

## The Evolution of Nagel's Panpsychism

0.

Nagel's 1979 article 'Panpsychism' is rightly credited with provoking the recent resurgence of interest in panpsychism, the thesis that in some sense mind exists throughout the natural world, after more than half a century of neglect. There is now a considerable and burgeoning literature on panpsychism and related positions, and though Strawson's 2006 target paper and volume has perhaps been a more proximal cause (see also Chalmers' discussion of panpsychism in his 1996, and Seager 2006), in none of these cases does the reasoning depart very far from Nagel's concise and powerful formulation. Thus Nagel is the true father of contemporary panpsychism.<sup>1</sup>

It is not so clear whether this offspring was ever desired in a wholehearted fashion, however. In his 1979 paper Nagel comes to the conclusion that panpsychism is no worse than any of the other 'hopelessly unacceptable' solutions to the problem of how minded creatures like us arise in a material world. And even in that paper there were signs that he was at least open to a similar, but apparently different, non-panpsychist, position. His recent work (2000, 2012) seems to make this trajectory of thought even clearer: he now appears explicitly to endorse not panpsychism but *neutral monism*.

In this paper I will trace the path of Nagel's thinking, from the reasons that led him to ambivalent embrace of panpsychism, to his current view. Having arrived at that view I will consider how to make best sense of it. Is it panpsychism, or not? And were the seeds of the view present all along?

1.

The problem Nagel confronts in 'Panpsychism' is one aspect of the 'mind-body problem', the broad question of how mind and matter interrelate, which is known as the *problem of consciousness*. Chalmers in a now famous phrase labels this the 'hard problem' - he claims that consciousness comprises the real challenge for the project to fit the mind into the material world that is so successfully described by science.<sup>2</sup>

What is consciousness, in this context? The term has a technical use that distinguishes it from such notions as political consciousness, consciousness in the sense of being awake, or self-consciousness (though the latter two notions are related). Consciousness as concerns the mind-body problem - or phenomenal

---

<sup>1</sup> For twentieth-century panpsychists that pre-date Nagel, see Whitehead 1925 and Hartshorne 1934.

<sup>2</sup> Chalmers 1996.

consciousness, as it is called<sup>3</sup> - is roughly equivalent to the idea of *sensation*. Though a lump of coal can be cold, and can be kicked by somebody, it feels nothing in either case. It is wholly inert and unfeeling. You are quite different: you would feel the cold, and you would feel pain as a result of the kick. This commonsense difference between the lump of coal and you can be put by saying that you have conscious mental states, or phenomenally conscious states in the jargon - also known simply as *experiences*. In other words you feel, you have sensations. In Nagel's phrase, there is something it is like to be you.<sup>4</sup> There is nothing it is like to be a rock, as far as we know. Those entities in the world that at least sometimes have sensations we say are phenomenally conscious. But for most purposes philosophers think about the conscious *states* of creatures - those states of themselves that they feel, in other words of which they are consciously aware. Comparatively few of the states of even conscious creatures are conscious: you do not feel the growth of the myelin sheaths around your neurons, or the condition of the enzymes in your stomach, for instance. By contrast a good deal of mental life is conscious: for example the visual images you are currently having of this page, and your accompanying thoughts. And those parts of your mind of which you are currently unaware, such as memories and long-term beliefs are - it is widely thought - at least capable of consciousness.<sup>5</sup>

What is the *problem* of consciousness? Think of the lump of coal again. It feels nothing, it is dead matter. Yet it is carbon, and we are made mostly of carbon. So how is it that when you mix carbon like that with enough water, nitrogen, salts, and some other ingredients, all of which are like the carbon in being insensate, and thereby build a functioning brain and organism out of them - a creature like one of us - that mixture decidedly *does* feel? How does sensation come out of the insensate? What magic trick is performed to make this happen? To conjure feeling out of the unfeeling seems one of the greatest jumps in all of Nature. Similar thoughts famously drove Descartes to posit the conscious mind as an immaterial bubble mysteriously tethered to the physical body. Cartesian dualism is out of favour these days in analytic circles, and most philosophers working on the problem of consciousness proceed on the assumption that we are material creatures. This trend has to do with the march of explanatory success in biology, notably neuroscience.<sup>6</sup> As a result, the conundrum of our 'material mentality'

---

<sup>3</sup> Block 1995.

<sup>4</sup> Nagel 1974.

<sup>5</sup> Philosophers disagree over which parts of mental life are capable of consciousness - for example, some believe (not consciously of course) that beliefs are never conscious. Others hold that occurrently believing a proposition is a paradigmatically conscious episode. I will largely ignore such controversies in this paper, by focusing on sensory states, which provide the standard examples of conscious mental states in the literature.

<sup>6</sup> See the Appendix of Papineau 2002 for an account of this reasoning.

must be faced squarely. But to say any more at this point would be to pre-empt premises in Nagel's argument for panpsychism. So let us turn next to those premises.

2.

Distilled,<sup>7</sup> the premises that carry Nagel to panpsychism, 'the view that the basic physical constituents of the universe have mental properties',<sup>8</sup> are the following quartet:

**P1. *Material composition:*** *Any living organism is wholly composed of a huge number of physical particles.*

*Explanation:* This is simply the denial of dualism, and an expression of adherence to the scientific orthodoxy about our nature. We are made of the matter we eat, and that originally came from stars. As Nagel says, 'If [star matter] were brought to earth, and grass were grown in it, and milk from a cow that ate that grass were drunk by a pregnant woman, then her child's brain would be partly composed of that matter.' (1979: 181) We are wholly material beings.

**P2. *Non-reductionism:*** *An organism's conscious experiences are not explained by any physical properties alone.*

*Explanation:* No physical account, an account in terms of the sorts of properties scientists measure and quantify, can explain why an organism has conscious states, or why those states are of the particular kind they are – why on a certain occasion you see red not blue, or feel pain instead of a tickle. This premise is vividly exemplified in the imaginary case of Jackson's Mary, the scientifically omniscient scientist who has never seen colour.<sup>9</sup> Jackson argues that Mary's complete scientific knowledge does not enable her to know what it is like to consciously experience red. She only finds out what it is like to experience red when she finally sees a red rose. Thus physical accounts cannot explain consciousness. Another expression of the same broad idea is the common thought that other people may systematically see colours different to those I see – perhaps they see green where I see red. Colourblindness of the familiar kinds is detectable due to physical and behavioural differences. But it appears conceivable that even someone physically identical to me, in the relevant respects, might see different colours. Such a 'colour inversion' would be

---

<sup>7</sup> I say 'distilled' because I have made minor alterations to some of Nagel's wording and phrasing, to aid clarity and ease of comprehension. See 1979: 181 for the original premises.

<sup>8</sup> 1979: 181.

<sup>9</sup> Jackson 1982. Antecedents of the argument can be found in Broad 1925 and Russell 1927a.

scientifically undetectable. This sense of the gap<sup>10</sup> that seems to exist between science's explanations and the nature of conscious experiences finds its most potent expression in the notion of a physical *zombie*,<sup>11</sup> a creature physically identical to a conscious human being who lacks experiences *altogether* - it feels nothing, despite manifesting complex behaviour that would lead you to believe otherwise, as well as all the requisite brain activity. The coherence of the zombie concept is meant to indicate that science cannot explain the fact that we are conscious at all.<sup>12, 13</sup>

**P3. *Realism about consciousness:*** *Consciousness belongs to the organism; it is a feature of its material constitution.*

*Explanation:* This premise affirms that consciousness is a real phenomenon, what some philosophers have doubted partly in view of the difficulties of explaining it, and that it is a property of the organism: the organism is conscious, not any 'soul' associated with it.

**P4. *Non-emergence:*** *All global properties of a complex system intelligibly derive from properties of its constituents, plus their arrangement.*

*Explanation:* Consider some liquid water at room temperature (to take a hackneyed example). The body of water is a complex system – a massive

---

<sup>10</sup> Levine 1983 influentially calls this the 'explanatory gap'.

<sup>11</sup> Chalmers 1996. See also Leibniz 1714 §17, Kirk 1974.

<sup>12</sup> It is worth noting that Nagel himself appeals to the way that the objective methods of science seem to omit what is distinctive about the subjectivity of consciousness, in making the case for the thesis of premise 2 (see e.g. his 1974), instead of appealing directly to the considerations I mention here (thank you to Leopold Stubenberg for raising this point). But these ideas are connected: one view might be that the inexplicability of consciousness by the physical, as expressed in the Mary story and the zombie thought experiment, is ultimately due to the elusiveness of subjective phenomena to the objective point of view.

<sup>13</sup> Premise 2 might seem unduly controversial; in particular it might seem to beg the question against the view that consciousness is purely physical. But this is not so, since many, perhaps most, contemporary physicalists actually accept that a physical explanation of consciousness is not possible. Physicalism, they emphasise, is primarily a *metaphysical* thesis about the nature of consciousness, not a theory about whether a satisfying physical *explanation* of consciousness can be given. Many of these theorists devote their energies to a physicalist explanation of why a physicalist explanation of consciousness is not possible, despite physicalism's truth. This position was not current at the time of 'Panpsychism', which may explain why Nagel does not consider it. Nonetheless, such 'type b physicalism', as Chalmers calls it (2003), faces serious objections, which Nagel would presumably endorse (e.g. Chalmers 2003). For examples of the position see Loar 1990, Papineau 2002. In effect premise 4 of Nagel's argument can be read as implicitly rejecting these positions, which deny an intelligible connection between an organism's physical matter and its consciousness.

arrangement of H<sub>2</sub>O molecules. The non-emergence claim is that all properties of the body can be transparently traced to, explained in terms of, the properties of its constituents. In other words every property of the water as a whole can be understood as a product of the properties of the molecules that form it, plus their interactions. Consider two examples of such properties: the water's liquidity, and its mass. That the water is liquid means it displays certain sorts of characteristic slopping and pouring behaviours; it is not solid, but not gaseous either. This is explicable in terms of the way the molecules are bonded – not so loosely that there are large gaps between them, as in a gas (i.e. steam), but loosely enough that they can slide around and over each other. That 'looseness' is ultimately explained in terms of the bipolar nature of each individual water molecule, and the implications this has for their inter-bonding. The mass of the water as a whole is explicable as the sum of the masses of its constituent molecules, whose masses are explained in turn by the masses of their constituents, down to fundamental particles. Roughly,<sup>14</sup> we need only add the masses of the fundamental constituents of the water to get its global mass. For there to be an 'emergent' property of the water, it would have to be a property not due in any clear way to the properties and behaviour of its constituents. It would be a property that popped out of nowhere: not just an explanatory, but a metaphysical, surprise. From premises 1, 3, and 4, it follows that the properties of those material systems that are human organisms must also be explicable in terms of the properties of the systems' material constituents, including the property of being conscious.

The two properties of water just considered, as well as being clearly non-emergent, are relevant in another way to Nagel's argument for panpsychism. In the case of liquidity, the system property is explained in terms of parts that do not have it: individual H<sub>2</sub>O molecules are not liquid. It is their group behaviour that constitutes liquidity. As regards the water's mass, however, the opposite is true: every material part of the water body has mass. Clearly, no massless component could contribute to the global mass of the body. This illustrates an important point for Nagel's argument – while some system properties can be explained in terms of parts that lack them, there are other, special, properties which seem to have to go 'all the way down': if a system possesses such a property, so must all the system's parts, down to the most tiny.<sup>15</sup> Nagel's argument is aimed at showing that consciousness is like mass in this respect: its presence at the macroscopic level of human organisms implies that it is fundamental.

---

<sup>14</sup> Because nuclear bonding involves energy, and some mass is thereby lost, the relation is not an entirely linear one.

<sup>15</sup> See also Strawson 2006: 15-16 on this point.

In sum: if creatures who are built of nothing but matter have conscious states, and these states are due to nothing but their material constitution, while the physical properties of that constituting matter – those properties measured by hard science – cannot explain consciousness, but still consciousness, like any other property of a complex system, *must* have an explanation in terms of the properties of the system's constituents, then it follows that there must be special non-physical properties of those material constituents that do suffice for consciousness. In other words there must be some secret properties of matter with a direct connection to consciousness, such that when you put matter together in the right way, as a brain (and perhaps body too), you get a conscious being. If the only property that would suffice is consciousness itself, then the argument shows that the material constituents of conscious organisms must themselves have conscious states – must feel something, have sensations, in other words. There would be things much tinier than organisms that have some form of inner mental life. There would be something it is like to be an electron!

So far the argument, if it works, shows that conscious organisms must have constituents with special non-physical properties that together create consciousness – those constituents must themselves be conscious, on the model of the explanation of macroscopic mass. But one or two more points need adding to get us all the way to Nagel's panpsychism, the claim that '[All] the basic physical constituents of the universe have mental properties'.<sup>16</sup> First is a point about the depth, and second a point about the breadth, of the special properties.

The question corresponding to the first point is whether the argument really requires that the very *smallest*, most fundamental,<sup>17</sup> material constituents – whatever are the basic subatomic particles, or Strings, for example – have the special non-physical properties, or whether it could be atoms, or molecules, or even bigger things like cells, say, that possess the key to consciousness: which would leave the more basic material constituents with only mundanely physical properties. Think of nature as a ladder with rungs: at the bottom are the smallest material things, electrons, quarks and so on (we currently think). On the next rung up are atoms; the next molecules, then proteins, then – eventually – composite material objects including living beings. The question is why consciousness must go all the way down to the lowest rung, as per Nagel's argument. Perhaps the argument shows that conscious beings require conscious constituents of some sort, but why not pick certain of their middle-sized parts,

---

<sup>16</sup> 1979: 181.

<sup>17</sup> Not all philosophers agree with Nagel's assumption that the universe's most fundamental level is also its most microscopic: a *priority monism* might see the universe itself as the basic metaphysical fact, with all facts about smaller things, including subatomic particles, dependent upon it (see e.g. Goff 2017). I will ignore this complication, and follow Nagel in linking metaphysical fundamentality with smallness.

like cells? Surely that is preferable to attributing consciousness to subatomic particles!

To see why Nagel's argument requires consciousness to exist at the most basic level of matter, consider the hypothesis that consciousness first arises at the cellular level: that it is the conglomeration of conscious cells that forms a human consciousness, but where the material constituents of these cells do not themselves have consciousness, only mundanely physical properties. Evidently we face the same puzzle as before in this situation, it has just been relocated to a lower level. Now instead of asking how a conscious *organism* could possibly derive its consciousness from purely physical ingredients, we must ask the same question of the conscious cell. And according to Nagel's argument, such a cell would require conscious components in order to explain the fact of its consciousness. It can be seen that this pattern will repeat for every sort of physical existent we consider above the basic level - every *complex* in other words. Nagel's argument will not be satisfied until we have ascribed consciousness to the fundamental material basis itself. This is one aspect of the parallel with mass.

But even this is still not quite Nagel's panpsychism, since could we not hold at this point that it is only the basic material constituents of *conscious organisms* that are required by his argument to have consciousness? That move would seem to do the least possible damage to the commonsense-scientific view of the world: if we must posit fundamental consciousness, then restrict it to the matter that makes up conscious beings, and keep the rest of the world's matter orthodoxly physical, i.e. non-mental. But once more the parallel with mass allows us to see why this restriction cannot be maintained. Given a massy object, like a table, one can in principle take matter from anywhere in the universe to make it up, to replace any mass it loses (imagine that a leg falls off). That possibility requires that all basic matter have mass. The same is true of a conscious organism: it can in principle be made of any matter in the universe; that matter only needs breaking down far enough and recombining into the right admixture of carbon, water, etc. (this was implicit in the explanation of premise 1).<sup>18</sup> Nagel does not include this principle, which we can call *fungibility*<sup>19</sup> among the premises of his argument, but he does add it afterwards,<sup>20</sup> and it is clearly needed for the full conclusion of *panpsychism* to follow - that conscious mentality is pervasive at the lowest material level. Once we see that a human consciousness can be made of any basic matter, it seems to follow by Nagel's argument that all such matter must itself have consciousness already.

---

<sup>18</sup> See Nagel's breathtaking list of the possible ingredients of a conscious being 1986:28.

<sup>19</sup> Coleman 2015: 67.

<sup>20</sup> 1979: 182.

This completes the explanation of Nagel's argument for panpsychism.

3.

There appears to be a loophole in Nagel's argument. The strict conclusion is that all fundamental matter possesses special non-physical properties, from which the conscious mentality of the organisms built of this matter transparently derives - just as the liquidity and mass of the Mediterranean derive from the properties of its basic material constituents. But unless we assume that *non-physical* implies *mental*, it need not follow that these special properties are mental. And if they are not mental properties, panpsychism is not entailed. The question, therefore, is whether there might exist some *third* kind of nature, neither mental nor physical, that belongs at the material world's fundamental level, and is capable of explaining consciousness.

It is a complication that expositions of his argument often ignore,<sup>21</sup> but Nagel is already alert to this possibility in 'Panpsychism'. Though no chain of purely physical explanations or inferences can end in something that could explain conscious mentality, Nagel maintains, it is conceivable that the mental explanatory chain and the physical explanatory chain each terminates in something that explains both - something that is therefore neither mental nor physical, considered just in itself.<sup>22</sup> On this view the fundamental constituents of matter would possess special properties that account both for the physical goings-on in the world - such events as the repulsions among electrons, atomic and molecular bonding, the building of a brain - and also for the fact that certain arrangements of matter have mental properties, including consciousness. This theory is not, it might seem, best described as panpsychism, but may better be termed a form of 'neutral monism' - the fundamental properties being of a single kind that is neutral between physicality and mentality. Neutral monism has a noble philosophical lineage, including among its proponents Ernst Mach, William James and Bertrand Russell. Still, the Nagel of 'Panpsychism' seems unwilling to join this distinguished list: He affirms that since, on his proposal, 'There would be properties of matter that were not physical [but] from which the mental properties of organic systems derived...[this] could still be called panpsychism.'<sup>23</sup>

---

<sup>21</sup> Van Cleve 1990 is a notable exception.

<sup>22</sup> P. 185. Also, on p.194 he proposes that 'the components out of which a [conscious] point of view is constructed would not themselves have to have [conscious] points of view', which at least suggests making consciousness out of the non-conscious (see Nagel 1974 for the close connection between consciousness and his notion of a 'point of view'). On the same page he talks of basic matter having not mental, but 'proto-mental', properties.

<sup>23</sup> 1979: 185.

Van Cleve objects to this piece of nomenclature,<sup>24</sup> and prefers C. D. Broad's term 'neutralism' for the aforementioned position. But he reasons that Nagel would be able to run a similarly structured argument, to the effect that neutral properties are no improvement on physical properties when it comes to explaining consciousness. (For, one might think, the problem Nagel's argument captures is that consciousness simply cannot derive from the *non-conscious*, and the neutral is no better off than the merely physical in that regard.) Thus Van Cleve agrees to ascribe full-blown panpsychism to Nagel: effectively, he holds that there *is* a gap in Nagel's argument, but it is one Nagel could close if he so wished.

I believe, however, that the situation is somewhat more complex, not to mention interesting, than Van Cleve suggests. Nagel evidently does not agree with Van Cleve that the same kind of gap would necessarily exist between the special neutral properties and consciousness as exists between physical properties and consciousness, since in subsequent work he has gone on to argue explicitly for neutral monism. Neither, though, does this mean that it is wholly inaccurate to characterise Nagel's early argument as one for panpsychism. Moreover, closer inspection will reveal that that argument does not after all feature a hole of the sort described. The fact that explains all these results is that Nagel appears to hold a rather subtle and sophisticated conception of the basic properties of matter, and of the connections these bear to mentality and physicality. This conception may ultimately warrant ascribing to him panpsychism *as well as* neutral monism, in some sense - the difference between the two may not be as clear-cut as is usually thought, in other words. His position and its evolution will require some careful probing in order to ascertain just what he now thinks, as well as whether he has always thought it.

4.

At times it seems clear that Nagel's position has developed into a thoroughgoing neutral monism, not least because he himself uses that term to describe it. Early in his recent monograph he says: 'the weight of the evidence favors some form of neutral monism over the traditional alternatives of materialism, idealism, and dualism'.<sup>25</sup> And a paper that precedes the monograph suggests the view that 'the fundamental constituents of the world, out of which everything is composed, are neither physical nor mental but something more basic',<sup>26</sup> going on explicitly to distinguish this suggestion from panpsychism. It might thus appear that Nagel's view simply changed between 1979 and the millennium: either he came to the firm conclusion that panpsychism was not after all needed to explain consciousness, or at least he came to think that the position Van Cleve called

---

<sup>24</sup> 1990: 217.

<sup>25</sup> 2012: 5.

<sup>26</sup> 2002: 231.

'neutralism' could not after all be described as panpsychism.

But, as Stubenberg notes,<sup>27</sup> this clean interpretation is muddied by certain other things Nagel says, not least by a tendency he sometimes evinces to characterise his current position as panpsychism. Later in the same monograph we get this decisive-sounding passage:

'Everything, living or not, is constituted from elements having a nature that is both physical and nonphysical—that is, capable of combining into mental wholes. So this reductive account can also be described as a form of panpsychism: all the elements of the physical world are also mental.'<sup>28</sup>

Thus the recent Nagel, even within the same text, characterises his theory in one place as neutral monism, in another place as panpsychism. These views have traditionally been taken to be mutually exclusive: they seem clearly to differ on the question of whether mentality is a fundamental property.<sup>29</sup> Is Nagel confused? Or does he perhaps vacillate, advocating neutral monism one moment and panpsychism the next? Or could it be that he defends a disjunction of the two? The following possibility seems far more likely, and demands further investigation: Nagel holds a univocal position which is simply resistant to classification according to the usual notions of panpsychism and neutral monism. The aim in the remainder is to get the metaphysics of this position as well understood as possible, using whatever clues we can find. At that stage we can broach the issue of finding the right name for the view, if such a name can be found.

5.

Setting matters of nomenclature aside for the moment, the difficulty of reconciling the passages cited above, and of ascribing a unitary view to Nagel on their basis, is that Nagel seems to say both that the world's fundamental constituents have mental and physical properties, but also that they are neither mental nor physical. It would appear on the face of it that these theses cannot consistently be held together: the basic constituents cannot be both mental and physical, as well as neither mental nor physical.

Stubenberg notes this difficulty, and attempts heroically to reconcile the apparently clashing Nagelian doctrines. He reports the following personal

---

<sup>27</sup> 2016: §5.4.

<sup>28</sup> 2012: 57.

<sup>29</sup> Connectedly, there is argument in the literature over whether theorists like James and Russell advocated panpsychism or neutral monism, or whether they held one view at one point and the other at another time – all indications that the theories are taken to exclude one another.

communication from Nagel about the puzzle:

the fundamental elements would be neither merely physical nor merely mental, but something that was necessarily both physical and mental, (or protomental); but since this necessary connection can't hold directly between the physical and the mental as we conceive them, it would require that the real character of these fundamental constituents be something more basic that accounts for their being both physical and (proto)mental.<sup>30</sup>

Stubenberg then proposes the following interpretation of Nagel's overall view:

The resulting picture is this. Described at the most fundamental level, the constituents of the world have properties that are neither mental nor physical. These neutral properties of every fundamental entity give rise to physical and mental (or protomental) properties. Thus each fundamental constituent is complex: it has mental (or protomental) properties, it has physical properties, and it has these two sets of properties as a necessary consequence of its having a third set of properties—the neutral properties.<sup>31</sup>

Stubenberg argues that this view best qualifies as a form of neutral monism, due to two factors: One is the foundational role of the neutral properties Nagel posits: these 'account for' the mental and the physical being of a fundamental element or constituent. The second factor is the impression one is given that the mental and physical features of the basic constituents involve nothing more than the instantiation of neutral properties. According to Nagel's personal communication 'the real character' of the fundamental constituents appears to be neutral, something more basic than mentality and physicality, and the constituents' mental and physical features are, in Stubenberg's phrase, 'no addition' to the being of their neutral properties. These factors combine to produce a picture of constituents whose essential nature is neutral, from which their mentality and physicality strictly, and necessarily, derive.

But I think this interpretation faces problems.<sup>32</sup> It appears that Stubenberg attributes to Nagel the view that a fundamental constituent has three sorts of nature – hence his statement that each constituent is 'complex'. Moreover, this tri-fold nature is according to Stubenberg a matter of how the constituent is in itself, not simply of how it behaves or how it interacts with other entities. Were it only a matter of such 'relational properties' then there might not seem to be any problematic clash of properties here. Consider Socrates' ring finger. Like most people's, it is short in comparison to his middle finger, but long in comparison to his little finger. It is thus long and short at the same time. But because these are properties it has essentially in relation to other things, and relative properties

---

<sup>30</sup> Stubenberg 2016, §5.4.

<sup>31</sup> *Ibid.*

<sup>32</sup> Stubenberg acknowledges this, both in his 2016 and in personal communication.

furthermore, there is no problematic complexity in the finger itself, as there would be if an object were said, for example, to be circular and square at the same time. The clashes we find worrisome are thus among the 'intrinsic' properties of things: those properties a thing possesses considered in itself and apart from its relations to other things. Put in these terms, we can say that on Stubenberg's interpretation of Nagel each fundamental constituent has a *tripartite intrinsic nature*: it is neutral, it is mental, and it is physical. Yet this ontological complexity seems to conflict directly with one of Nagel's statements of his view, that: 'This position is not equivalent to panpsychism. Panpsychism is, in effect, dualism all the way down. This is monism all the way down.'<sup>33</sup> A view on which each fundamental constituent is intrinsically mental, physical, *and* neutral can hardly be called a straightforward monism. Worse, Stubenberg's interpretation sees Nagel ascribing to each constituent what might appear to be an inconsistent set of intrinsic properties, in light of the opposition that both Nagel and Stubenberg acknowledge between the mental and the physical.

Another difficulty with Stubenberg's interpretation comes from his imputing to Nagel the view that the mental and physical parts of a constituent's nature are no addition in being to its neutral part; in other words, once the neutral nature is in place, the mental and physical natures are already accounted for. Usually such 'accounting for', or 'grounding' relations (as Stubenberg calls them),<sup>34</sup> obtain between different levels of being in the world, as where the properties of some set of smaller things is said to ground the properties of a big thing which they jointly compose. An example we have already encountered is the way the activity of non-liquid H<sub>2</sub>O molecules grounds - accounts for - the liquidity of the body of water they form. Another example would be the way panpsychists envisage the consciousness of a human to be grounded in the many micro-consciousnesses of her smallest material parts.<sup>35</sup> But Nagel's proposal, according to Stubenberg, is that one property (or aspect) of a single fundamental thing grounds other properties (or aspects) of that same thing's overall nature. So this is at the least a non-standard case of grounding. Grounding is, notably, a simultaneous relation: once the H<sub>2</sub>O molecules are interacting, there is no time lag required for the water's liquidity to obtain. That is because the molecular interaction, in a good sense, simply *is* the liquidity - that is all there is to it. Again this feature is difficult to understand when restricted to a single fundamental thing. As soon as the

---

<sup>33</sup> Nagel 2002: 231.

<sup>34</sup> The terminology of grounding is a recent and popular addition to analytic metaphysics. See Schaffer 2009 for an influential formulation.

<sup>35</sup> Though it standardly obtains between different levels, grounding need not run from the smaller to the larger: *cosmopsychists* hold that human consciousness is grounded in the single unified consciousness of the universe (Goff 2017, Shani 2015). More mundanely, we might think that the areas comprising the surface of an apple are red because the *apple* is red, not vice-versa.

constituent's neutral nature is in place, so are its mentality and physicality; that is what the idea of grounding entails. There is no easy model or analogy one can appeal to for a suitable understanding of this proposed relationship between the neutral, mental and physical natures of the fundamental constituent.

There is, to be sure, a familiar kind of case where grounding relations exist between the properties of a single thing: that is when that thing has a 'determinate' property that grounds its possessing the relevant 'determinable' property. To explain these terms, consider colours: Red is a determinable, which is to say that there are multiple ways of being red: a thing can be scarlet, or auburn, or vermillion, or crimson, or burgundy, etc. Each of these different specific ways of being red is known as a determinate of red. Red, in turn, is a determinate of the determinable colour: a more specific way of being coloured.<sup>36</sup> Now, when a thing has the determinate shade scarlet, this can be said to *ground* its being red, its also having that determinable property – for, clearly, nothing can be scarlet without being red. It is also true, of course, that nothing is red without being some specific shade of red. But, arguably, an item owes its being red to its being a specific shade of red, and not the other way around - since simply being red doesn't settle which specific shade of red something has. Thus it is the property of being scarlet that grounds the property of being red, and not vice-versa.

Does the determinate/determinable relation, as exemplified by colours and their shades, supply us a model with which to make sense of Stubenberg's construal of Nagel's position? The suggestion would be that the neutral property *n* of a fundamental constituent is a determinate, and grounds the constituent's mental and physical properties, which would be determinables related to *n* as red is related to scarlet. But this suggestion does not produce the right results. For one thing it fails to capture what is essential to Nagel's metaphysical proposal: the deep difference between the mental and the physical, which Nagel feels cannot be directly connected with one another given our conception of them. On the present proposal, the neutral property is as much a determinate of mentality as it is of physicality. But these two determinable properties are supposed to be very different to one another, irreconcilable in fact. It is therefore hard to see how a single specific property could be a determinate of both. It would be like saying that scarlet is a determinate not only of red but also of the *smell of food* – a determinable of another sort altogether. Determinates can fall under more than one determinable, admittedly, but not ones that clash in kind. And the problem that initiated Nagel's metaphysical project was the idea that mentality and

---

<sup>36</sup> Similarly, *shape* is a determinable of which *triangle* is a determinate. But triangle itself is also a determinable with determinates such as isosceles, equilateral, and so on. Thus being a determinable or a determinate is a relative affair.

physicality, taken on their faces, do clash. Just as scarlet cannot be a determinate red and a determinate food smell – for a colour cannot be a smell – *n* cannot be a determinate kind of mentality and a determinate kind of physicality, just in itself. A second problem with the model is that determinables arguably have no causal power beyond that of their determinates: whatever a red rag does to a bull is accomplished by its specific shade of red, to take a much discussed, if apocryphal, example.<sup>37</sup> This implies that mentality and physicality are without proper causal influence in the world – strictly, all the causation is done by the basic neutral properties. Thus mental and physical phenomena that we take to be paradigmatically causally efficacious – belief and desire pairs, pains, negative charge, mass – would be, qua mental and physical properties, no such thing. I doubt this is the result that Nagel was seeking. Lastly, there is no clear sense to be made of the suggestion that your present mental state, with all its specific qualitative character, is formed of a massive quantity of mental *determinables*. Your mental state is anything but determinable. Analogously, one cannot paint a room with only determinable colours.

Overall, it is not therefore a promising strategy for making sense of his position to ascribe to Nagel the view that a fundamental constituent has a complex, tripartite nature comprising of neutrality, mentality and physicality, when these are taken as fully occurrent intrinsic properties of the constituent.<sup>38</sup>

6.

Given that it is difficult to construe a fundamental constituent as having three intrinsic occurrent properties of such different kinds, at least in attempting to explain Nagel's avowedly *monist* position, there are two natural reactions to the failure of Stubenberg's interpretation, both of which deflate the status of the mental and physical properties. But neither is quite satisfactory, as we will see.

The first reaction is to ascribe to Nagel the view that 'the mental' and 'the physical' are really just *ways of thinking of* the neutral constituent, and not genuine parts of its nature. On this view being counted as mental or physical

---

<sup>37</sup> Gillett and Rives 2005.

<sup>38</sup> Stubenberg suggests that ascribing to Nagel a 'dual aspect' theory might help. But this does not in my view alter the substance of the difficulties: aspects are either mere ways of seeing a property (see below), or are properties themselves, neither option being an increase in the number of possibilities open to us. Even if aspects are something different from properties, we would face the conundrum of an allegedly neutral property whose nature comprises a mental and a physical aspect. In what sense would it be *neutral*, in that case, as opposed to being *mental and physical*? Would it perhaps have to have a third, neutral, aspect in addition to its mental and physical aspects? At that point we would seem merely to have swapped talking of *properties of the material constituent* for talking instead of *aspects of its nature*, but the substantive puzzles would remain intact for all that, even if they required rephrasing in terms of aspects.

would reflect mere classificatory practices, and would not answer to anything distinctive in the world itself. The terms 'mentality' and 'physicality' would be in that respect like statisticians' talk of 'the average person': there is no such person as the average person, though this phrase is used to pick out features of sets of real people. Perhaps, in this vein, we would want to call a fundamental constituent 'mental' when it was a relevant part of a conscious brain, and 'physical' when it was floating free in space and interacting with other constituents according to the laws of physics. But in both cases all we would really have is the neutral constituent and its behaviour. However, I think it is obvious that Nagel intends mentality and physicality to be more than *façons de parler*. Otherwise he could not possibly consider human conscious mentality to present such a deep and serious *metaphysical* as opposed to linguistic, puzzle. So this avenue of interpretation is not open.<sup>39</sup>

A connected view is something like that apparently held by Bertrand Russell during his neutral monist phase.<sup>40</sup> This is to construe the mentality and physicality of a fundamental constituent, perfectly neutral in itself, as relational or 'behavioural' properties, properties that only exist in its relations to other things - group effects, in effect. Being an uncle is an example of a relational property - it requires at least two things to exist, and involves both of them standing in a certain connection. Here the thought might be that what the neutral constituent's mentality consists in is its being part of a network of similar such constituents that jointly form a conscious brain. And what its physicality consists in is its interacting in physical ways - those ways physics describes - with other constituents. So it is physical in so far as it repels or attracts other fundamental constituents (perhaps our constituent is an electron, for instance), and it is mental in so far as it helps to constitute a conscious mind-brain. In this way, if the constituent does both jobs at once - as it must do if it is to form the relevant part of a conscious brain - it counts as mental and physical at the same time. Unlike on the previous interpretation, these are not mere ways of talking: the neutral constituent really is mental and physical, in these contexts, whether anyone thinks so or not. But all that those properties consist in is relations the constituent bears to other constituents. Its mentality and physicality amount to two very different sorts of relation it can stand in, akin to someone standing in the relations of being an uncle and also being the world's tallest pianist. Unlike Stubenberg's interpretation, this view does not require any complexity in the constituent itself. Russell compares the picture to a person's name figuring once in a geographical list, and again in an alphabetical list, in the telephone directory. Mentality and physicality, so construed, are simply two orders or contexts in

---

<sup>39</sup> Effectively this kind of view on the mind/body problem is already ruled out by the Realism premise of Nagel's original argument: this is the commitment to take conscious mentality wholly seriously as a real phenomenon.

<sup>40</sup> Russell 1927a, 1927b. See also Coleman 2015 for a similar proposal.

which a neutral constituent can figure.

But again this view is too deflationary for Nagel. It seems quite clear that he holds that a single fundamental constituent is itself, in some sense, mental and physical: these properties are part of its nature, regardless of what it does and whether or not it ever interacts with another constituent. This claim gains support from the fact that he is sometimes led to describe his position as panpsychism - there would be no need to do this were a constituent's mentality a matter only of its relations to other things. Russell's position, instead of taking mentality as fundamental, as panpsychism does, really reduces mentality to something else - relations among neutral constituents. Therefore we cannot model Nagel's position on Russell's neutral monism.<sup>41</sup>

7.

In this section I will attempt to offer a satisfactory substantive model for Nagel's position. But I should say that I am far from confident that Nagel would endorse it. It seems that he intends his theory to be extremely minimal, at this early stage of investigation into the material basis of consciousness.<sup>42</sup> It is strictly limited to assertion of the following bare logical possibility: that a fundamental constituent could possess a basic nature of a sort that ensured that it also had, of metaphysical necessity, a mental as well as a physical nature. Such a *basic* nature is called for as intermediary since the mental and the physical cannot be directly connected in the requisite necessary way. Nonetheless, the correlations we see between them at the macroscopic level - the way brain activity and conscious experiences systematically mirror one another - strongly suggest that they are necessarily tied together at a deep level.<sup>43</sup>

In a way I could stop there, since Nagel does. Stubenberg and I would be at fault for over-interpreting Nagel's view, offering analogies and examples in a bid to assign it more substantive content than it in fact has. It is really a logical skeleton of a view, and such a skeleton is presumably consistent with various - perhaps many - fleshed-out theories.

Yet it would not be at all satisfying to halt at this point. A thin theory of this kind does leave open multiple possible specific theories of consciousness. But presumably Nagel would agree that in fact only one theory can be adequate to the way things actually are - accurate, in other words. There is a right answer concerning how the mental and physical are related, he would surely concur. We

---

<sup>41</sup> Indeed we find an explicit rejection by Nagel of Russell's position on the grounds that 'both mental and physical properties are intrinsic' (2000: 447).

<sup>42</sup> After all, we have hardly been working on the problem for a couple of millennia.

<sup>43</sup> This description of his position is based upon private correspondence with Nagel.

ought to aim for that theory at all costs, and little is lost by aiming and missing. The more attempts we make to hit the target, surely, the better. Further, we have already seen that Nagel's theory is not so open-ended that we cannot discount some ways of fleshing it out - we found reasons to object to Stubenberg's way above, for example. This means that if we can propose a way of making the view more contentful - going beyond the bare logical possibility towards some kind of model, analogy, or example that would help us to understand it - a way that is not obviously incoherent and does not clash with other things we, or Nagel, believe, then we ought to consider that theory a serious contender for an explanation of consciousness. Finally, perhaps most importantly, it is not obvious on reflection that even Nagel's minimal theory helps us to make sense of the combination of theses we have been struggling to accommodate: that a basic material constituent is at once neither mental nor physical but also mental and physical. Nagel's skeletal thesis does not assist in this accommodation. What would help us to fit these assertions together is a positive concrete model, something that goes beyond Nagel's assertion of a bare logical possibility. For these reasons I will propose a possible way of spelling out Nagel's theory further.

What seems clear is that we want to understand how a constituent can be essentially neutral in its basic nature, yet also have mental and physical parts of its nature that derive directly from this core neutrality, in such a way as to make sense of calling it neither mental nor physical but also mental and physical *in itself*. Quite a challenge! It seems to me that the concept of a *disposition* may be of help here. Let us say that a certain vase is *fragile*; this means that it has the disposition to break if dropped from a great enough height onto a hard enough surface. This fragility is a property the vase has all the time, and, even though we explain or characterise that property in terms of its effects and interactions with other things (the drop, the height, the floor, the breaking), the vase is, itself, fragile whether or not these things ever happen. The vase on your mantelpiece is no less fragile for the fact that you may never drop it! In that sense its fragility is an intrinsic property, and occurrent - something the vase has at each moment. But it is also true that the vase's having that property is accounted for by - grounded in - its having another property, namely the particular molecular structure it possesses. It is the molecular structure that explains the vase's fragility - that is what *makes* it fragile. In a way the fragility can be seen as an aspect of the molecular structure, but it is also a perfectly good property in its own right. Moreover, fragility clearly plays a causal role in the breaking of the vase, if this occurs: the vase breaks because it is fragile. Even though the vase owes its fragility to the deeper property that grounds it, the molecular structure, this does not result in the fragility's causal power being screened off by that of

the molecular structure in the event of the vase breaking.<sup>44</sup>

The potential of this model should be apparent as regards our present topic. If we construe the fundamental constituent's mentality as a disposition it has, a disposition that is grounded in its neutral property, this has many pleasing results, which seem not only to accord with Nagel's various statements over the years, but also to deliver a plausible account of the relations among the constituent's properties. First, its mentality would be, like its neutrality, an intrinsic property of the constituent, albeit one grounded in - accounted for by - a deeper part of its nature. In the jargon, the neutral nature would provide the 'categorical ground' of the constituent's mental disposition (similarly, the vase's molecular structure categorically grounds its disposition to shatter if dropped from height). So this picture aptly captures the sense in which, for Nagel, the neutral property and the mental property are not on completely the same footing metaphysically, despite both being intrinsic to the fundamental constituent, which was the point where Stubenberg's interpretation met problems. The disposition/categorical property model does justice to Nagel's thesis that the constituent is, in its most fundamental nature, neutral, and its mental aspect derivative. Nor is there the same conflict that Stubenberg encountered between the constituent's being intrinsically neutral and its being intrinsically mental: if its mentality is a disposition it possesses, this does not imply any strong bifurcation in its basic nature. Still, dispositions are real, and they make things happen - they are perfectly good properties, by most philosophical lights. We can therefore say that the neutral constituent really is - also - mental, and that its mentality makes things happen. In the end, on this view, the vast interlocking pattern of interactions of the mental dispositions of trillions upon trillions of such constituents will be what is responsible for the generation of your consciousness, and for whatever effects your conscious mental states have upon the world. And though a disposition is strictly no addition in being to the categorical property that grounds it, still one can at least *imagine* the categorical property without the disposition: one can envisage an aspirin tablet of identical micro-constitution that lacks the disposition to dissolve, for example. In this sense, the neutral constituent's mentality represents some addition to its neutrality: it is a feature worthy of mention.<sup>45</sup>

Clearly we will want to say the same sort of thing about the constituent's physical nature, on the model - this is to be construed as another disposition, roughly to behave in the sorts of ways basic particles behave as recorded by

---

<sup>44</sup> If anything, the reverse is the case: in so far as the molecular structure causes the breaking, that is only because it grounds the vase's fragility, on a plausible view at least.

<sup>45</sup> Further to this point, it may well be that knowledge of the neutral property will not give one, *a priori*, knowledge of its disposition to mentality. Its mental tendency will thus be news.

physics. Again its physicality will be a real intrinsic property of the constituent, a property that makes things happen in its own right – such as subatomic repulsions and attractions. Since we identify dispositions by their effects, the result is secured that the mental disposition and the physical disposition are different properties of the constituent, since they cause rather different things to happen. One, eventually, causes consciousness (and its effects, e.g. conscious actions). The other, eventually, causes molecular bonding and the formation of rocks and organisms. But both are grounded in the neutral property. This feature, too - this way of framing the complex nature of the fundamental constituent - offers no obstacle, unlike the construal of the relationship on the determinable/determinate model. For a single categorical property can ground quite disparate dispositions. The vase's particular molecular structure is responsible not only for its disposition to shatter to bits, but is also what ensures that the vase is right now perfectly solid and of one piece. The electron's negative charge makes it *repel* one sort of thing (other electrons), and *attract* another (protons). There seems no reason, then, why the constituent's neutral nature could not ground its radically different mental and physical dispositions.

Can we now say what Nagel seems to want to say, namely that the constituent is neither mental nor physical yet both mental and physical? There seems to be a decent sense in which we can indeed say this. In its core being the constituent is neutral - that is its categorical nature. In its categorical nature, thus, the constituent is neither mental nor physical. Yet this nature grounds two radically disparate dispositions, one mental and one physical. In the sense of possessing the mental disposition, the constituent is, itself, mental - this is part of its overall intrinsic nature. It has a tendency to produce consciousness. And the same goes for its physical disposition – its tendency to produce physical interactions. So there is sense in the statement that the constituent is neither mental nor physical but also mental and physical. This is somewhat like saying of the vase that it is perfectly solid and stable but also fragile; or that an electron is repulsive as well as attractive. The position on offer lies somewhere between Stubenberg's proposal and Russell's view. Like Russell, we identify the dispositions, i.e. think about the constituent's mentality and physicality, in terms of what they cause in the world and how they relate to other things. But, as with Stubenberg's suggestion, these are not mere relational properties, but belong, intrinsically, to the constituent itself, regardless of what it ever does or causes, or what it does or does not relate with. Unlike Stubenberg, however, construing the constituent's mental and physical properties as dispositions allows us to capture their secondary status with respect to the neutral property, and their dependence upon it. That, at any rate, is my proposal for incarnating Nagel's skeleton-view as a living theory of consciousness.

Nagel's present position, if I am right about the most promising way of putting flesh on its logical bones, proves to be somewhat elusive with respect to the traditional categories of panpsychism and neutral monism. In a sense it is panpsychism, for the fundamental material constituents all have a disposition to mentality, to consciousness, as an intrinsic part of their nature. In a decent sense this means that they all possess mental properties, or perhaps proto-mental properties, to use the terminology Nagel sometimes favours. Whichever terminology we prefer, the material constituents have a part of their nature that is sufficiently closely tied to mentality as we know it for the view to merit the name 'panpsychism.' This could well be the reason behind the ruling Nagel made about the nomenclature of his position, which Van Cleve objected to. On the other hand, what grounds this mental disposition, what is ultimately there as regards the constituent's basic nature, is a property that is neither mental nor physical, but neutral. In that sense the view is neutral monism. We seem impelled to conclude that Nagel is propounding panpsychism at the same time as neutral monism! The labels and the difficulty of getting them both to stick to the position at once may not in the end be very important, of course. What seem more important are the observations that, first, if I am right, Nagel is offering a distinctive and interesting basic metaphysical picture aimed at explaining human consciousness, and, second, this seems to be more or less the same picture he has advocated ever since 'Panpsychism'. That we have achieved metaphysical as well as interpretational sense bodes well for the preceding discussion.<sup>46</sup>

### References

- Block, N. (1995) 'On a Confusion About a Function of Consciousness', *Behavioral and Brain Sciences* 18, 2: 227-247.
- Broad, C. D. (1925) *The Mind and its Place in Nature*, London: Routledge and Kegan Paul.
- Chalmers, D. J. (1996) *The Conscious Mind*, New York: Oxford University Press.
- (2003) 'Consciousness and its Place in Nature', in S. P. Stich and T. A. Warfield (eds.) *Blackwell Guide to the Philosophy of Mind*, pp. 102-42.
- Coleman, S. 'Neuro-Cosmology', in P. Coates and S. Coleman (eds.) *Phenomenal Qualities: Sense, Perception, and Consciousness*, Oxford: Oxford University Press, pp. 66-102.
- Gillett, C. and Rives, B. (2005), 'The Nonexistence of Determinables: Or, a World of Absolute Determinates as Default Hypothesis', *Noûs*, 39 (3): 483–504
- Goff, P. (2017) *Consciousness and Fundamental Reality*, Oxford: Oxford University Press.
- Hartshorne, C. (1934) *The Philosophy and Psychology of Sensation*, Chicago:

---

<sup>46</sup> I am grateful to Leopold Stubenberg and to Thomas Nagel for very helpful discussions in the preparation of this paper. Thanks also to an anonymous reviewer for some helpful suggestions.

- University of Chicago Press.
- Jackson, F. (1982) 'Epiphenomenal Qualia', *Philosophical Quarterly* 32: 127-136.
- Kirk, R. (1974) 'Zombies Vs. Materialists', *Proceedings of the Aristotelian Society* 48: 135-52.
- Leibniz, G. W. (1714) *The Monadology*
- Levine, J. (1983) 'Materialism and Qualia: The Explanatory Gap', *Pacific Philosophical Quarterly* 64: 354-61.
- Loar, B. (1990) 'Phenomenal States', *Philosophical Perspectives* 4: 81-108.
- Nagel, T. (1974) 'What is It Like to be a Bat?', *Philosophical Review* 83: 435-50.
- (1979) 'Panpsychism', in his *Mortal Questions*, Cambridge: Cambridge University Press.
- (1986) *The View from Nowhere*, New York: Oxford University Press.
- (2000) 'The Psychophysical Nexus', in P. Boghossian and C. Peacocke (eds.) *New Essays on the A Priori*, Oxford: Oxford University Press, pp. 433-471.
- (2002) *Concealment and Exposure*, Oxford: Oxford University Press.
- (2012) *Mind and Cosmos*, New York: Oxford University Press.
- Papineau, D. (2002) *Thinking about Consciousness*, Oxford: Oxford University Press.
- Russell, B. (1927a) *The Analysis of Matter*, London: George Allen and Unwin.
- (1927b) *An Outline of Philosophy*, London: George Allen and Unwin.
- Schaffer, J. (2009) 'On What Grounds What', in D. J. Chalmers, D. Manley & R. Wasserman (eds.) *Metametaphysics: New Essays on the Foundations of Ontology*, Oxford: Oxford University Press.
- Seager, W. (2006) 'The 'Intrinsic Nature' Argument for Panpsychism', *Journal of Consciousness Studies* 13, 10-11: 129-45.
- Shani, I. (2015) 'Cosmopsychism: A Holistic Approach to the Metaphysics of Experience', *Philosophical Papers* 44, 3: 389-437.
- Strawson, G. (2006) 'Realistic Monism: Why Physicalism Entails Panpsychism', *Journal of Consciousness Studies* 13, 10-11: 3-31.
- Stubenberg, L. (2016) 'Neutral Monism', in E. Zalta (ed.) *Stanford Encyclopaedia of Philosophy*. URL = <https://stanford.library.sydney.edu.au/entries/neutral-monism/>
- Van Cleve, J. (1990) "Mind-Dust or Magic?", *Philosophical Perspectives* 4: 215-226.
- Whitehead, A. N. (1925) *Science and The Modern World*, Cambridge: Cambridge University Press.