned, small conceits in the *Carne y Arena* experience: one is put in a cold room before going in, and is made to go barefoot in order to feel the actual sand under one's feet during the experience.

⁶⁵ For details see Sanchez-Vives, Slater et al (2018). For a helpful survey of this and other such experiments, see Ramirez (2018), and for a discussion of the psychological mechanisms at play, see Bertrand et al (2018).

attention, which is what they *see*. If what they see is an imperfect representation of what the real sufferers see, though, the explanatory weight falls on *where* they see it *from*, the perspective. So the question arises of what the relation is exactly between perspective-shifting and moral change.

It will be helpful to frame the rest of the discussion within Kenneth Walden's (2015) view that there are two fundamentally different ways art can shape our moral outlook. Sinnerbrink's earlier characterisation of film's power to 'sharpen our moral perception [... and...] enhance our understanding of moral-social complexity' fits straightforwardly Walden's characterisation of the most common such way: 'change by accretion in moral thought', which consists in 'the *acquisition* of new moral knowledge, the *refinement* of moral concepts and the capacity to apply our moral view in particular situations' (ibid., p. 283, my italics). Sinnerbrink's mechanism for that role of cinema was cinempathy, which it achieves by 'exercis[ing] our moral imaginations'. This is also how Walden characterises the accretionary effect of literature. Because literature typically plays 'an important role in moral imagination', he writes, it 'elicit[s] novel considered moral judgements [...] and in doing this we may take on *new* moral beliefs, *revise* some old ones, and reject others' (ibid., pp. 284-5, my italics).⁶⁶

The other, less common way art can impact our moral outlook is by restructuring it. Walden calls this, as opposed to change by accretion, change 'by reconfiguration'. Where an ethically accreting artwork engages with the propositional content of our moral framework (by refining, rejecting or advancing new beliefs), an ethically reconfiguring artwork engages instead with its structural features such that

⁶⁶ An illuminating account of how literature does this is offered by Antonia Peacocke's (2020) view that by calling forth past experiences, literature can help us focus on unnoticed phenomenal properties of them – and so gain new phenomenal concepts – without providing us with new experiences.

we can now ‘focus attention’ on aspects of reality that our prior framework occluded. Rather than showing audiences new contents, such an artwork takes what they had always seen from just one side, as it were, and shows it to them from a new side, thereby ‘mak[ing them] reflexively reckon with what has been pushed in front of their noses’ (ibid., p. 289).

One might think *Carne y Arena* does something along these lines. It shows us something we already knew from one side – the middle-class, news-reading side – from a new one: the very sufferer’s. In doing so, it allows us to represent to ourselves that phenomenon from that side, which ability we didn’t have before. Consider an analogous case: Jackson’s pre-liberation Mary reading Vladimir Nabokov’s *Transparent Things* (1972) and coming across the line: ‘he felt a roaring redness fill his head’. If we accept what Davis Lewis (1990) calls the ‘phenomenal information hypothesis’ to interpret Jackson’s experiment, the impossibility of Mary representing to herself this content is explained by her lacking the *phenomenal* concept ‘seeing red’ even if she already had the corresponding *theoretical* concept (Loar 1990). Alternatively, if we accept what Lewis calls the ‘ability hypothesis’ (his choice), that impossibility is explained by Mary’s lacking the ability to ‘visualize red’, which is Nemirow’s way (quoted by Lewis) of summarising what he spells out as the ‘ability to place oneself, at will, in a state representative of the experience [of seeing red]’.

Either hypothesis works for our case. The impossibility of a viewer representing to herself any content involving extreme suffering can be explained by her lacking the phenomenal concept of extreme suffering even if she has the corresponding theoretical concept. Alternatively, that impossibility can be explained by her lacking the ability to place herself, at will, in a state representative of the experience of extreme suffering. Phenomenal concepts refer to mental entities, to experiences or to their qualitative properties, and they are typically fixed, just like the

ability to put oneself in a state representative of an experience, when one has undergone an initial instance of it (Loar op cit., Lewis op cit.). Lewis' notorious example is the phenomenal concept of – or the ability to put himself in the state representative of – the experience of tasting Vegemite, which concept or ability Lewis never acquired because he never tried the reportedly peculiar spread (lest he 'spoil a good example!'). By assumption, the viewer has never undergone anything like extreme suffering either.⁶⁷ Traditional film goes to great lengths of aesthetic skill to restructure our moral thought into a novel perspective (a feat the harder the more alien that perspective is to one's default); meanwhile, González Iñárritu simply puts the perspective on our heads. Hence his success in making us 'feel subhuman' or 'reduced to cattle'.

Here's how these effects may be understood in Walden's light. While the accreting artwork's lesson consists in 'better grasping the extension of the moral concepts we use' (p. 284), the reconfiguring artwork's lesson consists in *wholly re-determining*, not just better grasping, the extension of those concepts. Haneke's films may well help to better grasp the extension of the concept 'extreme suffering' if the viewer, say, had never been exposed to extreme suffering even from a third-personal perspective. Meanwhile, the *de se* mode of presentation of suffering in *Carne y Arena* wholly determines a new extension for that concept along with a new extension, presumably, for 'morally wrong':⁶⁸ those extensions now include *this* experience that I

⁶⁷ For a similar argument against the imagination's power to give us understanding of the negative experiences of others that factors in, additionally, bodily constraints, see Mackenzie and Scully (2007).

⁶⁸ Such a way of categorizing *Carne y Arena* as a reconfiguring artwork relies on the distinction between phenomenal and theoretical concepts. It seems available to someone to argue, against this, that the concept *suffering* already was in the viewer's repertoire, and the experience just helped her to better grasp it. Because during the experience the viewer herself is suffering, however, the concept will *now* include her own state in its extension, and the extension-redetermining phenomenon is just part of Walden's characterisation of reconfiguring artworks.

am undergoing right now and which mirrors the immigrants': 'he is as I am, like this'.⁶⁹

That VR environments whose main contribution is to put the viewer in the virtual perspective of the target subject do enable her to use her phenomenal state as a sample for the target's despite the qualitative difference seems to be supported not just theoretically by the considerations discussed in Section 1.2 but also by psychological findings regarding the significance of perspective in VR. According to Edgar Ramirez (2018), the degree to which VR experiences are experienced '*as if* they were real' depends on two dimensions: *perspectival-fidelity* ('the degree to which a simulation accurately represents the subjective point-of-view of human observers') and *context-realism* ('the actual content of the simulation'). When it comes to the latter, 'photorealism appears to play only a small role in terms of the degree of context-realism in a simulation' (p. 7). We might cash out this requirement in terms of modal closeness to the actual world: 'a world in which subjects can perform superhuman actions is [...] likely to be experienced as less real than one where subjects are bound by the same rules as in our world' (p. 8). Extreme cinema's contents are all too close. And then, when it comes to perspectival-fidelity,

a simulation that presents a first-person point-of-view to the subject is higher in perspectival-fidelity than one presenting the same content from a third-person perspective (ibid.).

⁶⁹ I am stating the point, following Walden, that experiencing a novel point of view can have a 'reconfiguring' effect. But neither of us are giving an account of how exactly this happens. A promising one might be found in Elisabeth Camp's (2018) discussion of the way adopting a novel perspective (as a result of engagement with fiction) justifies novel moral judgements (about the fictional scenarios) in addition to causing novel emotions (about them). To summarise: on the one hand, new perspectives justify new moral evaluations, and on the other, they cause new emotions, which emotions in turn at least partially ground those new moral evaluations. A key role is played by a third element that Camp calls 'characterisations', a prototype-like mental representation that completes an 'intimate normative triangle among characterizations, emotions, and moral judgments' (Camp op cit., p. 82).

Phenomenology is complex, so Ramirez is wise to speak of experienced realness in terms of degree. Indeed, Mel Slater (2018) characterises VR systems' powers of immersion as a partial order (p. 432). On some accounts, traditional film itself counts as a VR system, albeit one on the least immersive end of the spectrum (Pan and Hamilton 2018). The point-of-view shot in traditional film, for example, may be taken to encode what Blanke and Metzinger (2009) call the 'weak first-person perspective', the geometrical property of a visual representation's having a 'point of projection functioning as its origin'.⁷⁰ What Blanke and Metzinger call the 'strong first-person perspective' encompasses, by contrast, the perceptual beliefs, emotions and other attitudes that come with experiencing, as an immersed 'epistemic subject', the world from a certain point of view (Blanke and Metzinger 2009, p. 8). The extent to which a VR system yields the strong first-person perspective – or the 'illusion of presence', as Slater calls it, or the 'embodied perspective', as Sanchez-Vives, Slater et al do – depends on how successfully the perceptual stimuli offered by the virtual environment prompts in the viewer a fine-grained imaginative experience of the target situation. This, again, is too complex a phenomenological state to regard as the result of this or that technical feature of the medium, although Sanchez-Vives, Slater et al suggest a threshold: the illusion of presence is 'a function *at least* of head-tracking' (Slater op cit., p. 432, my italics).

That psychological research has found head-tracking to be a threshold for the illusion of presence seems to support the contention made in Section 2.1 that there is a link between the subject's sense of agency and her ability to hold self-locating

⁷⁰ The immersiveness provided by the POV shot is often cashed out in terms of its allowing character identification: 'the point-of-view shot is often thought of as the locus of character identification in film' (Gaut op cit., p. 208). But authors who are sceptical about this connection (e.g. Gaut op cit., Smith 1995) have framed their scepticism in terms of an equivocation between the weak and strong first-person perspectives discussed here. The 'fallacy of POV', Smith writes, is the mistake of taking POV shots 'to represent, synechdochically, the entire mind of the character' (Smith 1995, p. 156).

thoughts. By being able to hold self-locating thoughts *about the represented experience*, the viewer seems able to use her state as a sample of the target subject's state – the immigrant's, in *Carne y Arena*; the abused woman's, in the Sanchez-Vives and Slater's experiment. The virtual and real experiences are qualitatively different, but prompted by the perceptual cues that are constitutive of the first-person perspective (action being directed towards the viewer, the visual field's point of projection depending on the direction of her gaze, and so on), the viewer's imagination is pushed to yield the target *de se* beliefs, 'I am being threatened', 'I can be shot/beat', and so on. Morally speaking, this is significant because experiencing a known phenomenon from a novel perspective allows one to form novel moral evaluations about it. Perhaps the viewer condemned extreme violence before, or perhaps she thought Latin migrants/women were deserving of such treatment. Either way, during the VR experience, the viewer may now include her own state in the same extension as that of the sufferer's in virtue of her state's standing as a sample of the sufferer's, which in turn allows her to morally condemn the phenomenon by pointing not just outwards, as Sinnerbrink's cinempathizer and Walton's 'sort of' empathizer might, but inwards, too.

Conclusion

Roger Ebert's optimism about the power of cinema would, not surprisingly, be taken up a decade later by advocates of VR. 'Virtual Reality can create the ultimate empathy machine', Chris Milk, the founder of a VR production company, said in a popular TED Talk in 2015. Not long after, film critics would celebrate *Carne y Arena* as 'the harrowing Cannes VR installation that Donald Trump needs to see' (Robey 2017). Such enthusiasm about a medium with the perceived power to make even Trump empathise with those he scorns is understandable. The aim of having

viewers understand, in Nemirow's sense, the experience of extreme suffering is one cinema has long harboured.

In this paper I've attempted to make headway towards explaining the success *Carne y Arena* has had in this regard and to point out some reasons to be wary of overblowing it. Its success can be explained, I have suggested, by the virtual environment's ability to steer the viewer towards a mental state that is sufficiently parallel to the immigrant's, so that it can be employed by her (the viewer) as a sample of the immigrant's mental state. What did this 'parallel' relation consist in? It did not consist, we saw, in phenomenological identity. Rather it consisted in the viewer's sharing the target subject's – the migrant's – visual perspective. Even in the presence of phenomenological dissimilarities, both visual and non-visual, between the virtual and real experiences, the sharing of a visual perspective allowed, in virtue of the viewer's enjoyment of a sense of agency, to elicit a *de se* truth about the suffering: it allowed her to index the experience to herself. Hence the viewer's ability to pass the Paul-Walton bar for empathy.

Having said this, the question arose of why filmmakers should strive to give the viewer the ability to hold the target *de se* thought given philosophers' insistence, in the aesthetics literature, that art *can* give epistemic access to suffering in a way that is ethically valuable, only via sympathy. The answer to this question amounted to an account of the moral significance of perspective-shifting. Walden's distinction provided the framework. Giving the viewer the chance to adopt a different perspective allows a 'restructuring' of her moral thinking, that is, it allows the extension of certain concepts to be redetermined for her by her new perspective, which in turn allows new moral evaluations. Even if a viewer already condemned violence against immigrants, for example, the novelty of her new moral evaluation lies in its including now a reference to the phenomenon the evaluator herself is

experiencing. So much seems to fit our data. *Carne y Arena*'s impact on film critics (as well as Sanchez-Vives and Slater's experiment's impact on domestic abusers) fit the description of the goals set by extreme filmmakers for themselves, the goal of offering their audiences first-person access to extreme violence in order to effect a moral change.

The framework in which to elucidate the aim of extreme cinema and the explanation, in that framework's terms, of why our test case seems to have succeeded constitutes the positive contribution of this paper. On the negative side, though, I also presented reasons to be wary of overblowing VR's power. Specifically, one should be wary of overstating its distinctiveness. I suggested that VR is not wholesale novel but continuous with other media. The connection between the sense of agency VR gives to the viewer and its ability to elicit *de se* thoughts is also present, albeit to a different extent, in first-person shooter video games; and the interplay between perception and imagination thanks to which the target *de se* thought is elicited is present in traditional film, too.

I want to close by pointing out the need for further reflection on some dimensions of the aesthetic of abusing the viewer now. First, one might be wary of overstating the moral significance of eliciting first-person imaginings of suffering. It is true that, as Langton (op cit.) tells us, philosophers have often placed a great moral weight on the motivational role of knowledge what-it-is-like of the experiences of others. Langton cites R. M. Hare, who wrote that '[knowing *what it is like for him*] is required for the full information which rationality in making moral judgements demands' (Hare 981, p. 92, in Langton op cit. p. 90), as well as Mark Johnston, who wrote that "the authority of affect" [...] can make the desire or action especially intelligible to the agent himself (Johnston 2001 pp. 187-8, in Langton op cit., p. 92). But as Langton herself remarks following Peter Goldie (2011) (and indeed

Bernard Williams (op cit.), although Williams goes unmentioned), one can imagine what an experience is like without imagining oneself in a particular subject's situation of experiencing it, much like, as mentioned earlier on, one can imagine a scene without imagining oneself – or anyone – seeing it (Langton op cit., pp. 87 and 94). One can have, in other words, a sense of the phenomenology of the experience without attaching it to any subject. Langton's example involves imagining swimming: just call to mind the properties of the sea, the movement of the waves, the taste of salty water, none of which necessarily involves imagining yourself – or anyone else – *being rocked by the waves, tasting the water*. Along similar lines, Kate Manne has complained that suffering's having a motivational role needn't mean it can't motivate anyone but the sufferer. 'The fact that *I* and not *you* feel *my* pain', she writes, only means that '[t]he motivational speaker lives in me, versus you. But [w]hy does this matter one way or the other, when it comes to the possibility of directly inspiring or motivating action? (Manne 2017, p. 19). After all, agents often fail to act on their own sufferings,⁷¹ just as they are often motivated by the sufferings of others and equally urgently. Consider 'a person's responding to his baby's cry of hunger' (Manne op cit., p. 19). This case brings up the point that parents hardly need to imagine themselves in their babies' shoes – or employ their state as a sample of their babies' – in order to be moved to act on their suffering.

So here is one first avenue for further research. The question of whether moral action requires empathy in Walton's and Paul's sense needs further examination. Secondly, even if one is sympathetic to extreme cinema's views, one should be mindful of issues that are beyond those internal to their project. Whether the experience González Inárritu offers is transformative or not is, as already

⁷¹ Due to *akrasia*, for instance, or to *asymbolia*, a pathology that robs you of your 'capacity to show care or concern for [your] own well-being' (cf. Cohen and Fulkerson 2017, §2.3).

mentioned, contingent across individuals;⁷² moreover, the question remains unanswered of how exactly the *Carne y Arena* viewer could engage in Haneke's dream of 'social and political change'. Even granting empathy for the immigrants is necessary for the viewer to feel a practical motivation, what action exactly can *Carne y Arena* viewers be expected to take after their experience? What concrete scenarios will their newly acquired *de se* knowledge – and their newly formed moral evaluations – lead to?

And finally, questions arise from cinema's stretching of its limits, if this is what is happening here. These questions concern not only the ethics of virtually abusing viewers (and whether the artistic context is relevant)⁷³ but also whether stepping into the shoes of those who suffer is ultimately a desirable – or even permissible – exercise for the privileged to engage in, for reasons concerning the ethical⁷⁴ and political⁷⁵ relations between those groups, and indeed, reasons concerning the ethics of this exercise's epistemic end.⁷⁶

In many ways, then, this discussion is incipient. But if Walden is right that 'it will be impossible to identify the [...] forms of artistic reconfiguration in advance of actually experiencing them' (op cit., p. 294), incipience is expected in exploring the implications of what we're just learning to identify: an audiovisual experience delivering a *de se* mode of presentation of suffering with the seeming potential to have unprecedented epistemic and moral effects.

⁷² The offenders in Sanchez-Vives, Slater et al's (2018) experiment, for example, unlike regular film-goers (and the experiment's own control subjects), seem to exhibit psychological deficits. Their experience was hence measurably more transformative than the control subjects' (ibid., p. 5) even if the affective response of both groups was similar, as evidenced further by a study where subjects were not violent and male but healthy and of both sexes (Sanchez-Vives, Slater et al 2020).

⁷³ See Parsons (2015), Madary and Metzinger (2016), Slater et al (2020b).

⁷⁴ For a worry along these lines, see Mackenzie and Scully (2007), pp. 345-6.

⁷⁵ See Srinivasan (2018), pp. 219-20.

⁷⁶ See Davis (2018).

Now, as a final note (and as a final twist), let me point out this. If extreme cinema is, as Weigel and others⁷⁷ have said, essentially European, then the fact that those who've come closest to achieving its goals are Mexican might mean Weigel is right that extreme cinema is, as a European tradition, effectively dead.^{78,79} But this last point may be a good thing. I, for one, don't mind calling whatever González Iñárritu is doing a different name than extreme cinema, nor do I mind the death of a genre if nationality was such a petty part of its definiens. On the other hand, the contingency of whether the virtual experience will be transformative, and to what extent, may only add to the richness of it. Consequently, however, whether González Iñárritu in fact achieves what has been suggested here ends up being an open question *you*, reader, will have to experience the piece to answer yourself.

References

- Agamben, G. 1999. *The Remnants of Auschwitz: the Witness and the Archive*. New York: Zone Books
- Alsmith, A. 2018. Perspectival structure and agentive self-location. In A. Alsmith and F. de Vignemont (eds.). *The Subject's Matter: Self-Consciousness and the Body*. Cambridge, MA: MIT Press
- Arendt, H. 1973. *The Origins of Totalitarianism*. New York: Harvest
- Barbot, B. and Kaufman, J. C. 2020. What makes immersive virtual reality the ultimate empathy machine? Discerning the underlying mechanisms of change. *Computers in Human Behavior* 111, doi:10.1016/j.chb.2020.106431

⁷⁷ Horeck and Kendall (2011).

⁷⁸ González Iñárritu's cinematographer is fellow Mexican Emmanuel Lubezki.

⁷⁹ Or does it mean extreme cinema is going *back* to its roots? Pérez (op cit.) characterises Haneke as *following* Luis Buñuel, who was Spanish-Mexican, in 'deliberately go[ing] beyond the limits of permissible displeasure'.

- Barnes, E. 2015. Social Identities and Transformative Experience. *Res Philosophica* 92(2), doi: 10.11612/resphil.2015.92.2.3
- Bertrand, P., et al. 2018. Learning Empathy Through Virtual Reality: Multiple Strategies for Training Empathy-Related Abilities Using Body Ownership Illusions in Embodied Virtual Reality. *Frontiers in Robotics and AI* 5(26), doi: 10.3389/frobt.2018.00026
- Bishop, B. 2017. Alejandro González Iñárritu's *Carne y Arena* proves that great virtual reality means going beyond the headset. *The Verge*, 8 July. Available at <https://www.theverge.com/2017/7/8/15941596/carne-y-arena-alejandro-inarritu-virtual-reality-installation-border> (accessed 24 July 2017)
- Blanke, O. and Metzinger, T. 2009. Full-body illusions and minimal phenomenal selfhood'. *Trends in Cognitive Sciences*, 13(1), 7–13.
- Blümlinger, C. 2010. Figures of Disgust. In R. Grundmann (ed.) *A Companion to Michael Haneke*. Oxford: Wiley-Blackwell
- Bradshaw, P. 2017. *Carne y Arena* review - dazzling virtual reality exhibit offers a fresh look at the refugee crisis. *The Guardian*, 22 May. Available at <https://www.theguardian.com/film/2017/may/22/carne-y-arena-review-inarritu-virtual-reality-refugee-cannes-2017> (accessed 24 July 2017)
- Brown, W. 2013. Violence in Extreme Cinema and the Ethics of Spectatorship. *Projections* 11, doi: 10.3167/proj.2013.070104
- Camp, E. 2018. Perspectives in Imaginative Engagement with Fiction. *Philosophical Perspectives* 31
- Chalmers, D. 2016. The Virtual and the Real. *Disputatio* 9(46), 10.1515/disp-2017-0009
- Cohen, J. and Fulkerson, M., 2017. Pain and rationality. In J. Corns (ed.), *The Routledge Handbook on the Philosophy of Pain*. London: Routledge

- Cumming, S., Greenberg, G., et al. Forthcoming. Showing Seeing in Film. *Ergo*
- Crane, T. 2006. Brentano's Concept of Intentional Inexistence. In Mark Textor (ed.), *The Austrian Contribution to Analytic Philosophy*. London: Routledge
- Davis, E. 2018. On epistemic appropriation. *Ethics* 128(4), doi: 10.1086/697490
- De la Peña, N. 2017. Curriculum Vitae. Available at http://www.immersivejournalism.com/wp-content/uploads/2010/05/Nonny_de_la_Pen%CC%83a_CV_Jan_2015.pdf (accessed 24 July 2017)
- Dougherty, T., et al. 2015. Expecting the unexpected. *Res Philosophica* 92(2), doi: 10.11612/resphil.2015.92.2.5
- Didi-Huberman, G. 2001. Images malgré tout. In: Chéroux, C. (dir.) *Mémoire des camps. Photographies des camps de concentration et d'extermination nazis (1933-1999)*. Paris: Marval
- Ebert, R. 2005. Ebert's Walk of Fame remarks. Available at <https://www.rogerebert.com/roger-ebert/eberts-walk-of-fame-remarks>
- Evans, G. 1980. Things Without the Mind. In *Collected Papers*. Oxford: Clarendon Press
- Farago, J. 2017. Inárritu's 'Carne y Arena' Virtual Reality Simulates a Harrowing Border Trek. *The New York Times*, 17 May. Available at <https://www.nytimes.com/2017/05/17/arts/design/alejandro-gonzalez-inarritu-carne-y-arena-virtual-reality-cannes.html> (accessed 24 July 2017)
- Fodor, J. 1998. *Concepts: Where Cognitive Science Went Wrong*. New York: Oxford University Press
- Fraser, R. 2018. The Ethics of Metaphor. *Ethics* 128 (4), doi: 10.1086/697448

- Frey, M. 2014. The Ethics of Extreme Cinema. In J. Choi and M. Frey (eds.). *Cine-Ethics: Ethical Dimensions of Film Theory, Practice and Spectatorship*. New York: Routledge
- Frey, M. 2016. *Extreme Cinema: The Transgressive Rhetoric of Today's Art Film Culture*. New Jersey: Rutgers University Press
- Friend, S. 2007. The Pleasures of Documentary Tragedy. *British Journal of Aesthetics* 47
- Gaut, B. 2012. *A Philosophy of Cinematic Art*. Cambridge: CUP
- Gendler, T. 2007. Alief and Belief. *Journal of Philosophy* 105(10), pp. 634-663
- Gleiberman, O. 2017. Cannes Virtual Reality Review: Alejandro G. Iñárritu's 'Carne y Arena'. *Variety* 20 May, 2017. Available at <http://variety.com/2017/film/reviews/carne-y-arena-review-alejandro-g-inarritu-1202438293/> (accessed 24 July 2017)
- Grønstad, A. 2012. *Screening the Unwatchable*. London: Palgrave Macmillan
- Gustafsson, H. 2014. Archaeology after Auschwitz. In Grønstad, A. and Gustafsson, H. (eds.), *Cinema and Agamben: Ethics, Biopolitics and the Moving Image*. London: Bloomsbury
- Hall, R. J. 2008. If it itches, scratch!, *Australasian Journal of Philosophy*, 86(4): 525–535
- Hammond, M. M. 2017. *Empathy and the Psychology of Literary Modernism*. Edinburgh: EUP
- Horeck, T. and Kendall, T. (eds.) 2011. *The new extremism in cinema: from France to Europe*. Edinburgh: Edinburgh University Press
- Jackson, F. 1986. What Mary Didn't Know. *Journal of Philosophy* 83(5), pp. 291-295

- Jackson, M. B. 2018. Justification by Imagination. In F. Macpherson and F. Dorsch (eds.), *Perceptual Imagination and Perceptual Memory*. Oxford: Oxford University Press
- Khaldi, T. 2017. Event: virtual reality seen by Iñárritu with Carne y Arena. *Festival de Cannes*, 22 May. Available at <http://www.festival-cannes.com/en/festival/actualites/articles/event-virtual-reality-seen-by-inarritu-with-carne-y-arena> (accessed 24 July 2017)
- Kosslyn, S. 1994. *Image and Brain: The Resolution of the Imagery Debate*. Cambridge: MIT Press
- Kozhevnikov, M. and Dhond, R.P. 2012. Understanding immersivity: image generation and transformation processes in 3D immersive environments. *Front. Psychology* 3:284. doi: 10.3389/fpsyg.2012.00284
- Lang, B. 2017. How Alejandro G. Iñárritu Used Virtual Reality to Tackle Illegal Immigration. *Variety*, 17 May. Available at <http://variety.com/2017/film/markets-festivals/alejandro-g-inarritu-virtual-reality-1202430264/> (accessed 24 July 2017)
- Langton, R. 2014. Projected Love. In *Understanding Love: Philosophy, Film, and Fiction*, Oxford: OUP
- Loar, B. 1990. Phenomenal states. *Philosophical Perspectives* 4, doi: 10.2307/2214188
- Lewis, D. 1990. What experience teaches. In W.G. Lycan (ed.). *Mind and Cognition*. Oxford: Blackwell
- Lübecker, N. 2015. *The Feel-Bad Film*. Edinburgh: Edinburgh University Press
- Mackenzie, C. and Scully, J. L. 2007. Moral Imagination, Disability and Embodiment. *Journal of Applied Philosophy* 24 (4), pp. 334-351

- Madary, M. and Metzinger, T. 2016. Real Virtuality: A Code of Ethical Conduct. Recommendations for Good Scientific Practice and the Consumers of VR-Technology. *Frontiers in Robotics and AI* 3, doi: 10.3389/frobt.2016.00003
- Mitchell, J. Forthcoming. Self-Locating Content in Visual Experience and the ‘Here-Replacement’ Account. *Journal of Philosophy*
- Ordóñez Angulo, E. 2017. Transformative Suffering. *Los Angeles Review of Books*, 11 August. Available at <https://lareviewofbooks.org/article/transformative-suffering>
- Paul, L.A. 2014. *Transformative Experience*. Oxford: Oxford University Press
- Paul, L.A. 2015. Précis of Transformative Experience. *Philosophy and Phenomenological Research* 91(3), doi: 10.1111/phpr.12249
- Paul, L.A. 2018. De se preferences and empathy for future selves. *Philosophical Perspectives* 31
- Pan, X., & Hamilton. 2018. Why and how to use virtual reality to study human social interaction: The challenges of exploring a new research landscape. *British Journal of Psychology*. 109(3), 1–23.
- Parsons, T. D. 2015. Virtual Reality for Enhanced Ecological Validity and Experimental Control in the Clinical, Affective and Social Neurosciences. *Front. Hum. Neurosci.* 9, 660
- Pérez, G. 2012. Bourgeois nightmares. *London Review of Books* 34(23)
- Putnam, H. 1975. The meaning of “meaning”. In *Mind, Language and Reality*. Cambridge: Cambridge University Press
- Ramírez, E. 2018. Ecological and ethical issues in virtual reality research: A call for increased scrutiny. *Philosophical Psychology*, doi: 10.1080/09515089.2018.1532073

- Ramírez, E. and LaBarge, S. 2018. Real moral problems in the use of virtual reality. *Ethics and Information Technology*, doi:10.1007/s10676-018-9473-5
- Robey, T. 2017. Carne y Arena: the harrowing Cannes VR installation that Donald Trump needs to see. *The Telegraph* 19 May, available at <http://www.telegraph.co.uk/films/0/alejandro-inarritus-carne-y-arena-harrowing-cannes-vr-installation/> (accessed 24 July 2017)
- Robjant, D. 2012. Learning of Pains. *Wittgenstein-Studien* 3, doi: 10.1515/wgst.2012.261
- Sanchez-Vives, M. 2018. The Transformational Power of VR. *TEDxIESEBarcelona*. Accessible at https://www.youtube.com/watch?v=CUBXS_5aaLE
- Sanchez-Vives, M., Slater, M., et al. 2018. Offenders become the victim in virtual reality: impact of changing perspective in domestic violence. *Nature Scientific Reports* 8, doi: 10.1038/s41598-018-19987-7
- Sanchez-Vives, M., Slater, M., et al. 2020. First person virtual embodiment modulates cortical network that encodes the bodily self and its surrounding space during the experience of domestic violence. *ENEURO*.0263-19.2019, doi: <https://doi.org/10.1523/ENEURO.0263-19.2019>
- Saxton, L. 2008a. Close Encounters with Distant Suffering: Michael Haneke's Disarming Visions. In Kate Ince (ed.), *Auteurism from Assayas to Ozon: Five Directors*. Manchester: Manchester University Press
- Saxton, L. 2008b. Ethics, spectatorship and the spectacle of suffering. In L. Downing and L. Saxton (eds). *Film and Ethics*. London: Routledge
- Saxton, L. 2008c. *Haunted Images: Film, Ethics, Testimony and the Holocaust*. London: Wallflower
- Searle, J. 1992. *The Rediscovery of the Mind*. Cambridge, MA: MIT Press,

- Sexton, D. 2017. Cannes 2017: Carne y Arena, film review – Hot ticket to the revolution. *Evening Standard*, 26 May. Available at <http://www.standard.co.uk/goingout/film/cannes-2017-carney-arena-film-review-hot-ticket-to-the-revolution-a3549796.html> (accessed 24 July 2017)
- Sicart, M. 2011. *The Ethics of Videogames*. Cambridge, MA: MIT Press
- Shoemaker, S. 1968. Self-Reference and Self-Awareness. *Journal of Philosophy* 65: 555–67
- Sinnerbrink, R. 2011. A Post-Humanist Moralism. *Angelaki* 16(4), doi: <http://dx.doi.org/10.1080/0969725X.2011.641350>
- Sinnerbrink, R. 2015. *Cinematic Ethics: Exploring Ethical Experience Through Film*. London: Routledge
- Slater, M. 2018. Immersion and the illusion of presence in virtual reality. *British Journal of Psychology*, 109, doi:10.1111/bjop.12305
- Slater, M. and Sanchez-Vives, M. 2018. Enhancing Our Lives with immersive Virtual Reality. *Front. Robot. AI* 3:74. doi: 10.3389/frobt.2016.00074
- Slater, M. et al. 2020a. Embodied Perspective as a Victim of Sexual Harassment in Virtual Reality Reduces Action Conformity in a Later Milgram Obedience Scenario. *Sci Rep* 10, 6207
- Slater, M. et al. 2020b. The Ethics of Realism in Virtual and Augmented Reality. *Front. Virtual Real.* 1:1. doi: 10.3389/frvir.2020.00001
- Smith, A. 1759/1976. *The Theory of Moral Sentiments*. Glasgow Edition of the Works and Correspondence of Adam Smith, Vol. i. Oxford: Oxford University Press
- Smith, M. 1995. *Engaging Characters: Fiction, Emotion, and the Cinema*. Oxford: Clarendon Press

- Stear, Nils-Hennes. 2017. In *Other Shoes: Music, Metaphor, Empathy, Existence*.
British Journal of Aesthetics, doi: 10.1093/aesthj/ayw067
- Srinivasan, A. 2018. The Ineffable and the Ethical. *Philosophy and Phenomenological Research* 96(1), 215-223
- Strawson, P.F. 1959. *Individuals*. London: Methuen
- Thompson, A. 2017. Why Alejandro González Iñárritu is the Director Who Finally Got VR Right — Cannes 2017. *IndieWire*, 20 May. Available at <http://www.indiewire.com/2017/05/alejandro-gonzalez-inarritu-carne-y-arena-cannes-vr-1201819096/> (accessed 24 July 2017)
- Thornton, S. 2018. Solipsism and the Problem of Other Minds. *The Internet Encyclopedia of Philosophy*, ISSN 2161-0002, <http://www.iep.utm.edu>
- Utichi, J. 2017. The Birth Of An Art Form: How Alejandro G. Iñárritu And Emmanuel Lubezki Learned To Master Virtual Reality – Cannes. *Deadline*, 21 May. Available at <http://deadline.com/2017/05/alejandro-g-inarritu-emmanuel-lubezki-carne-y-arena-virtual-reality-cannes-1202099184/> (accessed 24 July 2017)
- Vendler, Z 1984. Vicarious Experience. In *The Matter of Minds*. Oxford: Clarendon Press
- Walden, K. 2015. Art and moral revolution. *The Journal of Aesthetics and Art Criticism* 73(3), doi: 10.1111/jaac.12173
- Walton, K. 1978. Fearing Fictions. *Journal of Philosophy* 75(1), pp. 5-27.
- Walton, K. 2015. Empathy, Imagination, and Phenomenal Concepts. In *Other Shoes: Music, Metaphor, Empathy, Existence*. Oxford: Oxford University Press
- Weigel, M. 2013. Sadomodernism. *N Plus One Magazine* 16
- Williams, B. 1973. Imagination and the self. In *Problems of the Self*. Cambridge: CUP

- Williamson, T. 200). *The Philosophy of Philosophy*. Oxford: Blackwell
- Williamson, T. 2016. Knowing by Imagining. In Kind, A. and Kung, P. (eds.).
Knowledge Through Imagination. Oxford: OUP
- Wheatley, C. 2009. *Michael Haneke's Cinema: the Ethic of the Image*. New York:
Berghahn
- Wittgenstein, L. 1953. *Philosophical Investigations*. Oxford : Basil Blackwell
- Wollheim, R. 1984. *The Thread of Life*. Cambridge: CUP
- Yablo, S. 2002. Coulda, Woulda, Shoulda. In T. S. Gendler and J. Hawthorne
(Eds.), *Conceivability and Possibility*. Oxford: Oxford University Press
- Zeitchik, S. 2017. Cannes 2017: Alejandro Iñárritu's virtual reality project takes film
to new frontiers—and questions. *Los Angeles Times*, 21 May. Available at
[http://www.latimes.com/entertainment/movies/la-et-mn-cannes-vr-inarritu-
20170521-htmstory.html](http://www.latimes.com/entertainment/movies/la-et-mn-cannes-vr-inarritu-20170521-htmstory.html) (accessed 24 July 2017)

Seeing the infinite

Abstract

According to the consensus in the epistemology of mathematics, experience provides 'no impression from which the idea of the infinite may be derived'. Here I argue against this consensus.

I shut my eyes – I opened them. Then I saw the Aleph.
I arrive now at the ineffable core of my story. [...]
How, then, can I translate into words the limitless Aleph,
which my floundering mind can scarcely encompass?

– Jorge Luis Borges

Introduction

In the notoriously obscure Section 66 of his *The Nature and the Meaning of Numbers* (1888/2012),⁸⁰ Richard Dedekind attempts a proof of the existence of an infinite ‘system’.⁸¹ Much has been written about this passage – about the presence of metaphysics in a work of sober mathematics; about the unsuccessful, it has been largely thought, ontological argument in it.⁸² But the way Dedekind puts his question gives rise to further reflection. ‘Does such a system exist at all in our realm of thoughts?’, he asks. The system Dedekind claims to be infinite and existing is a mental entity, his *Gedankenwelt*, his ‘realm of thoughts’. Hence, the proof is at once a metaphysical and epistemological project. It is geared toward showing both that infinite systems exist and that we are aware of them.

The theorem goes like this:

⁸⁰ The original, perhaps better-known title is *Was sind und was sollen die Zahlen?*. I quote the title of the translation with which I’m working.

⁸¹ ‘System’ is Dedekind’s pre-Cantor term. As it would be argued later, it wasn’t clear he could have used ‘set’, as the existence of a universal set gave rise to paradox; but it wasn’t clear he meant just ‘class’, either. Following the interpretive literature, I’ll stick to Dedekind’s own term when discussing his theorem.

⁸² For reactions, see e.g. Sieg and Schlimm (2005); Ferreirós (2011), §2; Klev (2018), esp. pp. 262-72; and McCarty (1995).

Theorem: there exist infinite systems.

Proof: my own realm of thoughts – or thought-world – is the totality, S , of things that can be an object of my thinking, and it is infinite. If we let a stand for an element of S , then the thought, b , that ‘ a can be thought about’, is itself an element of S . Let $\varphi(x)$ be a mapping from any element x in S to an element ‘ x can be thought about’. Its image, S' , the set of all output values the mapping function may produce, that is, is a proper part of S , because there are elements in S , e.g. my own self, that are not contained in S' . Finally, if we let a and b be distinct elements of S , then their images a' and b' are also distinct; hence the mapping φ is on the whole distinct. Hence S is infinite.⁸³

The proof is meant to show that Dedekind’s thought-world satisfies his own definition – proposed in Section 64, at the beginning of Part V, ‘The Finite and Infinite’ – of what an infinite system is:

Definition: A system S is infinite whenever there is a mapping f from S to S such that it is injective but not surjective.⁸⁴

For every element in S , there is exactly one element the mapping function can output, that is, the function is 1-1; hence the injection.⁸⁵ Yet there is also an element that is not the output of any such mapping: my own ego, myself, not a

⁸³ My gloss of the theorem, and my reading of Dedekind generally, owes much to Klev (op cit.).

⁸⁴ I’m glossing here, again guided by Klev (op cit.), over Dedekind’s notion of ‘similarity’ between systems, which is the term he employs in the definition.

⁸⁵ As Klev (op cit., pp. 65-6) explains, drawing from other writings by Dedekind, Dedekind simply assumes certain identity conditions of thoughts: ‘[i]f s and t are different elements of S , then the thoughts $J(s)$ and $J(t)$ also differ, since what they are about differs: whereas $J(s)$ is about (*handelt von*) s , $J(t)$ is about t ’.

thought about anything. Hence the failure of surjection.⁸⁶ So the system S meets Dedekind's definition. And it is easy to see in what sense this renders it infinite. If the thought 'o can be thought about' is itself an object of thought, then there is a further thought corresponding to it, "o can be thought about" can be thought about'; and so on. For all Dedekind tells us, the first element that gives rise to thoughts of this form is the self. The structure of the thought-world may thus be represented as in Figure 1, where the outer circle is S , the thought-world, and inner circle, S' , the image of the mapping $\varphi(x)$, and each closed rectangle an object of thought:

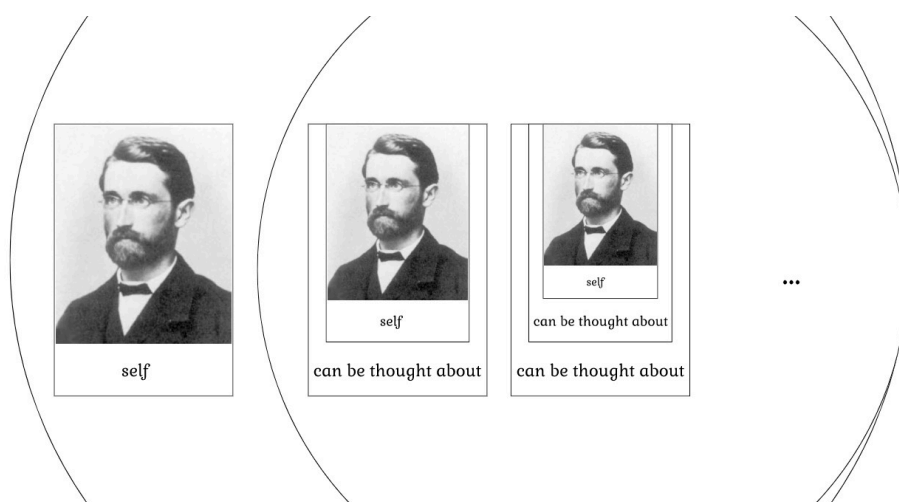


Figure 1 (picture of Dedekind from public domain).

Following my suggestion that the proof does not only have ontological but also epistemological import, one might note a salient issue with the proof: the notion of awareness of the self. Dedekind's reliance on it is problematic because the

⁸⁶ The simplest argument for this is that the self is an object of thought but not itself a thought at all; 'not a judgement', that is, 'with a determinate subject and predicate expressible in a sentence' (ibid.).

knowledge of such a thing, it might be thought, is empirical, whereas knowledge of the sort of entity Dedekind is targeting, an infinite system, is most likely not. Dedekind himself called his proof, albeit focusing on its ontological import, a ‘*logical proof of existence*’ (quoted in Klev 2018, p. 259, my italics). There is a case to be made, to be sure, as Anston Klev (op cit.) points out, that Dedekind is in fact thinking of a logical notion when he speaks of a ‘self’: not a flesh-and-bone entity or even a soul but ‘what Kant calls “the logical subject of thinking”’, in which case the object of which awareness grounds awareness of infinity for Dedekind is, just as the target entity, ‘an ens rationis, [...] “the object of a concept to which there does not correspond any intuition”’ (Klev op cit.).

Whether such a Kantian defence of Dedekind is feasible I will not adjudge. (McCarty 1995 attempts it; Klev is sceptical.) What I want to point out is that the issue of grounding our awareness of infinity is so fraught that what appears to be the most notable attempt to elucidate it found no echo after its author’s awkward failure, such as it was. Klev voices the impression, indeed, that it is only ‘in the form of the Axiom of Infinity that Dedekind’s proof survives today’ (Klev op cit., p. 265). Based on his investigations of Ernst Zermelo’s writings (both his correspondence and published work), Klev tells us that, convinced that the project was untenable, ‘[i]nstead of attempting to prove the existence of an infinite set’, or again, to elucidate our knowledge of it, ‘Zermelo postulates an Axiom of Infinity’, axioms being the kinds of truths we take as primitive, as explaining rather than to be explained.

Consider Zermelo’s axiom:

(Axiom of Infinity) There exists a set Z such that $\emptyset \in Z$ and $\{a\} \in Z$ whenever $a \in Z$.

‘The inspiration of Dedekind’s proof for the formulation of this axiom should be obvious’, Klev writes,

the empty set takes the place of the self and the mapping $a \mapsto \{a\}$ – which clearly does not have \emptyset in its range – the place of Dedekind’s mapping $\varphi(x)$. Zermelo indeed says that the axiom “essentially stems from Mr. R. Dedekind” (Klev op cit., pp. 265-6).

The empty set being a safely non-empirical notion, this axiom is, if not a vindication of Dedekind’s theorem, at least an expression of its conclusion that is free of ‘the intrusion of non-mathematical concepts’, as Charles Parsons complained about the theorem; an expression that is free of that ‘wildly non-mathematical item’, the representation of the self, as George Boolos saw it.⁸⁷

But why take the route, one might wonder, of simply asserting that (we know) there exist infinite sets rather than grinding on investigating how it is that (we know) they exist? Well, the natural first step toward *that* seems to be, in any case, what Zermelo did, to homogenise the facts in question: to make all the objects involved – or all the ‘items’, or all the ‘concepts’ involved, as Boolos and Parsons say – non-empirical. But that being done, the question Dedekind was tackling (how could an infinite system exist in our realm of thoughts?) loses its specialness. It becomes absorbed by the broader literature on a priori knowledge. Which is, perhaps, just as well. Surely it would have been a non-starter to go the other way around? Surely it would have been stranger still to ask the question not in fully ‘ens rationis’ terms, not in terms of ‘concepts to which there do not correspond any intuitions’,

⁸⁷ Both criticisms quoted in Klev (op cit.), p. 262.

but in terms of whether not just the starting item (Dedekind's self or Zermelo's empty set, any first item of an ω -sequence) but indeed the target entity, an infinite set (an entire ω -sequence!) is empirically knowable. This would have been stranger still because the infinite seems to be in even worse standing than the self when it comes to the sorts of things about which empirical knowledge is possible. The self is controversial in this respect; the infinite is not even so. As we will shortly see, the consensus seems to be that there is simply no experiencing infinity. In this paper I'll attempt just such a non-starter. As a segue into how, a word will be in order on the poor chances the infinite has been given in the literature to be empirically known, or as I'll say henceforth, known 'by experience'. It will be helpful to change gears now and ground ourselves in the context of a contemporary debate, which focuses not so much on metaphysics – as Dedekind (thought he) did – but on the role knowledge of infinity plays in justifying the edifice of mathematical knowledge.

The debate I'll introduce now is close in spirit to that famously articulated by Paul Benacerraf (1973) when he asked after our justification for believing mathematical truths. The difference is that infinity takes centre stage here. Here it is, at the core of Fraser MacBride's (2008) scepticism about Stewart Shapiro's (1997, 2011) epistemology of mathematics: '[m]athematical novices are inevitably beset', MacBride writes, 'by a poverty of stimulus which prevents them from ever extracting, or extrapolating, the infinite from what is given (structurally or otherwise) in experience' (p. 161). The debate concerns whether Shapiro's epistemology provides a satisfactory account of how someone who doesn't already know 'the principles which are employed by mathematicians to infer the general from the particular' (MacBride *op. cit.*, p. 159) can justifiably form a belief such as, 'all the natural numbers have a distinct successor'.

Shapiro's epistemology handles the matter of justification by appealing to the coherence of mathematical theories, where our ability to coherently talk about mathematical objects is itself evidence (to know) that they exist (Shapiro 2011, p. 135).⁸⁸ But the question is how a mathematical novice, who has no theory yet, might be warranted in making the jump from her beliefs about particular objects – '1 has a successor', '2 has a successor', '3 has a successor' – to a belief about *all* the objects in the relevant set. Shapiro is happy to attribute this jump to 'just having a hunch' or perhaps to 'innate knowledge' (Shapiro op cit., p. 140). Neither is satisfactory to MacBride. Perhaps hunches or innate knowledge can be reliable methods for belief formation, but exactly the contents of such beliefs 'match' the reality they are about seems in need of being spelled out. Shapiro is undisturbed. 'The right response', he writes borrowing Crispin Wright's words, '[...] is not to conclude that the acquisition of genuine warrant is impossible, but rather to insist that it does not require this elusive kind of security' (Wright quoted in Shapiro op. cit., p. 144).

It must be said, to MacBride's credit, that the question of warrant for generality is not a niche one. A similar worry arises in the context of postulating diagrammatic reasoning as a source of mathematical knowledge. Suppose we arrive at the justified belief, via working with a diagram of a right-angled triangle, that Pythagoras's Theorem holds of that triangle. How do we know that the theorem holds of *all* right-angled triangles? The problem is that diagrams, Marcus Giaquinto writes, can 'tempt one to make *unwarranted generalizations*, as one's thinking may too easily depend in an unnoticed way on a feature represented in the diagram that is not common to all members of the class one is thinking about' (Giaquinto 2007, p. 77, my italics).

⁸⁸ One might think, on a similar vein, that the Dedekind-Peano axioms generally can be arrived at by means of an abductive argument, since they are what explains the entire body of arithmetical truths.

Yet the problem as it arises in the MacBride-Shapiro exchange is not whether but *how* novices might be ‘tempted’, in the first place, to make unwarranted generalizations. This is not just a question about ‘infer[ring] the general from the particular’, which phrasing could suggest the problem of induction; rather, the problematic generality here involves, crucially, an infinite domain. In this paper I will tackle the problem as it arises in this exchange. Embedding my discussion in the MacBride-Shapiro debate is useful because both MacBride and Shapiro accept that one way of having epistemic warrant to form the target general belief is by doing it through a less controversial method than a hunch: sense experience. If we are able to assuage MacBride’s concerns, perhaps a similar line of thought can illuminate other cases of seemingly unwarranted generalizations, such as the case Giaquinto brings out about diagrammatic learning.

Touching, as it does, on disparate questions in philosophy of mind and epistemology, the dialectic that follows will be elaborate, so I highlight from the start the key idea behind everything. In this paper I want to provide something of an empirical counterpart to Dedekind’s argument that awareness of infinity is possible. To do that, I’ll rely chiefly on two ideas: that a concrete sequence can be potentially, if not actually, infinite, and that such a potentially infinite sequence can be experienced in virtue of perceptual experience’s ability to present, even at an instant, an extended – e.g., sequential – phenomenon. My proposal will thus implicitly channel Dedekind but also, more explicitly, Aristotle, for whom infinity was to be understood in terms of modality and temporality, the two notions underpinning my proposal. The upcoming case being one of perceptually acquired knowledge of infinity, it will constitute an answer to MacBride’s challenge, the one Shapiro can’t (won’t?) provide. The gap is, again, in clarifying how a mathematical novice can be justified in holding a belief that quantifies over all the elements of an infinite set. By

the end of the paper, I hope to have convinced the reader a mathematical novice can be *perceptually* justified in holding such a belief.

The paper is structured around a particular test case, and it divides into six sections. In the first section, I elaborate on the discussion within which I'll frame the paper, and argue that knowledge of infinity is at its centre. In the second section, I introduce our test case. In the third section, I introduce the potential vs. actual infinity distinction, which, as I will suggest in the fourth section, will be helpful in understanding the epistemic state of the subject of our test case. In the fifth section I will spell out the proposal that our subject's epistemic achievement was to acquire perceptual knowledge of the axioms of arithmetic. Having at that point drawn the lessons wanted from this test case, in the sixth and final section I'll show how a Wittgensteinian worry that threatens an imagination-based epistemology may be overcome by my perception-based proposal.

1. Knowledge of small and infinite structures

Let me start by showing, by way of background, that at the heart of the contemporary debate I've presented is indeed the issue that was troubling Dedekind: the issue of knowledge of infinity. This is a piece of knowledge any epistemology of mathematics should want to grapple with, but in Shapiro's, it seems to have less weight than that of small, finite numbers.

Shapiro proposes, in short, a 'stratified epistemology' of mathematics. In Shapiro's view, the subject matter of mathematics – what his epistemology tells us how we come to know – is structures. Structures 'exist objectively, independently of the community of mathematicians and scientists, their minds, languages, forms of life, etc.' (Shapiro *op cit.*, p. 130). Because structures are abstract objects consisting

in places that stand in certain relations to each other, the referents of singular terms in mathematical sentences are the places that, together with the relations that hold between them, constitute those structures. The subject matter of arithmetic, for example, is the structure of the natural numbers, and natural number terms refer to places in that structure, which could nevertheless be instantiated not only by the system of the natural number set and the successor relation but ‘any countably infinite system of objects that has a certain successor relation obeying certain principles’ (ibid).

So, for Shapiro, mathematical knowledge is knowledge about the abstract entities structures are. The question of how we acquire knowledge of *that* is answered by him in steps.⁸⁹ The first step is a mental process he calls ‘pattern recognition’: ‘[a] subject observes one or more systems of objects arranged in various ways, and abstracts a pattern, or structure, from the systems’ (ibid., p. 136).⁹⁰ You observe a pair of apples and a pair of pears, or trios of them, and recognise a pattern common to the pairs or trios; thereby, you acquire knowledge of the structures consisting in the cardinal numbers 2 and 4.

The next step is the process Shapiro calls ‘projection’:

The subject mentally arranges the first few cardinality structures, say, and realizes that they themselves exhibit a pattern. Each such pattern seems to be extendable to a larger one, by adding a place. The subject then projects this pattern of patterns far beyond those hitherto encountered via simple pattern

⁸⁹ Here we are just concerned with (MacBride’s criticism of) the first two for reasons to be expounded shortly.

⁹⁰ A system is a set of objects and relations that hold between them which is capable of exemplifying a structure. This can be put formally as an ordered pair $\langle A, R \rangle$, where A is a set and R is the relation that holds between A ’s members, which in the case under discussion, the set of natural numbers, if such entities exist, is the succession relation.

recognition. This yields knowledge of large finite structures, such as the cardinal-9,422 structure, and eventually knowledge of the natural number structure itself (ibid.).

Projection, then, is meant to allow the formation of, first, a singular thought about some number that pattern recognition has previously given us knowledge of: the belief that *it* is extendable ‘by the addition of a next longest pattern’ (MacBride op. cit., p. 159), that is, the belief that *it* has a successor; and, second, the formation of a general thought about all such numbers: that *they all* have their distinct successor.⁹¹ So MacBride’s challenge is, to repeat, that it is not clear how a mathematical novice can ‘pass’ to the second thought from having the first, how she will ‘transit’ from the first, particular kind of epistemic state to the second, general kind (p. 160). But his scepticism is actually more pervasive: already at the stage of pattern recognition he thinks that ‘an illuminating philosophical description of the process remains to be given’ (ibid.). Even granting that, however (he says), the problem with Shapiro’s account of projection is that someone who entertained the general thought ‘all natural numbers have a successor’ from having the singular thoughts ‘2 has a successor’, ‘3 has a successor’, ..., would be already exhibiting ‘knowledge of the principles which are employed by mathematicians to infer the general from the particular’ (ibid.). Indeed, ‘[i]n the absence of any grasp of these principles, there can be no assurance that the features displayed by a given finite structure are representative of the features characteristic of the infinite structure of which it is an initial fragment’ (ibid.). Some such principles – principles that allow one to infer the general from the particular – are the Dedekind-Peano axioms of

⁹¹ The generality here ranges over the class of cardinals to which projection-known cardinals belong, that is, finite cardinals.

arithmetic: statements describing the infinite structure of the natural numbers. At the heart of the matter, then, I suggest, is the problem of knowledge of infinity.

The challenge MacBride raises for Shapiro is twofold: although it starts by asking how you know the structures denoted by the numerals '2' and '3' from seeing pairs and trios of things ('please offer an illuminating philosophical account of that'), its punch is to ask, even if that query is satisfied, how you know the *infinite* structure denoted by the numeral ' \aleph_0 ' from knowing the finite structures denoted by '1', '2', '3', etc. As noted earlier, Shapiro will be satisfied with his own answer to this whereas MacBride won't be. Shapiro's answer is basically to concede that the subject is not exactly justified in forming the general thought that all natural numbers have a successor but to say that it doesn't matter. Whether the confidence with which the novice forms such a belief comes from a hunch or from innate knowledge, Shapiro is at any rate prepared to call this belief whatever MacBride likes if he (MacBride) doesn't accept it as knowledge. 'The critic can call [it] a hypothesis (for what that is worth)' (ibid., p. 142). For Shapiro, the justification of this general thought does not come from grounding it on something else but from 'recogniz[ing] its role in our intellectual enterprise' (ibid.).

So here we reach an impasse. Indeed, even an account that might satisfy what I called the first, less pressing aspect of MacBride's access problem, the demand for an illuminating account of how we acquire knowledge of finite structures in the first place, would fail to satisfy his second worry. But let us pause and oblige him anyway. Such an account is provided by Marcus Giaquinto (2001, 2007, 2012). Giaquinto basically develops what Shapiro calls pattern recognition as a general method to acquire knowledge of universals, but he calls it, instead, 'abstraction'. Giaquinto takes up Russell's distinction between knowledge by description and knowledge by acquaintance, which terms he uses interchangeably with knowledge via theory and

knowledge via experience, but amends Russell's characterisation of the two to claim that one can count as knowing *something* via experience not only when being presented with *it* directly, as when *it* is a concrete object, but also when being presented with an *instance of it*, as when *it* is a universal (2007, p. 215; 2012, p. 503). So one counts as knowing whiteness via experience from being presented with a white horse and with white sugar and from having then performed the process of abstraction, the 'mental elimination of irrelevant properties when thinking of a body [or plurality of bodies]' (2012, p. 501)⁹² which yields direct awareness of just the salient property they share: in this case, the colour white.⁹³ When it comes to the universals that structures are, one counts as knowing one via experience if one has directly perceived two systems instantiating it (say, again, a pair of apples and a pair of pears) and if then one has performed the process of abstraction that yields direct awareness of the property common to the two systems – in this case, the structure they share.

If we may linger on this a little: drawing on Giaquinto's detailed account of the notions of structure and of isomorphism, the phenomenon of two systems sharing a structure (2007, pp. 214-229), we might say that the process of abstraction achieves, in formal terms, the detection of an order-preserving correlation between two sets. A set here is just the collection of a system's elements, e.g. the two apples or

⁹² Let the term 'perform' here not mislead the reader. For Giaquinto (as for Russell, his inspiration), abstraction is not a mental act but a mental process as involuntary as sense perception itself (Giaquinto 2012, p. 501).

⁹³ Geach (1957) objects to something like this account by complaining that it assumes what it claims to explain. If the property in question is salient to the subject, has she not already grasped it? I will not be concerned with this worry because, as Marshall (2017) notes, for the case of numbers Giaquinto can appeal to a particular capacity for perceptually knowing cardinalities: the famous 'number sense'. Marshall himself thinks this still doesn't help Giaquinto because the epistemic bar he's set for himself (full acquaintance with numbers) is too high; but he presents an alternative that would serve Shapiro's needs just as well (op cit., pp. 53-5). The point here is that MacBride's demand for a clarification on how perception can yield knowledge of small finite cardinalities (the first step in Shapiro's stratified epistemology) can be met.

the two pears. A *structured* set, more specifically, is a set considered under specific relations, functions or constants. The set of natural numbers, for example, is structured under the successor function and contains the constant zero. Two structured sets share their structure, or are isomorphic, whenever there is an order-preserving correlation between one structured set and the other.^{94,95} This yields an identity: if the condition is met, the structure of the first set *is* the structure of the other. The structure of the set of natural numbers considered under the successor function, for example, is identical with the structure of the set of finite von Neumann ordinals considered under the membership relation. (Call any system that instantiates that structure an ω -sequence.) Now, the sets of natural numbers and of von Neumann ordinals are abstract, but of course concrete collections can form structured sets and be isomorphic too. The structure of the set consisting of a cell and two generations of cells formed from the initial one via mitosis considered under the ‘x is a parent of y’ relation can be identical with the structure of the set consisting in

⁹⁴ For a formal definition of ‘order-preserving correlation’ please see Giaquinto (2007), pp. 215-16. The details are ignored for present purposes.

⁹⁵ Here I am employing Giaquinto’s view on identity of structures, which might be too stringent for Shapiro’s purposes. Under the stated condition, the structure of the natural numbers with zero, successor, addition and multiplication is distinct from the structure of the natural numbers with zero, successor, addition, multiplication and exponentiation, since this is a new relation. So which of them does the numeral ‘ \aleph_0 ’ refer to? Here is where Shapiro’s third step in his stratified epistemology – the stage of ‘description’ – becomes relevant. What is needed for determinate reference – for referring to at most one structure, regardless of whether the systems that instantiate it can be multiple – comes in only at this stage; it is the notion of categoricity. A theory is categorical when all of its models are isomorphic. Now, this third step is beyond the scope of the worry we are focusing on here, which is the worry of awareness of general principles, not of whether the theory constituted by those principles is categorical. It must be said, however, that MacBride’s worries about this third stage circle back to the issue of unwarranted generalisations: for Shapiro, the notion of a theory’s being categorical is ‘best explicated, if not defined, in set-theoretic terms’, and set theory is as vulnerable as arithmetic to the problem of how one is justified in generalizing: it includes the axiom of infinity, after all. This is why, in MacBride’s view, ‘Shapiro’s epistemic strategy is guilty of vicious circularity’ (MacBride op cit., p. 162). But perhaps what I say is relevant to the possibility of being justified in generalising in set theory, too.

the nodes of a hand-drawn diagram considered under the ‘there is a hand-drawn arrow from x to y ’ relation.⁹⁶ It is systems like these, concrete and observable, that allow for Giaquinto’s proposed process of abstraction. When a subject detects an order-preserving correlation between the sets consisting in the elements of two systems she perceives considered under their respective relations, the subject is effectively *abstracting* (‘extracting, extrapolating, teasing out’⁹⁷) from those two concrete entities – those two collections – a third: the abstract entity – the structure – instantiated by them.

Something like this, incidentally, is what authors in the literature on abstractionism in the philosophy of mathematics have suggested ‘principles of abstraction’ state: the identity between two *abstracta* given a certain condition. Borrowing Ebert and Rossberg’s (2016, pp. 3-4) way of putting it: ‘[t]he general form of an abstraction principle can be symbolized like this:

$$\S\alpha=\S\beta \leftrightarrow \alpha\sim\beta$$

Here “ \S ” is a [singular-]term-forming operator applicable to expressions of the type of α and β , and \sim is an equivalence relation on entities denoted by expressions of that type. [...] The *abstracta* denoted by the terms featuring in the identity statement on the left are taken to be introduced, in some sense, by the abstraction principle, giving the equivalence on the right-hand side conceptual priority over them’. Following this schema let me suggest, then, that when a subject performs Giaquinto’s abstraction to gain knowledge of a structure, she conforms to something like the following principle:

⁹⁶ Both examples are Giaquinto’s (ibid).

⁹⁷ These three phrasings are Salmon’s (2018) way to introduce abstraction principles.

The structure of the system	if and	there is a bijection c between sets
$\langle A, R \rangle$ is the structure of the	only if	A and B such that for all x, y in
system $\langle B, S \rangle$		A , xRy iff $c(x)Sc(y)$

where the *abstracta* on the left-hand side, the structures, are ‘introduced’ in the content of the subject’s mental state, as it were, whenever the subject recognises an order-preserving correlation between the members and relations of two sets that she’s observed and which must hence be in the content of her mental state for the abstraction process to take place.⁹⁸ Putting abstraction in these terms shows why in Giaquinto’s account only *finite* structures can be known via experience: because

⁹⁸ Precisely because, as I’ve phrased it, this principle mirrors Gottlob Frege’s (1884/1953, §64) original example (viz. the direction of line A is the direction of line B iff A and B are parallel), it is not meant as a definition of ‘structure’ but simply as the specification of the semantic contents of whole sentences where that term occurs. Now that is also called ‘contextual’ or ‘implicit’ definition in the literature, and is thought to be faulty qua (reference-fixing) definition due to its liability to what people have called ‘the Julius Caesar problem’. This problem, in our case, would be the issue that even if our principle successfully specifies that the content of a sentence about structures, e.g. ‘the structure of the set of natural numbers is Julius Caesar’, comes down to ‘there is an order-preserving correlation between the set of natural numbers and Julius Caesar’, it (our principle) fails to determine whether that is true—and, in failing to do so, it fails to determine the references of the terms for [structure]’ (Dummett 1991, pp. 156-7). Here we won’t be concerned with that problem because we’re not interested in fixing the reference of abstract object terms or picking out their identity conditions but simply in making clear how an abstract entity’s (a structure’s, in this case) featuring in the content of a subject’s state can arise from other two entities’ (two sets’, in this case) featuring in it, which is exactly what our principle, if I got it right, helps to specify. As Frege himself remarks of his own example: even though we don’t get a ‘demarcated concept of direction’ from this, ‘we have in our definition the means to recognize this object [the direction of A] when it should occur in another guise as the direction of B ’ (ibid. §66, my emphasis). On a related note, though, it helps perhaps to say that just as one can define ‘parallel lines’ independently so as to avoid that Frege’s principle be taken, contrary to his intention, as a definition – even a contextual definition – of ‘parallel’ in terms of ‘direction’ (cf. Salmon op. cit., pp. 1637-8), so too we find in Giaquinto an independent definition of ‘order-preserving correlation’ that allows us to translate his view on structure into a principle of this form. Although, again, I do not intend it as a definition, I do intend, like Frege, to put the priority (which contents determine or introduce which) on the right-hand side, which is why I make this second clarification (and why I won’t propose an analogue of Frege’s strategy of identifying abstracta with equivalence classes in view of the Julius Caesar problem to indeed provide, after all, a reference-fixing definition).

knowledge of a structure arises only from prior knowledge of the systems whose elements form the sets on the right-hand side of the principle, and perceptual knowledge of a system is only possible if the system is finite – and finitely *small*, at that. As Giaquinto writes: ‘most structures, even most finite structures, are too big and too complicated for visual cognition’ (2007, p. 236). It is the same point Shapiro makes when he says of his label for the same process that ‘[a]t most, pattern recognition accounts for knowledge of small finite structures’ (2011, p. 136).⁹⁹ Now, Giaquinto and Shapiro offer different accounts of the way we *can* know a bigger structure from knowing two isomorphic sets. Shapiro’s way, as mentioned, is projection, where we might not *know* in MacBride’s warranted sense but suspect or have a hunch or otherwise form beliefs about bigger structures from knowing smaller ones, which eventually leads to forming a belief about the infinite structure of the natural numbers. Shapiro mentions in this same vein Parsons’ (2007) account,¹⁰⁰ which we won’t discuss (yet) but is akin to Shapiro’s in that the step from knowledge of a finite cardinality to knowledge of the infinite is justified by a purported faculty of intuition – and one suspects that, in MacBride’s book, Parsons’ intuition wouldn’t be any more secure than Shapiro’s hunch. By contrast, for Giaquinto, if the knowledge of sets is not perceptual, then it will be theoretical, in which case we’ll know the infinite structures abstracted from those sets via theory or by description as well. Thus, we might know the infinite structure corresponding to the cardinal \aleph_0 by the definite description ‘the structure common to all models of Σ ’, if we let Σ be a set of sentences constituting a version of the Dedekind-Peano axioms for arithmetic (Giaquinto 2007, pp. 215-6).

⁹⁹ The possibility of seeing a small but infinitely divisible object is not discussed, as infinity is understood in non-modal terms by these authors.

¹⁰⁰ See especially ch. 5-6, and for a more recent defence, Jeshion (2014).

Now, one clarification is in order before moving on. Giaquinto is not very explicit about this (indeed, he acknowledges the point only in footnotes (cf. 2007 pp. 237-8, fn. 13 and 14), but by ‘the set of Dedekind-Peano axioms’ he means ‘the set of *second-order* Dedekind-Peano axioms’. He seems to regard this as the natural way of viewing knowledge of them. ‘Of course certain second-order assumptions are built in to the underlying conceptions of space, time, and motion here’, he writes about assuming a second-order version of the fifth axiom, the axiom of induction, in his discussion of knowledge of structure, ‘but these are our natural conceptions’ (ibid.). Whether this is so or not (whether it is ‘natural’ to assume the second-order version of the axioms), Giaquinto’s motivation to do so is that his explanandum here is knowledge, specifically, of the natural number structure. This requires the second-order axioms because the first-order ones are not categorical: they are true of structures other than the natural number structure as well as of it. Indeed, Giaquinto acknowledges that the knowledge he is targeting, ‘knowledge [of the structure of models of such-&-such axioms,] requires knowing that the axiom set is categorical, meaning that all its models are isomorphic’ (op cit. pp. 228 and 238). I will join Giaquinto in meaning second-order whenever I talk about the Dedekind-Peano axioms, and for similar reasons. I have advertised the subject of this paper to be knowledge of infinity, but I’ve also shown that one rare attempt to make sense of it, Dedekind’s theorem, took as its target object a particular kind of infinite set, an ω -sequence, the kind Giaquinto’s discussion takes as target object of knowledge, too. So, although categoricity isn’t needed for infinity, I will aim for these authors’ specific kind of structure as well (which motivates, as will become clear, Section 5.2). Perhaps our ‘natural conceptions’, as Giaquinto says, of space, time and such make it so that the natural structure to think of when thinking of infinity is an ω -sequence,

as he and Dedekind seem to have thought. If so, it would be nice if perception turned out to yield knowledge of exactly that structure.

So, to sum up. I have suggested that at the heart of MacBride's challenge is the issue of knowledge of infinity. MacBride demanded first an illuminating philosophical account of how we know small, finite structures; I have appealed to Giaquinto's views for that. But the more pressing demand was to explain how we know infinite structures, in particular the structure of the natural numbers, which we count as knowing when we know that all naturals have a successor. The answer that we know this structure in the same way we know any structure that falls outside the scope of acquaintance, i.e. via theory, would fail to satisfy MacBride, however, who wants an answer that works for the mathematical novice – so, of course, assumes no theory. Of the two epistemic methods proposed by Giaquinto only acquaintance would have satisfied MacBride, but again, we saw that knowledge of structures by acquaintance only applies to finite, small structures. Hence, if there is one thing all three authors seem to agree on, it is that, as Shapiro says (quoting MacBride's channelling of Hume), 'experience provides no corresponding impression from which the idea of the infinite may be derived' (Shapiro 2011, p. 142).¹⁰¹

¹⁰¹ From the way I have phrased the problem in these last few lines, it may not be clear why larger or more sophisticated structures than those corresponding to the first small cardinals do not pose equally serious epistemological problems to infinite structures. The first thing to say is that in this paper we're not examining the epistemological problem of infinite structures per se but the problem of warrant for jumping from beliefs about the first small natural numbers to a belief about all of them, which is helpfully framed by the twofold challenge MacBride poses to Shapiro, and the first step of which I have argued to be illuminated by Giaquinto's view. Another thing to say is that Shapiro will presumably want to account for knowledge of larger or more sophisticated structures, infinite or not, by appeal to categoricity (that is, by appeal to the third step in his stratified epistemology). As noted already, MacBride's worry here would circle back to the problem of warrant for generalisation, because the axioms of set theory needed by Shapiro to 'explicate, if not define', the notion of categoricity are not free from the epistemological worry about generality. So, if we focus on the generality problem by examining the case of how a mathematical novice can in fact extract or extrapolate the infinite from what is given in experience, then the epistemic warrant we might attribute to this novice could trickle down

2. An encounter with infinity

This brings us to the main thesis of this paper, which is that, contrary to what the philosophers above suppose, knowledge of infinity by acquaintance is indeed possible in something like Giaquinto's sense. Let me present, without further ado, the episode that will constitute my case study. It is drawn from an anecdote by science writer Sarah Scoles (2016):

When I first encountered infinity, I was a four-year-old in the Latter-day Saints temple in Atlanta, Georgia. My parents, my sister and I were [there for] a ceremony meant to unite us as a family forever. I didn't really understand what 'forever' meant. [...] My sister and I knelt on the floor with our elbows on a pedestal, and my parents arranged themselves the same way across from us. Behind them was a mirror, and behind us was a mirror. The two surfaces reflected each other's images back and forth, creating infinite reproductions of our family together. I looked into the images and was amazed to see that they never ended; they just got smaller and smaller. At some point I could no longer discern the individual reflections, but I intuitively grasped that I was merely running into the limits of my vision. The reflections kept going and going. 'Oh,' I thought, with a chill of understanding. 'Forever.'

What we have in the passage is not, of course, a theoretical description of a phenomenon, let alone an explanation of it. What we have is a subject's first-person

to those theories – about finite but complex structures, or infinite and bigger than the natural number structure – she will later come to know in the third stage.

report of an experience she underwent as a younger self. She's telling us that this experience gave her understanding of something she hadn't understood theoretically before, the temporal variety of infinity ('forever'), so her anecdote involves also what seems the acquisition of perceptual knowledge, or at least the epistemic upshot of that perceptual experience. My suggestion will be that young Sarah's perceptual experience gave her justification to make a generalisation of the kind mathematical novices seem typically unequipped to make. But let us start by examining the phenomenology of the experience in some detail.

Suppose we begin the inspection of Sarah's experience after she's been brought to the centre of the mirror arrangement. Suppose her eyes are closed in that process (her dad is carrying her there; they tell her to close her eyes and expect a surprise). When she finally opens them, she is suddenly confronted with a system of what seem to her to be endless reflections, as roughly illustrated below by Figure 2. Call this point in time t_0 .

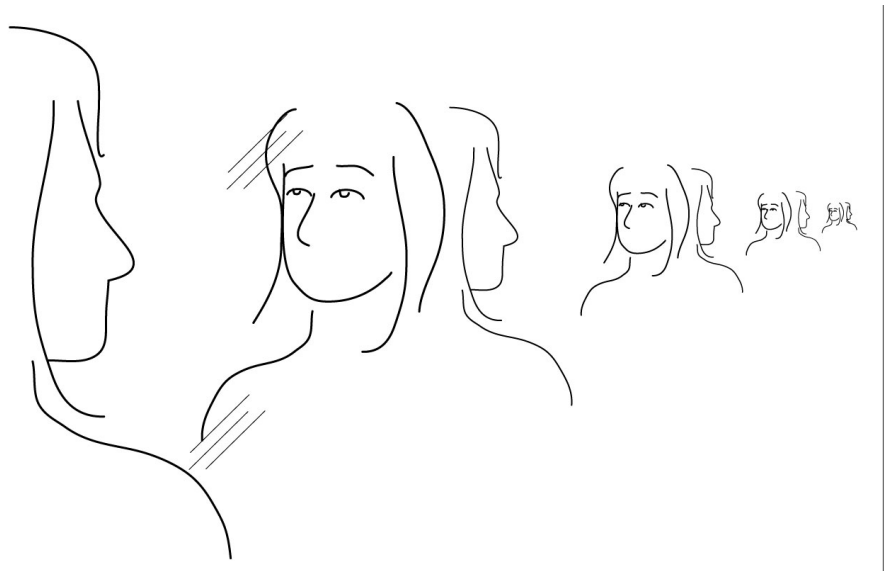


Figure 2.

Astounded, Sarah takes a second to absorb what she is seeing. But then astonishment gives way to curiosity and she moves a little to the right. Call that point t_1 . Because mirrors present a window-like behaviour in visual experience (Mac Cumhaill 2011), Sarah's experience has a specular experience-like feature, too: 'elasticity'. This means that, exactly as it would happen were she seeing through a window, her experience exhibits 'characteristic patterns of expansion and contraction that cue awareness [*as*] of the presence of objects and the empty space that is "outside"' (Mac Cumhaill *op. cit.*, p. 492). So what she sees stretches and contracts a little in her visual field just like empty space and present objects would on the other side of a window. To get a clearer sense of this, contrast seeing a landscape through a window and seeing one in a picture. Where in the former case the space between the landscape's elements – say, the trees – is perceived to be continuous with the space that separates you, the viewer, from what you see, in the former case it is not. (In Figures 3 and 4, the spaces a and b are continuous in the first case but not the second.) So, in the experience of seeing through a window, the space seen 'in' the window behaves just as the space around you. Because specular experience replicates window experience, the space seen 'in' a mirror behaves just like this, too.

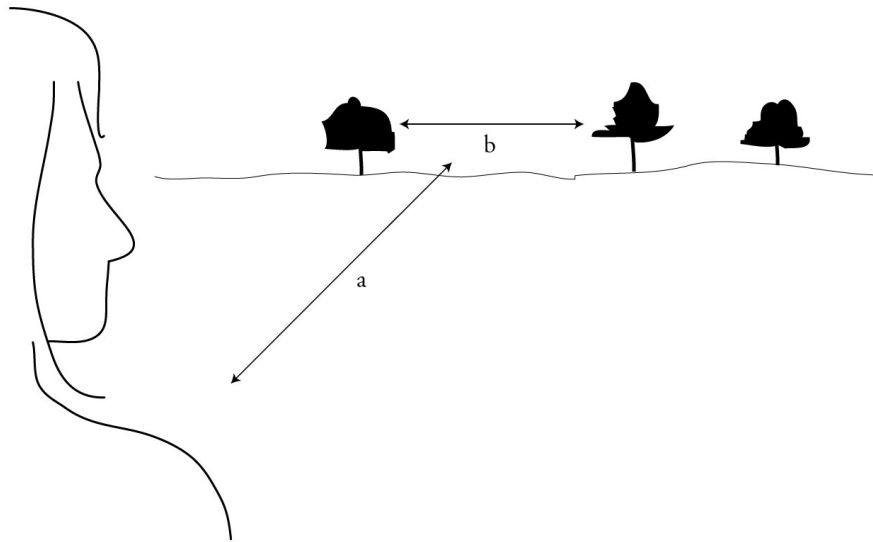


Figure 3.

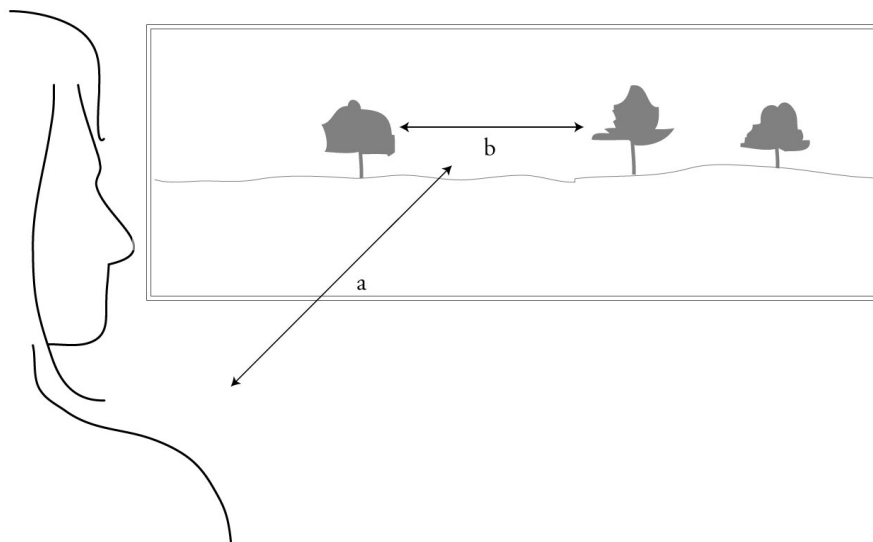


Figure 4.

Now, a related feature of the experience of seeing through a window is that when we move to either side, objects that were previously occluded come into sight.

Likewise, the system of n reflections becomes for Sarah a system of $n+1$ at t_1 . Then, when she moves further still, it becomes one of $n+2$ at t_2 , then $n+3$ at t_3 , $n+4$ at t_4 , and so on (Figure 5).¹⁰²



Figure 5.

Perhaps she reaches finally an awkward position at which the angles of the mirrors don't let her see more reflections because then she'd start to see behind the mirror, or because she'd bump into someone outside the centre of the arrangement, or because the reflections grow fuzzier. At any rate, she understands this makes seeing one more reflection impossible as a matter of fact rather than as a matter of principle. If the platform where she stands were rotating, or if the mirrors rotated with her as she moves left or right, or if the optical properties of the system were different, she

¹⁰² A feel for the elasticity phenomenon and the coming of more reflections into sight can be found in infinity mirror videos, for example, here: <https://youtu.be/3Ll4IbdVhOQ>

would continue to see one more reflection, over and over again, forming a circle with the rotating arrangement of mirrors – ‘forever’.

What is happening here? One striking feature of the experience is that, unlike in Giaquinto’s and Shapiro’s cases of acquaintance with small, finite structures, the subject is not actually being presented with the cardinality she seems to think she is – she is not, in other words, *actually* being presented with infinitely many reflections. It is a physical fact – a fact of physical optics – that there is only a finite number of reflections visible in a mirror that is placed in front of another mirror (Gbur 2011). Yet this is what it seems like to her. Even as her adult, science-writer self, the subject describes the experience in terms of there actually being an infinity of reflections: ‘I looked into the images and was amazed to see that they never ended; they just got smaller and smaller’. One wonders if the writer indeed knows the physical fact just mentioned (nothing in the piece, which is about the power of ‘mathematical experiences’, as she calls them, over discursive media, suggests she does know it). The thought naturally arises that there is something like an illusion of actual infinity occurring, then. At the same time, the temporal vocabulary of the report (the reflections ‘never ended’, ‘kept going and going’) suggests a belief that the infinity being presented was not just spatial – a falsity – but also connected to time.

I will put my interpretation of the case along these two dimensions. On the one hand, the mirror set-up presented young Sarah with what seemed to her like an actual infinity of reflections. This was an illusion because there was no actual spatial infinity there. On the other hand, there seemed to be a grasp that time had something to do with the number or reflections being infinite. There seemed to be a belief, in other words, that the number of reflections extended not just in space, as per the illusion, but also in time.

Let us work with the hypothesis that the illusion of actual infinity enabled Sarah to grasp just that: actual infinity, even if what she saw was not actually infinite but, as I'll shortly suggest, potentially so. I will defend this hypothesis by way of defending three ideas. First, I will spell out the suggestion that even if the system of reflections cannot be characterised as actually infinite, it can be characterised as potentially infinite (Section 3). Secondly, drawing on the Aristotelian connection between time and modality, I will suggest that the temporal vocabulary in Sarah's report expresses a grasp of said potential infinity; specifically, I will ascribe to her a belief of the form, 'for every reflection, there can always be a distinct successor' (Section 4). These two first ideas should do justice to the fact that what even though Sarah's experience was not of actual infinity, there was a sense – a temporal/modal sense – in which she really did experience infinity. Thirdly, I will discuss how the experience enabled our subject to grasp the structure of the natural numbers, as described, as we've seen, by the Dedekind-Peano axioms, which involve no modal vocabulary. The key will be to remind ourselves that the target kind of knowledge, the knowledge that MacBride was anxious to understand possession of in the case of mathematical novices, is general or universal knowledge, that which quantifies over all the members of an infinite set. Thus, in Section 5 I will argue that the experience enabled Sarah to conceive of a model for the Dedekind-Peano axioms, which ability is what typically justifies subjects in holding a general or universal belief about all the elements of an infinite set.

3. Potential vs. actual infinity

Ever since Cantor's pioneering work in the formal study of infinity (1883, 1887), the concept of infinite size has come with the embedded notion of *completed* size.

Mathematically, after all, a set has its cardinality as a fixed property. But this is a relatively recent development. Apart from a few exceptions,¹⁰³ before Cantor, and ever since Aristotle, the prevalent notion of infinity came with the embedded notion of potentiality rather than of actuality. We might put this in slightly anachronistic jargon by saying that for a number to be infinite was for it to correspond to the size of a set with potentially infinite members rather than with a fixed size.¹⁰⁴ Thus, Aristotle wrote, ‘generally the infinite is as follows: there is always another and another to be taken’ (*Physics* 3.6, 206a27- 29).¹⁰⁵ This characterisation contrasts with today’s Cantorian concept of the infinite, which does not rely on temporal vocabulary such as Aristotle’s ‘always’, or on dispositional vocabulary, such as Aristotle’s ‘to be taken’. For Cantor, an infinite set being infinite simpliciter means it is actually infinite. This part of why Giaquinto says ‘we cannot experience an entire instance of the natural number structure’: not only is actual infinity as a matter of fact uninstantiated by concrete systems such as the reflections in the mirrors; our perceptual capacities are, in any case, finite.

The thought that actually infinite structures simply can’t be instantiated by finite concrete systems – let alone be objects of perception – is somewhat Platonic in spirit. In his rejection of Plato’s view that mathematical objects are independent from empirical reality, and in his view that infinity exists only potentially, never actually, Aristotle effectively allowed our beliefs about the infinite to be empirically justified

¹⁰³ Duns Scotus and Gregory of Rimini are just two examples of a few actualists on the margins. See Moore (2001), pp. 45-55.

¹⁰⁴ This is anachronistic because, as Moore (op. cit., pp. 36-7) points out, Aristotle endorsed only the idea of infinite magnitude, not of number, but I’m applying his process-based reasoning to our discussion of infinite number for purposes of the dialectic. Also, pre-Cantor, people didn’t think in terms of sets.

¹⁰⁵ I remain neutral on the question whether Aristotle located potential infinity in the process, as Hintikka (1966) suggests, or in the structure of the magnitude to which the process merely bears witness, as Lear (1979) suggests. I hope everything I say is compatible with either exegesis.

by the sorts of things we can, in fact, do: bisecting and extending lines, he thought, are in principle never-ending processes. Now, both an Aristotelian mathematician in the process of extending a line and Sarah are, unlike mathematical abstracta, concrete beings embedded in time. What they are experiencing and whatever beliefs arise from that experience must be understood in dynamic rather than in static terms. Aristotle could not have come up with his view had he existed only at some point in time rather than experiencing, in a temporally extended fashion, indefinitely extensible processes. Picture him by a river in Athens, brooding over Platonic forms. ‘Sure, because the world is finite, the amount of water in this river is finite too’, we find him thinking. ‘But there is nothing in principle to stop this finite amount of water from running the course of the river over and over again. And if I were immortal, I would be free to sit here and watch it do that *forever*’.

As Adrian Moore points out, the connection of Aristotle’s actual/potential infinity distinction with time, and hence with lived experience, was not fortuitous:

For Aristotle, the infinite was the untraversable. But traversal takes time. So there is no making sense of the claim that something is untraversable save with respect to the whole of time (Moore 2001, p. 40).

This contrasts with processes the length of which you can in fact ‘traverse’, such as, contrary to Zeno, the process of Achilles reaching the finish line in the race against the tortoise. For the river’s running of its course to be infinite is thus for one not to be able to traverse the length of its duration. But notice that this ability is to be understood as a matter of principle. There is nothing in the river’s running of its course itself, or, as Jonathan Lear (1979) puts it, nothing in ‘the structure of [its] magnitude’ (p. 193) that impedes its being untraversable. Only Aristotle’s mortality – and perhaps his capacity for boredom – prevent him from witnessing this going on

forever. Hence, his potential/actual infinity distinction is to be understood *essentially* in temporal terms:

The actual infinite is that whose infinitude exists, or is given, at some point in time. The potential infinite is that whose infinitude exists, or is given, *over* time; it is never wholly present (ibid.).

Belief in the latter is what Aristotle argued was empirically justified and the former what wasn't.¹⁰⁶ My suggestion is that Sarah's experience, while seeming to her, falsely, to be of *actual* infinity, can be regarded to be of genuine infinity in the Aristotelian sense. There is nothing in the structure of the system of reflections itself that puts any limits on the process of getting the next element into sight: reaching a final reflection is impossible, in other words. The system of reflections is, as we might now say, 'untraversable'.

Interpreting our case this way would be of little use if we didn't have a cogent account of the notion of potential infinity, but recent work by Shapiro himself and Øystein Linnebo (2017) offers one such account. The details are intricate, but for our purposes it will be enough to put their suggestion as follows. Contemporary – Cantorian – mathematicians' qualms about the concept of potential – Aristotelian – infinity can be summarised in one complaint, the complaint that “a clear meaning has never been given to” the phrase “x is potentially infinite”.¹⁰⁷ This makes Linnebo and Shapiro's task straightforward: to offer simply an explication – making logical sense or defending the coherence – of potential infinity. This will involve a 'single

¹⁰⁶ Indeed, Moore argues that, for Aristotle, time is not just connected with the potential/actual infinity distinction but modality generally: 'Aristotle believed that questions of possibility and impossibility were themselves intimately connected with time, so that asking whether or not something was possible was akin to asking whether or not it would be so—at some time' (ibid.).

¹⁰⁷ Karl-George Niebergall quoted in Linnebo and Shapiro op. cit., p. 166.

controversial claim' (op. cit., p . 158): that the non-modal language of contemporary mathematics is not fully explicit. According to this view, when one says the natural numbers form a (countable) set that is infinite, one *really* means they form a (countable) set that is potentially infinite.¹⁰⁸ To which the Cantorian mathematician might insist that she doesn't understand what 'potentially infinite' means. Here Linnebo and Shapiro assume that to understand a concept is to know what inferences the concept allows, to know its logical behaviour; hence, their task amounts to telling us what logical operations a potentialist view allows. Do we get to quantify over all natural numbers, for example, if they purportedly constitute a domain of not fixed but modally variant size? Because, again, the potentialist's position is not that their view about the infinity of the natural numbers is correct whereas a Cantorian view is not but instead that a modal expression of infinity is merely more explicit, the inferences the potentialist's proposed view will allow will have to match the ordinary mathematician's. Here is where the substance of Linnebo and Shapiro's work lies: in their proof that the entailment relations that obtain in a potentialist language correspond to the entailment relations that obtain in ordinary mathematics when translated into its non-modal language.¹⁰⁹

For our purposes, the upshot of their argument is that it entitles us to make sense of the claim that the system of reflections Sarah saw is potentially infinite. Linnebo and Shapiro illustrate their explication of potential infinity with a hypothetical debate on the structure I claim Sarah grasped perceptually, the structure

¹⁰⁸ From here on, I'll omit this qualification because I'm only concerned with the smallest countable infinity.

¹⁰⁹ As Linnebo and Shapiro point out, '[t]he only difference between actualism and [...] potentialism [...] detected so far is a philosophical one, concerned with the presence or absence of an implicit modal aspect of ordinary mathematical quantification' (p. 176). Further, more 'technical' differences emerge when the authors extend their account to handle uncountable infinities, e.g. the infinity of real numbers. We will not be concerned with those here, but see their op cit., §§7-8.

of the natural numbers. The actualist character is played by an adherent of classical infinitary mathematics and the potentialist by a critic, an intuitionist, say. For an actualist, to say that the set of natural numbers is infinite is to say that every natural number has a successor, which she would express by:

$$\forall m \exists n \text{ Succ}(m,n)$$

For the potentialist, to say that the set of natural numbers is infinite is to say that, at any point in time at which we have finitely many natural numbers, we can always generate or construct – there can exist – one more. Which she would express by:

$$\Box \forall m \Diamond \exists n \text{ Succ}(m,n)$$

The potentialist's view, then, is that when the actualist – i.e. the ordinary mathematician – says that ' $\forall m \exists n \text{ Succ}(m,n)$ ', she *really* means that ' $\Box \forall m \Diamond \exists n \text{ Succ}(m,n)$ ', or that *that* is what she would say were she to use a fully explicit form of expression.

So here's an informal gloss on how the claim that the system of reflections is potentially infinite can be made sense of in Linnebo and Shapiro's framework. Let the reflections stand for the mathematical objects of a potentialist theory's model. Because the theory contains modal operators, the model for it contains possible worlds. For the system of reflections to be potentially infinite is for any world w_0 containing n reflections to access another world w_1 containing $n+1$ reflections.¹¹⁰ At

¹¹⁰ This involves, of course, an actual infinity of possible worlds. But what is being explicated here is the potential infinity of mathematical objects. Presumably, what we're employing for such explication, possible worlds semantics (which is done in a meta-language without modal operators), can appeal to an actual infinity. Linnebo and Shapiro anticipate the worry that appeal to possible worlds semantics might seem illegitimate – since, for a potentialist critic, there exist no actual infinities – by clarifying that 'we understand [the contemporary heuristic

w_0 , the n reflections exist actually whereas the extra reflection exists potentially. Then, at w_1 , the process of mathematical construction – which here consists of the process of the subject moving a bit further to get one more reflection into sight – has resulted in that extra reflection being actualised and another still coming into potential existence because now w_1 accesses w_2 , and so on. Because we can generalise with our quantifiers across all accessible possible worlds rather than just over a single world, though, already at w_0 we can make the claim that for every reflection there can always be a succeeding reflection, which is to say that the set is, in effect, potentially infinite.

4. Knowledge of potential infinity

4.1 Modal knowledge and experience

If the suggestion to regard the system of reflections as potentially infinite sounds plausible, now we need a bridge that accounts for the subject's knowledge of that property of it. To say that Sarah perceptually learned that for any reflection there could always be a successor is to make a claim about perception's role in the acquisition of knowledge that is not only general but also (and just as problematically) modal. A word on the epistemology of modality is thus in order.

Following the logic of the last section's last paragraph, I will put the problem as follows. The infinity of the system was said to be explicable by quantification over elements of it that exist in worlds accessible from the world in which we're making

of possible worlds] as only heuristic, as a manner of speaking. Our official theory is formulated in the modal language, with (one or both of) the modal operators as primitive. The modal language is rock bottom, not explained or defined in terms of anything else' (op cit., p. 170).

such a quantification. But according to orthodoxy, what we can know by perception is only what's actual – what exists at *our* world.¹¹¹ Furthermore, knowledge by acquaintance is, as Bertrand Russell wrote, 'the converse of the relation of object and subject which constitutes presentation' (1911, p. 108), so for something to be knowable by perception it must not only exist in our world but also be, paradigmatically, *present* before us. As Crane and French summarise the point, 'it seems that you can only see or hear or touch what is there' (2017, §1.1.2; §3.4).

The problem of modal knowledge in our case concerns both possibility and necessity, however. Our target is knowledge of a necessary possibility. So we need to account both for knowledge of what can be the case but is not and for knowledge of what is not just the case but must be. One way to getting started with this problem is to consider the role of another faculty that can also yield visual imagery, the imagination, in the acquisition of modal knowledge; then ask whether perception plays a role in the imagination's performance of it. One of the most popular recent views on modal epistemology, that which explains modal knowledge in terms of counterfactual knowledge, offers especially ample room for such a twofold – perceptive-imaginative – mechanism on account of its naturalistic spirit, i.e. its resistance to appeal to any cognitive capacities other than those which are 'used in ordinary life, perhaps trained, developed, and systematically applied in various special ways' (Williamson 2007a, p. 136). This is the view pioneered by Bob Stalnaker (1968) and David Lewis (1973), and most recently defended by Tim Williamson (2007a, 2007b, 2016).

To illustrate both how counterfactual knowledge is equivalent to modal knowledge and how it is arrived at in everyday life, Williamson himself has offered a

¹¹¹ Cf. e.g. Chalmers 1996, p. 137; Salmon 2005, p. 254; Hale 2013, p. 252; but the thought may be traceable to Ancient Greece, to Plato's view that only the intellect gave us access to the realm of Forms.

sketch of how perception and the imagination interact to deliver knowledge of counterfactual scenarios in a way that also is intended to persuade us that this is, indeed, the way in which the imagination thrives – by interacting with perception. To develop the point, Williamson starts by drawing our attention to the fact that just like the attention, the imagination can operate in both voluntary and involuntary ways. To illustrate the involuntary mode, he considers a group of early humans who are distracted and might be instantly pulled to imagine the dangers of the forest they're entering when they suddenly perceive the sounds of wolves (Williamson 2016, p. 115). This is a scenario the imagination could conjure up on its own but is *pushed* to by an occurrent perception. In this case, the proposition learned can be represented as a subjunctive conditional: 'if we were to enter the forest, we would stumble upon wolves'. To illustrate the voluntary mode, consider a case where perception *enables* the imagination to conjure up a scenario it wouldn't be able to do otherwise. A hiker might want to imagine jumping a stream that is on a path he wants to take in order to strategise whether to jump it or find an alternative path instead. This imaginative exercise, however, requires actually going to the stream to gather some necessary perceptual knowledge, knowledge about 'the width of the stream, [...] the awkwardness of the place from which he would have to launch himself, the slipperiness of the rocks on which he would have to land, how tired he is, and so on' (Williamson op cit., p. 116).¹¹² The proposition learned in this case might be represented as: 'if I jumped from this slippery spot, then I would fall'.

The above suggests a pattern of perception-imagination interaction in Sarah's case. Because at t_2 , say, Sarah sees more reflections than she saw at t_1 , the knowledge of what happens when she moved from t_1 to t_2 *enables* her to imagine what will happen when she next moves from t_2 to t_3 , namely, that another reflection will come

¹¹² This is a modified version of Williamson's hunter example.

into view. Thus she could, if she wanted, voluntarily represent this scenario to herself at t_2 : ‘if I moved a little further, another reflection would come into view’. But from her report, it appears that such an imagining was not a choice. Much in the way the involuntary mode of the imagination of the early humans was activated by the sounds of wolves, Sarah’s imagination may have been *pushed* at t_2 , by what she experienced during t_1 - t_2 , to represent what would happen during t_2 - t_3 – and so on, and so on. Yet the modal knowledge acquired by Sarah involves something more than possibility. It is not just knowledge that something would happen but that, necessarily, it would, or in the above, Aristotelian way of putting it: ‘ $\Box\forall m\Diamond\exists n$ Succ(m,n)’.

This now goes beyond the scenarios Williamson suggested were delivered by the imagination as pushed/enabled by perception. Both the early humans’ and the hiker’s cases are ones in which perception played an essential role in the acquisition of knowledge of possibilities *primarily*, not of necessities. This point bears elaboration. There is a sense in which learning whether it is possible or not to jump is ‘easy’ knowledge of a necessity: to know that ‘ $\neg\Diamond p$ ’ is to know that ‘ $\Box\neg p$ ’. But to think that this already gives us what we need – an argument, that is, to the effect that perceptual knowledge of necessity is possible – would be a mistake. The mistake is pretty easy to make because it *is* true that knowledge that ‘ $\Box q$ ’ could be ascribed to someone when they know that ‘ $\neg\Diamond p$ ’ (along with other truths: that $\neg\Diamond p$ is equivalent to $\Box\neg p$, for example, and that $\neg p$ is equivalent to $\neg p$).¹¹³ This is so even if the method of learning $\neg\Diamond p$ is perceptual, which some have thought precludes the second, ‘ $\Box q$ ’-knowledge ascription (Dretske 2003, 2005). Perceptually knowing that ‘ $\neg\Diamond p$ ’ can indeed, however, be construed to entail perceptually knowing that ‘ $\Box q$ ’ if, for instance, we construe the same fact that was presented to the subject as carrying the

¹¹³ This is true given certain closure principles. See Hawthorne (2005).

information that $\neg\Diamond p$ and that $\Box q$, or if, alternatively, we construe perceiving that $\neg\Diamond p$ as putting the subject in a position to perceive, indirectly, that $\Box q$.¹¹⁴

Now, perhaps the issue is not the relation between \Diamond and \Box , actually, but the kind of modality in play. When the hiker forms the perceptual belief ‘I can’t jump’, the modality is highly restricted: the modal verb – ‘can’t’ – expresses immediate practical possibility, just like ‘need’ in ‘I need to get up’ after a long nap expresses the corresponding restricted necessity. The kind of modality expressed by the modal operator in ‘ $\Box\forall m\Diamond\exists n \text{ Succ}(m,n)$ ’ is not like this. Both are objective rather than epistemic, yes, but one can construe knowledge of the kind the hiker learns as knowledge about herself, about her own abilities relative to the environment, whereas the knowledge Sarah learns is knowledge about the structure of reflections, which is knowledge about the world – and, ultimately, knowledge about a mathematical structure. It is knowledge about something entirely independent of herself, in other words. The difference between the hiker’s and Sarah’s experience shows up in a difference in their phenomenology. Ruling out a possibility by means of perception, as the hiker does, feels differently from the way it feels to learn, by the same method, that something must be the case. Where the hiker senses an *incompatibility*, I want to suggest, between what he perceives – the river’s width, the slipperiness of the spot – and a scenario in which he has jumped, Sarah senses the positive force of a necessity. The experience gives her the knowledge that necessarily, for any reflection there can be a successor. The hiker’s experience of incompatibility seems to be related to the involvement of negation in a sentence expressing what he’s learned – that ‘I *cannot* jump’. The idea that sentential negation is grounded in the phenomenology of incompatibility has received attention recently (Berto 2015, Berto and Restall 2019,

¹¹⁴ These are the general shapes Luper (2020, §3.2) gives to responses to the view that perceptual knowledge, as every distinct mode of access to knowledge, is not closed but open, as Dretske complained.

Kinkaid 2020), but to see the significance of this here, consider the converse situation from our hiker's. To see that $\Diamond p$ (that one can jump) is, one might think, to see that $\neg\Box\neg p$ (that it is not necessary not to jump, so not necessary to find an alternative route). So one might be tempted to think that seeing that one can jump is also to have an experience of a negated necessity. But the phenomenology of seeing that one can jump differs starkly from the phenomenology of experiencing an incompatibility, as in our unable hiker's case. In the positive case, the phenomenology involves an excitement over the nearby world, the world where one has jumped, coming into sight, the bridge towards it materialising out of thin air as the accessibility relation from one's current world to *that* one. Such a phenomenology emerges from the visualisation of the target positive scenario, I want to suggest, not from the mere awareness of the absence, in the actual world, of obstacles for getting there.

Sarah's experience shares something with this positive phenomenology. Notice her description of a 'chill of understanding'. What we need, if this sounds right, is an account that explains Sarah's perceptual knowledge of a positive necessity, of the necessity of a positive scenario. Now, in our case, that scenario itself involves a possibility, and indeed, knowledge of possibilities has been the primary focus of work exploring whether perception can deliver modal knowledge directly, especially by examining cases where perception tells us about our abilities to interact with the environment, where both our abilities and what makes the environment receptive to them are viewed as dispositions.¹¹⁵

¹¹⁵ Nanay (2011, 2012) argues that dispositional properties of objects can feature in the contents of perception; Strohminger (2015) argues that we can have perceptual knowledge of possibilities but she sees the modal propositions in question as part of the subject's knowledge-that, which may or may not be grounded in perceptual representation. Incidentally, Strohminger's account also allows perceptual knowledge of necessities in something like the hiker's case. To know by perception that one is not able to climb a tree, say, is to know that in no physically possible world one finds oneself up there by one's own means. The point that inferential knowledge is also achievable by perception allows

Let me put what we've learned in this section in terms of constraints. First, we have learned that what we want is stronger than what the literature has considered. Our target is knowledge of a necessity arrived at via an experience with a positive phenomenology, not a negative one, and it is knowledge of a necessity which embeds possibility (which will render extant accounts helpful, after all). Secondly, we have learned that although, as a first pass, it seems plausible to map Sarah's experience onto Williamson's patterns of perceptually-guided imagination, the fact that Sarah's grasp was instantaneous rather than the result of an extended exercise (suppose this, then what would follow?), in addition to the precedent found in the literature that perception can indeed yield some modal knowledge, suggests that Sarah's grasp of potential infinity, if informed by both perception and the imagination, differs from Williamson's cases in that it was instantaneous. (Indeed, Sarah's report suggests that she ascribes her grasp of infinity to perception alone.) So the immediate task is to explain how perception might yield knowledge of a necessary proposition.

4.2 Modal knowledge, time, and experience

To begin to sketch an answer, notice that there's been a conflation of modal and temporal vocabulary in my discussion. In my paraphrase of the proposition ' $\Box \forall m \Diamond \exists n \text{ Succ}(m,n)$ ', 'for any successor, there can always be a successor', 'always' stands for 'necessarily'. This is not a coincidence. I have already suggested viewing

Strohminger's view to deliver less restricted necessities also: e.g. if one can see that two things are distinct and one accepts the Principle of the Necessity of Distinctness, one may also see that it is metaphysically necessary that those two things are distinct (pp. 368-9). Although Strohminger doesn't seem to see a difference between the phenomenology of such a purported experience of metaphysical necessity and one of physical possibility or impossibility, knowledge of possibilities does take centre stage for her, too. More on Strohminger shortly.

Sarah's experience through an Aristotelian lens. Following Aristotle again, I want to suggest now that the temporal profile of Sarah's experience is key to understanding both modal aspects of the target knowledge – the necessity involved and what is known to be necessary, a possibility.

There are roughly three main ways of understanding the claim that experience is temporally extended (Dainton 2018). One is in terms of the causal processes involved in temporal experience. Even as I stand here looking at my screen, light takes time to travel from and around it to my eyes so that I can enter the state of seeing it. Assuming, as Simon Prosser (2016) puts it, that 'causal influence takes time to travel', one can hardly deny experience is temporally extended in this sense. This is not the way of understanding the temporality of experience that interests us. The other two senses are interrelated (or can be). One is the sense in which experience can be regarded as extended, regardless of the causal processes that underlie it, and the other is the sense in which experience seems to present us with temporal extendedness, which may or may not imply that it itself is extended.

Our interest is in the last of these issues: on the ability of perceptual experience to present us with temporally extended phenomena. Part of the puzzling nature of Sarah's case arises from the fact that she seems to have grasped at first glance, instantaneously ("Oh," I thought, with a chill of understanding'), something that is best understood in dynamic than in static terms: the potential infinity – which is to say, the *untraversability*, a property logically connected with time – of the system of reflections. A similar puzzling asymmetry, that between the seeming *instantaneousness* of an experience and the seeming *extendedness* of its content, can be found in the philosophy of temporal experience itself. 'Philosophical puzzlement [about our perceptual experience of temporal properties such as succession, persistence and change] arises as a consequence of assuming that, if one experiences

succession or temporal structure at all, then one experiences it at a moment', Ian Phillips writes (2008, p. 176). The puzzle is dissolved if that assumption is shown to be false. Phillips' argument to that effect would take us too far afield, but its thrust can be felt intuitively. Consider and contrast two auditory experiences, the experience of hearing a staccato G played on a piano in the context of a sequence (C major, broken triad played staccato) and in isolation. The following two pieces of score represent the two scenarios:



Figure 6 (from Phillips op cit., p. 178).

Both are cases of hearing G at time t_2 . But in the first case and not the second, that note is experienced as following from – as being part of a succession of – more notes. ‘One would be failing to characterise fully how things were for you at $[t_2]$ if one only mentioned the fact that one was hearing a G note’ (Phillips op cit., p. 178). I claimed earlier the system of reflections Sarah sees bears a property that should be understood in Aristotelian terms, in terms of its being always possible for there to be a successor for any reflection. This property was suggested to be explicable as its being possible to generalise, from the world of evaluation, over the objects existing in worlds accessed from *that* world. Thus, if such a generalisation was what Sarah was able to make at the instant of her perceptual grasp, the instant of her ‘chill

of understanding’, one might be tempted to represent her instantaneous awareness of that infinity of successive objects thus:

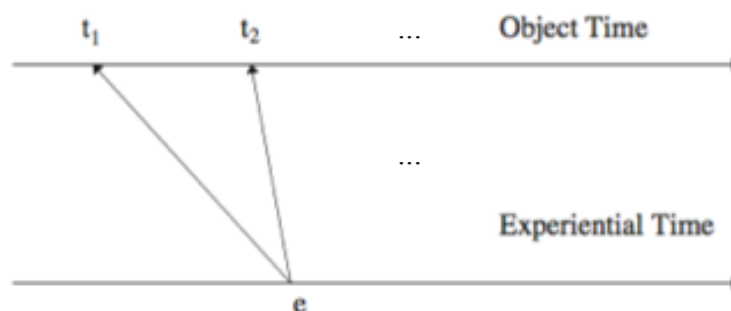


Figure 7 (adapted from Phillips op cit., p. 183).

The three dots over the object timeline represent the ‘always’ bit from my phrasing of the target proposition. The three dots below it represent Sarah’s awareness of the infinity of objects even from the one instant – ‘e’ – represented on the experiential timeline. This diagram is how Phillips represents the view he thinks is mistaken. Experience does not present us with temporal phenomena at one single moment. For our purposes, we can summarise the upshot of his view thus. If one thinks, as many do in the literature, that to make sense of the experience of succession we need to appeal not just to occurrent perception but also to memory, then one will want to posit ‘constitutive relations between past and present experience’ (p. 197). If one instead thinks that we are conscious, at any given time, of both past and present objects, then one will want to posit, as the theoretical unit in which to locate present experience, not a point-like moment but rather an extended ‘interval of experience standing in constitutive relation to present experience’ (ibid.). In either case, and returning to the musical example, the consequence is that ‘if one has not been having an auditory experience in some *temporal window around the*

current moment, one will simply not be in a position *currently* to have an auditory experience at all' (p. 198, my italics).

Let us assume, as per Phillips' view, that present experience is not really constituted by moments but by temporal 'windows', intervals, and that our freezing an experience in time and focusing on moments is just a theoretical device. Even if this is so, the artificial moment we're examining in Sarah's case is different from Phillips' musical case. Phillips is focusing on the moment at which the subject experiences the end of a succession; we are focusing on the moment at which a subject does not experience the end or the beginning of any succession but rather *that* there is a succession to be experienced at all. Unlike in Phillips' case, our moment or interval is still isolated from the experience of succession, which Sarah does undergo after her perceptual grasp that such an experience was there to be had, that is, once she started interacting with the system of reflections as represented in Figure 2 above. So how might we explain such an instantaneous grasp if 'the temporal window around our current moment' is not, as in Phillips' case, filled with the successive objects (musical notes in his case, reflections in ours) that constitute the succession in question?

Here's where understanding the experience as a necessary *possibility* pays off. As Margot Strohming (2015) points out, one way of appreciating the point that one can have instantaneous perceptual knowledge of modal facts is to realise that 'knowledge that can be obtained by inference can sometimes also be obtained by perception' (p. 368). When we're learning how to recognise an object as a member of its kind (say, a bird as an oriole), we often arrive at the target piece of knowledge 'by inferring it from some further claims (to simplify, that the bird has plumage with certain markings, and that the bird is an oriole if it has those markings)' (ibid.).

Eventually, we will be able to know this without relying on inferences, purely by perception.

Now, perhaps Sarah could have arrived at our target knowledge of necessity by inferential means. She could have studied the properties of systems whose elements are related by the succession relation, and come to the conclusion that a contradiction would follow from denying that all such elements have a successor. The fact that she came to know this instantaneously, though, clearly does not follow Strohminger's model of perceptual skill honing, where what is knowable inferentially later becomes knowable by perception. But Strohminger herself tell us that a perceptually available piece of knowledge needn't be always be preceded by a process of inferential learning. I can come to know 'I can reach the mug' by perception without this being a case of perceptual refinement of a long-practised ability to recognise when I can and cannot reach for things by making calculations based on a measurement of my arm's length, the distance between objects and myself, and so on. There may have been a process of learning behind my ability to perceptually know this, but chances are that this learning did not involve conscious inferences but rather motor-perceptual development of the kind infants undergo early on.

Sarah's experience at the moment of her grasp may not be surrounded by a temporal window filled with successive objects. But the four years she's lived are enough for her to have had experiences that *are* surrounded by such kind of succession-filled temporal windows. Just as the perceptual stimuli I receive at the point at which I see a gigantic tree with no branches below the 3 metre mark are enough for me to judge 'I can't climb that tree' even if I've never seen that tree before (I have encountered less challenging trees and know what it is to climb them), so are the perceptual stimuli Sarah receives at the point at which she first glances at the mirror arrangement enough for her to judge that the system of reflections is

untraversable (she has encountered mirror reflections before, as well as traversable processes: long Christmas dinners have seemed *almost* untraversable before she can open the gifts waiting under the tree).

Let me bring these considerations back to our concern, Sarah's perceptual grasp that ' $\neg\forall m\Diamond\exists n \text{ Succ}(m,n)$ '. Defending the possibility of an instantaneous acquaintance with temporal phenomena in virtue of having repeatedly enjoyed *extended* acquaintances with them beforehand would be of little use without an explanation of Sarah's grasp that what she's seeing in that ceremony is one of such cases. Her understanding that there could always be a successor for any reflection is underpinned, I want to suggest, by her grasp of the iterability of certain relations. Children are generally able to recognise what they see on smooth surfaces as a reflection of themselves rather than as another child by 24 months of age.¹¹⁶ Let's call this accomplishment their understanding of the reflection-of relation. It involves an understanding that there is one object and not two (of the same kind) in the relevant domain of quantification, and that they are seeing the iteration of that object's availability for being perceived. This is indeed one way of understanding Maarten Steenhagen's (2017) objection to Clare Mac Cumhaill's (2011) view that mirrors produce an illusion, the illusion that there are objects (and space) where there is just a surface. Steenhagen suggests understanding mirrors not as producing an image of what's in front of them, which would involve positing the existence of a further object from the physical one (the image), but as simply allowing us see things by looking in more than one direction. Thus mirrors do not create an image – and so not an illusory image – at all.¹¹⁷ Whatever the correct view is of the metaphysics of

¹¹⁶ This is so despite behavioural responses that might mislead one to think some children of this age do not pass the self-recognition test. See Broesch, Callaghan et al (2010).

¹¹⁷ The argument relies on a distinction between real and virtual – i.e. 'non-existent' – images that is irrelevant for present purposes. See Steenhagen (op. cit.), pp. 1234-5.

what we see in mirrors, both accounts of specular experience support my suggestion. Whether what children understand by age two should be regarded as the relation between an image and what it is an image of *or* as the relation between an object and a perspective from which it's visible that is distinct from where it is located, either relation is captured by our ordinary usage of the 'reflection of' predicate, and both involve the notion of repetition – of an appearance (as instantiated by an object and its image), in the one case, of a place from which something is visible, in the other. The key insight is, again, that children by age two understand that there aren't two children but one in the suitably restricted domain of quantification.

The notion of repetition embedded in the notion of the 'reflection-of' relation is what underpins, I submit, the possibility involved in the learnt proposition. And that this repetition can happen indefinitely is what underpins the necessity that embeds the former. Not all experiences of relations involve a sense of repetition, of course: consider an experience that involves the relation 'in front of'. And not all experiences of relations that involve repetition involve the sense of indefinite iteration: consider a music listener's extended experience of successive, qualitatively identical notes, similar to Phillip's example, where the temporal window covers the succession's end. Further, not all experiences of relation that involve a sense of indefinite iteration involve repetition. Consider Aristotle's experience of segmenting a line in halves again and again, where the 'half of' relation is, if anything, something like the converse of the 'reflection of' relation. As this last example shows (and we will later see), the 'reflection' of relation needn't be the only one that underpins an understanding of indefinite succession. But in Sarah's case, it does seem to be what underpins such an understanding. Her grasp that there are always a successor for every reflection is due to her grasp of the 'reflection of' relation (that's her grasp of possibility); her grasp that there is *always* a successor for

any reflection is due to her grasp of this relation's indefinite iterability (that's her grasp of necessity).

Now, if my judgement 'I can't climb the tree' is not a hunch but a piece of perceptual knowledge, there is no reason to regard Sarah's judgement that 'for any reflection there can always be a successor' as enjoying an inferior epistemic standing. Both climbing trees and reflections succeeding one after the other are phenomena that are themselves not instantaneous but extended; yet, the tree's unclimbability and the potential infinity of the system can be instantaneously evident. And just as inferentially available modal knowledge is sometimes perceptually available, so does modal knowledge that we achieve by the imagination seem to be perceptually available, too. I did not have to imagine climbing the tree to arrive at my modal judgement in the way the hiker did have to imagine crossing the river to arrive at hers. The same can be said, I think, for the case of Sarah. She did not have to imagine the experience she'd undergo a moment later in order to grasp that the system of reflections would afford it. Nor does it seem like her perception at that initial instant pushed her to imagine the subsequent experience in the way the early humans were in fact pushed to imagine wolves by the perception of their howls.

A little elaboration on this latter point might be helpful. According to a perception-guided imagination proposal of the kind sketched earlier, what Sarah actually sees at some moment – the n reflections she is actually presented with at t_n – *pushes* her to imagine, still at t_n , the reflections she will actually see an instant later, at t_{n+1} . Now we are in a position to see that to divide her experience in this way would be like dividing the experience of the music listener. 'Our experience of temporal phenomena cannot be understood if we attempt to break experience down into instantaneous slices', Phillips writes (p. 176). This is most readily evident in cases where the temporal phenomenon in question is experienced as it unfolds. What

Sarah's case shows, however, is that this is not the only way temporal phenomena can be experienced. Normally, we do experience trees' climbability *as we interact* with trees, and we do 'experience properties logically connected to time such as change in volume, pitch, speed, tone etc.' (Phillips op cit., p. 177) *as they occupy* a temporal window around the current moment. But we can also, on occasion, experience trees' climbability and properties logically connected to time instantaneously, albeit in a modally qualified way. I *can* climb the tree; there *can* always be a successor.¹¹⁸ This is just the way Sarah experienced succession: instantaneously. And this may also explain, I want to further suggest, the fact that at she experienced the sequence of reflections as actually infinite. Because we typically experience succession extendedly, it is perhaps natural that something that is in fact modal in nature – viz. the potential infinity of the system of reflections – would be experienced as non-modal – as actual – in an atypically instantaneous experience (atypical in that we typically experience properties logically connected to time in an extended way).

4.3 *An interlude*

The above may have struck the reader as a bit of an excursus into a tangential matter, the matter of the temporality of experience. If the length of that discussion obscures its goal, let me highlight, before moving on, that the key idea I wanted to bring out

¹¹⁸ The exact mechanisms through which perception gave Sarah the knowledge that for any reflection there could always be a successor are outside the scope of this paper. Because, on Williamson's imagination-based view, modal/counterfactual knowledge is arrived at by means of 'offline predictive mechanisms' that take into account much of 'one's background knowledge and beliefs' (2007b, pp. 152–153), we can speculate that online predictive mechanisms of the kind that have been theorized as underlying perception will be at work in Sarah's case. It is perhaps useful to note, on this subject, that appeal to such online predictive mechanisms needn't commit one to a view on which perception yields only indirect epistemic justification – it needn't entail, that is, that the justification of our perceptual beliefs is derived from the justification of the premises of the unconscious inferences that underlie the relevant predictions. See Drayson (2018).

from it is that one can have an experience of *untraversability*, of there being *always more*, while being a finite subject, finite in the sense of being mortal and in the sense of having limited perceptual capacities. That such an experience is possible is, I take it, a surprising result; hence the need to unpack how an experience can be regarded as presenting an infinite sequence.

This unpacking involved characterising the experience as a peculiar one in two ways: first, in that it was instantaneous rather than extended (whereas experiences of modal phenomena tend to be extended), and second, in that it was an illusion rather than veridical. Sarah's experience being peculiar in these two ways may not be a coincidence. Consider, on the one hand, that a non-illusory experience of actual infinity is difficult to imagine, because actual infinity does not empirically exist (and if it does, e.g. as instantiated by the universe, it cannot be present in its entirety to a human subject). So the only way to experience actual infinity may be in a kind of non-veridical way. And then, on the other hand, a non-instantaneous experience of potential infinity may not have yielded the epistemic goods Sarah's experience did; it may not have yielded, that is, a belief about actual infinity. What might a non-instantaneous experience of potential infinity be like? The answer is, I'd submit, the kind of experiences I painted Aristotle as having. Sitting by a river looking at clouds reflected on the water succeed one another; seeing a finite line of 3 marks being turned by the teacher into one of 4, of 5, and so on. This kind of cases seem uncondusive to the sort of belief that Sarah's experience, which was *as of* actual infinity, elicited. It is not clear, in other words, that typical experiences of potential infinity – typical in being veridical and in being extended, as experiences of modal phenomena tend to be – would elicit the sort of thought mathematical novices are not justified, according to MacBride, to hold: what we have called a general belief. Perhaps the Humean consensus is right, in this sense, that there exist no *veridical*

experiences ‘from which the idea of the infinite can be derived’, since there are no actual infinities to veridically experience.

And now a final point. That Sarah could have believed the infinity of reflections to be actual while grasping the modal nature of the sequence – while grasping, that is, that another reflection *could always* come into view – may be accommodated in several ways. Perhaps there was a cognitive difference, a difference in kind of cognitive state, between the modal knowledge she acquired and her mistaken belief of actual infinity, the former being an ‘alief’, say, and the latter a belief;¹¹⁹ or the former a belief and the latter an ‘intellectual seeming’.¹²⁰ Or perhaps the former, being related (if not, as argued earlier, identical) to Sarah’s knowledge of her capacity to move and view further reflections, i.e. of the reflections being viewable for her, should be regarded as encoded in some kind of non-propositional conceptual state, as some theorists view the cognitive states that explain action,¹²¹ or indeed a different, non-conceptual state altogether.¹²² This turns on the outcomes of debates such as the one around which kind of mental states there are and that between intellectualism vs. anti-intellectualism about action, both of which are, to be sure, beyond the scope of this paper. It does seem, however, that a subject like Sarah would grasp that another reflection could always come into view, as is suggested by interaction with the mirror arrangement, and that, given her ignorance of the physical facts, she nevertheless would take there to be actually infinite reflections before her, as is suggested by her own report.

¹¹⁹ Cf. Gendler (2008).

¹²⁰ Cf. Gow (2019, 2020).

¹²¹ Cf. Bengson and Moffett (2012), Bengson (2016).

¹²² That action is always explained by propositional mental states is called the intellectualist view (see Stanley and Williamson (2001, 2016)); anti-intellectualism is the view that intellectualism is false (see Brownstein (2014), Brownstein and Michaelson (2016), Michaelson (ms.)).

So, to conclude the interlude. I have argued in this section that an experience of untraversability is possible at least in the way explained here, namely, by experiencing potential infinity in an instantaneous way and by experiencing it *as* actual. This seems promising as a way of tackling the heart of the MacBride-Shapiro debate, which I argued in Section 1 is the issue of knowledge of infinity. Knowledge of infinity, we also saw there, can be understood as knowledge of ‘the principles which are employed by mathematicians to infer the general from the particular’. How Sarah’s experience gave her knowledge of some such principles is the topic of the next section.

5. Generality

5.1 Positive generality

Our interest in Sarah’s perceptual access to potential infinity in the foregoing section responded to our interest in her access to knowledge of the kind some philosophers are even more sceptical about: general knowledge. The problematic thing in our original puzzle was not the modal operator(s) in the target proposition, which cropped up later, but the universal quantifier in it. Differently put, the problem was the possibility of coming to know anything about infinite structures by perception.

Let me take a step back to show how I mean to exploit the above discussion in Sarah’s case. To form a belief about the structure of the natural numbers is to form a belief about infinity, that is, a belief of the kind mathematical novices are, on MacBride’s view, unwarranted to form. Yet Sarah did form such a belief, and if I am right that she formed it on the basis of perceptual experience, not a hunch, then that belief was epistemically justified (assuming perception is a reliable or safe method).

Compare another mathematical novice who is next in line for the ceremony and forms, prior to undergoing the experience, *from a hunch*, the belief, ‘there is a potentially infinite number of reflections in that arrangement’. What is it that post-experience Sarah and mathematical experts share, epistemically, that her fellow four-year-old does not?

My suggestion is that they are both able to conceive of a model of Dedekind-Peano arithmetic. It is a natural thought that to be justified in believing a mathematical theory, one should have the ability to *conceive* of a model for that theory. Conceivability is largely taken to be a source of (corrigible) justification, after all, and not just in mathematics. Furthermore, it is natural in many cases to cash out the ability to conceive the truth of some proposition or set of propositions in visual terms. Such an ability, Gendler and Hawthorne write, for instance, ‘enables us to represent scenarios to ourselves using words or concepts or *sensory images*, scenarios that purport to involve actual or non-actual things in actual or non-actual configurations’ (Gendler and Hawthorne 2002, p. 1, my italics). The representational aspect of the conceivability capacity is key. What would it mean – what would it *look like*, a mathematical novice might wonder – for *all* natural numbers to have a distinct successor, which is to say, for there to be *no end* to their set? This is the question Sarah’s experience answered for her.

So the target theory, the theory for which the ability to conceive of a model is our target, is a theory about the structure of the natural numbers: in this case, Dedekind-Peano arithmetic. Here are its five axioms:

(PA1) Zero is a natural number.

(PA2) Zero is not the successor of any natural number.

(PA3) Every natural number has a successor that is also a natural number.

(PA4) Different natural numbers have different successors.

(PA5) For any property P , if $P(0)$ holds, and if, whenever $P(n)$ holds, then $P(n + 1)$ holds, then $P(n)$ holds for all natural numbers.

Perhaps a mathematical expert acquires the ability to conceive of a model for (PA1) – (PA5) in accordance with Shapiro’s stratified epistemology, which ends with the stage of ‘description’, a fully theoretical stage.¹²³ But how might a visual experience allow a novice to conceive of such a model? To get us started in answering this, consider Charles Parsons’ (op cit.) imagination-based proposal, for the exposition of which I will rely on Robin Jeshion’s recent defence of the view (2014).¹²⁴ The thought is that one might be able to conceive of a model for arithmetic without knowing, as per MacBride’s demand, ‘the principles which are employed by mathematicians to infer the general from the particular’ from the get-go by imagining a similarly infinite set, a sequence of stroke-strings ‘|, ||, |||, ...’. We can unproblematically ‘intuit’ this sequences’s properties just as we intuit those of other ‘quasi-concrete’ objects such as ‘letters, words, sounds and shapes’ (Jeshion 2014, p. 330). Such an intuitive episode can ground, ‘*in a certain sense*, our knowledge of the first four of the Dedekind-Peano axioms’, that is, ‘our knowledge of the infinitude of the natural numbers’ (Jeshion op cit., pp. 327 and 336).¹²⁵

Here is roughly how this works. After introducing a language L of which the stroke ‘|’ is the only symbol and the well-formed expressions are strings such as ‘|, ||, |||, ...’, Parsons basically interprets the string formed by a single stroke as zero and

¹²³ But her justification for believing in the truth of (PA1) - (PA5) would still be, according to MacBride, questionable. See mention of the description stage in a footnote above.

¹²⁴ I should say I don’t aim to argue against Parsons himself. Rather, I want to employ the view – again, as most succinctly presented by Jeshion – for necessary contrast.

¹²⁵ The fifth axiom, the axiom of induction, cannot be grounded in mathematical intuition on Parsons’ view, but we will attempt to tackle it shortly.

the operation of adding one stroke to the right as the successor operation. This renders the sequence '|, ||, |||, ...' isomorphic to the set of the natural numbers. To know by intuition that each string can be extended by one stroke more is much like knowing that each natural number has a distinct successor, which constitutes, Parsons thinks, 'the weakest expression of the idea that our "language" is potentially infinite' (1980, p. 105). Knowing this by intuition about L amounts to knowing the following, then:

(PA1') | is a stroke string.

(PA2') | is not the successor of any stroke string.

(PA3') Every stroke string has a successor that is also a stroke string.

(PA4') Different stroke strings have different successors.

Seeing how intuitive awareness of the stroke strings yields awareness of the propositions (PA1') and (PA2') is, Jeshion says, 'straightforward'; '[t]he axiom that presents the most complexities is (PA3')' (op cit., pp. 336-7). That is because (PA3') amounts to the general thought we saw McBride demand a non-theoretical warrant for, which very worry Bob Hale and Crispin Wright press against Parsons' view as well: '[t]he problem, of course, is to see how, following Parsons' intuitive route, knowledge of general truths about intuited objects [...] can be achieved' (2002, p. 106). Jeshion's and Parsons' answers to this challenge rely on imagination's – the basis for intuition – capacity to deliver 'vague', i.e. 'arbitrary', i.e. non-particular mental representations (of the stroke strings), which non-particularity Hale and Wright regard as insufficient for the requisite generality (of knowledge about all the

stroke strings), or in other words, as insufficient to justify genuine universal quantification (over all stroke strings).^{126,127}

The problem seems to be that the non-particularity of the deliverances of the imagination may not amount to *general* but rather *vague* or *indeterminate* deliverances. Thus: unwarranted generalisation. As James Page (1993) has observed, it is not obvious, first, how Parsons's subject is entitled to regard the final string as different in length from the first string she imagines, seeing as their lengths are, as the properties of imagined items usually are, indeterminate. Yet suppose the strings are not imagined but perceived (as marks on paper, say). She draws a stroke-string and then draws another, and another. In that case, what allows the subject to make the generalisation involved in (PA3') is her intuition that the structure of the type of string is irrelevant to its being possible to add another stroke. This renders her formation of the belief expressed by (PA3') analogous to the case of someone who forms the belief that Pythagoras's Theorem holds of any right-angled triangle on the basis of discovering, via diagrammatic reasoning, that it holds of the particular diagram with which she's working. The problem then returns: how is the subject warranted in believing that the properties she's attributing to that particular stroke-string (that particular right-angled-triangle) hold of all stroke-strings (all right-angled-triangles)?

Perhaps the issue is, at bottom, that '[t]he notion of intuition seems murky and obscure, and is widely deemed a creature of darkness', as Jeshion herself writes (2014, p. 330). It would be better, in this light, if a subject were endowed with the

¹²⁶ This is just one of their diagnosed problems with it. See Hale and Wright (op. cit.), pp. 108-11.

¹²⁷ Arbitrariness of content is defined fairly intuitively by Parsons. We imagine an arbitrary string much in the way we imagine a large crowd: it is not part of the content of such a mental state that the string/crowd consists exactly of n strokes/people (Parsons 2008, pp. 173-4).

ability to conceive of a model of arithmetic not by somehow intuiting the properties of an infinite system but by being perceptually presented with them. Following Parsons' strategy of reinterpreting the Dedekind-Peano axioms to capture intuitive knowledge about the potential infinity of imagined strokes, we can reinterpret those axioms to capture the perceptual knowledge Sarah has of the potential infinity of the perceived reflections.

Let the first reflection seen by Sarah stand for zero and her seeing of every successive reflection to stand for the successor operation. Then, by Jeshion's (presentation of Parsons') demonstration of how awareness of those two things yields awareness of the PA' truths (Jeshion op. cit., p. 336), we have that Sarah has come to grasp that:

(PA1'') [Demonstrative referring to first reflection] is a reflection.

(PA2'') [Demonstrative referring to first reflection] is not the successor of any reflection.

(PA3'') Every reflection has a successor that is also a reflection.

(PA4'') Different reflections have different successors.

If this seems right, then by grasping the truths (PA1'') - (PA4''), Sarah also grasped what Jeshion thinks is the hardest – but not impossible, on her view – truth for Parsons to justify intuitive knowledge of, the stroke-string analogue of the Axiom of Infinity of set theory:

(Axiom of Infinity') There are infinitely many distinct strings of stroke (ibid., p. 337),

or, in Sarah's case:

(Axiom of Infinity'') There are infinitely many distinct reflections,

which knowledge of is sufficient for knowledge of the general thought (PA3”), as knowing (PA3”) is necessary for knowledge of *it*,¹²⁸ and which is exactly what Sarah is telling us she came to understand:

I looked into the images and was amazed to see that they never ended; [...] the reflections kept going and going.

One final clarification is in order. (PA3”) and (Axiom of Infinity”) are actually the result of translating what Sarah comes to perceptually grasp, the Aristotelian thought of the form ‘ $\Box\forall m\Diamond\exists n \text{ Succ}(m,n)$ ’, into the non-modal language of ordinary mathematics. We are entitled to this translation for two reasons. First, one of the main results of Linnebo and Shapiro’s work is, as mentioned, the argument that an entailment obtains in the non-modal language of ordinary mathematics whenever its modal counterpart obtains in the potentialist’s language – which is, in other words, to say that ‘classical first-order Dedekind-Peano arithmetic [...] is equivalent [to a potentialist translation of that theory]’ (Linnebo and Shapiro op. cit., p. 16). Secondly, though, even if what Sarah perceptually grasped was a modal proposition (that there could always be another reflection) because a non-modal one wasn’t available for her to grasp, since she wasn’t been presented with an actual infinity, or infinity simpliciter, it did *seem* to her that she was being presented with such an actual, not a potential thing, an actual ω -sequence. So the Dedekind-Peano analogues we should attribute to her a belief in are indeed non-modal, as expressed by (PA1”) – (PA4”), and indeed, as I’ll argue shortly, something like PA5. It appears,

¹²⁸ Along with two other truths: in the case of the stroke language, that ‘we can indefinitely iterate the operation of adding one new string’, and ‘the new string of strokes obtained from adding a new stroke is in fact new’ (Jeshion op cit., p. 336, fn. 14). In our case, those two truths are: ‘we can indefinitely iterate the operation of bringing one new reflection into view’ and ‘the new reflection brought into sight is in fact new’.

after all, that she understood the relevant infinite structure could be instantiated by systems other than the system of endless reflections it seemed to her she was seeing: a system, for example, of endless instants of time. ‘Oh,’ I thought, with a chill of understanding. ‘Forever.’

5.2 *Negative generality*

As the reader will have noticed, I have ignored, in the foregoing discussion, the fifth axiom. This is in part because that discussion followed Parsons’s, which ignores it too. But, as Hale and Wright (op cit., p. 108) remark on this feature of Parsons’s view, full knowledge of arithmetic cannot be accounted for without it. Why is this so? What does knowledge of the fifth axiom add?

The object I’m arguing Sarah acquired knowledge of, and which we saw in the beginning Dedekind target too, is the structure instantiated by any ω -sequence. This is the structure described by the Dedekind-Peano axioms. The axioms we have discussed so far, the first four of them, are true of all ω -sequences, but they are also true of others. If we interpret occurrences of ‘number’ in the axioms as meaning not ‘natural number’ but ‘natural number or integer’, say, and interpret ‘successor’ as denoting the operation of adding one, then the axioms are also true of the structured set formed by the union of the set of integers and the set of natural numbers (‘zero’ denoting a natural number). This is not an ω -sequence because there is no first member, and there are numbers in the sequence – all the integers – that are not reachable from zero by applying the ‘successor’ operation a finite number of times. Because such an interpretation is unintended, the model of the four axioms it generates is a non-standard model of arithmetic. Here is where the fifth axiom comes in. It ensures that the structure described by the theory is the structure of all *and only*

ω -sequences. It does so by ruling out unintended interpretations like the one just proposed. Consider what it says again: if a predicate is true of zero, the axiom tells us, and if it is true of the successor of a number n whenever it is true of n , then it is true of any number n . This describes something like a ‘domino’ effect: once a statement is true of the first member of the sequence, the rest will follow. This axiom comes out false on an interpretation where the extension of ‘number’ includes the integers, because there is no such first member in that sequence, no ‘base case’ from which to carry out the mathematical induction, in other words; from any given member, there are members that are not reachable by iterating the ‘successor’ operation. So there is no domino effect in such a case. The domino effect of the axiom has a reverse side, then. By ensuring that only domino-apt sequences – only sequences that have a first member to ‘push’ and whose members are all related by the relation of succession; only ω -sequences, that is – are described by the theory, it rules out other – non-standard – models. Call this the ‘limitative’ role or effect of the fifth axiom.

Appreciating this effect allows us to see that knowledge of the fifth axiom is, in a sense, negative knowledge, knowledge of the kind one has when one knows the edges of things. In the process of honing in on the object Dedekind was targeting, the fifth axiom tells us when to stop, when to look no further. The first four axioms tell us what to look for, yes: a set C ordered under an indefinitely iterable function S that contains an element 0 which is not $S(x)$ for any x in C , and such that, for any x in C , $S(x)$ is in C , and such that if $S(x) = S(y)$, then $x = y$. This already implies that C has infinitely many elements, since by those axioms the sequence $0, S(0), S(S(0)), \dots$ cannot contain repetitions. Then the fifth axiom tells us that all members of the sequence should be reachable by the operation S from the one member that is not $S(x)$ for any x ; that C should be, in other words, domino-apt. It tells us to look *no further* than an ω -sequence, the smallest infinity. To non-fortuitously conceive of an

ω -sequence, then – to conceive of an infinite set that doesn't just *happen* to be the size of the set of the natural numbers – is to conform not just to the first four axioms but also to the fifth.¹²⁹

I argued that our subject acquired, in her experience, the ability to conceive of a model of arithmetic. I can now refine that statement: she became able to conceive, specifically, of a *standard* model of it. In this, she seems to differ from a subject who exercises the imagination as per Parsons' account. Parsons explicitly recoiled from arguing that the intuition can deliver knowledge of the fifth axiom along with knowledge of the other four. It can be helpful to see why. 'Induction as a general principle has an essentially higher-order character,' he writes in his (1986), 'and for that reason it seems evident that it cannot be intuitively known (p. 227). It is a point acknowledged elsewhere. 'For Brouwer the general notion of induction [...] belongs to "mathematics of the second order", whereas concrete additions and multiplications belong to "mathematics of the first order"', Mark van Atten writes in his comparative study of Parsons' and L. E. J. Brouwer's views on that principle (2015, pp. 9-10). The difference between the first- and second-order 'characters' of the first four and fifth axioms seems to indicate they generate different epistemological worries. MacBride's worry was about the entitlement a novice has to generalise over all objects within an infinite set. The first four axioms are truths about such objects: terms for them simply occupy the subject positions of those sentences. The fifth axiom is, by contrast, a truth about those objects' properties. The subject position of (the various clauses of) the sentence expressing the fifth axiom are occupied by a property term, 'P', not by reference to any first-order object. Here is that sentence again:

¹²⁹ I'm indebted in this section to Tim Williamson for pointing me towards the limitative role of mathematical induction in the first place.

(PA5) For any property P , if P holds of zero, and whenever P holds of any natural number, P holds of its successor, then P holds for all natural numbers.

The reason a truth like (PA5) cannot be intuitively known is that, on Parsons' view, what we intuit are only first-order objects: the strokes, on his proposal, with which our imagination forms strings; and intuition of objects grounds intuitive propositional knowledge. We intuit that p if p can be seen to be true on the basis of our intuitions of the object(s) p is about (Parsons 2008, p. 171). Where (PA1) – (PA4) are truths about objects, and their string-stroke equivalents about objects we can intuit, (PA5) is a truth about *all* the properties that meet a certain condition relative to those objects. The objects that would need to be intuited here seem to be not first-order entities but sets of them.¹³⁰

A proposal on which the fifth axiom, or an analogue of it, is known perceptually seems bound to run into a similar problem. Such a proposal would *prima facie* have to deal with the fraught question of whether properties can themselves be perceptually known, and so truths about them known directly, as opposed to known as part of the process of perceiving their bearers. But the discussion of the limitative effect of the fifth axiom allows us to bypass this complication and appeal instead to the ability to perceptually acquire negative knowledge in order to sketch a way in which our subject did acquire something like the knowledge the fifth axiom expresses, and so, not just perceptual knowledge of an infinite sequence but specifically of an ω -sequence, the very kind Dedekind was targeting.

¹³⁰ One might worry that this is an unduly object-centred view of intuition. Feliz Mühlhölzer (2010 p. 270) does so, and thinks it results in a neglect of the role of intuition in ordinary mathematical practice, in which what is thought to be intuitive is most often proofs. Mühlhölzer's views on the epistemic accessibility of proofs seem to permeate his discussion of surveryability, as we'll see shortly.

So, what is it to perceptually acquire negative knowledge? Earlier on I hinted at an answer: to experience a negative fact, I said, is to experience incompatibility. This is at least an idea that has been discussed recently (cf., again, Berto 2015, Berto and Restall 2019, Kinkaid 2020). Attached to this proposal seems to be the idea that there is a phenomenology to negative knowledge acquisition: even though $\Diamond p$ is equivalent to $\neg \Box \neg p$, I suggested, knowing by perception that $\Diamond p$ feels different from learning that $\neg \Box \neg p$, as the difference between scenarios in which a hiker learns that he *can* jump and that it simply isn't necessary to find an alternative suggested. From this, one might think that learning negative facts *needs* to come with some sort of negative phenomenology, or at least a phenomenology differing from what I called experiencing a 'positive force'. But not all negative facts are made, and so felt, equal. The kind of negative knowledge I am discussing right now is, specifically, limitative knowledge; knowledge I likened to knowing the edges or limits of things. Knowledge of limits is derivative on knowledge of what one knows the limits of, and so the phenomenology of acquiring it plausibly derivative of the phenomenology of acquiring the latter. Consider Stephen Barker and Mark Jago's (2012, p. 134) discussion of edge perception:

[T]he perception of edges [...] is an important component of human object recognition. [In] the analogous case of computer vision, [...] one might think [...] that an image can be represented as positive information only. This is not the case, however. The value representing the intensity of each pixel contains both positive and negative information: a pixel's intensity value says that the pixel is at least this bright but *no brighter*. The numerical value associated with each pixel encodes both a positive and a negative fact. We

suspect that a similar story holds of edge-detection in human perception (emphasis original).

Regardless of whether Barker and Jago's analogy between computer and human vision is right, the insight that perception of edges is a key part of the perception of what the edge is of is plausible enough, and it is helpful. Recall the question that arises from Hale and Wright's discussion of how, in a proposal like Parsons's, knowledge of arithmetic cannot be accounted for without knowledge of the fifth axiom. How might a subject be said to be able to conceive of a model of arithmetic if she doesn't know what structure exactly she is meant to be able to recognise among different systems now, and also, channelling David Lewis (1990), to imagine and remember systems as instantiating? How does she know, in other words, that the structure she is to recognise, imagine and remember is that of the shape $\| \dots$, which meets the conditions of there being a first member and all the other members being reachable from it by application of the successor operation, rather than a structure of the shape $\dots \| \dots$, say, which does not meet those conditions? The first four axioms are true of both kinds of models, after all. Hale and Wright's question can be rephrased, in light of the limitative role of the fifth axiom, as the question of how knowledge of arithmetic can be accounted for without knowledge of the *limits* of what the theory describes, knowledge that the theory describes a structure like the one of the natural numbers and no bigger (to the exclusion, then, of the structure of the union of the set of natural numbers with the set of integers, say).

If one role of the fifth axiom is to let us know that arithmetic describes no bigger set than *that*, no bigger sequence than an ω -sequence, then the epistemic need being pointed out as unmet here is much like the one fulfilled by our capacity of edge detection, which consists in the knowledge that this or that object extends thus far

and *no further*. The problem with Parsons' view from this perspective might be that, on his account, intuition works by constructing the sequence of strokes in the imagination, such that no edge detection is possible in that experience because detection is something one can do, or fail to do, with respect to what is presented to one, not what one constructs. The equivalent in the experience of Parsons' subject might be limit setting, not detection. And limit setting is an ability that is entailed, to be sure, by the ability to conceive of the target model in the first place.¹³¹ So this is one reason to think perceptual knowledge of arithmetic is possible in a way intuition-based knowledge may not be.

How may we account for such knowledge in the case of our subject? My suggestion is that we needn't spell out how a self-standing analogue of the fifth axiom became known to her in the way we have done so for analogues of the previous four axioms. The relevant epistemic need being met by the fifth axiom is, unlike in the case of the first four, derivative, insofar as knowledge of limits is. Now, it is true that the fifth axiom is not logically entailed by the first four, whence the need to account for it, but that is not the sense in which knowledge of it is derivative. Rather, the fact that our subject became able, upon her experience of an ω -sequence, to conceive of that structure points to the possibility that the limitative knowledge a mathematical novice requires to imagine an ω -sequence was simply built into the experience, i.e. into the presentation of the sequence of reflections to our subject, just as knowledge of the edges of objects comes built-in when we are presented with objects (but not

¹³¹ I draw the contrast between what is presented to one and what one constructs, rather than what one imagines, because there is a sense in which one can indeed notice features of scenarios that one conjures up in the mind's eye, as Peacocke (2020) suggests, for example, occurs in literary experience. The experience of Parsons' subject is patently not like this, however. His experience is meant to be one of forming a string of strokes in a way that seems to be active rather than receptive, as in the case of literary and perceptual experience.

when we have to construct them in our mind's eye, in which case we have to set their limits ourselves).

The reason I'm calling this subsection 'negative generality' is that the knowledge I am claiming Sarah acquired along with knowledge of analogues of the first four axioms is, like them, general knowledge, even though it is negative in a sense, too. To appreciate this, consider one way to express the negative general fact in question. The fact is the answer to the question: what is excluded as a natural number? Let the constant N below stand for membership into the set of natural numbers. According to the first four axioms, N is true of zero and of the successor of any natural number, starting with *it*:

- (i) $N(0)$, and
- (ii) $\forall x(N(x) \rightarrow N(S(x)))$

This roughly summarises what it is to be a natural number according to the first four axioms. The extra, limitative knowledge we need is that nothing else besides zero and its successor, and its successor, and its successor... is part of the set. If we let P denote the property of meeting any of the above two conditions (the condition of being zero or the successor of a natural number), we might represent this limitative knowledge thus:

- (iii) $\forall x(\neg P(x) \rightarrow \neg N(x))$

Nothing that does not have the property of being the kind of object specified by (i) and (ii) is part of the set of natural numbers. Again, this statement is not entailed by the previous two; and indeed it comes out true in a non-standard model

with the positive and negative integers added and N interpreted as true of all of the elements in that set. Holding the meaning of N fixed, though, that is, presupposing a grasp of natural numberhood, this at least illustrates the kind of knowledge Sarah counts as having merely in virtue of being aware of a sequence of which the first member counts as part of the sequence and of which the rest of the members are those objects which succeed that first as per the relevant iterable relation, and of which nothing else is part.¹³² That nothing besides the first reflection and the reflections that followed it is part of the sequence was simply evident to her.

6. The importance of a glance (and back to Dedekind)

Let me take stock. In the first section, I argued that at the heart of the debate around general knowledge, knowledge of the kind it is puzzling a mathematical novice would enjoy given her ignorance of ‘the principles which are employed by mathematicians to infer the general from the particular’, is knowledge of infinity. The issue was especially pressing since leading epistemologies such as Shapiro’s seem to build knowledge of the infinite upon perception of the small and finite. The challenge was

¹³² This ‘illustration’ is drawn from a discussion in Peregrin (ms.). Peregrin is driven by a similar thought as I am in this section, the thought that the epistemic upshot of the induction axiom is limitative, but this leads him to argue for a kind of formal equivalence between the negative proposition, (iii), and that axiom. His argument, unfortunately, ignores the fact mentioned, that (iii) can be true of non-standard models if the meaning of N is not kept fixed, which means that, if (iii) is meant to be *equivalent* to the axiom, its truth cannot presuppose what the five axioms are jointly meant to describe, the structure of the natural numbers. Here, though, I am not trying to convince the reader that a certain piece of negative knowledge is *equivalent* to knowledge of the fifth axiom. I am just trying to exploit the fact that in perception, awareness of things tends to come with awareness of their limits, such that the function fulfilled by the fifth axiom in the context of describing the structure of the natural numbers has no counterpart in the context of our subject’s perceptual experience of a system instantiating that structure. For, having perceived the relevant ω -sequence, there is no further proposition for our subject to discover in the way that, having specified the first four axioms, we *do* need to further specify the fifth one in describing the natural number structure, in order to ensure that we are describing an ω -sequence.

to account for a novice's justification of the jump in a way that didn't appeal to a hunch. In the second section, I introduced a test case: a mathematical novice's seeming acquisition of just that kind of general knowledge by the very means Shapiro thinks we come to know about small and finite structures: visual perception. It was a seeming perceptual encounter with an ω -sequence. In the third section, I defended my calling this sequence 'infinite' by taking on board the idea that infinity can be made sense of in potential terms, not only actual ones, such that the system of reflections could be viewed as potentially infinite. In the fourth section, I argued that the subject of our test case acquired just such a piece of modal knowledge, knowledge of potential infinity, which we put as a fact of the form ' $\Box\forall m\Diamond\exists n \text{ Succ}(m,n)$ '. Following Aristotle, I called the experience in which this proposition was learnt an experience of 'untraversability' or of 'there being always more'. It was suggested, furthermore, that experiencing this potential infinity not as such but as *actual* helped our subject to grasp just that, actual infinity, which grasp was understood, in the fifth section, as the ability to conceive of a model of Dedekind-Peano arithmetic. This involved a separate treatment of the first four and the fifth axioms. The first four axioms already imply infinity; however, they are compatible not just with standard models of arithmetic but also with non-standard models. Hence, to show that our subject became acquainted with the structure instantiated by the natural numbers (i.e. what we call the standard model), I had to show that she acquired knowledge of the fifth axiom, which excludes as members of the described sequence – thus ensuring the sequence is indeed an ω -sequence – any non-standard numbers, e.g. the negative and positive integers. I suggested that a perceptual experience is indeed amenable to knowledge of the fifth axiom in a way that the experience of constructing strings of strokes in the imagination is not. This was because the knowledge in question is limitative, which is a kind of knowledge perception often

does yield. (The clarification was made here that while being, in a sense, knowledge of a negative fact, this needn't involve a negative phenomenology because it is part of the perception of what the limit is of.) The upshot of this all is that a mathematical novice like our subject can be justified in believing something about *all* the elements of an ω -sequence by forming, upon a perceptual experience, beliefs with content analogous to 'the principles which mathematicians employ to infer the general from the particular'.

The above comes down to the suggestion that, contrary to the Humean consensus, experience *can* provide an 'impression from which the idea of the infinite may be derived'. Now, before closing, let me bring out an issue that connects this proposal back to the anecdote I discussed in the introduction. Actually, though, Dedekind's Section 66 is no mere anecdote. If I am right that it has an epistemological upshot, then something already in Dedekind's theorem prefigures a problem threatening an intuition-based account like Parsons's. To see the problem, notice that Dedekind's view does not involve the construction of the target infinite system, as Parsons' account does. Dedekind wants to show that there simply exists, in a time-independent way (at least relative to the subject's life), an infinite thought-world. There is a temporal asymmetry between the ways in which Dedekind and an intuition-based view account for knowledge of infinity. The intuition-based account involves an extended imaginary exercise; Dedekind's does not. But the latter involves of course an inferential achievement. Is there a difference in epistemic value between accounting for knowledge of infinity by involving an extended process as such that of imaginatively constructing strokes and accounting for it without involving one? The difference isn't so much a matter of time but of whether the epistemic achievement consists in having 'surveyed' an object, as Wittgenstein would put it.

I'll elaborate. The notion of surveyability seems to be a highly debated matter in Wittgenstein scholarship (cf. e.g. Mühlhölzer 2005, Majetschak 2016, Floyd 2021); and it seems relevant not only to his philosophy of mathematics but also to his view on anthropology and aesthetics (Ben-Moshe 2021).¹³³ This very fact, I want to suggest, the fact that it has been shown to be consequential to such diverse experiences as one of mathematical understanding and one of art, points to the relevance of perception to it. Without going into Wittgenstein exegesis, let me briefly sketch what surveyability consists in and how it relates to the present issue. The literature treats surveyability as a property of representations and 'to survey' as a success term, as something epistemic agents do to representations. Thus representations are surveyable or unsurveyable and subjects either survey a representation or not. (Here I'll treat mental representations, such as the one in Parsons' subjects case, as apt for this qualification.)

Nir Ben-Moshe (op cit.) locates one succinct phrasing of Wittgenstein's view of what a 'surveyable representation' is in the following passage:

The concept of surveyable representation is of fundamental importance for us. It denotes the form of our representation, the way we see things. [...]

This surveyable representation brings about the understanding which consists precisely in the fact that we 'see the connections'. Hence the importance of

¹³³ There is a debate on whether the notion of 'surveyability' Wittgenstein employed in his *Remarks on the Foundations of Mathematics* (1956) is the same as the one employed in earlier works, especially his *Remarks on Frazer's Golden Bough* (1931), which has been taken to present his view on philosophical method (Majetschak 2016) but Ben-Moshe (2021) additionally relates, specifically, to his aesthetics. See Majetschak (2016, esp. p. 74) for a defence of the view that the two notions are indeed the same.

finding *connecting links* (Wittgenstein 1993, quoted in Ben-Moshe op cit. §2, emphasis original).¹³⁴

The contrast here is with the notion of ‘causal-hypothetical’ representation. Where a causal-hypothetical representation of a phenomenon conveys its ‘origins’, a surveyable representation conveys its ‘inner nature’, that is, the ‘internal relations’ and ‘formal connections’ that give it its shape (all terms in quotes are Wittgenstein’s). Consider illustrating ‘an internal relation of a circle to an ellipse by gradually converting an ellipse into a circle’. Such an illustration does not tell us that ‘a certain ellipse actually, historically, had originated from a circle’; it simply tells us that certain ‘internal relations’ or ‘formal connections’ hold between circles and ellipses (ibid.). There are many attributes to each of these two kinds of representations,¹³⁵ but one Stefan Majetschak (2016) draws our attention to is the ‘generality’, as opposed to ‘specificity’, of surveyable representations. A surveyable representation is, in Majetschak’s words, ‘*a picture of the general connections*, that is a picture of structural relationships between things or facts’ (Majetschak op cit., p. 72, emphasis original). This dimension of a surveyable representation seems to be key to its significance for mathematics. Where evidence in the natural sciences tends to come from empirical observation, that is, experiments, which yield specific data, evidence in mathematics needs to be the opposite: not particular but general. Evidence in mathematics comes in the form of proofs. So proofs need to show ‘general connections’. To grasp a

¹³⁴ Translations of Wittgenstein sometimes have ‘perspicuous’ rather than ‘surveyable’ (see e.g. passages in Ben-Moshe 2021 and Floyd 2021), but here I’ll employ ‘surveyable’, since this is the term of choice in the explicitly exegetical literature (Majetschak 2016, Mühlhölzer 2005). Needless to say, the perceptual aspect of that term is also fitting to my purposes.

¹³⁵ And what attributes one sees as distinguishing them may depend on what the context is. Again, Ben-Moshe’s (2021) argument that ‘surveyability’ is consequential to Wittgenstein’s aesthetics takes notice of aspects of it that differ from aspects discussed by e.g. Majetschak (2016) and Mühlhölzer (2005), who focus on its significance for mathematics.

surveyable representation such as the representation of the formal connections between a circle and an ellipse is to know something about all circles and all ellipses, to be able to extrapolate what one learns about what one ‘surveys’ to claims about other, structurally similar cases: other circles, other ellipses.

Consider another example, a surveyable representation of the fact that $3 > 2$:

[t]he pattern	L	R
	x	0
	x	0
		0

may be regarded as a demonstration of $3 > 2$, a demonstration which is as good as any geometrical demonstration. (AWL 1979: 170)

Figure 8 (from Majetschak op cit., p. 76).

‘As a surveyable representation’, Majetschak tells us, ‘[this pattern] shows the connections which are to be proved and lets the aspects that matter be seen *at one glance*’ (p. 76, my italics). Majetschak’s discussion takes pains to emphasise the perceptual dimension of the epistemic achievement that surveyable representations allow. ‘Wittgenstein wishes to say that in every proof it is ultimately a matter of what the “characteristic visual shape” (RFM 1978: III, 11) allows to appear before our eyes’; ‘the “visual criterion” (AWL 1979: 168) of whether we see certain connections in a constellation of signs or not is the ultimate one that we have’ (ibid.).¹³⁶

It is curious that a seeming perceptual epistemic achievement yields in this case the opposite of what perception is largely taken to yield knowledge of: the general rather than the particular. This is why the contrast with a distinct form of

¹³⁶ Indeed, as Mühlhölzer notes, Wittgenstein’s original German term for ‘surveyable’ has *sehen*, ‘to see’, at its root: “*über-sicht-lich*”, “*über-seh-bar*” (Mühlhölzer 2005, p. 68).

representation, one that does make particularities salient, is helpful. Why are causal-hypothetical representations not surveyable? ‘For reasons of length or extent’, Majetschak writes, ‘they do not involve the aspect that needs to be made visible in a *single overview*’ (p. 78, my italics). The thought seems to be that ‘the aspect that needs to be made visible in a single overview’, the aspect of the phenomenon that can be extrapolated to other phenomena, is its general ‘visual shape’, which is visible only in a synoptic view of it.¹³⁷

Now, because the kind of representation most often discussed as surveyable in Wittgenstein’s sense is mathematical proofs (and because the work in which the notion plays its most central role is in the *Remarks on the Foundations of Mathematics*), one might think the visual vocabulary employed is metaphorical. Proofs that one cannot take in ‘at one glance’ are the order of the day. This is something like Feliz Mühlhölzer’s (2005) view: that Wittgenstein’s vocabulary is meant to convey not something about the actual experience of understanding proofs but about what it means to understand a proof. In short, the emphasis on having the object in view at once, and on its relevant aspect being a general ‘visual shape’, as constituted by ‘formal connections’, comes down to the view that ‘the surveyability of a proof consists in its possibility of *reproduction*’ (Mühlhölzer op cit., p. 59, emphasis original). Three conditions come with this. That ‘the reproduction of a proof is the sort of reproduction of a *picture*’, that ‘this reproduction must be an *easy task*’, and that ‘we must be able to decide with *certainty* whether the reproduction produces the same proof’ (ibid., all emphases original). Mühlhölzer illustrates what these conditions come down to with a negative example:

¹³⁷ There is also mention of a relation between the notion of synopsis and the ‘answer to a puzzle’ (Ben-Moshe op cit., p. 7).

[What these conditions exclude] are not very long and complicated proofs per se, but proofs that are unsurveyable in the manner in which they, e.g., represent numbers. This would be the case, say, if we represented the natural number n by a string of n strokes, [...] such that for very big n it is not an ‘easy task’—and it may even be a practically impossible task: for human beings anyway [...]—, to reproduce the sign representing n (Mühlhölzer op cit., p. 61).

Mühlhölzer’s example brings to mind one problem with Parsons’s account: that limited as she is, his subject is unable to survey arbitrarily long symbols. Let me spell this out a little. Picture Parsons’s subject mid-task, holding a string of strokes in mind already. The project of grounding her knowledge that every stroke has a distinct successor relies on the possibility of her imagining an arbitrarily long string’s being extended by one stroke yielding a *distinct* arbitrarily long string. And not any distinct such string: a string that remains the same except for the extra stroke, rather than, say, shrink by one stroke at the beginning. Hale and Wright’s concern about the arbitrary nature of the mental representation here – their concern that while failing to be particular this representation does not yield the requisite generality – might be put in Wittgensteinian terms by saying that such representation is not surveyable to the subject: it does not seem that the imagined string’s visual shape – its structure – is immediately clear and recognisable in such a way that the subject knows – clearly and immediately – that she can indefinitely extend it in the required way, in the way characteristic of ω -sequences. Will the next imagined string remain the same at the beginning rather than lose its first member? That limited subjects like us are unable to survey arbitrarily long symbols is not something exclusive to

intuition, though. That inability would be the same if the exercise consisted in drawing the strokes on paper and visually perceiving them.

If we focus on *mental* representations, the requirement that the subject be able to reproduce the sequence *easily*, in the way *pictures* are reproduced, and with *certainty*, sounds much like the ability Lewis (1990) attributed to subjects who undergo new experiences: the ability to recognise, imagine and remember the experience. Indeed, the suggestion came up earlier that we regard *our* subject as able, unlike Parsons's, to recognise, imagine and remember the sequence as of the shape |||... and not ...|||...,¹³⁸ say, which Parsons's subject would seem unable to do. But note that this inability is not just because of the arbitrary character of her mental representation; it's also because a certainty of the kind involved in Wittgenstein's requirement – the certainty that a reproduction not be ...|||... whereas the original is |||... – involves, as discussed earlier, knowledge of the fifth axiom, which Parsons's subject is at any rate ignorant of, even if she did count as intuitively knowing the first four, because what she knows does not exclude the 'reproduction' – the next imagined string – being of the shape ...|||....

Contrast this with a representation of Dedekind's thought-world, as depicted in Figure 1 (repeated here as Figure 9):

¹³⁸ The dots imply infinity.

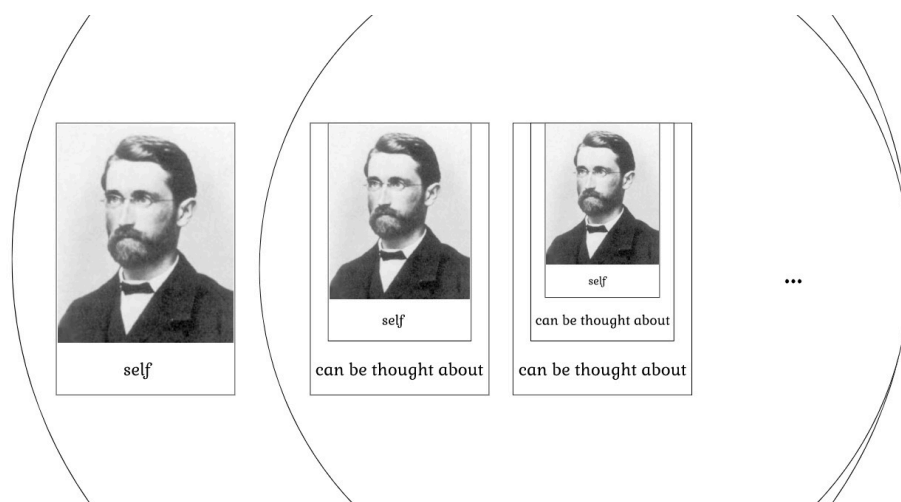


Figure 9.

Once you have in mind the structure of the thought-world, you would seem able to reproduce it in the way Wittgenstein says you should because you know it begins with a member that does not belong to the $\varphi(x)$ mapping – the self, which is not itself a thought about anything – and know that the rest of the members follow from it by the relevant iterable relation. To lose that first member would be to lose the entire sequence; there is no reproducing it without reproducing the first member. Whether the entire theorem is a surveyable representation or not – whether proofs in general are surveyable – is not a question I want to tackle here; let us focus on a mental representation of the kind one deploys when one thinks of Dedekind’s thought-world. If perception is representational, I want to suggest, then Sarah’s mental representation also seems to have been surveyable. If it is not, then we can regard the sequence of reflections that results from the arrangement of mirrors as the surveyable representation (albeit a three-dimensional one) itself. The represented object is, at any rate, like Dedekind’s in Section 66: an ω -sequence. The phenomenology of Sarah’s experience agrees with the visual vocabulary of taking the representation in ‘in one glance’, of grasping the ‘visual shape’ of what she sees, and of

grasping the ‘internal relations’ that hold the elements of the sequence together. It was clear to Sarah that there was a first member – the first reflection – that did not disappear as further elements came into view, and on which the rest of the members followed by the relevant iterable relation, just as it is clear, for a subject who thinks of Dedekind’s thought-world, that the first member is the self. And the worry that truths one knows to hold of a particular representation – e.g. a particular imagined string of strokes – may not be immediately extrapolated to claims about other structurally similar cases (what Majetschak called the ‘general’ as opposed to ‘specific’ nature of surveyable representations) seems not to have been a problem for Sarah. Sarah patently did extrapolate what she learned about the sequence of reflections to a different kind of infinite sequence: a sequence of moments in time. One might say that she did so because the knowledge she acquired was not of that particular system of reflections – or not *only* of it – but of the ‘formal connections’ that bring the elements of any ω -sequence together, as instantiated by mirror reflections *or* instants of time, beginning with a first moment – the moment she was born into her family, say – and ‘forever’. The problem of surveyability that threatens a subject like Parsons’s did not threaten Sarah.

To close, at the risk of repetition, by revisiting an earlier case. The child who sees the gigantic tree doesn’t need to go through the steps of trying to climb it to perceptually know that she can, or cannot, climb it. If we may stretch Wittgenstein’s notion to make the world and not just representations the object of surveyability, perhaps we can say that the child is able to ‘survey’ the tree and know, in a glance, the modal fact ‘I can/cannot climb this’. Possibilities are not themselves visible entities, but as we saw in Section 4, this is not an impediment for acknowledging perception to be a reliable or safe method for modal knowledge sometimes. Similarly, limited subjects like us are unable to survey arbitrarily long symbols, either in the

mind's eye or by perceiving them on a piece of paper. Yet this is not an impediment to acknowledge that we can indeed survey *potentially* infinite sequences and acquire perceptual knowledge of a modal fact like ' $\Box\forall m\Diamond\exists n \text{ Succ}(m,n)$ ', even if this gives such a sequence the appearance of actual infinity.

Conclusion

The present proposal is clearly closer to Dedekind's spirit than to Parsons's, even though only the latter focuses on visual imagery. Dedekind secured the non-arbitrariness of his representation – and so, unwittingly, its surveyability – by underpinning it with awareness of a first, primitive entity, something that cannot disappear in the construction of arbitrary symbols because it grounds the entire structure: awareness of the self. Dedekind did not choose a system that was merely characterised by any old iterable relation, such as the function that maps any object to a thought about *it*. Rather, if I am right that there was epistemological agenda here, he specifically chose a sequence that his readers would likely already be in possession of, their own thought-world. Yet the very intent of using non-mathematical concepts is what critics saw as a flaw. My proposal has reversed Dedekind's strategy and explored the possibility of thoroughly perceptual access to just such that kind of sequence. I have replaced the self with a mathematical novice of the kind MacBride envisaged, one with no prior knowledge of the principles that allow one to infer the general from the particular, and I've replaced the loaded notion of the self with a simpler, empirical notion that is close enough: a *physical* reflection of oneself. In doing this, the proposal does what I called, in the introduction, a non-starter. It has gone the other way around from Zermelo and broached the issue not in fully a priori terms, not starting with an a priori concept, the empty set, and seeing

what one can build upon it, but in fully empirical terms, asking whether one can have a visual grasp not just of the first element of an infinite sequence but also of the infinite sequence itself.

Now, this might all seem a little pointless if our aim had just been to give an idiosyncratic explication of Scoles' anecdote, but let me point out a consequence of the discussion. As Hale and Wright point out, it is widely assumed that in providing an epistemology of mathematics, '[t]wo broad approaches seem possible: *intuitional* and *intellectual*' (Hale and Wright op. cit., p. 104). Parsons' and Shapiro's views on knowledge of infinity might be roughly put in the first box, Giaquinto's in the second.¹³⁹ By pressing his challenge, MacBride seemed to demand a third approach. How can we know that all the natural numbers have a successor other than by a hunch or intuition *or* by already knowing the principles that allow inferring the general from the particular? In grounding our subject's knowledge of the above truths about the structure of the natural numbers on neither intuition nor theory but on perceptual experience, I have, in effect, challenged that wide assumption. Strictly speaking, the idea of perception being a route to mathematical knowledge had been explored by Shapiro and Giaquinto themselves before, insofar as knowledge of small numbers counts as mathematical knowledge; however, in suggesting that perceptual access to knowledge of the natural number structure is possible, I have pushed for a stronger view, thereby tackling what I claimed was the heart of MacBride's challenge: the issue of grasp of infinity.

Before closing, some reflection is in order on why it should make a difference that this grasp is delivered by experience and not the intuitional or intellectual

¹³⁹ I'm referring to the hunch and/or innate knowledge we saw Shapiro suggest accounting for the jump to the general thought. He also offers another way in earlier work, 'characterisation', which is straightforwardly intellectual (Hale and Wright (op. cit., p. 112-3) call it a 'canonical axiomatic description'), but which is not relevant to our discussion.

approach. Earlier, I phrased the gap MacBride pointed out in Shapiro's epistemology as the need to clarify how a mathematical novice can be justified in holding a belief that quantifies over all the elements of an infinite set. Shapiro's suggestion that the particular-to-general jump could be grounded in innate knowledge or a hunch was not unsatisfactory to MacBride because those two cannot justify; innate knowledge and hunches may well be safe or reliable methods for acquiring knowledge, after all. His dissatisfaction, as I read him, was that the general-belief-forming process was not clearly laid out by Shapiro. Indeed, this dissatisfaction also applied to the particular-belief-forming process, which Shapiro did think involve perception: MacBride demanded an illuminating philosophical account of how we know small, finite structures, and form beliefs about particular natural numbers in this way. Giaquinto helped with this. On his view, we form beliefs about small natural numbers by abstraction from our perceptual knowledge of small collections. Hunches may well be as safe or reliable an epistemic method as perception, but perception seems to be more theoretically tractable. (Following Russell, Giaquinto thinks it is, anyway. We have a way, 'abstraction', to think about how perception may yield particular natural number beliefs.) Giaquinto's view did not help with the more pressing of MacBride's demands, though, which to see how the structuralist framework explained how we know infinite structures, in particular the structure of the natural numbers. The reason Giaquinto did not help here is that his answer to this further question relies on theory, which is what a mathematical novice, by hypothesis, does not have. Another answer, not considered in the introduction along with Giaquinto's but mentioned in the latter sections was Parson's, according to whom we can know the natural number structure by intuition. Now, this did not presuppose theory, hence was not vulnerable to MacBride's mathematical novice objection. But we saw it had problems of its own. One important source of those problems was the indeterminate,

vague or arbitrary nature of intuition's deliverances. Page, Hale and Wright voiced concerns about how they, while being non-particular, fell short of the requisite generality. I raised a further issue, too, that an arbitrarily long string of strokes, tokened in the mind's eye or on a concrete drawing on paper, is not surveyable. A perceptual approach is invulnerable to these objections: it does not presuppose theory on the part of the subject, nor is it indeterminate (or vague or arbitrary), or unsurveyable.

But to say that perception can play a role in a mathematical novice's formation of a belief quantifying over all the elements of an infinite set – a set, specifically, structured like the natural numbers – is not to say that the intellect and intuition cannot play such a role, nor that the above problems with the intellectual and intuitional approaches deprive the intellect or intuition of their justificatory properties. Those problems deprive *us* of clarity with respect to how intuition and the intellect may work. A holist like Shapiro may be comfortable with such unclarity. The justification those of us who didn't grow up Mormons enjoy for our non-perceptually-formed beliefs about all the natural numbers would, for him, not be threatened. All I wanted in this paper was to convince the reader that a mathematical novice can be perceptually justified in holding such a general belief, not that she cannot be *otherwise* justified.

Now, my strategy to do that was to make the case for a visual experience *as of* actual infinity, which I unpacked by relying on the modal nature of the sequence of reflections and the ability of perception to present to us, even in a non-extended way, sequential phenomena. Does this experience having an illusory aspect threaten the subject's justification? I think it does not. The illusion concerned the actuality of the infinity being presented; accordingly, the subject was not claimed to have had a veridical perception of actual infinity. She was not justified in believing anything like

the proposition, 'I have an actual infinity before me'. The beliefs we're seeking justification for are those related to the particular-to-general jump. So, the relevant propositions are the following. Because reflection on the mirror set-up suggests our subject became aware of the potential nature of the sequence of reflections, the first proposition I brought up was the Aristotelian one, ' $\Box \forall m \Diamond \exists n \text{ Succ}(m,n)$ ', which may turn out to be encoded by something else than a conscious belief. But in detecting the relation that held between reflections, our subject detected the way the system was structured, which is the way the natural numbers are structured. Structural properties are attributed to a system regardless of its modal nature; it is this that allows Linnebo and Shapiro to argue for an equivalence between modal and non-modal formulations of arithmetic. Perception itself gave our subject access to the way the system of reflections was structured; hence, her belief in the propositions that describe such a structure, the Dedekind-Peano axioms, is justified by perception. Thus Sarah enjoyed not one but two improbable things: a visual phenomenology as of actual infinity (which we have long thought impossible, since actual infinities do not exist), and the acquisition of genuine knowledge of the axioms that describe it (not just a false belief that there was an actual infinity before her).¹⁴⁰

True: the knowledge and phenomenology discussed in our test case involved rather peculiar circumstances. So perhaps the interest of it lies in the questions it opens up. Is there some epistemological value in taking seriously MacBride's challenge and refuse to rely simply, as Shapiro does, on a holistic view of justification? To cling, that is, to the demand for a more tractable knowledge formation method for mathematical novices than a hunch? If so, is there a role for

¹⁴⁰ Another way to see how the illusory dimension of the experience does not threaten our subject's justification to hold the relevant beliefs is to see its epistemic upshot as an ability: the ability to conceive of a model for the Dedekind-Peano axioms, which ability is what typically justifies subjects in holding a general or universal belief about all the elements of an infinite set.

more ordinary cases of perceptual experience in our acquisition of knowledge about mathematical reality? I suspect there is a positive answer to these questions. This is but the beginning of a research avenue for future work.

References

- Aristotle. 1942. *The Basic Works of Aristotle*. NY: Random House
- Ben-Moshe, N. 2021. On Wittgenstein's Notion of a Surveyable Representation: Rituals, Aesthetics, and Aspect-Perception. *Australasian Journal of Philosophy*, DOI: 10.1080/00048402.2021.1952458
- Bengson, J. and Moffett, M. A.. 2012. Two Conceptions of Mind and Action: Knowledge How and the Philosophical Theory of Intelligence. In *Knowing How: Essays on Knowledge, Mind, and Action*. Oxford: Oxford University Press
- van Benthem, J. 2006. Epistemic logic and epistemology: the state of their affairs. *Philosophical Studies* 128: 49–76
- Bengson, J. 2016. Practical perception and intelligent action. *Philosophical Issues*, 26, doi: 10.1111/phs.12081
- Bengson, J. and Moffett, M. 2012. Nonpropositional Intellectualism. In *Knowing How: Essays on Knowledge, Mind, and Action*. Oxford: OUP
- Davidson, D. 1980. *Essays on Actions and Events*. Oxford: OUP
- Broad, C. D. 1923. *Scientific Thought*. London: Kegan Paul
- Broesch, Callaghan et al. 2010. Cultural Variations in Children's Mirror Self-Recognition. *Journal of Cross-Cultural Psychology* XX(X), doi: 10.1177/0022022110381114

- Brownstein, M. 2014. Rationalizing flow: agency in skilled unreflective action. *Philos Stud* 168:545–568
- Brownstein, M. and Michaelson, E. 2015. Doing without believing: Intellectualism, knowledge-how, and belief-attribution. *Synthese* 193, 2815–2836
- Buijsman, S. 2020. Two roads to the successor axiom. *Synthese* 197, pp. 1241–1261
- Burgess, J. and Rosen, G. 1997. *A Subject with No Object: Strategies for Nominalistic Interpretation of Mathematics*. Oxford: OUP
- Cantor, G. 1874. Über eine Eigenschaft des Inbegriffes aller reellen algebraischen Zahlen. *Journal für die reine und angewandte Mathematik* 77, 258–62.
- Chalmers, D. 2002. Does conceivability entail possibility?. In T. Gendler and J. Hawthorne (eds.). *Conceivability and Possibility*. Oxford: Oxford University Press
- Crane, T. and French, C. 2017. The Problem of Perception. In Edward N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy (Winter 2016 Edition)*
- Davidson, D. 1970. Action and reaction. *Inquiry* 13, pp. 137–48
- Davies, J. E. 2017. Towards a Theory of Singular Thought About Abstract Mathematical Objects. *Synthese*, doi: 10.1007/s11229-017-1644-0
- Dretske, F. 1969. *Seeing and Knowing*. London: Routledge and Kegan Paul
- Dummett, M. 1978. The philosophical basis of intuitionistic logic. In *Truth and Other Enigmas*. Cambridge, MA: Harvard University Press, pp. 215–247
- Dummett, M. 1991. *Frege: Philosophy of Mathematics*. Cambridge, MA: Harvard University Press
- Drayson, Z. 2018. Direct perception and the predictive mind. *Philos Stud* 175, 3145–316. <https://doi.org/10.1007/s11098-017-0999-x>
- Ebert, A. and Rossberg, M., eds. 2016. *Abstractionism: Essays in the Philosophy of Mathematics*. Oxford: OUP

- Fodor, J. 1990. A theory of content II. In Jerry Fodor, *A Theory of Content*.
Cambridge: MIT Press
- Fodor, J. 1994. Concepts: A potboiler. *Cognition*, 50, 95–113
- Floyd, J. 2021. *Wittgenstein's Philosophy of Mathematics*. Cambridge: CUP
- Frege, G. 1884/1953. Translated by J.L. Austin. *The Foundations of Arithmetic*.
Oxford: Blackwell
- Gbur, G. 2011. Infinity is weird... even in infinity mirrors! Skulls in the Stars: The
intersection of physics, optics, history and pulp fiction. Available at
[https://skullsinthestars.com/2011/07/30/infinity-is-weird-even-in-infinity-
mirrors/](https://skullsinthestars.com/2011/07/30/infinity-is-weird-even-in-infinity-mirrors/)
- Geach, P. 1957. *Mental Acts*. London: Routledge
- Gendler, T. S., and Hawthorne, J. (eds.). 2002. *Conceivability and Possibility*. New
York: Oxford University Press
- Giaquinto, M. 2001. Knowing Numbers. *Journal of Philosophy*, xcvi
- Giaquinto, M. 2007. *Visual thinking in mathematics*. Oxford: OUP
- Giaquinto, M. 2012. Russell on Knowledge of Universals by Acquaintance.
Philosophy, 87, doi:10.1017/S0031819112000381
- Ginsborg, H. 2006. Empirical concepts and the content of experience. *European
Journal of Philosophy*, 14(3), doi: 10.1111/j.1468-0378.2006.00230.x
- Gow, L. 2019. Everything is clear: All perceptual experiences are transparent.
European Journal of Philosophy, 27, 412–425.
- Gow, L. 2020. A new theory of absence experience. *European Journal of Philosophy*,
29, 168–181.
- Hale, B., and Wright, C. 2002. Benacerraf's dilemma revisited. *European Journal of
Philosophy*, 10, 101–129.

- Hawthorne, J. 2005. The Case for Closure. In Steup, M. and Sosa, E. (eds.), *Contemporary Debates in Epistemology*. Malden, MA: Blackwell.
- Higginbotham, J. 1998. Conceptual Competence. *Philosophical Issues*, 9, pp.149-152
- Hintikka, J. 1962. *Knowledge and Belief*. Ithaca, NY: Cornell University Press
- Hintikka, J. 1966. Aristotelian Infinity. *Philosophical Review*, 75, pp. 197-218
- Hintikka, J. 1969. Semantics for propositional attitudes. In J.W. Davis, D.J. Hockney & W.K. Wilson (eds.), *Philosophical Logic*. Dordrecht: Reidel, pp. 21–45
- Hintikka, J. 1970. Objects of knowledge and belief: acquaintances and public figures. *Journal of Philosophy*, 67, pp. 869–83
- Iaquinto, S. and Spolaore, G. Forthcoming. Outline for a logic of knowledge of acquaintance. *Analysis*, doi: 10.1093/analys/any005
- James, W. 1890/1983. *The Principles of Psychology*. Cambridge, MA: Harvard University Press
- Jackson, F. 1986. What Mary Didn't Know. *Journal of Philosophy*, 83(5), pp. 291-295
- Jeshion, R. 2010. Acquaintance, semantic instrumentalism, and cognitivism. In R. Jeshion (ed.), *New essays on singular thought*. Oxford: Oxford University Press, pp. 105–140
- Jeshion, R. 2017. Intuiting the infinite. *Philosophical Studies*, 171, doi: 10.1007/s11098-013-0274-8
- Klev, A. 2018. A Road Map of Dedekind's Theorem 66. *HOPOS The Journal of the International Society for the History of Philosophy of Science* 8, doi: 10.1086/698660
- Korbmacher, J. and Schiemer, G. 2017. What are Structural Properties? *Philosophia Mathematica*, nkx011, doi: 10.1093/philmat/nkx011

- Lear, J. 1979. Aristotelian Infinity. *Proceedings of the Aristotelian Society*, 80, 187–210
- Lewis, D. 1990. What experience teaches. In W.G. Lycan (ed.). *Mind and cognition*.
Oxford: Blackwell
- Linnebo, Ø. 2008. Structuralism and the Notion of Dependence. *Philosophical Quarterly* Vol. 58, No. 230, pp. 59-79
- Linnebo, Ø. 2013. The potential hierarchy of sets. *The Review of Symbolic Logic*, 6(2), 205-228, doi:10.1017/S1755020313000014
- Linnebo, Ø. and Shapiro, S. 2017. Actual and Potential Infinity. *Nous*, doi: 10.1111/nous.12208
- Loar, B. 1990. Phenomenal States. *Philosophical Perspectives*, 4, 81-108, doi: 10.2307/2214188
- Luper, S. 2020. Epistemic closure. *Stanford Encyclopedia of Philosophy*
- Michaelson, E. Ms. Iffy Know How
- MacBride, F. 2008. Can Ante Rem Structuralism Solve the Access Problem? *The Philosophical Quarterly*, 58, pp.155-64.
- Mac Cumhaill, C. 2011. Specular space. *Proceedings of the Aristotelian Society*, cxi(3), 487–495. doi:10.1111/j.1467-9264.2011.00320.x
- Maddy, P. 1980. Perception and Mathematical Intuition. *The Philosophical Review* 89 (2), pp. 163-196
- Maddy, P. 1997. *Naturalism in Mathematics*. Oxford: OUP
- Maddy, P. 2007. *Second Philosophy: a Naturalistic Method*. Oxford: OUP
- Majetschak, S. 2016. Survey and Surveyability: Remarks on two central notions in Wittgenstein's later philosophy. *Wittgenstein-Studien* 7, doi: 10.1515/witt-2016-0106
- Marshall, O. 2017. Giaquinto on Acquaintance with Numbers. *Journal of Philosophy* 114 (1):43-55

- Mühlhölzer, F. 2005. "A Mathematical Proof Must Be Surveyable" What Wittgenstein Meant By This And What It Implies. *Grazer Philosophische Studien* 71, pp. 57–86
- Mühlhölzer, F. 2010. Mathematical Intuition and Natural Numbers: A Critical Discussion. *Erkenntnis* 73, doi: 10.1007/s10670-010-9225-3
- Moore, A. W. 2001. *The Infinite*. Oxford: OUP
- Papineau, D. 2002. *Thinking about Consciousness*, Oxford: Oxford University Press
- Papineau, D. 2009. Phenomenal and perceptual consciousness. In Torin Alter & Sven Walter (eds.), *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*. Oxford: Oxford University Press, pp. 111--144
- Parsons, C. 1980. Mathematical Intuition. *Proceedings of the Aristotelian Society*, 80, pp. 145-68.
- Parsons, C. 1986. Intuition in Constructive Mathematics. In *Language, Mind, and Logic*. Cambridge: Cambridge University Press, pp. 211-229.
- Parsons, C. 1990. The structuralist view of mathematical objects. *Synthese* 84, pp. 303–346
- Parsons, C. 2007. *Mathematical Thought and its Objects*. Cambridge: CUP
- Peacocke, C. 1992. *A Study of Concepts*. Cambridge: MIT Press
- Peacocke, A. 2020. How literature expands your imagination. *Philosophy and Phenomenological Research*
- Pehkonen, E., et al. 2006. Infinity of numbers: how students understand it. In Novotná, J., Moraová, H., Krátká, M. & Stehlíková, N. (eds.). *Proceedings 30th Conference of the International Group for the Psychology of Mathematics Education*, 4, pp. 345-352
- Peregrin, J. Ms. Logic and nothing else. <https://philpapers.org/rec/PERH-3>

- Piazza, M. et al. 2011. Subitizing reflects visuo-spatial object individuation capacity. *Cognition* 121, pp.147-153
- Prior, A. 1957. *Time and Modality*. Oxford: Clarendon Press
- Prosser, S. 2016. Is Experience Temporally Extended? In *Experiencing Time*. Oxford: Oxford University Press
- Pryor, J. 2000. The Skeptic and the Dogmatist. *Noûs* 34: 517–49.
- Railo, H. et al. 2008. The role of attention in subitizing. *Cognition* 107, pp.82-104
- Railo, H. and Hannula-Sormunen, M. 2012. Subitizing. *Encyclopedia of the Sciences of Learning*, pp. 3233-3235
- Resnik, M. 1996. Structural relativity. *Philosophia Mathematica* (3) 3, 83–99
- Rips, L. and Asmuth, J. 2007. Mathematical Induction and Induction in Mathematics. In A. Feeney and E. Heit (Eds.), *Inductive Reasoning*. Cambridge: Cambridge University Press
- Russell, B. 1911. Knowledge by acquaintance and knowledge by description. *Proceedings of the Aristotelian Society*, 11(5), 108–28.
- Salmon, N. 2018. Julius Caesar and the numbers. *Philosophical Studies* 175, doi: 10.1007/s11098-017-0927-0
- Scoles, S. 2016. How thinking about infinity changes kids' brains on math. *Aeon*. Available at <https://aeon.co/ideas/how-thinking-about-infinity-changes-kids-brains-on-math>
- Shapiro, S. 1997. *Philosophy of Mathematics: Structure and Ontology*. Oxford: Oxford University Press
- Shapiro, S. 2011. Epistemology of Mathematics: What are the Questions? What count as Answers? *The Philosophical Quarterly*, 61, doi: 10.1111/j.1467-9213.2010.658.x

- Shapiro, S. and Wright, C. 2006. All Things Indefinitely Extensible. In Rayo and Uzquiano (eds.), *Absolute Generality*. Oxford: OUP
- Steenhagen, M. 2017. False reflections. *Philosophical Studies*, 174, doi: 10.1007/s11098-016-0752-x
- Soteriou, M. 2013. *The Mind's Construction*. Oxford: OUP
- Soysal, Zeynep. 2017. Why is the universe of sets not a set? *Synthese*, doi: 10.1007/s11229-017-1513-x
- Stanley, J. & Williamson, T. 2001. Knowing How. *Journal of Philosophy* 98 (8):411-444
- Stanley, J. & Williamson, T. 2016. Skill. *Noûs* 51:4, 713–726
- Tall, D. 1980. Mathematical intuition, with special reference to limiting processes. *Proceedings of the Fourth International Congress on Mathematical Education*. Berkeley, CA. 170176. Available at: <http://www.warwick.ac.uk/staff/david.tall>
- Tall, D. 2001. A child thinking about infinity. *Journal of Mathematical Behavior*, 20, pp.7-19
- Weiskopf, D. and Bechtel, W. 2004. Remarks on Fodor on Having Concepts. *Mind and Language* 19, 48-56
- Williamson, T. 2007a. Philosophical Knowledge and Knowledge of Counterfactuals. In C. Beyer and A. Burri (eds), *Philosophical Knowledge: its Possibility and Scope*. Amsterdam: Rodopi
- Williamson, T. 2007b. *The Philosophy of Philosophy*. Malden, MA: Wiley-Blackwell
- Yablo, S. 2002. Coulda, Woulda, Shoulda. In T. Gendler and J. Hawthorne (eds.). *Conceivability and Possibility*. Oxford: Oxford University Press

What do we see when we see absences?

Abstract

A growing body of literature about the phenomenological and epistemic role of structural features of experience, as well as on the topic of absence experience itself, point toward the view that absence experience is non-veridical. Here I attempt to challenge that result.

What is a ghost? A terrible event condemned to repeat itself? An instant of pain, perhaps. Something dead that appears at times alive. An emotion suspended in time, like a blurry photograph, like an insect trapped in amber.

– Guillermo del Toro

Introduction

Nobody believes little Carlos when he says, in Guillermo del Toro's *The Devil's Backbone* (2001), that he has been seeing a ghost at night. These are the nineteen-forties, and this, a group of Spanish republicans. Grown-ups believe in science here. They are fighting, in fact, the old mindset that made people believe royals had divine rights, that magical healing worked (the head of the group is a physician), and that there exist such things as ghosts. But what if ghosts were not viewed by these sceptics to be substance-like, yet immaterial things? Not actually dead people walking but the living's own evocation of the past? Not bodily but happening or fact-like entities, or as the narrator puts it, 'terrible events'? Would the sceptics be willing to accept the existence of ghosts in that case? Would they be able, like Carlos, to *see* them?

It takes a conceptual shift of this kind for the unwilling to see what is haunting them in this ghost story. The present paper attempts a shift more or less like that. I want to convince the reader to believe me when I say that I can see something like ghosts – that I can see absences. To say this is to say that my visual experience of them is veridical. But what does it mean for a perceptual experience to be veridical? The answer is usually given in contrast with two other notions: those of hallucinatory and illusory experience. If someone is having a visual experience as of a

white horse, it may be that she is really seeing a white horse (in which case the experience is veridical), that she is seeing a black horse (in which case it is illusory), or that she is seeing no object at all (in which case it is a hallucination). At a minimum, for an experience to be neither illusory nor hallucinatory, the presence of the perceived object before the subject seems to be required; and the presence of an object entails, of course, its existence. Given this, there are objects that we know we don't, cannot, perceive: objects such that, if they appear in any experience, then it's likely to be non-veridical. Experiences as of unicorns, say, or as of pink elephants. Unicorns do not exist; they are the object of hallucination. Elephants do but they are typically not pink; pink elephants are likely objects of illusion. But there are cases that aren't so clear-cut. Absences are one such case. There is no consensus on whether our ontology should accept them as worldly constituents, and if it should, whether they should be recognised as having observable properties such that we could see them. I will show that most prominent views on the matter of absence experience lean toward diagnosing it (explicitly or implicitly, and for different reasons) as non-veridical. This tendency can be linked to a couple of independent views: that absence perception is a species of object perception (where the dichotomy is sharp between this and another kind of perceptual state, fact perception), and that there are features of absence experience that are explained not by appeal to what is in view but to the subject herself. The problem with treating absence experience as non-veridical is, however, that it doesn't do justice to ordinary discourse about it, according to which we take ourselves to be enjoying veridical experiences whenever we see absences. A view that honours the relevance of facts about the subject while dispensing with the assumption that absence experience has to fit either the object or the fact-perception model would open the possibility of regarding absence experience as something other than hallucinatory or illusory, i.e. as veridical, because there would be no posited

object, non-existent or misperceived, to non-veridically see. The project is ambitious, and in the length of this paper I cannot carry it out in full detail. My aim is to make the idea that absences are veridical at least intelligible, to make room for it, to show that they are on similar footing as other entities we do unproblematically recognise as seen. My aim is, in other words, to effect a shift in the reader that opens her up to regarding absence experience in the way sceptics regarded, after their shift in perspective, Carlos's experience of ghosts.

1. Absence experience as non-veridical

1.1

Let me start by addressing the point that we ordinarily take ourselves to enjoy veridical experiences whenever we see absences, and that when we report such experiences, we mean to say something true. This is likely an intuitive point, but to bring it home, consider an absence sceptic. A motivation behind such a sceptic's view can be found in Clare Mac Cumhaill's (2018) way of bringing out the puzzling nature of absence experience. Mac Cumhaill starts from the thought that if we really do see absences, there must be a distinctive phenomenology to the experience. But 'can there be "something it's like" to perceive absences? Emboldened by philosophical theory, the answer might seem plain: "no"' (op cit., p. 34). The 'philosophical theory' that drives this scepticism stems from the thought that experience is 'transparent' or 'diaphanous', that is, devoid of any intrinsic qualities, qualities that may be attributable to the experience itself rather to what it is an experience of. 'The thought is that, in attempting to discover by introspection the presence of such

qualities, one only ever manages to “see through” the experience to the worldly things and features that one’s experience is of (op cit., p. 35). If introspection reveals only mind-independent objects and properties, then, ‘there being no sensible properties for introspection to alight upon when absences are putatively perceived, can it be granted that there is, after all, ‘something it’s like’ to perceive absence?’ (ibid.).

Mac Cumhaill goes on to answer the question *almost* in the positive (why ‘almost’ will become clear later). For now, let us stay with the sceptic who remains convinced that there is nothing it is like to see absences, since there are no such things out there for us to see. Whenever we report a seeing of absence, this view might go, we are speaking non-literally. When Sartre says he sees Pierre’s absence from the café, say, he is literally saying a falsehood.¹⁴¹ There is no such thing as Pierre’s absence in the café for Sartre to see. Instead, that report should be taken as shorthand or metaphor, or some other truth-value-shifting device, for a negative, true statement: that (while he sees other things) Sartre does not see Pierre there. Similarly, when someone unfamiliar with Swiss cheese says, astonished, that they see holes in the cheese, they should be taken to mean that they don’t see cheese in certain spherical locations in an otherwise cheese-filled region. And when someone says they see total darkness, they should be taken to mean that they don’t see any light in that place.

As Mac Cumhaill notes, this line of thought is given voice by Sartre himself:

¹⁴¹ Assuming a view of definite descriptions on which ‘the absence of Pierre’ asserts or entails the existence of such an entity (Russell 1905). On a competing view on which definite descriptions presuppose it rather than assert it (Strawson 1950), Sartre’s statement may have an undefined truth value, or as Strawson says, generate a ‘truth-value gap’. I pass over this complication and treat such statements as false because, unlike the uses of definite descriptions in which, as Strawson pointed out, one does not mean to commit oneself ontologically (e.g. ‘the Greek god of Thunder is bearded’), in discourse about absence perception we do typically mean to refer to something we have seen.

I look at the room, the patrons, and I say, “He is not here”. Is there an intuition of Pierre’s absence, or does negation indeed enter in only with judgment? At first sight it seems absurd to speak here of intuition since to be exact there could not be an intuition of nothing and since the absence of Pierre is this nothing. Popular consciousness, however, bears witness to this intuition. Do we not say, for example, “I suddenly saw that he was not there”. Is this just a matter of misplacing (*déplacement*) the negation? (Sartre 1943, quoted in Mac Cumhaill op cit., p. 37).

When describing an experience of absence, should we put the negation on the side of the subject – describe it as the subject’s failure to see something – rather than on the side of the scene – describe it as some negative thing, a privation, being perceived? Might the fact that ‘popular consciousness bears witness’ to the latter be evidence that in experiencing absences we are all victims of collective illusion, or worse, collective hallucination? Perhaps we are. But this would call for a way of explaining how it is that we seem to routinely enjoy experiences that we report to be about something, an absence, that isn’t really there.

A brief look at absence perception discourse will both illustrate this last point and give us an overview of the varieties of absence experience.¹⁴² Let me propose three paradigm cases, and list them by order of their seeming contributions to general phenomenology. Seeing holes, seeing absences of individuals, and seeing total darkness. (A fourth case will emerge below.) The case of seeing each of these generates different issues, but they all seem to be determinate cases of one determinable type, seeing absence. Seeing the absence of some substance in a region that is surrounded by the presence of it (holes), seeing the absence of individuals

¹⁴² In what follows I focus on visual experience, but ultimately my proposal is meant to be generalisable.

from times and places (e.g. the absence of Sartre's friend Pierre from the café), and seeing the absence of light (darkness).¹⁴³

The contribution of seeing holes to the character of our experience tends to be limited: seeing the material around the hole – its 'host' – tends to contribute more. In their limited phenomenological impact, holes seem akin to ordinary particular objects. Indeed, they can be viewed, like them, to have volume, size, and a precise spatiotemporal location. I can think 'I see something big there' while pointing at the hole's location, or think, 'that is a big one!', demonstratively referring to the hole. Now, there are salient differences between holes and darkness. Whereas the relation between darkness and absence is straightforward, that holes of is not. One might intuitively identify holes with absences ('holes [...] are] absences of matter': Sorensen 2015, p. 542), such that to see a hole *is* to see an absence. Or one might, rather, explain holes by reference to absences, such that to see a hole is not just to see an absence: perhaps holes are not identical to but rather 'result from the absence of matter' (Lewis and Lewis 1970, p. 260); perhaps a hole is a place where an absence is (Finn 2018). At any rate, there seems to be a correlation between experiencing the absence that is contained in, identical to, or part of what produces the hole and experiencing the hole's 'host'. In what follows, I want to stay clear of issues in the metaphysics of holes and focus on the absence to which a hole is related in any of these ways. However absence-related, holes are arguably as vulnerable as absences to the suspicions of a parsimony-advocating sceptic.

Consider now the other end of the spectrum, where our phenomenology seems is determined by our experience of the absence: seeing total darkness.

¹⁴³ One might think that because some of these cases have more obvious 'phenomenal reality' than others, as O'Shaughnessy puts it, our experience of them should count as experience of positive entities. Total darkness, for instance, can be regarded in this way, in contrast with the experience of silence. Because I want to cling to the intuition that they are all experiences of absences, my proposal will treat them equally as such.

Darkness, too, is attributed observable properties. If I turn the lights off, I can see total darkness and its apparent colour, black, upon which I seem epistemically entitled to believe that darkness is black. What's more, darkness can be seen in a photograph, just like light and illuminated things can be. If the photograph in question is not totally but just partially dark, I am aware of the partial absence of light either in the photographed scene, if the photo is not underexposed, or the photographic image itself, if it is. If the photo is totally dark (because it looks totally black), then I am aware of the total absence of light, either in the original scene or in the image. And darkness has an observable extension. It can fill an entire room, just like light can, or it can concentrate in some corners of it and dissipate towards an open window. In this, darkness appears to be akin to light itself or water – that is, to non-particular stuff.

Now, most phenomenologically arresting of all might be the absence of individuals. Take the following excerpt from Wisława Szymborska's 'Cat in an Empty Apartment':

Die – you can't do that to a cat.
Since what can a cat do
in an *empty* apartment?
Climb the walls?
Rub up against the furniture?
Nothing seems different here
but nothing is the same.
Nothing's been moved
but there's more space.
[...]

Someone was always, always here,
then suddenly disappeared
and stubbornly stays disappeared.

The absence of the cat's owner is felt by the cat to be extended, much like the absence of light can be. It can fill the entire apartment. Without the owner, there is more space, so much, in fact, that the apartment is felt to be empty. More speculatively, the absence seems to have properties of the kind we experience persons as having. In the cat's case, the owner's absence is felt to be stubborn. Sartre may have been similarly impacted by Pierre's absence, just like Pierre himself could affect him – Sartre may have been annoyed or struck or saddened by it.

But absences of individuals, like holes, raise identity questions. Do the other patrons see Pierre's absence? If one is inclined to answer that they don't, then the question arises whether that is because Pierre's absence is identical to the absence of many other things, such that Sartre is not really seeing Pierre's absence simpliciter but rather seeing that entity – which is identical to the absence of many people or things – *as* the absence of Pierre. An account of the identity conditions for absences is likely linked to an account of what ontological category they belong to. For example, if absences of individuals are identified with facts (Jago and Barker 2011), it seems easier to regard Pierre's absence as distinct from Simone's absence.¹⁴⁴ If absences are concrete entities, though, that is, tangible objects of some kind, then it is harder to see how several distinct such entities – perhaps infinitely many – could all fit in one small café.

¹⁴⁴ This is so even if facts are regarded as spatiotemporal entities, as in Jago and Barker's (2011) view, because the constituents of each fact differ – Pierre partly constitutes the first fact, and Simone the second.

Eventually, my proposal will not give an account of the identity conditions of absences, but it will dovetail with the absences of Pierre and Simone being distinct, and not fit with a view on which they are identical and one only sees that absence *as* the absence of one or another friend or lover. To elaborate on this a bit: the latter kind of view might be motivated by the following thought. Imagine Sartre sits and turns to the empty chair beside him. The absence he sees might be thought to be located roughly in the same region Pierre would occupy if he'd been sitting on that chair. But the exact same spatiotemporal region, which Sartre sees as empty of Pierre, may be seen by someone else (or Sartre himself in different circumstances) as empty of Simone. A view on which the empty region is the mind-independent absence Sartre sees would have trouble accommodating a case in which the appointment was not at a café but at the Montmartre funicular, say, at some given time, and in which Sartre sees Pierre's absence in seeing that none of the other occupants is his friend. The car is packed, and there is no empty spot to identify with the absence here.

So, while total darkness, holes and the absences of individuals seem all determinates of the determinable absence, each raises different ontological issues. In this, though, they are as diverse as present entities can be. Indeed, this very variety seems to be perceptible: as my brief survey shows, the variety of predicates we employ in reporting our experience of absences is as wide as the variety of predicates we employ in reporting our experience of present things. The fact that even cats perceive absences, assuming Szymborska is right, speaks to the scope of the 'popular consciousness' Sartre concedes, for a moment, could be mistaken. For popular consciousness to be mistaken about absence perception would mean for perceptual reports and predications like the above to be false. Sentences like 'total darkness is black', 'that hole is big' and 'I can see Pierre's absence' would all be false because there would be, on this view, no such entities out there that we could see. Rather,

those beliefs would be explained by our seeing an almost all-light-absorbing room, stuff surrounding a certain region, and people other than Pierre at the café, then forming beliefs about those scenarios that, while seeming to refer to some *thing* that is black, big, and seen, actually are just about something not black, not big, not seen failing to present in the context.

So, what does it come to, exactly, to say that absence experience is not veridical, and that speech in which absences are reported as seen and in which predications are made of them is false? I'll start by offering a brief summary of two recent ways of answering this question. One of them is explicitly a non-veridical view of absence experience; the other, as I will show, ends up predicting it to be so as well.

1.2

Stephen Mumford's (forthcoming, §7) view is straightforward: it is the view that absence experience is 'an illusion, albeit a useful one, of the mind'. His reason to claim so is straightforward too: '[w]hat is not there cannot be seen. Only presences – what is there – can be seen. Yet absences *seem* experientially accessible. They have a phenomenology, even if an elusive one' (my italics). On Mumford's view, it is an adaptive advantage of our cognitive system that we experience darkness, holes and people's absences (he considers all three kinds) as constituents of the world along with light, substances and people themselves. Take a group of our ancestors as they're walking through a jungle. In order to detect a hole that lies before them, it would be more efficacious for them to experience the hole substantially, as an *object* to be avoided in the way other objects are, e.g. predators, rather than having to deduce from what they veridically see – the positive, corresponding earth-lining – that they would fall if they walk on. 'And yet this [experiencing the absence] must be an

illusion', Mumford writes. 'What is not there cannot be seen. Only presences – what is there – can be seen' (Mumford op cit., §7.5).

The standard schema characterises illusion as seeing an F as a G: seeing something that doesn't look in any way, for example, as positively black appearing, in total darkness' case; or perhaps seeing a region of space unoccupied by material as occupied by something, a hole; etc. Mumford doesn't tell us how his view fits this schema, and so not how exactly something that is F – non-existent or not-present – appears, on his view, to perceivers as G – existent or present, and appearing to have observable properties. His view that to experience absences is to experience something that 'is not there' seems more like a view of it as hallucination. But the key to Mumford's use of the term 'illusion' – and the core of his view altogether – is his reliance on Daniel Dennett's notion of 'user illusion'.

Mumford summarises the idea in this way:

The sort of illusion that is in mind is like the user illusion we have when we operate a computer. The computer's workings are understood by programmers in terms of machine code and physically are a system of on-off switches. They also have a user-friendly interface for those of us who understand none of what lies beneath. Dennett's suggestion is that there is something similar occurring with the minds that we use in order to navigate our way around the world (ibid.).

The illusion consists not in seeing an F as a G but in experiencing ourselves to be seeing some *thing*, Pierre's absence, for example, and believing ourselves to be referring to it when we say things like 'I can see Peter's absence', which in fact does not report a seeing of that object – that absence – but is instead the result of an

inference, ‘an automatic and direct non-deductive inference drawn from what is seen to what is absent’ (ibid.). The inference is non-deductive because ‘[t]here is no deductive path from what is seen to what is absent’, since the presence of anything anywhere is basically compatible with the presence of something else there. For example: inferring the fact that Pierre is absent from the fact that Paul, Jean, the waiters, etc. are present in the same place is not warranted. Pierre could be hiding behind Paul at any time, or behind the viewer. So a further fact from the facts the viewer knows by perception would be needed: ‘[t]he additional required fact is negative: that there are no more facts’. But of course, ‘[a] general or totality fact is another negative, and how is this general absence known?’

Mumford’s emphasis on the non-deductive nature of the inference helps to bolster his view that generally speaking, the mind engages in a trade-off between epistemic security and rapidity in its delivery of beliefs about the world as we navigate it. The beliefs about absences we form during putative absence experience are problematic twice over: first, because they refer to entities that aren’t actually there for us to see; second, because even understood as the result of an inferential process, the relevant inference does not enjoy the status of deduction. For the trade-off to yield its returns, then, we need to *experience* such beliefs as being just as safe as the beliefs we form about an apple that we have seen, touched, and eaten; we need to experience them, that is, as being about an object with which we’re perceptually acquainted like any other. Only in this way could such beliefs guide action in the way necessary to survive by, avoiding, for example, walking over holes, relying on people who aren’t there, failing to slow down and trust the rest of our senses when an expanse of darkness blocks our sight.

Mumford’s view is geared toward respecting data that other possible views would seem to conflict with. A view on which we veridically see absences, for

instance, would seem to conflict with a causal condition on perception, which states that ‘to qualify as a genuine perception, the object perceived must be the cause of the perception’ (Mumford op cit., §7.3.3).¹⁴⁵ Absences being, if existent, incapable of bouncing off light that can strike your retina, pressing itself onto your body, or affect a change on your tongue, they cannot be objects of genuine perception, on this view.¹⁴⁶

A slight digression may be in order at this point. That absences are unfit to enter perceptual relations in virtue of their causal inertness seems too significant a point to list among others against their perceptibility. This is, of course, a source of epistemological scepticism in the case of other causally inert entities too, for example, abstract objects. The case of absences is special among these because, unlike with abstract objects, we do ordinarily take ourselves to see absences, such that to say otherwise is to contradict common sense. But that is the cost, a sceptic might say, of subscribing to ‘naturalism’ in philosophy, by which term we are to understand the view that reality, including perceptual phenomena, is causally closed. The naturalistic spirit thus drives theorists to require a physical, causal contact between perceivers and perceived things. A visual experience lacking a physical contact with the object seemingly presented risks revealing itself as, say, a veridical hallucination, as in H.P. Grice’s (1961) thought experiment in which a subject’s experience matches the scene before him – there is a pillar *there* – without the relevant object, the pillar *there*, causing the experience – because there is, unbeknownst to the subject, a mirror in

¹⁴⁵ Mumford attributes this view to Farennikova (2013), according to whom absences are genuinely seen because absence experience is explained by the subject’s expectations and expectations penetrate perception.

¹⁴⁶ Although it isn’t part of my dialectic, the causal condition is also, on Mumford’s view, a problem for Martin and Dokic’s (2015) ‘meta-cognitive’ view, according to which absence experience is explained in neither perceptual nor cognitive – or belief-involving – terms but rather by the feeling of surprise that accompanies it. ‘If it is problematic to account for how an absence can cause a perception’, Mumford writes, ‘is it not equally problematic to understand how an absence can cause a metacognitive feeling of surprise?’ (ibid.).

front of him that reflects a distinct but qualitatively identical pillar, located not *there* but *over there*. Channelling this view, Roy Sorensen (2015 pp. 554-5) commends theories in favour the perceptibility of causally inert objects that acknowledge this challenge, for example the view that holes are perceptible only mediately, through the perception of a material thing that does enter causal relations, the hole's host (Casati and Varzi 1994), and warns that if one is to defend the perceptibility of non-hole absences, one faces a similar challenge: '[t]hose committed to the perception of absences appear committed to causation by absences'.

The proposal I'll eventually make would *prima facie* stumble upon Sorensen's challenge. The way to deal with it will not involve defending absence causation¹⁴⁷ but rather pointing out that Sorensen makes a further assumption: that those committed to the perception of absences are committed to a view of them as ordinary objects, such that if ordinary objects of perception strike our retinas, absence enthusiasts need to show absences do that, too. My proposal will take a different view of absences, and thus will not face this challenge. For now, though, it bears noting that there are, anyway, various ways to complicate, rather than meet, the causal demand as it stands. One first way might be to offer a view of perception on which the relevant causal arrow doesn't stem from the seen object. If one thinks that to see an object is to receive information about it through a visual channel (Dretske 1983), then there are cases in which this occurs without the seen thing being the cause of the information transfer but rather being caused by the same event that causes the transfer. Call this a 'common cause' theory of perception (Ganson 2020). Another way might be to offer a more sophisticated version of the explanandum, such that what is to be explained is not one's perceptual contact with an object (which seems to have the causal condition built-in) but one's acquisition of perceptual knowledge about objects.

¹⁴⁷ On that subject see Schaffer (2000), Jago and Barker (op cit.).

Here, too, there are cases where a demonstrative thought, say, is made available by perception but in which the relevant perceptual link is not *caused* by the object referred to by the demonstrative term. A first example shouldn't be too controversial: consider two stable workers who come upon an empty stall: 'where is that horse?', one asks; 'That horse is out pasturing', the other responds. The thoughts were prompted by perception and the demonstrative term refers to an existent object; yet, the perceptual link that enabled the thoughts did not have the relevant object at the other end.¹⁴⁸ A second example is perhaps less obvious but also plausible. Consider a young student who exclaims, upon first encountering an irregular figure in his geometry textbook: 'that shape is weird!'. The thought is not about the marks on the page but about an abstract object, the same thing the teacher takes up reference to when she replies to the student: 'yes, that shape is *irregular*'. In both cases, the object referred to is not perceptually available to the subject in the way objects in ordinary demonstrative contexts are; nevertheless, there are ways to make sense of the idea that there is a perceptual *link* to them here, albeit one that is more complicated than the kind a simple causal interaction between an object's surface and the subject's retina usually establish.¹⁴⁹

To return to our dialectic: Mumford seems unaware of this kind of theoretical possibilities. He takes it as a datum that absences cannot be perceived because they cannot cause a perception. In fact, he writes: 'perceptual cases bring home the sheer empirical implausibility of causation by absence' (ibid.). So absence experience shouldn't be treated in perceptual terms. And on the other hand, a view

¹⁴⁸ For one account of the first case according to which perception is involved in making salient a relation that holds between the demonstrative's index and the object of reference, see Georgi (2012). In ordinary cases, that relation is identity, in others, such as this case, it will be a more complex one. See also Nunberg (1993) on deferred reference.

¹⁴⁹ For an account according to which the role of perception is to establish a link between the subject and the (abstract) object *via* the latter's concrete tokens, see Juvshik (2018).

on which absence experience is explained purely terms of a belief that is formed upon the experience of present things is at odds with the fact that there is a perceptual(-seeming) phenomenology to absence experience, such that we are disposed to treat absences, upon (purportedly) experiencing them, as objects of thought and speech and about which certain actions are to be taken, just as we do with veridically perceived objects.¹⁵⁰

The Denettian ‘user illusion’ dimension of Mumford’s account deals with both perception and belief-based dangers. To say that absence experience is an illusion is to deny that it is a veridical perception;¹⁵¹ to say that it is a useful illusion allows one to accommodate the phenomenology, akin to that of veridical perception, which is no doubt useful, into the account.

In short, for Mumford, ‘absences are experienced perceptually but only because of a mechanism where an inference is presented phenomenally’ (op cit., §7.6). Return to the group of early humans. They infer, from the earth-lining they veridically see, that they will fall if they walk on (because no ground will be there to support them). It just so happens that this inference lacks the phenomenology of conscious reasoning – because it is not – but has one resembling sensory experience, and it results in an automatically formed belief with no negation in its logical form – so with nothing like the form ‘there isn’t any ground in that location’ – but rather a belief asserting or presupposing the existence of something with a different, yet equally existing nature (apparently) as the rest of the things in the scene, a hole, which everyone in the group can jointly pick out because it *seems* to them to be there,

¹⁵⁰ As Mumford himself notes, the ‘cognitive’ view does not have clear-cut representatives, but it is a theory both he and Gow (2020) – as we will shortly see – consider as occupying logical space and thus in need of consideration.

¹⁵¹ Where being a ‘veridical perception’ is incompatible with being a ‘veridical hallucination’.

available for demonstrative reference and, crucially, action, namely, the action of being circumvented.

Now, Mumford's view, as brief as I've offered a summary of it here, may or may not be right. It would seem that at least some empirical evidence is needed to conclude that animal cognitive systems have indeed developed, as a matter of adaptative advantage, the ability to delude the animal itself – at his personal, conscious level – into believing there exist things with observable properties such as colour and shape that we call absences but which aren't really there. My aim in putting Mumford's view on the table, though, is not to directly challenge it for now but to show its kinship to another recent view that does not explicitly diagnose absence experience as non-veridical. Then I will show how this diagnosis, explicit or not, might be pushed on us by independent, more general facts about perceptual experience. (Eventually I'll offer, as I've said, a way to resist it.)

1.3

Mumford's strategy to present his view in a way that specifically avoids the pitfalls he identifies as threatening other possible views is also found in recent work by Laura Gow (2020). Like Mumford, Gow divides the logical space between 'perceptual' and 'cognitive' views: views on which absence experience is explained in terms of a perceptual state and views on which it is explained in terms of a cognitive one, by which Gow means, specifically, a belief. Gow's main reason to reject what she calls the 'strong' perceptual view, the view that we literally see absences, is similarly naturalistic in spirit: it is a commitment to the causal condition on perception. Only Gow's wariness is broader. It is theoretically costly to 'tak[e] a realist stance towards absences and permit[...] them a role in the perceptual process', she writes (op cit., p.

172), not just because of the burden of explaining how absence causation might work but also because of reasons to do with ontological parsimony.¹⁵² And her main reason to avoid what she calls the ‘weak’ perceptual view, according to which absence experience is a genuinely visual phenomenon but only *as of* something out there, not really *of* it, is that it renders the experience non-veridical. That regarding absence experience to be non-veridical helps with the issue of ontological/causal problem is something she thinks is self-evident:

Veridical visual perception involves objects reflecting photons which then interact with the photoreceptors which compose the retina, and veridical auditory perception involves vibrating objects causing sound waves which enter our ears, and so forth. It hardly needs to be said, but absences do not, of course, reflect photons, nor do they produce sound waves by vibrating (Gow *op cit.*, p. 171).

Yet the cost of non-veridicalism, Gow writes, is to deny that ‘absence experiences “get the world right”’ (p. 172). This is a problem because Gow thinks absence experiences do yield knowledge of the world. The point needs a little elaboration. Gow would likely say that, strictly speaking, the kind of discourse about absence experience discussed earlier is false: ‘[a]bsences do not seem to have colour properties, for example. Nor do they seem to have timbre, volume, pitch, motion, sweetness, sourness and so forth’ (p. 173). However, for absence experience to ‘get the world right’ needn’t mean for our attributions of observable properties to

¹⁵² Gow doesn’t actually put this point in terms of parsimony. She writes, rather, that ‘anyone with physicalist leanings would be wise to avoid a theory which posits such entities’ [as absences] (Gow *op cit.*, p. 170). I’ve decided to interpret her in this way to simplify things and avoid the issue of what it means for something to be a physical entity, which question she herself doesn’t address.

absences to be right; it just means for information about the failure of things to be present in certain contexts to be carried to the perceiving subject during or by the experience. How absence experience could get the world right if absences do not exist – and if they exist, do not have the requisite causal or observable properties – can therefore only be explained, for Gow, by a kind of cognitive account, albeit not one that simply regards absence experience as a combination of a perception of present entities plus a belief or judgement about what is not there. Such a view, which would analyse the experience of a gone laptop as ‘[t]he perceptual phenomenology (of the empty table top, say) plus the belief (that one's laptop is no longer there)’ (p. 175) faces the following ‘decisive’ objection:

[A]bsence experience exhibits resilience to a change of belief. If you are told that your laptop is still on the table and a magician is playing an elaborate trick with mirrors, you will still experience the absence of your laptop even though you now believe that your laptop is not absent. A proponent of the belief/judgement version of the cognitive view must therefore allow that a subject can simultaneously (and consciously) believe or judge *p* and *not p* (op cit., p. 176).

This problem gives Gow her cue. On her view, absence experience is indeed best explained by a non-perceptual state that is belief-like in that it bears propositional content, but not belief-like in that it's not consciously available for reasoning, for example, and would not typically be endorsed by the subject. She calls this state an ‘intellectual seeming’. Intellectual seemings ‘have a belief-like or judgement-like phenomenology, and the same direction of fit as beliefs and judgements. [...] Consequently, their content can be true or false/accurate or

inaccurate depending on whether the world really is the way it is represented by the intellectual seeming' (ibid.). Gow refers us to extant literature on intellectual seemings for a thorough characterisation,¹⁵³ and relies on examples for her own exposition:

[A] bat and a ball cost \$1.10 and the bat is \$1 more expensive than the ball—how much does the ball cost? The answer most people give to this question is '10 cents', which is incorrect—the correct answer is 5 cents. Interestingly, even when we are familiar with the puzzle it may still seem to us that the answer is 10 cents. If so, this would be an intellectual seeming. It cannot be a belief, since we know and believe that the correct answer is 5 cents (ibid.).

Just as, when presented with the puzzle, it immediately seems to most people that the answer is '10 cents', so too when presented with an absence, it immediately seems to me that there is something out there, an absence. But absences are appropriate contents only of beliefs or intellectual seemings; they cannot be, for the rehearsed reasons, contents of experience (assuming a representational framework) or constituents of it (assuming a non-representational one). And just as the intellectual seeming with the content '10 cents' might be explained in a different way from an intellectual seeming with the content 'heavier objects fall faster', say (another of Gow's examples), on account of there being different background beliefs or intellectual seemings about how subtraction and physics work, so might different intellectual seemings of absences be explained differently. A mismatch between what is perceived and the subject's expectations, understood propositionally (as holding

¹⁵³ She highlights Huemer (2007), Reiland (2015) and Tucker (2013), but mentions other relevant work. See her op cit., pp. 169 and 173.

some attitude expressible as ‘expecting that...’), might explain a few cases;¹⁵⁴ a failure of the subject’s body schema to update itself after a bodily change might explain other cases, as when one expects one’s tooth to be absent (because one just saw the dentist remove it) and is nevertheless perceptually struck by the experience of feeling the gap it left.¹⁵⁵ Moreover, some experiences of absence might have a non-perceptual dimension that can be accommodated by a further, ‘meta-cognitive’ view, one which explains the feeling of surprise that often – but not always – accompanies experiences of absence.^{156,157}

The kinship should be clear by now between Gow’s and Mumford’s views. Although Gow sees the result that absence experience is non-veridical as an argument against the weak variety of the perceptual view, her own view ends up, in a sense, yielding a similar consequence. On her view, again, absence experience is not to be explained in visual terms, veridical or non-veridical, at all. But then perceptual reports about absences, that is, sentences in which absences are predicated observable properties and/or are referred to by the complement of perceptual verbs, would turn out to be false. It is just not true that ‘darkness is black’, that ‘I can see a big hole’ or that ‘I can see Pierre’s absence’, on Gow’s view. There are no such perceptual states as those reported. Presumably, Gow would say that these propositions are the result of taking the relevant intellectual seemings at face value, just as a careless member of the

¹⁵⁴ Cf. Farennikova (2013).

¹⁵⁵ Cf. Cavendon-Taylor (2017).

¹⁵⁶ Cf. Martin and Dokic (2013).

¹⁵⁷ Because, as mentioned earlier, Martin and Dokic’s (op cit.) view may help to explain something at the ‘meta’ level (the meta-cognitive or meta-perceptual level, that is), I take it to be compatible with both views and leave it largely undiscussed. Yet as other (esp. Cavendon-Taylor op cit. and Gow 2018) have pointed out, the feeling of surprise is contingent. Just like Szymborska’s cat, a bereaved person might in fact experience something like the converse. They might suffer to see their expectations of the absence confirmed time and time again when entering new rooms in the house, or visiting once shared spaces.

public might count as doing if they simply endorsed their intellectual seeming that ‘the ball costs 10 cents’ without thinking their answer through.

The form of illusionism I’m attributing to Gow is less straightforward than Mumford’s. It is, as we saw, part of the schema of perceptual illusion that it is consistent between subjects (because the misleading appearance is indeed a property of the object itself); meanwhile, intellectual seemings do not seem to be so consistent. On the one hand, it is only ‘most people’ who experience the ‘illusion’ that the ball costs 10 cent or that heavier objects fall faster;¹⁵⁸ on the other, Gow illustrates her view with false intellectual seemings, but that doesn’t entail that all such seemings are false (whereas all visual illusions, by definition, are). Nonetheless, there is a prediction of systematic *wrongness* of some kind for perceptual reports about absences on Gow’s account. Even if we gain knowledge – ‘get the world right’ – during absence experience, such knowledge cannot be regarded as being about such a thing as an absence. Perhaps sentences about absences should then be paraphrased in an absence-nominalistically acceptable way for the knowledge in question to be expressed truthfully. Not ‘I see holes in the cheese’ but ‘I see that the cheese is perforated’, things like that.¹⁵⁹ Viewed as experiences *about absences*, though, Gow’s account sides with Mumford’s. Where on Mumford’s view absence perceivers undergo an illusion in taking themselves to enjoy veridical experiences when they attribute colour, size, and visibility to darkness, holes, and people’s absences, on Gow’s view subjects might count as experiencing something like an intellectual illusion whenever the ball *seems* to them to cost 10 cents, heavier objects *seem* to them to fall faster, and absences to be there.

¹⁵⁸ Gow does not cite empirical evidence for the claim that ‘most people’ experience this. Anecdotally, I don’t experience this intellectual seeming.

¹⁵⁹ But see Lewis and Lewis (op cit.), Bricker (2016), §4.2.

2. A step back: on the structure and objects of experience

2.1

One reason to focus on the above two as samples of current work on our subject matter is that, as we've seen, they position themselves as surmounting the obstacles of extant alternatives. But another is that there are independent reasons from the above authors' to support non-veridicality about absence experience, which need to be considered separately. I am referring to work by Sorensen (2008), Clare Mac Cumhaill (2015, 2018) and Dominic Alford-Duguid (ms.), all three of which examine the phenomenological and epistemic upshot of certain invariant – or 'structural' – features of visual experience and in doing so provide support for non-veridicality.

Sorensen's is probably the most discussed work on all matters of darkness and shadows. His point of departure is that experiencing darkness, while subjectively indiscriminable from the state of not seeing anything, is a perceptual state like any other, not to be identified with the state it is indiscriminable with. This comes with the consequence that, assuming we can see a uniformly illuminated white wall and fail to believe that we do (because we lack the relevant concepts; because we think we see pure white light, or whatever), we can see total darkness and fail to believe that we do. Darkness is, like any present entity, something we can non-epistemically see.

This point might make it seem like Sorensen will sympathise with a veridical view of seeing darkness, but it's not clear he does. Consider Sorensen's observation, which he makes channelling vision scientist Leo Hurvich, that 'the visual system does not treat darkness as a privation of light' even though that's what it *is* (2008, p. 138); '[t]he culprit', Sorensen writes, 'is our pre-Newtonian visual system' (p. 250). The

result is the visual system's responding with a totally black experience to total darkness just as it would respond with a totally white experience to a uniformly illuminated white wall if pressed up against one's nose.¹⁶⁰ So our experience of total darkness is not to be understood as a perceptual encounter with a black thing. If Sorensen is right that the culprit of our 'mistreatment' of the absence of light is our pre-Newtonian visual system, it is plausible that some other creature's visual system would treat the absence of light as the privation it is. It is difficult, evidently, to imagine such a creature's phenomenology, but if the mark of our mistreatment of light privation is that it manifests in a positive experience of blackness, then that other creature's 'correct' treatment of it may manifest in a different response, not in blackness, as *that* is the result of our visual system's 'mistake'. Such a line of thought seems conducive to predicting, like the views above, an attribution like 'total darkness is black' to be false.

Now, Sorensen doesn't claim, like Mumford, that seeing darkness is an illusion. In fact, he does make this diagnosis for a peculiar kind of case:

An astronaut with his back to the sun would see space as black even though she is bathed in much light, [but] because darkness is the absence of light, the astronaut's black experience is a *false* representation of darkness [...] the astronaut is experiencing an illusion' (Sorensen op cit., pp. 250-1, emphasis original).

Does this mean Sorensen regards ordinary, attic-variety experiences of absences as veridical? I'll refrain from exegesis here. My point is to note that Sorensen's work is a precedent of the inclination to account for absence experience in

¹⁶⁰ The example is Dretske's (1969, pp. 26-7).

terms of invariant facts about the subject's visual system rather than about what she sees. Mac Cumhaill's and Alford-Duguid's are just two of the most recent authors to follow this path, albeit focusing on a different kind of invariant fact about the subject: not contingent facts resulting from human evolution but facts about the structure of experience itself.¹⁶¹

Structural features of experience play a role in other instances of experiencing absence. Take the absence of a specific individual in a context – Pierre's absence in the café; the absence of Gow's laptop from her table. Mac Cumhaill (2018) has recently developed a Sartrean view of cases like this. Mac Cumhaill starts from a point of sympathy for a theory that is, at least, compatible with realism about absences, which compatibility can seem to be at odds with the insight that Sartre's cognitive attitudes, specifically his expectation to find Pierre at the café, must have some explanatory role in his experience of Pierre's absence. Mac Cumhaill's solution is to shift the object of Sartre's perception. On her view, '[i]t can be granted that absences are ways the world is. Nonetheless, it is insisted that what Sartre perceives is not an absence but an *absential location*, the bounds of which are circumscribed attitudinally – by reference to Sartre's interests and intentions' (Mac Cumhaill 2018, p. 41). The distinction is between absences, which are 'ways the world is at a particular locale' but don't have 'sensible properties [... that may help them] ground there being "anything it's like" to perceive them', and absential locations, which are 'places at which the application of the predicate 'is absent' to an absentee is true when evaluated relative to those places' (op cit., p. 36). Places being sensible, concrete entities, absential locations can be perceived just like any other place is.

¹⁶¹ As Mac Cumhaill (2015, p. 688) points out, this body of literature seems to have been sparked by M.G.F. Martin's (1992, 1993) comparative analysis of structural features of visual and bodily/tactual experience. See also Richardson (2010), Soteriou (2013) and Phillips (2013).

This new explanandum is less problematic. The question of how seeing absential locations is possible does not have the ontological burden seeing absence perception has. A view that is ontologically silent – silent on issues to do with what kind of things are perceptible – and focuses merely on the subject-side of the phenomenon becomes, then, acceptable. Mac Cumhaill’s proposal is that Sartre has already offered such a view. In perception, Mac Cumhaill tells us that Sartre writes, “there is always the construction of a figure on a ground”, where, notably, the ground is seen ‘in addition’ to the figure’ (p. 39). Just as thoughts are often thought to have a distinctive structure (a subject-predicate structure, say), so does perception have a distinctive figure-ground structure. In absence perception, ‘one intuits *only* the ground’ (ibid.). The café may play the role of the figure for someone who is looking for a place to sit; for Sartre, the café – the ‘mirrors, patrons and rattling saucers’ – all blend into an undifferentiated ground onto which a figure fails to appear. What explains the experience of absence is not, then, anything that hits Sartre’s retina. But the explanation remains perceptual because the figure-ground structure is a fact about the perceptual apparatus. Which (again) still makes room for attitudes’ playing an explanatory role. As Mac Cumhaill hastens to remark, ‘low-level implicit expectations [of the kind appealed to by Farennikova] are likely to govern explanatorily relevant patterns of perceptual activity’ (p. 45). Expectations might determine what will count as figure and what as ground.

Finally, a third reason to think structural features are relevant to the issue at hand involves spatial awareness. As Mac Cumhaill herself (2015) and, more recently, Alford-Duguid (ms.) have argued, reflection on the phenomenon of seeing empty space reveals aspects of perceptual experience – such as its phenomenology and the range of beliefs it can justify – that cannot be explained purely in terms of what is

perceived.¹⁶² Consider a perceiver's sense that the space she sees is part of a larger space. This 'sense' is a matter of visual phenomenology but also something that can figure in the content of a belief. On Alford-Duguid's view, this belief – expressible as 'the region in view is a sub-region of a larger space' – is justified by the perceiver's awareness of her visual field's boundedness, which very notion implies an outer realm, something beyond what the bounds are of. One curious feature of this belief is that while it is not a priori (one has to open one's eyes in order to enjoy visual awareness of the boundedness of the visual field), the belief is immune to the kind of defeaters to which perceptual beliefs are typically vulnerable, such as reliable testimony that one is suffering a hallucination, say. Alford-Duguid's way to explain this difference takes after the way P.F. Strawson (1950) puts the difference between asserting and presupposing that *p*. 'Visual experience perceptually presupposes that the region in view is a sub-region of a larger space, and does so because it involves awareness of the spatial sensory field as bounded by our sensory limitations' (Alford-Duguid op cit., §4.2). The belief that there is a red apple in front of you loses its justification ('is defeated') when you learn the apple is white (the red appearance is just a trick of the light). This kind of justification is dubbed by Alford-Duguid 'presentational': it pertains to beliefs about what is presented to the subject. The belief that the region in view is a sub-region of a larger space cannot be defeated in this way, because the belief is not about anything that is presented to the subject. One might make a similar point with regards to the contribution of presentational and structural features of experience to *phenomenology*. If there is something it is like to be aware of the boundedness of one's visual field, then this aspect of visual

¹⁶² Mac Cumhaill eventually challenges this, but on her view (as in her (2018) work on absential locations, incidentally), the explanandum turns out to be different from the one with which she started. She ends up re-casting the phenomenon of seeing empty space as one of seeing a positive thing, a region with a look in its own right, rather than as a species of absence experience.

phenomenology is not contributed by any presented objects.¹⁶³ A consequence of this latter point is that the notion of ‘diaphaneity’ is threatened.¹⁶⁴ This was the notion, recall, that experience is ‘transparent’ or ‘diaphanous’, that is, devoid of any qualities that may be attributable to the experience itself rather to what whatever objects it is an experience of. Take a hypothetical case of experiencing emptiness, as in the all-white scenarios from the *Matrix* films that are meant to represent a blank canvas in which an experience simulation can be run.¹⁶⁵ There is, by hypothesis, nothing the experience is of in this case. This blank canvas – called ‘the Construct’ in the films – is not only visually empty; it is supposed to be gravity-less, odourless; to be devoid, in short, of all kinds of circumstances of which we ordinarily gain knowledge by being perceptually open to the world. Characters are able to stand, as though in an Earth-like gravitational setting, by a further specification, and similarly see themselves and each other thanks to the presence of light, and so on. But they know there is in fact nothing there.¹⁶⁶

Suppose you find yourself in the Construct. Your experience might resemble that of Sorensen’s astronaut if she found herself in a genuinely light-devoid space, floating in a space that is all-black. There is something it is like to find yourself in this case. The white look of blank space helps to imagine being aware of the bounds of your visual field. There is nothing being presented to you; yet, you still can have

¹⁶³ The notion that perception of limits or bounds generally presents a counterexample to the thought that the phenomenology of experience depends entirely on presented objects can be found elsewhere (Barker and Jago 2012, p. 134). More on this shortly.

¹⁶⁴ To renegotiate the view is indeed one of the aims of Mac Cumhaill (2015). See esp. pp. 190-2.

¹⁶⁵ See <https://matrix.fandom.com/wiki/Construct>

¹⁶⁶ Not even they themselves are there, according to the fiction, because the Matrix is a simulation. I assume here, not uncontroversially perhaps, that if it is possible to imagine a scenario where one’s embodied self, not just one’s virtual avatar, is floating in a blank void, perhaps having learned to live without oxygen, then that is a logically possible scenario. I ask the reader to follow me in assuming this and in supposing such a modified version of the scenario in the main text.

one perceptually justified belief, the belief that there is emptiness around you, where the phenomenology of that experience cannot be attributed to any object it is of, since there isn't one. In this (at least logically possible) scenario, the character of the experience is thus explained purely in terms of a structural feature of it, not in terms of anything in view.

To wrap up. In the previous section, I presented views that predict reports of absence experience – reports that predicate observable properties of them and/or treat them as the referents of complements of perceptual verbs – to be false, since they predict absence experience itself to be non-veridical. In the last few pages, I've presented work that, while silent on *that* issue, seem to indirectly support non-veridicality. The experience of darkness can be explained without appeal to anything expansive and black; it can be explained by appeal just to the pre-Newtonian nature of the visual system. Experiencing the absence of individuals can be explained by appeal to nothing that is stubbornly absent, as Szyborska says; it can be explained by pointing out that perceptual states have a distinctive figure/ground structure, and that in (what we call) absence experience, what is seen is just a figureless ground. And the experience of emptiness can be explained by appeal to no further entity beyond the perceiver herself; it can be explained simply by pointing out that visual experience gives us awareness of a larger region than the one bounded by the visual field, even when that region is empty.

2.2

The result of the above survey is pessimistic for a veridicalist about absence experience. The literature seems to favour the view that absence experience can be accounted for without the need to posit an entity with which the absence perceiver

stands in a perceptual relation. So the bits of discourse discussed earlier on are false.

Is there anything else a realist about absences could cling to?

One might start by noting a pattern in the structural explanations just discussed. Visual experience systematically gives us both positive and negative knowledge, knowledge not only about what the experience is of but also about what it is *not* of. And absence experience is not the only instance of this phenomenon. Consider Stephen Barker and Mark Jago's (2012, p. 134) discussion of edge perception:

[T]he perception of edges [...] is an important component of human object recognition. [In] the analogous case of computer vision, [...] one might think [...] that an image can be represented as positive information only. This is not the case, however. The value representing the intensity of each pixel contains both positive and negative information: a pixel's intensity value says that the pixel is at least this bright but *no brighter*. The numerical value associated with each pixel encodes both a positive and a negative fact. We suspect that a similar story holds of edge-detection in human perception (emphasis original).

Regardless of whether Barker and Jago's analogy between computer and human vision is right, the insight that perception of things comes with perception of their edges is plausible enough. (Indeed, one might read Dretske's (1969) view that what is seen must be 'visually differentiated from its immediate environment' as saying that the perception of x's edges enables the perception of x.) If this is right, then edge or boundary perception too is an invariant feature of visual experience. Experience is simply such that seeing things typically comes with seeing their limits.

And in cases where one does not see the limits of what one sees, as in the case of seeing a uniformly illuminated white wall pressed against one's nose, then one is aware of one's own sensory limits at least. So one enjoys awareness of some limit or other in having any visual experience. Note that this further structural feature of experience may help to explain hole perception. Holes can be regarded as the inner limits of their material hosts. Indeed, a kinship between holes and edges is seen by Ian Phillips in the auditory case: '[i]f we think of pauses as auditory "holes," we can think of such phenomena [as hearing a single sound cease] as auditory "edges"' (2013, p. 341).

Generally speaking, then, just by opening our eyes, we put ourselves in a position to gain both positive and negative knowledge, knowledge that light is not present, that Pierre or Simone aren't, that nothing is being presented (the Construct), or that x – a seen object, or our own visual field – extends up to here and *no further*. We might follow Alford-Duguid in calling these bits of negative knowledge 'structurally', not 'presentationally', justified, where both varieties of justification count as perceptual. That the knowledge acquired in absence experience is non-presentational is fitting; by definition, the purported experienced entity is not a presented object.

But I said I'd give a defence of veridicalism a go. Didn't I just bolster the adversary's point? Granting – no, having myself made the case – that all our examples can benefit from a structuralist explanation, I want to point out now that they leave an aspect of the experience unaccounted for: the contribution of absences to the phenomenology of experience. Let me make this point by analogy. Take a flower that reflects wavelengths into the near-ultraviolet. Our experiencing the petals as blue-ish whereas hummingbirds do as ultraviolet-ish can be explained by appeal to structural differences in our visual systems. Neither we nor the hummingbirds are experiencing

an illusion; both ways of appearing are objective features of the mind-independent flower. The phenomenon of variation between subjects' experience of one and the same thing comes in handy when philosophers want to illustrate the point that perceptual experience involves not merely perceiving objects but doing so in certain ways. As M.G.F. Martin writes, '[w]hy cannot the *ways* in which things are presented make a difference to what the experience is like in addition to what is perceived?' (Martin 1998, p. 175). One natural way to accommodate this thought is by regarding perception to be a three-place relation where the subject and the object are the main two relata and the third element is something like the 'standpoint' from which the subject relates to the object (Campbell 2009), which standpoint may include facts like the number of the subject's cones. Alternatively, the third relatum might be constituted by the conditions of viewing (the light being cool or warm, say) relative to which the object will appear in one way or another (Brewer 2011). As has been recently argued, though, the acknowledgement that objects can appear differently to subjects needn't push one to see perception as including a further, third element than the subject and the object themselves (French and Phillips 2020). The different ways in which objects are presented to subjects needn't be, in other words, reified. Acknowledging, precisely, that 'there is no unique way of perceiving [an object]', say a red car, leads one to the thought that 'nothing other than the car and its redness need be presented to [a subject]' in the first place (French and Phillips *op cit.*, p. 7). The objects we are presented with in perception can shape the character of our experience in one way or another depending of a variety of factors like the lighting and our number of cones. This is just part of what it is for an object to appear in certain ways and for a subject to be sensitive to certain ways objects have of being experienced.

Pairs of subjects like humans and hummingbirds – call them ‘contrast pairs’ – help to bring out the contribution of what we see to the phenomenology of our experience. If what we see does not make a phenomenological contribution, there would be no difference in the character of our experience between differently sensitised subjects. Consider our experience of total darkness and that enjoyed by the creature whose visual system responds to it by eliciting what we would call ‘whiteness’. Here we have a pair of subjects who, like the pair of humans and hummingbirds, have a different phenomenology in response to one and the same thing. Both cases have a shape we might characterise by saying that ‘ S_1 experiences some thing x in some way w_1 and S_2 in some way w_2 ’. But if absences do not exist or are not really there, we are missing the x element that is treated in a pre-Newtonian and post-Newtonian way, respectively, by us and the post-Newtonian creature. That absence experience can vary in the way ultra-violet-flower experience can is supported by our ability to think up similar contrasting pairs for other varieties of absence. Consider a Matrix programmer who makes a mistake and doesn’t code for emptiness to look white but red for some particular Construct user. Neo experiences emptiness as white; Trinity as red.¹⁶⁷ Consider the absence of ink in a circular region of a two-dimensional, otherwise black image, and imagine a robot whose vision is not programmed to encode negative information as Barker and Jago tell us computer vision typically is, and so does not detect a hole– or the inner limits of the expanse of black ink – but rather a white round object. Finally, imagine Sartre is accompanied to the café by a creature who has the power to see the shape of individuals who were present minutes before, perhaps because she sees the molecules left by their scent in the way we see shooting stars’ wakes. Suppose that by the time Sartre gets there,

¹⁶⁷ In image processing software, absence of content tends not to be represented as white – which is, as in the Construct, a further specification – but as a pattern of white and grey squares.

Pierre has only just left. Then Sartre and the creature will experience the absence differently.¹⁶⁸

If theoretical parsimony drives us to accept a picture of perception on which the ways in which reality is perceived are not further entities to count among the subject and object but rather to be associated with the presented objects (to be regarded as attributes of them, as ways *they* have of appearing), then a non-veredicalist about absences, driven, presumably, to their view (at least in part) by parsimony, will end up missing what the relevant ways of appearing in the above cases are *of*. But these ways of appearing are what explains intersubjective phenomenological differences. Perhaps absences appearing to be ontologically unnecessary in an account of absence experience is partly explained by an intuition like the one I have phrased, borrowing Alford-Duguid's words, by saying that the knowledge acquired in absence experience has a structural, not presentational, justification. We've already acknowledged this intuition seems, admittedly, correct. The relevant bits of knowledge do not hinge, by definition, on any object being *presented* to the subject. Yet this doesn't mean there is *nothing* the subject is perceptually aware of, which would be true in cases of absence of awareness of anything. The Construct user's case is the clearest one in this respect. The knowledge in question depends on no presented objects, yes;¹⁶⁹ still, there is perceptual

¹⁶⁸ This scenario turns on experiencing Pierre's absence in a way that cannot be generalised to, for example, a case where Pierre wasn't there. But other thought experiments can fix this. Imagine Sartre's café is not the Deux Magots in Paris but the canteen in the space station from Andrei Tarkovsky's *Solaris* (1972), where a nearby planet's strange atmosphere makes it so that when a person misses someone, the absentee materialises and becomes visible. Here the contrast pair is Paris Sartre and space station Sartre, both of whom experience the same thing – Pierre's absence – differently.

¹⁶⁹ It is not clear an objector could say the presented object in this case is one's own body. Alford-Duguid's view rests on work by Richardson (2010) and Soteriou (2013) who start, precisely, by contrasting our awareness of the region in view as bounded, which we have merely by awareness of our sensory limitations, with the awareness we might have of the limits of our body. '[T]he limits of bodily awareness do not strike us as merely sensory; they

awareness being enjoyed here, and there is something it is like to be in such a state. For first timers, as shown in the film, it is quite striking.

So what does this phenomenology hinge on? There is another intuition in the background of this gap, an intuition about the sort of things we can perceive. It can be traced to a view on which there is a rough dichotomy between object and fact perception, a dichotomy perhaps not explicitly stated but often assumed (Dretske 1969, 1999, 2000, etc.; Williamson 2000). Let me elaborate. A natural way to check whether a perceptual report expresses one or the other kind of perceptual state is to focus on the complement of the perceptual verb. If the complement is – or is amenable – to a sentential complement, then the reported state is one of fact perception. For example: ‘I see the limits of my visual field’ has a noun phrase as complement, but it seems amenable to a rephrase like: ‘I see that my visual field is bounded’. This seems fitting. Such a perceptual state makes reference to no presented objects, and the perceptual belief in question is, as we’ve seen, non-presentationally justified. Compare: ‘I think I saw Jon crying, but I’m not sure if he was crying or actually laughing’. The uncertainty precludes rephrasing the bit ‘I saw Jon crying’ as ‘I saw that Jon was crying’, which would indicate a case of *fact* perception. Given the dichotomy, ‘Jon crying’ refers to an *object* of perception. Luckily, on this view, objects are a broad church. Here is Dretske (1999, p. 121):

[T]oken events, states, and conditions are spatio-temporal particulars which are (like apples and stars) distinct from both the facts and properties from

seem due to the boundaries of an object—the body’, Alford-Duguid writes. Why not so in the case of the boundedness of the visual field? The response might be that the contours of the visual field do not seem to be perceived as associated with anything bodily – as the contours of our eye sockets, say. In half-closing our eyes, we *are* aware of our eyelids, yes, but this now counts as seeing something within our visual field rather than as modifying its boundaries.

which I distinguish objects. Events and conditions have a (temporal) beginning and an end. Properties do not. Neither do facts. As Dostoyevsky put it, a person's suffering (an event or condition) ends, but the fact that the person suffers endures forever.

If the *fact* that a person suffers has no beginning or end, it cannot be non-epistemically seen. Only entities with a spatiotemporal location – 'spatiotemporal particulars' – can be seen like this. The reason is that seeing things non-epistemically is a primitive, non-belief-involving state, one to be understood as part of the causally closed, non-minded course of reality. It is a process or event as primitive, in Dretske's words, as that of stepping on things (2000, p. 101).

A similar ontological category to Dretske's one of events, states, and conditions is that of 'situation'. Here is Williamson (2000, p. 38):

There is a distinction between seeing that A and seeing a situation in which A. One difference is that only the former requires the perceiver to grasp the proposition that A. A normal observer in normal conditions who has no concept of chess can see a situation in which Olga is playing chess, by looking in the right direction, but cannot see *that* Olga is playing chess, because he does not know what he sees to be a situation in which Olga is playing chess.

When coupled with a that clause, 'see' is what Williamson (op cit., p. 34) calls a *factive* mental state operator. This is why seeing 'that A' counts as *fact-seeing*. But seeing 'a situation in which A' seems to fall under what Dretske calls object-seeing. Situations seem to have, along with states, events and conditions, a place in the list of entities that count as *objects* of awareness in Dretske's sense even though

they're not ordinary material objects. Hence they cannot be put together with factive mental state operators such as 'see that'. Dretske requires objects to be spatiotemporal particulars, and he assumes that events, for instance, are. Williamson seems to assume the same thing about situations. So only spatiotemporal particulars can be referred to by the complement of a perceptual report that expresses non-epistemic seeing, which is to say, again, that only they can be objects of perception.

So much seems intuitive enough. The intuition dovetails with Mumford and Gow's assurance that absences cannot be veridically experienced: only spatiotemporal particulars bounce light off, vibrate, etc. Perhaps absence experience can be made sense of in terms of *fact* perception, according to this view (recall, indeed, Mumford and Gow are cognitivists), just like the experience of the boundedness of our visual field can be made sense of in terms of seeing *that* one's visual field is bounded. Neither of those involves an object being presented. Yet to say as much and leave it at that would be to ignore the possible difference in phenomenology between a pre-Newtonian and a post-Newtonian subjects' experience of darkness. (Ditto for the rest of the contrast pairs.) This is because it is objects to which we typically attribute the *ways* of appearing that explain such differences. And it precludes absence seeing being non-epistemic, too, which it intuitively *can* be.¹⁷⁰ These two problems are interrelated. If we posit an object of perception out there filling the other place of the two-place relation, we can, as per the object/fact dichotomy, unproblematically regard it as non-epistemically seen. And we can associate the different *ways* of appearing that explain intersubjective differences with that object.

Absences are not spatiotemporal particulars, so this schema of explanation is not available in a case of absence experience. Might there be a way to give absences,

¹⁷⁰ Except for Gow's (op cit.), I haven't encountered any view to the contrary. For a classic discussion that absence perception can be non-epistemic, see Sorensen (2008), pp. 246-47.

as ‘ways the world is’, to borrow Mac Cumhaill’s words, a role in perceptual experience?

3. Ways the world is

3.1

It will be no surprise, given the amount of argument collected so far to the contrary, that my attempt at defending the veridicality of absence experience will seem a bit radical. In one of the sections of the work that inspires this attempt, Mark Johnston himself asks, ‘is this madness?’

The idea Johnston fears will be received as mad is that presence – ‘the variety of ways in which real or ostensible items, be they objects, qualities or whatever, disclose some aspect of their nature’ (Johnston 2004, p. 233) – is independent of the mind, ‘objective’, in Johnston’s phrase. By ‘presence’ Johnston means something like what Russell (1910-1911, p. 108) meant by ‘presentation’, which yields the thought that presence or presentation is a relation between an object and a subject, and so that presence ‘is fundamentally presence to something, a self or subject of experience’ (Johnston *op. cit.*, p. 257).

On this familiar view, because an object is present to a subject only when and where the subject is, presence is a ‘local phenomenon’ (p. 234), and it is brought about by a mind’s entering a perceptual relation with reality.¹⁷¹ This predicts

¹⁷¹Two clarifications: Johnston swerves between calling our side of the relation ‘mind’ and ‘consciousness’; and he thinks presence can be also an intellectual phenomenon, not just perceptual, as when propositions are ‘THERE, [...] available as objects of attention’ (p. 234).

presence to cease to exist whenever ‘the last individual consciousness’ does. It is minds that are the ‘Producers of Presence’ (tongue-in-cheek capitals Johnston’s). The mad proposal is that minds are not producers but rather ‘Samplers of Presence’. Reality is already there, available to be sensed by us. Facts about a subject do contribute to the explanation of her experience of reality, but they do because they illuminate the subject’s particular viewing conditions. In the case, for example, of looking up from one’s desk and seeing, through the window, ‘one’s dogs running in the front yard [...] the perceptual experience is of the dogs and their running being present in a certain way’ (p. 233). This ‘certain way’ includes the dogs looking oblong rather than square, if they are running sideways rather than towards the window; it includes their fur looking bright and reddish rather than matte and brownish, if the sun is out. All these are, in Johnston’s Fregean phrase, ‘modes of presentation’ of what one sees. But what one sees has those modes of presentation built-in. ‘[A]ll the possible ways of [...] sensing each such thing come into being with the things themselves, whether or not there are any individual minds to sample these modes of presentation, i.e. to access them in individual mental acts’ (p. 235).

Where for Frege a mode of presentation is something like a description an object uniquely satisfies, such that it (the mode of presentation) is an entity associated with the thinking subject, something *she* contributes to the phenomenon of reference, for Johnston a mode of presentation ‘*is just* [...] some item or other in the world, presenting in a certain manner’ (p. 246). Johnston speaks of modes of presentation being ‘perspectival’. One might read this to mean more than just the geometrical feature of a scene looking a particular way from a particular point of view, because the notion of ‘point of view’ seems to include, for him, more than a spatial position. ‘Each one of us finds him or herself at the center of an arena of

I focus on the perceptual case, which he also thinks is ‘the best way to bring presence into view’ (ibid.).

presence and action', Johnston writes (p. 260). His examples do involve mostly perspectival – in the geometrical sense – differences among the ways in which one and the same thing appears to differently-positioned subjects, but it is natural to include other differences in ways of appearing. A creature with a different number of cones in their visual system will be sensitive to different changes from the bright-and-reddish/matte-and-brownish change *I*, in my conditions, might be sensitive to when it gets cloudy in the dog scene. Someone with more training than I would be sensitive to a taste I am not when sampling wine. Johnston deals with this phenomenon by positing a distinction between the *properties* some item has, say the surface reflectance profile of an object (which we refer to via the colour words), and the *qualities* we are perceptually aware of when sensing the item in a particular occasion (which we also talk about using those words). A green thing can have a blue-ish *quality* in certain conditions and for certain subjects, even though it does not have the *property* of being blue.

This distinction between properties and qualities seems idiosyncratic, but it is important to the idea of presence being 'sampled', not 'produced', so let me put the point in the more familiar terms of things having similar appearances or looks (Martin 2010). On such a view, to us and in certain appropriate conditions, the flower will have a look that is similar to that of paradigmatic blue things. Looks ascriptions like '*o* looks blue' do not pick out one particular property borne by all blue-looking things; rather, there are various, distinct ways of looking, for example one borne by a blue flower and one by an ultra-violet flower, which are qualitatively similar. Viewing the blue look of the ultraviolet flower as distinct from the blue look of a bluebell accounts for the intuitive intransitivity of the similarity relation. A looking similar to B and B to C does not entail A looking similar to C. This intransitivity wouldn't be respected if we ascribed one and the same look to all blue-

looking things. So this view, like Johnston's, posits a two-layered kind of perceptual contact with presented objects. Items have basic properties, say shape and colour, and in virtue of having these, items have looks. We don't, as it were, perceptually access looks *rather than* the things themselves and their basic properties; that's not what the two-layered view of perceptual contact means. What it means is that there is, as we might pre-theoretically think, a distinction between appearance and reality, both of which are aspects of what we see. A stick can be actually straight but look bent if half-submerged. The bent look is similar to the look of actually bent things; in that, the stick strikes us as having something in common with them. But its straightness also plays a role in our experience. The stick wouldn't look bent if it weren't straight, after all.

The point that presence is objective can be put in these terms by saying that both being and appearance – both being straight and looking bent – are objective features of things. Whether or not we accept Johnston's quality/property distinction, we can accept the point that appearances don't come into existence when reality is perceived. Appearances, or the ways reality is present, discloses itself, are independent of our experience. The fact that appearance is 'perspectival' (in the comprehensive sense of the term) doesn't contradict this; it doesn't entail it is somehow mind-dependent. Reality itself has several 'modes of presentation', in Johnston's terms, or 'looks', in Martin's, or 'appearances', more generally; and different subjects in different conditions will access different such modes, looks, or appearances. That a red apple has a green-ish look is still true of it even when there is no colour-blind

subject in the context to access that look.¹⁷² In Johnston's terms, we might say the apple is present in that way nevertheless.¹⁷³

My proposal that absence is no different from presence in Johnston's sense. Just as we sample, not produce presence, so too we sample absence. A lesson from the discussion so far in this section is that it is a pervasive feature of experience that we access reality in different *ways*, because reality is readily present to us those ways. So perhaps we access absences in different *ways*, because different things fail to be present in different ways, too. It is natural to extend the view that Johnston samples his dogs running in the front yard in a particular way, whatever 'dogs running' refers to (a state, an event, or fact), to a view on which the dogs' being absent is also something he could sample, and also in different ways. If the dogs' presence is 'the variety of ways in which [the dogs] disclose some aspect of their nature', the dogs' absence would be 'the variety of ways in which [the dogs] fail to disclose any aspect of their nature'. Johnston might access one of the ways the dogs disclose some aspect of their nature, and Johnston's friend, who is colour-blind, another such way. Similarly, the two of them access different ways for the dogs to fail to disclose any aspect of themselves. To Johnston, those dogs' absence might feel overwhelming; it might come with a deafening silence, which he has to get out of the house to stop sensing. (I'm thinking of a dog-bereaved Johnston.) To Johnston's friend, who disliked them, the absence might feel quite differently. But both are just sampling different ways in

¹⁷² That such a look-ascription is true in that context is compatible with its being non-assertible in that context, because no one is sensitive to it there. Cf. Martin (op cit.).

¹⁷³ As Umrao Sethi (forthcoming) notes, this might mean that, for all we know, items have potentially infinite ways of being experienced, only some of which we access, in certain conditions. This plenitude is less problematic than it seems if one remembers appearances are not to be counted as further elements to the object and its properties themselves. An object has ways of being experienced, which are relational properties, in virtue of its basic, non-relational ones.

which the dogs fail to disclose themselves, which is just as much an objective, sampleable way for the world to be as the dogs being present is.

Now, suppose Johnston liked one dog better than the other, and conversely his friend. Suppose Johnston looks up at the front yard and sees one dog only. Suppose, too, the dogs are almost identical twins. Perhaps before he looks closely, the absence of his favourite dog, Pierre, will look to Johnston similarly to the way the absence of the other dog, Paul, looks to Johnston's friend. They both look up at the same time and feel the same surge of dread upon seeing the one-dogged yard (what happened to Pierre/Paul? Did he stay at the vet? I'd swear they were in the garden together!). This means Paul's absence has a way of being experienced that is similar to Pierre's absence, just like red and orange objects have similar ways of looking.

Another situation. Only Pierre is absent from the house. Because Pierre is Johnston's favourite, Johnston knows what Pierre's little paws running to him sound like on the parquet; Johnston's friend doesn't. So, when they both come into the house, only Johnston recognises Pierre's absence, because only he knows that the paws they hear approaching do not include Pierre's. This, again, doesn't mean Pierre's absence is somehow subjective to Johnson. The absence is there, available to be detected by anyone who is exposed to it. Except Johnston's particular make-up is to Pierre's absence as a hummingbird's to ultraviolet, or a sommelier's to a complex wine. All three – the absence, ultraviolet and corked wine – are there for anyone to sample; only I, who don't know Pierre, lack a fourth cone and am useless with wine, will fail to experience the qualities of those things to which Johnston, hummingbirds and sommeliers *are* sensitive, even when I'm also presented with the worldly entities they are qualities of.

Now go back to the point that absence experience is recalcitrant, i.e. that it remains even when one's beliefs are incompatible with it. This seemed to speak to

Mumford's and Gow's view of it as an illusion. Just as the illusion that the Müller-Lyer arrows are different lengths remains even when one learns they are the same length, so too the experience of one's laptop's absence remains even when one learns the laptop is there, and a magician is just playing a mirror trick. This phenomenon would be explained, on my proposal, by the two ways of the world to be – the laptop being there and its being absent – having similar ways of being experienced. Again, this means they can be experienced similarly in different sets of viewing conditions, just like Pierre's absence feels similarly to Johnston as Paul's absence does to his friend, in one of my examples above. Gow's laptop's being there – that way for the world to be – looks to her, tricked by the magician, similarly to how someone's laptop's being absent looks to someone who isn't being tricked and was in fact robbed of it.

Consider now the case of total darkness. Light's being absent looks to humans black. Such seems to be the result of our possessing a pre-Newtonian visual system. But it might look different to a creature with a post-Newtonian visual system. Both are ways of total darkness' being experienced. Light's failure to disclose itself feels different to subjects with different kinds of visual systems. And someone might have access to no ways of a certain absence to be experienced at all. Consider our Construct user. She is aware of the emptiness around her, we had said, in virtue of her awareness of the boundedness of her visual field. This made it seem like her experience was fully explicable without appeal to anything being presented to her, but now, on the present view, that subject's ability to be aware of her sensory limits helps to make her *sensible* to the emptiness around her, just like the sommelier's training helps to sensitise her to the subtleties of wine. So a creature with no such ability – with no ability to be aware of her sensory limits – would likely not be sensitive to the emptiness around her in the Construct. Even though the belief that

the region in view is a sub-region of a larger space is not defeated by the typical defeaters of perceptual beliefs, it is still not an a priori belief. If a subject had been born in the Construct but never opened her eyes, or had been born without them, she wouldn't have been able to acquire that belief visually. So the emptiness would have been there, but she could have failed to visually sense it. If Martin (1998) is right that one does not have perceptual awareness of empty space through the body, as when, for instance, one holds out one's hands, then she would not sense the emptiness by touch either. (For all she knows, a stack of books may stand between her hands at any time.)

3.2

A slogan for this proposal might be that in absence experience, we *are* perceptually aware of something, of a way the world is. What we are aware of in such cases may not an object in the discussed, Dretske-Williamson sense, yet it does contribute to the character of the experience in a certain way, just like present objects can do. This proposal hinges on denying that perceptual awareness is necessarily of an object, which naturally dovetails with Johnston's denial that presence is necessarily to a subject. The compatibility of the two denials can be seen as follows. If objects' ways of being experienced are there for us to sample, then those ways can fail to be sampled by some subject. Such is part of the concept of availability. Similarly, a subject who is perceptually open to the world can fail to exploit such openness with regard to some object, to fail to sample its presence, because the subject does not come into contact with the object or because the object has no way of being experienced that the subject is sensitive to, i.e. that the subject is equipped to

sample.¹⁷⁴ Yet, as the examples of experiences discussed in Section 1 show, in some occasions where presence fails to be sampled, perceptual openness is exploited by the sampling of an absence itself.

Let me unpack this paragraph by taking, in turn, the denial that perceptual awareness is necessarily of an object and how it is that in some occasions, the openness unexploited by presence is exploited in sampling an absence itself. Following Phillips (op cit.), we can cash out the rejection of the view that perceptual awareness is necessarily of an object in two – mutually compatible – ways. The first is as challenging diaphaneity, understood, at least, as the view that the character of experience depends entirely on the presented objects. As Phillips puts the target view, ‘[i]f experience can entirely be analyzed in terms of its objects, then where there is no object, there is no experience’ (Phillips op cit., p. 345). His own argument to reject this, which he employs to defend the idea that absence can be experienced in audition (i.e. that silence can be experienced), is to say that, analogously to how sight makes us aware of a visual field independently of any objects standing there, in audition we can be aware of a temporal extension independently of any sounds extending over it. A subject whose experience consists solely in this soundless temporal extension differs from a deaf subject in that in such an experience she, unlike the latter, is perceptually open for a sound to make itself present.¹⁷⁵ Now, the other, compatible way of denying that perceptual awareness is necessarily of an object is by rejecting that there are sharp boundaries between perceptual modalities. The following goes beyond Phillips’ point, but the thought occurs to one naturally from reading him that if a deaf person develops a hypersensitive sense of touch that allows

¹⁷⁴ This might be the situation in our experience of total darkness. There might indeed be some light present in the region of space perceived, but not enough for human vision to register perceptually.

¹⁷⁵ Phillips (op cit.), pp. 344-57.

her to be aware of sounds by their vibrations, then she, too, will be able to experience silence. So the hearing and the deaf may experience one and the same thing – the same absence – differently, which is to say that one and the same absence will have those two ways of being experienced. It does so in virtue of what it is an absence of. Part of the nature of sounds is that they can fail to disclose themselves in different ways to differently sensitised subjects: in silent and in vibrationless ways, respectively.

This example takes me to the second thing I wanted to address: that sometimes, the openness unexploited by presence can be exploited by sampling an absence itself. Here is where structural features of experience come in. In addition to having, as we saw, an epistemic role in our experiences of absence, structural features have a role in explaining character. The two roles are related. The kind of knowledge we gain in absence experience, I suggested before, can be regarded as negative knowledge. It is knowledge about what is not presented to us in experience. This knowledge does not have, by definition, a ‘presentational’ justification, to use Alford-Duguid’s phrase: it is precisely knowledge about what is not presented. But we saw that it has a different, still perceptual kind of justification: ‘structural’. The negative knowledge in question tends to be specific, though; it tends to encode information about what the experience is not of. Let me illustrate this point by showing how different structural features help in our acquisition of knowledge about different kinds of absences in the examples we have discussed. One structural feature of experience is that it makes us aware of a temporal extension (which functions as something like an auditory field). This feature sensitises us to the absence of *auditory objects* (but it has no role in our sensitivity to the absence of visual objects). Another structural feature is that in experience we detect the edges of materials. This helps to explain our awareness of the absence (of material) located in, or identified with, holes. Another feature of our visual system is that it is pre-Newtonian. This helps to

explain our awareness of darkness. And the feature of having a figure/ground structure helps to explain our awareness of the absence of individuals, which can fail to appear on a ground. Finally, the feature of visual experience that during it we are aware of the boundedness of the visual field helps to explain our awareness of a more general absence, the absence of any visual objects (as in the Construct). So, generally, the structural features of experience do, as predicted by the literature, have a role in explaining our phenomenology during absence experience; yet this phenomenology is not exhaustively explained by citing those structural features. An absence has to be cited as being sampled or sensed. The structural features of experience are just part of what makes us perceptually receptive to them.

3.3

I would like to stress two things before closing. One is that this proposal does not simply replace what fulfils the role of object of perception. As I suggested in Section 2, popular views like Dretske's, while indeed employing the notion of 'object' formally, take a stance on what can actually fulfil the role. Following similar assumptions, authors have taken absences not to be possible objects of perception. Even granting absences exist, one realism-friendly view, Mac Cumhaill's, held that only spatiotemporal particulars, such as absential locations, can fulfil that role. Here I've suggested that absences *themselves*, which may in fact fail to meet the assumed criteria for objecthood, can be veridically experienced. I've remained neutral as to how exactly to make sense of such a state, whether to regard it as a state of object or fact perception (assuming again those are the candidates). The more general, and admittedly vaguer phrase that absence is 'a way the world is' is flexible. But so is the notion of presence.

We do know a little about these ways the world is, however. Johnston tells us that one aspect of the view he opposes is that presence is a ‘local phenomenon’; but by that Johnston means presence doesn’t come into being when an item is presented to a subject. His point is that things’ being present was a phenomenon already happening there. In this sense, then, presence *is* in fact local. So much is admitted by Mac Cumhaill for the converse of presence (absences are ‘ways the world is at a particular locale’; only these local phenomena, she tells us, have no sensible properties. Now, on Johnston’s quality/property view, it is qualities, not properties, that explain the character of our experience of items, such that Mac Cumhaill’s remark doesn’t obviously threaten entities without sensible properties. (To repeat: the flower’s blue-ish quality explains our experience of it and its ultra-violet-ish quality explains the hummingbird’s. The flower’s property of having a certain reflectance profile explains *those* qualities at best.) This can sound as though our perceptual contact with reality is mediated, and that we access just the quality layer of reality, not the layer of objects and properties themselves. To avoid getting ourselves into exegetical questions about Johnston’s view, we can just rely, again, on the more familiar view on which things can have similar appearances or looks. This view, while not supposing such a mediated nature of perceptual contact, nevertheless posits a two-layered structure in the way appearances or looks are explained. Our experience of reality involves a perceptual relation to objects, which, in addition to non-relational properties like colour and shape, have, in virtue of those colours and shapes, certain relational properties, appearances or looks, which they display – but don’t have or cease to have – relative to certain viewing conditions. If our experience of absence is also an experience of reality, then we should expect this two-layered structure of experience to pervade it. And so it does. There is no object in such cases

with which to associate looks or appearances, to be sure, but in this the case of absence experience is not alone. Not only objects have ways of being experienced.

The best non-absence example might be particular property instantiations, or ‘tropes’. Take the sentence ‘John saw the beauty of the rock formation’ (Moltmann 2013, p. 51). Beauty characterises, on the present view, a way something has of being experienced. In this case, that something is ‘the rock formation’, which Friederike Moltmann argues is not an event, state or situation but a trope. Yet tropes are abstract, not concrete, particulars. So tropes fail to meet the standard of objecthood. They, like absences, fail to have sensible properties such as colour and shape. However, a trope can be experienced as beautiful, such that ascribing beauty to a trope needn’t be regarded as false. There may be other examples of items to which we ordinarily ascribe qualities in reports but lack sensible properties. Redness – the universal redness – is abstract, yet it’s perfectly felicitous for Nabokov to write a phrase like: ‘he felt a roaring redness fill his head’. And the referent of ‘Jon crying’, from our earlier example, can be an ugly thing to see. Now, I don’t intend to weigh in on what the right ontological category for ‘the rock formation’ (or ‘Jon crying’, etc.) is. I just want to point out that there are all kinds of entities that may also fail to be sensible in the traditional sense – because they may be abstract or causally inert, say – and yet seem to contribute, with their ways of being experienced, to the character of our experiences. Our attribution of qualities to those things in perceptual discourse is a consequence of it. Absences, as ways that the world is, are just another such kind of item.

The second thing I want to stress is that this view is intended to cover, but not to be limited to, visual perception, but it *is* intended *not* to cover what authors call ‘cognitive phenomenology’, at least as deployed when explaining a particular kind of experience, the experience of holding (non-perceptual) propositional

attitudes, most notably beliefs. As the reader may have noticed, I have phrased the phenomenon in question as ‘seeing’, ‘perceiving’, and ‘experiencing’ absence. This may seem like a cheat, since authors who hold an illusion view still think we have an experience of absence, although it is not explained by appeal to anything we perceive. Take, for example, Gow’s (op cit.) example of experiencing the absence of your best friend’s text message congratulating you on your birthday. Among the rest of the text messages, you clearly have a sense of their one being missing. Suppose that we are not talking about the experience you have as you go through your day but rather when you get home at night, having forgotten your phone, and check the text messages. My intuition is that experiencing the absence of your best friend’s name among the other names – you don’t even open the messages; you just scroll through the names on the notifications screen – has an important perceptual component. What does your friend’s failure to send a message look like (to you)? Like this! And it has a similar a way of being experienced as another way for the world to be: your friend’s having sent the message and unwittingly turned her phone off before the message could go through. You could have, of course, learned by testimony that your friend didn’t send a message (someone at home saw the notifications and told you), and held the belief that her message wasn’t there. But the phenomenology of this experience is starkly different from the one of going through the notifications yourself, tired and alone at night, the screen glowing in the dark, and failing to find your best friend’s name. Many of the cases discussed here, as this one, rely on multimodal perception, and many will need to be explained partly by appeal of cognitive attitudes (expectations, again, or body schemas). This does not mean that in being exposed to absences, you are enjoying any less of a perceptual experience than you would be when exposed to your dogs running in the yard, or a beautiful rock forming. All those have different ways of being experienced, and depending on

your viewing – or, more generally, experiencing – conditions, which may well include your cognitive attitudes, you will access one such way.

Conclusion

How does this all vindicate absence experience discourse? Well, on the illusion view, reports like ‘I see something black’ when seeing total darkness, ‘I see something stubborn’ when seeing Pierre’s absence, and ‘I see something big’ when seeing holes are all false. On my view, blackness, stubbornness and bigness are bona fide attributes of absences, available for us to sense, in the appropriate conditions, because they follow from the way those different absent things fail to disclose themselves. Light fails to disclose itself in a way we experience by seeing total blackness (but some other creature might as seeing white); Pierre fails to disclose himself in a way that Johnston experiences as sad (but Johnston’s friend, for whom Pierre was annoying, might experience as relieving); ink fails to disclose itself in a circular region on an otherwise black page in a way that we experience as a hole (but some other creature might experience as a white object). Blackness, sadness and bigness are objectively attributes of those entities just like a blue look is objectively an attribute of the ultraviolet-reflecting flower (even if a hummingbird experiences it as violet-ish), sadness of a piece of music (even if someone experiences it as dull), and thickness of the very piece of paper on which the absence of ink is. Now, here I am perhaps assuming too much about the objective ways of being experienced of present things themselves. I am assuming sadness somehow lies in the music, for instance, and stubbornness in people, just as those can be attributes of absences. Let my view take the form of a conditional, then. If there is a way to make sense of my experiencing music as sad by attributing that way-of-being-experienced to the music itself, then there is a way to

make sense of experiencing darkness as black, holes as big, and loved ones' absences as stubborn by attributing those way-of-being-experienced to the absences themselves.

This shows two ways in which the view needs refining. One is the issue of how exactly absences can be experienced in any way if they have no basic sensible properties; another is the issue of how they can be experienced as stubborn or sad, if they can be experienced at all. Both are, however, issues that pertain to other kinds of entities too, most notably tropes. At any rate, what I've tried to do here is persuade the reader that such refining is even worth pursuing, i.e. that there is theoretical room at all for a veridicalist view of absence experience. I hope also to have conveyed the spirit of my motivation. The variety of ways in which absences are experienced seems as wide as the variety of ways in which present items are, and in perceptual reports we record both kinds accordingly. This is just a pervasive feature of our experience of the world. If absences are *there*, available to be sampled in different ways and become 'the topics of thought and talk', as Johnston says, then they contribute, via those different ways of being experienced, to our general phenomenology just as red things (via their now orange, now redd-ish appearance) do. The key to accord absences a place in the realm of entities we see may thus be to stop thinking of them as objects that are required by 'philosophical theory', but fail to, strike our retinas. And perhaps to stop regarding sense experience as just as plain a phenomenon as stepping on things, too.

References

Alford-Duguid, D. Ms. On the Epistemic Significance of Perceptual Structure

- Barker, S. and Jago, M. 2012. Being Positive About Negative Facts. *Philosophy and Phenomenological Research*, 85: 117–138. doi:10.1111/j.1933-1592.2010.00479.x
- Brogaard, B. 2018. *Seeing and Saying*. Oxford: OUP
- Dretske, F. 1969. *Seeing and Knowing*. Chicago: University Of Chicago Press
- Dretske, F. 1981. *Knowledge and the Flow of Information*. Cambridge, MA: The MIT Press
- Dretske, F. 1994. If you can't make one, you don't know how it works. *Midwest Studies in Philosophy* 19, 468-482
- Dretske, F. 1999. The Mind's Awareness of Itself. *Philosophical Studies*, 195: 103–124.
- Drestke, F. 2000. *Perception, Knowledge, and Belief*. Cambridge: CUP
- Farennikova, A. 2013. Seeing Absence. *Philosophical Studies* 166 (3), pp. 429-454
- Farennikova, A. 2015. Perception of Absence and Penetration From Expectation. *Review of Philosophy and Psychology* 6 (4), pp. 621-640
- Finn, S. 2018. Is a hole a real thing, or just a place where something isn't? *Aeon*, <https://aeon.co/ideas/is-a-hole-a-real-thing-or-just-a-place-where-something-isnt>
- Ganson, T. 2020. An Alternative to the Causal Theory of Perception. *Australasian Journal of Philosophy*, doi:10.1080/00048402.2020.1836008
- Gow, L. 2020. A new theory of absence experience. *European Journal of Philosophy* 29, doi: 10.1111/ejop.12577
- Juvshik, T. 2018. Abstract Objects, Causal Efficacy, and Causal Exclusion. *Erkenntnis* 83 (4):805-827
- Johnston, M. 2007. Objective mind and the objectivity of our minds. *Philosophy and Phenomenological Research* 75/2: 233-268

- Lewis, D and Lewis, S. 1970. Holes. *Australasian Journal of Philosophy* 48(2), 206-212, doi: 10.1080/00048407012341181
- Martin, M. G. F. 1998. Setting Things Before the Mind. In Anthony O’Hear (ed), *Contemporary Issues in the Philosophy of Mind*. Cambridge: CUP
- Martin, M. G. F. 2010. What’s in a look? In B. Nanay (ed.), *Perceiving the World*. New York: Oxford University Press
- Martin, J.-R. and Dokic, J. 2013. Seeing Absence or Absence of Seeing? *Thought: A Journal of Philosophy* 2, doi: 10.1002/tht3.72
- Mac Cumhaill, C. 2015. Perceiving Immaterial Paths. *Philosophy and Phenomenological Research* XC, doi: 10.1111/phpr.12037
- Mac Cumhaill, C. 2018. Absential Locations and the Figureless Ground. *Sartre Studies International* 24, doi:10.3167/ssi.2018.240104
- Moltmann, F. 2013. *Abstract Objects and the Semantics of Natural Language*. Oxford: OUP
- Mumford, S. Forthcoming. *Absence and Nothing*. Oxford: OUP
- Nunberg, G. 1993. Indexicality and deixis. *Linguistics and Philosophy* 16, pp. 1-43
- Phillips, I. 2018. No more than meets the eye. In T. Crowther and C. Mac Cumhaill (eds.), *Perceptual Ephemera*. Oxford: OUP
- Quine, W. V. O. 1948. On What There Is. *The Review of Metaphysics*
- Quine, W. V. O. 1960. *Word and Object*. Cambridge, MA: MIT Press
- Russell, B. 1905. On Denoting. In R. C. Marsh (ed.), *Logic and Knowledge*. London: Allen and Unwin.
- Sartre, J-P. 1943. *Being and Nothingness*. New York: Philosophical Library
- Sethi, U. Forthcoming. Diaphaneity and the Ways Things Appear. In Otavio Bueno & Jan Voshoolz (Eds). *Essays on Markus Gabriel’s New Realism*. Synthese Library, Springer

- Sorensen, R. 2008. *Seeing Dark Things*. Oxford: OUP
- Steenhagen, M. 2017. False reflections. *Philosophical Studies*, 174, doi:
10.1007/s11098-016-0752-x
- Strawson, P.F. 1950. On Referring. *Mind* 59: 320–344
- Williamson, T. 1990/2013. *Identity and discrimination*. Oxford: Blackwell
- Williamson, T. 2000. *Knowledge and its Limits*. Oxford: OUP