Keith E. Turausky
ket_austin@utexas.edu
University of Texas at Austin
Graduate Program

Contemporary Panpsychism: Its Varieties and Vagaries

Abstract

After spending the better part of a century in philosophical exile, panpsychism has recently seen something of a resurgence in the philosophy of mind. With an eye toward the so-called Hard Problem (i.e., the putative emergence of the mental from the physical), contemporary panpsychism holds that consciousness *simpliciter*—not self-consciousness, emotion, desire, belief, cognition, or even really awareness as known in the human case, but *bare subjective experience*—is a fundamental force of our universe, intrinsic to literally everything. The view remains highly controversial, accepted reluctantly even by some of its most prominent advocates. As such, panpsychism has not yet been robustly developed as a theory of mind. Indeed, it might not be apparent to all observers that there are at least two, and arguably three, distinct strains of panpsychist thought in philosophy today. Some views hold that consciousness is as inherent in matter as mass, while others are more restrictive, limiting subjective experience to either informational/computational or biological systems. Herein, I shall examine the similarities and differences among these views, as well as arguments for and against each. Special consideration will be given to what I deem the Chalmers/Strawson Divide: a cluster of important conceptual tensions between two of the most influential voices in the nascent panpsychist "movement." Largely unexplored in the literature, these tensions microcosmically reflect larger debates in the philosophy of mind, and may pose a significant challenge to the project of establishing a consistent and viable theory of panpsychism. Finally, I will offer some tentative suggestions for the development of such a theory.

1. Panpsychist Motivations: "When You Have Eliminated the Impossible…"

David Lewis famously noted "the incredulous stare" as perhaps the most common response to his theory of modal realism (according to which so-called *possible worlds*

really do exist).[1] Yet it seems fair to say such slack-jawed incredulity is an even more common response to *panpsychism*, the philosophical hypothesis that mind (roughly speaking) is everywhere. Indeed, while modal realism is taken rather seriously even by those who refuse to follow Lewis's logic to its literal conclusion, panpsychism by contrast has, as Pierfrancesco Basile notes, tended either to be "marginalized" or, "[when] taken notice of… treated with scorn and ridiculed."[2] Only over the past decade or so have a few brave thinkers—most notably Galen Strawson, David Skrbina, and (less enthusiastically, but no less influentially) David Chalmers—succeeded in reviving serious discussion of this long-neglected hypothesis.

Their relative success in this, one must suspect, is due in no small measure to their shared appeal to the "neutral monist" panpsychism advocated by Bertrand Russell, whose imprimatur in the analytic age resists glib dismissal.[3] Nevertheless, the reigning consensus among contemporary philosophers of mind was surely voiced when, in response to Strawson's provocative 2006 article "Realistic Monism: Why Physicalism Entails Panpsychism," Colin McGinn trenchantly deemed panpsychism "a complete myth, a comfortable piece of utter balderdash."[4] And while McGinn seems to find panpsychism roughly as harmless as a child's belief in Santa Claus, others (including Chalmers, who is openly *sympathetic* to the idea!) have somewhat more ominously noted the "threat of panpsychism."[5]

A much-needed counterpoint in this dialectic has been provided by Skrbina, who "often acts as a ruthless Minority Whip in the field of panpsychist studies,"[6] encouraging the sympathetic-but-timid to overcome their fear of ostracism and, as it were, come out of the

---

[1] David Lewis, *On the Plurality of Worlds* (Oxford: Blackwell, 2003), 133.
[2] Pierfrancesco Basile, "Back to Whitehead? Galen Strawson and the Rediscovery of Panpsychism," in *Mind that Abides: Panpsychism in the New Millenium*, ed. David Skrbina (Philadelphia: John Benjamins Publishing Company, 2009), 179–180.
[3] A very similar line of argument was advanced by early twentieth-century astrophysicist Arthur Eddington, as will later be discussed.
[4] Colin McGinn, "Hard Questions," in *Consciousness and its Place in Nature: Does Physicalism Entail Panpsychism?*, by Galen Strawson et al., ed. Anthony Freeman (Charlottesville, VA: Imprint Academic, 2006), 93.
[5] George Molnar, *Power: A Study in Metaphysics*, ed. Stephen Mumford (Oxford: Oxford University Press, 2003), 70; but see also David J. Chalmers, *The Conscious Mind: In Search of a Fundamental Theory* (New York: Oxford University Press, 1996), 154.
[6] Graham Harman, "Zero-Person and the Psyche," in Skrbina 2009, 278.

closet as panpsychists. Particularly admirable is the extensive work Skrbina has done to establish the vast (and vastly underappreciated) influence of panpsychism throughout the history of Western philosophy.[7] Yet even this recovered pedigree leaves many wholly unmoved; as Peter Simons has quipped, "that great philosophers such as Leibniz or Whitehead have been panpsychists is insufficient recommendation: everyone makes mistakes."[8] Clever enough, of course, but arguably misguided in more than one way: as Basile points out, although "no panpsychist bases his case upon an argument from authority… it does seem that a theory with such credentials deserves a fair hearing."[9] I would further note that the "incredulous stare"—that common and cocksure kneejerk conviction, implicit in McGinn and Simons's barbs, that no right-minded philosopher could *possibly* take panpsychism seriously—is itself merely an *argumentum ad populum*, a fallacy of the same genus as Skrbina's alleged argument from authority. *Tu quoque*, the panpsychist might thus respond (supposing, that is, that the panpsychist indeed had nothing more than authority to which to appeal!).

In any case, as Chalmers notes, the fact remains that panpsychism "is often regarded as outrageous, or even crazy."[10] Faced as they are with such unyielding skepticism, those friendly to panpsychism tend naturally to assume a rhetorical position redolent of Sherlock Holmes's famed dictum: "When you have eliminated the impossible, whatever remains, however improbable, must be the truth." That is to say that, at least so far as the literature goes, the motivation for panpsychism tends to be based less on positive arguments *per se* than on counterfactuals and, especially, negative arguments against the prevailing theory of mind—i.e., materialism in all its sundry guises. Because all materialist efforts fail so spectacularly to account for (or, in some cases, even acknowledge) the existence of consciousness, the argument goes, panpsychism is the

[7] See, e.g., David Skrbina, *Panpsychism in the West* (Cambridge, MA: The MIT Press, 2005)—and this is to say nothing of the far more *prima facie* prominent tradition of panpsychism in the East!
[8] Peter Simons, "The Seeds of Experience," in Freeman 2006, 146. As previously mentioned, Russell and Eddington are apparently to be counted among the "mistaken." One might add to the list—as Skrbina has—philosophical luminaries ranging from Xenophanes to Spinoza to William James (not to mention that curiously prescient martyr to science and philosophy, the "Academic of no Academy," Giordano Bruno; see Skrbina 2005 as well as Ingrid D. Rowland, *Giordano Bruno: Philosopher/Heretic* [Chicago: University of Chicago Press, 2009]).
[9] Basile in Skrbina 2009, 179.
[10] Chalmers 1996, 293.

only option left standing in the philosophy of mind. In this sense, the panpsychist hypothesis is indeed something of a *reductio ad absurdum*, but not in the way its critics would prefer; rather, panpsychism emerges as the only viable resolution to a *reductio* implicit in the materialist hypothesis. This sort of argument was articulated most clearly and forcefully in Galen Strawson's seminal "Realistic Monism." Target paper of the *Journal of Consciousness Studies* collection *Consciousness and its Place in Nature: Does Physicalism Entail Panpsychism?*[11] and heralded by the cheerleading Skrbina as a "soon-to-be classic,"[12] it will be the focus of the following section.

2. Strawson's "Realistic Monism"

Strawson begins his argument in "Realistic Monism" by asking—and then answering—the question "What does physicalism involve?":

> Well, one thing is absolutely clear. You're certainly not a realistic physicalist, you're not a real physicalist, if you deny the existence of the phenomenon whose existence is more certain than the existence of anything else: experience…. Full recognition of the reality of experience…. is the obligatory starting point for any theory that can legitimately claim to be 'naturalistic' because experience is itself the fundamental given natural fact; it is a very old point that there is nothing more certain than the existence of experience.[13]

Strawson therefore has no patience for those, like Daniel Dennett, who "are prepared to deny the existence of experience," and he dismisses them with Nietzschean aplomb: eliminativism is, for Strawson, "the strangest thing that has ever happened in the whole history of human thought, not just the whole history of philosophy," and he shamefacedly laments as "grievous" that "[i]t falls, unfortunately, to philosophy, not religion, to reveal the deepest woo-woo of the human mind."[14] Next to the eliminativist's denial of experience, Strawson opines, "every known religious belief is only a little less sensible

[11] Note that this collection (henceforth, Freeman 2006) is not to be confused with David Chalmers's influential 2002 article of the same name (also cited herein as it appears as a chapter in Chalmers's 2010 *The Character of Consciousness*); both pay titular homage to C. D. Broad's 1925 classic, *The Mind and its Place in Nature*.

[12] Skrbina 2009, xiv. Indeed, in *Mind That Abides*, Skrbina's collection of essays by contemporary panpsychist sympathizers, Skrbina positions "Realistic Monism" as the keynote article, preceded only by Skrbina's own condensed overview of "Panpsychism in History." In some respects, one might say, Skrbina has played T. H. Huxley to Strawson's Darwin (though the soft-spoken Strawson, at least in print, comes off as something of his own "bulldog," as the quotations to follow in this paper should amply illustrate).

[13] Galen Strawson, "Realistic Monism: Why Physicalism Entails Panpsychism," in Freeman 2006, 4.

[14] Strawson in Freeman 2006, 4–5.

than the belief that grass is green"—and in fact, he adds in a sardonic footnote, "religious believers are in infinitely better shape, epistemologically, than the Dennettians."[15]

If nothing else, all this anti-eliminativist sentiment should make abundantly clear that, to borrow Chalmers's phrasing, Strawson *takes consciousness seriously*. And like Chalmers, Strawson also takes science—a.k.a. physicalism and/or naturalism—seriously. But needless to say, Strawson is at pains to distinguish Dennettian "physicSalism"[16] (the aforementioned "deepest woo-woo of the human mind") from his preferred thesis of "real physicalism," namely:

> [RP] experience is a real concrete phenomenon and every real concrete phenomenon is physical.[17]

On Strawson's view, no sober-minded physicalist can deny RP, for "they cannot deny that when you put physical stuff together in the way in which it is put together in brains like ours, it constitutes—is—experience like ours; all by itself. All by itself: there is on their own physicalist view nothing else, nothing non-physical, involved."[18] Indeed, Strawson suspects (and Chalmers's own "informal surveys"[19] have suggested) that "at least some of those who call themselves physicalists are realistic physicalists—real realists about experiential phenomena."[20] And Strawson himself endorses RP, of course—yet it is important to note his third-person remove in the recent quotations: "*they* cannot deny," "some who call *themselves*." Strawson sets himself apart from other physicalists, and even other "realistic physicalists," by identifying a tension between RP and a second thesis to which "many—perhaps most—of those who call themselves physicalists or materialists… are committed."[21] This thesis Strawson calls "NE" (short for "non-experiential[ism]"), and he states it as follows:

---

[15] Strawson in Freeman 2006, 6f7.
[16] Strawson in Freeman 2006, 4. The "physicSal" thus comprises the putatively non-experiential aspects of reality, i.e. what is traditionally called the *material*.
[17] Strawson in Freeman 2006, 12.
[18] Ibid.
[19] "Informal surveys suggest that the numbers run two or three to one in favor" of the view that "there is something further that needs explaining" about consciousness once all the physical facts are in (Chalmers 1996, xiii).
[20] Strawson in Freeman 2006, 12.
[21] Strawson in Freeman 2006, 11.

> [NE] physical stuff is, in itself, in its fundamental nature, something wholly and utterly non-experiential.[22]

The thesis seems innocuous enough, and Strawson gamely concedes that not only most philosophers but "the vast majority of human beings"[23] endorse it. But there is a problem afoot—the Hard Problem itself, in fact. For how *can* one simultaneously espouse RP and NE? That is, if "the fundamental given natural fact" is *experience itself*, how can it also be true that the "fundamental nature" of "every real concrete phenomenon" is "wholly and utterly non-experiential"? Strawson here notably inverts the traditional presentation of the mind–body problem: instead of asking how consciousness could possibly arise from wholly unconscious matter, he asks how matter could possibly be wholly unconscious *given that* consciousness exists. The Dennetts of the world will surely cry foul at this inversion (one can already hear the objection: "Why take consciousness as given?"), but as we have seen, Strawson does not suffer eliminativists gladly, and so it is to the rest of us "wannabe materialists" (to use Terence Horgan's term[24]) that he directs his question: how can NE be true *given that* RP is true?

The go-to answer involves some sort of "emergence" of experience from the non-experiential, but Strawson is well aware that no satisfactory account of such radical emergence has yet found its way into the literature. And he thinks he knows why: the very idea is "incoherent," despite having "acquired an air of plausibility (or at least possibility)… simply because it has been appealed to many times in the face of a seeming mystery."[25] The appeal to emergence most commonly takes the form of an argument from analogy with liquidity—a phenomenon that arises from collections of, e.g., water molecules without being characteristic either of water molecules themselves or of the

---

[22] Ibid.
[23] Strawson in Freeman 2006, 12.
[24] See Terence (Terry) Horgan, "Materialism, Minimal Emergentism, and the Hard Problem of Consciousness," in *The Waning of Materialism*, eds. George Bealer and Robert Koons (Oxford: Oxford University Press, 2010) 309–329. Note, however, that despite deeming himself a "wannabe materialist," Horgan is strongly critical of the most prominent contemporary brand of materialism, specifically what he and John Tienson call "new wave materialism" (the position Chalmers classifies as type-B materialism; see his "Consciousness and its Place in Nature" and Horgan and Tienson's "Deconstructing New Wave Materialism"). Horgan's views on type-F monism—and pan(proto)psychism in particular—remain to be seen, but Strawson's argument that physicalism entails panpsychism certainly seems intended to offer safe harbor to Horgan-style "wannabe materialists."
[25] Strawson in Freeman 2006, 12.

atoms and smaller particles of which they are composed. Though Strawson sympathizes with the analogy ("how, we may think, could *this* arise from individual non-liquid molecules?"), he ultimately finds it misleading. His critique invokes the principle of "total dependence" upon which he thinks all cases of true emergence necessarily hinge:

> It seems plain that there must be a fundamental sense in which any emergent phenomenon, say Y, is wholly dependent on that which it emerges from, say X. It seems, in fact, that this must be true by definition of 'emergent'; for if there is not this total dependence then it will not be true after all, not true without qualification, to say that Y is emergent from X. For in this case at least some part or aspect of Y will have to hail from somewhere else and will therefore not be emergent from X. Plainly this is not how it is with liquidity.[26]

That is, no "part or aspect" of liquidity hails from elsewhere than (or is in any other way *independent* of) water molecules and their constituent atomic and subatomic particles: large-scale liquidity is unambiguously and exhaustively *dependent* on small-scale physics. As such, both liquidity and its non-liquid antecedents are unambiguously and exhaustively *physical*: the relevant observed and explained facts about these states reduce solely to the "shape-size-mass-charge-etc. phenomena" that Strawson abbreviates as "'P' phenomena."[27] This homogeneity makes it fairly easily for us to overcome any initial skepticism and see that "liquid phenomena (which are wholly P phenomena) are emergent properties of wholly non-liquid phenomena (which are wholly P phenomena)"—in both cases, "we move wholly within a completely conceptually homogeneous (non-hetereogenous) set of notions."[28] In a crucial sense, we understand that nothing truly *new* has emerged here: the P phenomena were there the whole time, and what's there now is nothing other than P phenomena. The emergence of liquidity isn't really mysterious, then, and it certainly isn't "brute." And thus, for Strawson, the analogy from liquidity to consciousness is a nonstarter:

> You can get liquidity from non-liquid molecules as easily as you can get a cricket team from eleven things that are not cricket teams. In God's physics, it would have to be just as plain how you get experiential phenomena from wholly non-experiential phenomena. But this is what boggles the human mind.[29]

[26] Strawson in Freeman 2006, 14.
[27] Strawson in Freeman 2006, 13.
[28] Strawson in Freeman 2006, 15.
[29] Ibid.

Non-eliminativist "wannabe materialists" who respect this boggling epistemic gap—roughly, those categorized by Chalmers as type-B materialists[30]—must concede that when it comes to the mind–body problem, we most certainly do *not* "move wholly within a completely conceptually homogeneous (non-hetereogenous) set of notions." Indeed, type-B views tend not merely to concede but to *turn upon* the conceptual heterogeneity of the physical and the experiential—see, e.g., Davidson, Loar, and McLaughlin. The insurmountability of the epistemic gap is *itself* presented as Occamic evidence against the existence of an ontological gap: it is because our phenomenal and (neuro-)physical concepts are so disparate, the type-B materialist argues, that we find an insurmountable epistemic gap and wrongly infer an ontological gap. But then the analogy from liquidity, where we find *no* insurmountable epistemic gap, simply won't do, even on a type-B materialist account. As Strawson puts it, "we need an analogy on a wholly different scale,"[31] and he helpfully obliges by considering both the putative emergence of the extended from the unextended and of the spatial from the non-spatial:

> Suppose someone proposes that there are real, concrete, intrinsically, irreducibly and wholly non-spatial phenomena ('wholly non-S phenomena'), and that when they stand in certain wholly non-spatial relations they give rise to or constitute real, concrete, intrinsically and irreducibly spatial phenomena ('S phenomena'), these being emergent features of wholly non-S phenomena. Those who claim to find no difficulty in the idea that genuinely unextended concrete entities can give rise to or constitute genuinely extended concrete entities may like to consider this case separately, because they presumably take it that their putative mathematical-point entities are at least spatial entities, at least in the sense of being spatially located. My hope is that even if they think they can make sense of the emergence of the extended from the unextended, they won't think this about the more radical case of the emergence of the spatial from the non-spatial.[32]

By Strawson's reckoning, the claim that the experiential emerges from the non-experiential is as strong—yet as unjustified and ultimately incoherent—as the claim that the spatial emerges from the non-spatial. This is radical or "brute" emergence, which is "by definition a miracle every time it occurs, for it is true by hypothesis that in brute

[30] Eliminativism being the type-A materialist view; see David J. Chalmers, *The Character of Consciousness* (New York: Oxford University Press, 2010) [Chalmers's influential 2002 article "Consciousness and its Place in Nature" appears as chapter 5, 103–140, in that book].
[31] Strawson in Freeman 2006, 15.
[32] Strawson in Freeman 2006, 17.

emergence there is absolutely nothing about X, the emerged-from, in virtue of which Y, the emerger, emerges from it."[33] In fact, as Strawson sees it, the notion of "brute emergence" is itself a contradiction in terms:

> X has to have something—indeed everything—to do with it [i.e., Y]. That's what emerging is (that's how liquidity arises out of non-liquid phenomena)…. It cannot be brute. Otherwise it will be intelligible to suppose… even that concrete phenomena can emerge from wholly abstract phenomena. Brutality rules out nothing. If emergence can be brute, then it is fully intelligible to suppose that non-physical soul-stuff can arise out of physical stuff—in which case we can't rule out the possibility of Cartesian egos *even if we are physicalists*.[34]

In light of the latter possibility, brute emergence (whether or not it is, in the final analysis, self-contradictory) clearly represents a Pandora's Box for the "wannabe materialist." Worse, as Strawson sees it, Pandora's Box has here become the very cornerstone of a philosophy ripe for deconstruction.[35] Though not quite as scathing toward emergentism as toward eliminativism, Strawson again pulls no punches:

> How did the notion of brute emergence ever gain currency? By one of the most lethal processes of theory formation, or term formation, that there is. The notion of brute emergence marks a position that seemingly has to exist if one accepts both RP (or, more simply, the reality of experience) and NE [the fundamental non-experientiality of matter]. And since many are irredeemably committed to both RP and NE, the notion of brute emergence comes to feel substantial to them by a kind of reflected, holographical energy. *It has to be there, given these unquestioned premisses, so it is felt to be real.*[36]

RP remains unquestioned (unquestionable, really), for Strawson; after all, he takes the reality of experience as "the fundamental given natural fact." But if a commitment to

---

[33] Strawson in Freeman 2006, 18. He goes on to note an interesting dilemma: "given the standard assumption that the emergence of Y from X entails the 'supervenience' of Y on X… it turns out to be a strictly lawlike miracle. But a miracle is by definition a violation of a law of nature!"

[34] Strawson in Freeman 2006, 19; italics in original.

[35] Though Strawson delivers, as McGinn puts it, a laudably "eloquent and uncompromising rejection of the idea of brute emergence," (McGinn in Freeman 2006, 90), Strawson is not as alone as he suggests in noticing the problem. Chalmers, in his anti-reductionist efforts, has come up against the issue (see, e.g., Chalmers 1996, 129–130), while the *Stanford Encyclopedia of Philosophy*'s entry on panpsychism notes that "not only is there a problem simply in accounting for the emergence of something so distinctive as consciousness from mere matter, it is surprisingly difficult to articulate a form of emergentism that does not threaten to make the emergent features causally impotent or epiphenomenal. This is not the place to discuss the difficulties of all the varieties of emergentism, but they seem serious…." [William Seager and Sean Allen-Hermanson, "Panpsychism," *The Stanford Encyclopedia of Philosophy (Fall 2010 Edition)*, ed. Edward N. Zalta, plato.stanford.edu/archives/fall2010/entries/panpsychism].

[36] Strawson in Freeman 2006, 18–19; italics mine.

both RP and NE requires a commitment to brute emergence, and if brute emergence is incoherent (if not, strictly speaking, metaphysically impossible), *why not question NE?* If we don't, Strawson warns, we necessarily take on "(a) a commitment to something—wholly and essentially non-experiential stuff—for which there is *absolutely no evidence whatever*, along with (b) the wholly unnecessary (and incoherent) burden of brute emergence, otherwise known as magic."[37] Therefore, he concludes, we real physicalists (a.k.a. non-eliminativists) should reject NE: we should deny that "physical stuff is, in itself, in its fundamental nature, something wholly and utterly non-experiential." The stage is thus set for panpsychism—not as a presupposition, it should be noted, but as a logical conclusion. For to reject NE and its concomitant doctrine of brute emergence is to acknowledge that:

> [I]f experience like ours… emerges from something that is not experience like ours… then that something must already be experiential in some sense or other. It must already be somehow experiential in its essential and fundamental nature, however primitively or strangely or (to us) incomprehensibly…. Given that everything concrete is physical, and that everything physical is constituted out of physical ultimates, and that experience is part of concrete reality, it seems the only reasonable position, more than just an 'inference to the best explanation'. Which is not to say that it is easy to accept in the current intellectual climate.[38]

Intellectual climate be damned, Strawson feels he has discharged his Holmesian duty in eliminating the impossible; whatever remains, however improbable, must be true. At this point, Strawson briefly entertains *micropsychism*—the hypothesis that "some but not all physical ultimates are experiential"[39]—but ultimately finds it dissatisfying, akin to "the idea that some but not all physical ultimates are spatio-temporal." Indeed, he notes, with its "radical heterogeneity at the very bottom of things," micropsychism appears to be "a form of dualism."[40] Therefore, with eliminativism, "wannabe"/type-B materialism, and

---

[37] Strawson in Freeman 2006, 20; italics in original. Our purported lack of evidence for non-experiential matter derives from Strawson's Russell-Eddington line of argumentation, but as Chalmers has also appealed extensively to Russell on this point, I will reserve further consideration of it for the subsequent discussion of Chalmers's pan(proto)psychist arguments.

[38] Strawson in Freeman 2006, 24.

[39] Strawson in Freeman 2006, 25. NOTE STRAWSON'S REVISED (?) POSITION.

[40] Ibid. And Chalmers would seem to agree, given his speculation in "Consciousness and its Place in Nature" on viable ways to cash out "type-D" substance dualism and on the close relation between type-D and type-F views.

micropsychism off the table (and with "Cartesian"[41] substance dualism and Berkeleyan idealism never even invited to the table!), Strawson sees no reasonable alternative: *all* physical ultimates must be experiential "in some sense or other," so panpsychism *must* be true (in some sense or other). "This sounded crazy to me for a long time," he readily admits, "but I am quite used to it now that I know that there is no alternative short of 'substance dualism'…. Real physicalism, realistic physicalism, entails panpsychism, and whatever problems are raised by this fact are problems a real physicalist must face."[42]

A later section will consider the very real problems panpsychism both raises and faces (beside, of course, the incredulous stare), as well as some potential responses. For now, however, let us focus on the relevant contributions of David Chalmers, whose arguments leading to a panpsychist conclusion differ from Strawson's in some important ways.

3. Chalmers's Type-F Pan(proto)psychism (w/Detour into Biological Panpsychism)

Chalmers, unlike Strawson, is no Nietzschean fire-breather. "I know that the debate between the deflationary and inflationary views of consciousness [and] the argument between the likes of Dennett and the likes of me is sometimes cast as an ideological battle," he writes, but "I learned long ago that I am not much of an ideological crusader… and I am only occasionally inclined to rhetorical extremes."[43] Perhaps more to the point, Chalmers wisely hedges his philosophical bets—to the occasional frustration of his opponents, advocates, and hero-seeking acolytes alike. As he coyly notes of the *non-reductive* ("inflationary") type-D substance dualist, type-E epiphenomenalist, and type-F neutral/dual-aspect monist views he defends against the *reductive* ("deflationary") opposition views A through C, "My own loyalties are fairly evenly spread among these

[41] Strawson prefers scare-quotes here, as he believes Descartes has been quite unfairly represented by "anti-Cartesian" monists such as Dennett, who are in fact more in thrall to "Cartesian" thinking than was Descartes himself. Strawson mentions this in "Realistic Monism," but says much more on the topic in "Panpsychism? Reply to Commentators with a Celebration of Descartes," the concluding entry in Freeman 2006. Interestingly, at least some of Strawson's arguments seem to have been anticipated by Filip Radovic of Göteborg University, whose poster presentation at *Toward a Science of Consciousness 1998*, "Towards a Proper Monism," (1) claimed that "materialists are in a way implicitly accepting the original distinction proposed by Descartes," (2) distinguished between "false" and "proper" monism, and (3) invoked Spinoza's metaphysics as a better conceptual model than Russell's "neutral monism," with which, as will be seen, Strawson also takes issue; see www.phil.gu.se/posters/prop.html.

[42] Strawson in Freeman 2006, 25–26.

[43] Chalmers 2010, xv.

three views, depending on the day of the week."[44] Nevertheless, he grants, "I think that in some ways, the type-F view is the most appealing," although "this sense is largely grounded in aesthetic considerations whose force is unclear."[45] These aesthetic considerations, it seems, are roughly as follows:

> (1) Type-D substance dualism does not respect the causal closure of the physical.
> (2) Type-E epiphenomenalism denies our profound natural intuition that consciousness *does something*.
> (3) Type-F neutral/dual-aspect monism (a) respects the causal closure of the physical, (b) preserves the possibility of some causal role for consciousness,[46] and furthermore (c) holds "the promise of integrating phenomenal and physical properties very tightly in the natural world… [delivering] a deeply integrated and elegant view of nature."[47]

Thus, as Chalmers puts it, while Type-F monism "arguably fits the letter of materialism, it shares the spirit of anti-materialism."[48] All this sounds admirably elegant indeed—but then, what exactly *is* Type-F monism? "One could give the view in its most general form the name *panprotopsychism*," Chalmers writes, "with either protophenomenal or phenomenal properties underlying all of physical reality."[49] The former, protophenomenal case, he says, "can be seen as a sort of neutral monism," while the latter case, with fully/genuinely (i.e., non-"proto") phenomenal properties "ubiquitous at the fundamental level," could be "characterized… as a sort of panpsychism."[50]

Characteristically, Chalmers maintains a studied agnosticism between protophenomenal neutral monism (i.e., panprotopsychism in the non-general sense) and full-fledged panpsychism—but even pan*proto*psychism is a bridge too far for many philosophers.[51] Perhaps to soften the blow with a mild appeal to authority, Chalmers introduces the type-

---

[44] Chalmers 2010, xvii.
[45] Chalmers 2010, 138.
[46] Interestingly, the question of epiphenomenalism aside, Chalmers and (especially) Strawson lean toward incompatibilist determinism when it comes to the question of free will. That consciousness could play a causal role without being "free" does not seem *prima facie* self-contradictory, but no doubt deserves more extensive exploration than it has yet been given, or will be given here.
[47] Chalmers 2010, 133.
[48] Chalmers 2010, 134.
[49] Ibid.
[50] Ibid.
[51] By contrast, as will be seen in the following section, Strawson considers any Chalmers-style "proto" hypothesis to be a question-begging hedge—a crucial difference in views between the two philosophers.

F position as "Russellian monism"[52]; after all, one popular route to type-F pan(proto)psychism comes via some comments made by Bertrand Russell and Arthur Eddington, two quite estimable thinkers (though again, the argument for panpsychism does not hinge on Russell and Eddington's authority so much as on the astuteness of their observations). As Chalmers summarizes:

> Russell pointed out that physics characterizes physical entities and properties by their relations to one another and to us…. At the same time, physics says nothing about the intrinsic nature of these entities and properties. Where we have relations and dispositions, we expect some underlying intrinsic properties that ground the dispositions…. [H]owever, physics is silent about the intrinsic nature…. of fundamental physical systems[.][53]

Coincidentally, physics also appears silent about phenomenal consciousness (type-A materialist views again notwithstanding). What's more, phenomenal consciousness certainly seems to be at least *part* of the "intrinsic nature" of thinking beings such as ourselves. The implication of this dual mystery is clear enough: perhaps phenomenal consciousness *is* (at least part of) the intrinsic nature of "fundamental physical systems." In other words, perhaps *mind is everywhere* (the literal meaning of "panpsychism"). To quote Russell directly: "The physical world is only known as regards certain abstract features of its space-time structure—features which, because of their abstractness, do not suffice to show whether the physical world is, or is not, different in intrinsic character from the world of mind."[54] Or, as Eddington queried, "what knowledge have we of the nature of atoms that renders it at all incongruous that they should constitute a thinking object? ...It seems rather silly to prefer to attach it to something of a so-called 'concrete' nature inconsistent with thought, and then to wonder where the thought comes from."[55] (Nowadays, of course, we'd talk about quarks or strings instead of atoms, but whatever "ultimates" we may consider, the point remains the same.)

This sort of argument turns on the very definition of science, or in any case a definition held dear by many: that science (at least "hard" science, which of course includes the

[52] Chalmers 2010, 133.
[53] Ibid.
[54] Bertrand Russell, *Human Knowledge: Its Scope and Limits* (London: Routledge 1948/1992), 240; quoted by Strawson in Freeman 2006, 20.
[55] Arthur Eddington, *The Nature of the Physical World* (New York: Macmillan, 1928), 258–260; quoted by Strawson in Freeman 2006, 10.

study of fundamental physics) must deal exclusively in third-person (i.e., "intersubjectively verifiable"[56]) evidence. Especially for those, like Chalmers, who accept the conceivability—and, by extension, metaphysical possibility—of "zombies" and "zombie worlds" (identical to humans and our world from a third-person point of view but utterly devoid of first-person experience), the Russell-Eddington line of reasoning is naturally compelling. For in effect, Russell and Eddington observe that our actual world is, from the point of view of science, *indistinguishable from a zombie world*. As even the resolutely anti-panpsychist McGinn concedes, "physics does indeed leave the intrinsic nature of matter unspecified.... it tells us what the ultimates *do*, particularly their dynamic properties, but it doesn't tell us what they *are*. Physics is a kind of functionalist theory of material reality."[57] But one could drop any talk of zombies and/or functionalism and simply note that solipsism—the hypothesis that *only I* am conscious—is, however distastefully improbable, generally accepted to be *logically possible*, insofar as one cannot, via third-person scientific means, conclusively establish that anyone (even oneself) is conscious. Like it or not, then, first-person experience (i.e., consciousness *itself*) is, for each of us and all of us, the only source of anything approaching justified true belief when it comes to the presence of consciousness in the universe.

So, in fact, science *necessarily* lacks a "consciousness detector" (a piquant phrase Chalmers introduced whilst waving around a hairdryer meant to stand in for the proposed device). Rather, our scientific knowledge about the "outside world"—indeed, *any* knowledge we have about the "outside world"—is limited to surfaces, structures, and functions. Science, by method and design, tells us nothing about what David Armstrong called the "stuffing for matter."[58] Nor, simultaneously (and again, by method and design), does science tell us anything about phenomenal consciousness: as Chalmers notes, "if it were not for our direct evidence in the first-person case, the hypothesis [that conscious experience exists] would seem unwarranted, almost mystical, perhaps."[59] Taking these points together, Eddington concludes the following (using "pointer readings" as a stand-in for "scientific observations and measurements"):

---

[56] Space forbids an analysis of the significance of "subjective" in the phrase "intersubjective verifiability."
[57] McGinn in Freeman 2006, 93.
[58] David M. Armstrong, *A Materialist Theory of the Mind* (New York: Routledge, 1968/1993), 282.
[59] Chalmers 1996, 5.

> In one case—namely, for the pointer readings of my own brain—I have an insight which is not limited to the evidence of the pointer readings. That insight shows that they are attached to a background of consciousness… I may expect that the background of other pointer readings in physics is *of a nature continuous with that revealed to me in this way*… [though] I do not suppose that it always has the more specialized attributes of consciousness."[60]

The Russell-Eddington route to panpsychism suggests we could solve two mysteries for the price of one, or at least consolidate them into a single mystery: perhaps phenomenal consciousness, for all its material inexplicability, *is* the "stuffing for matter." Chalmers, like Strawson, notes that "the idea sounds wild at first, but on reflection it becomes less so. After all, we really have *no idea* about the intrinsic properties of the physical. Their nature is up for grabs, and phenomenal properties seem as likely a candidate as any other."[61] It is important here to note that Strawson too draws heavily on the Russell-Eddington line; in fact, in combination with his previously detailed objections to NE (the thesis of the fundamental non-experientiality of matter), it constitutes more or less the whole of his case in "Realistic Monism." We shall see in the subsequent section that Strawson deploys the Russell-Eddington line to somewhat different ends than Chalmers, but for now it will be worth exploring another crucial path by which Chalmers has, however tentatively, reached his own pan(proto)psychist conclusions.

It might be called the Darwinian argument for panpsychism: given that humans have phenomenal experience, it seems a fair bet that most "higher mammals" (e.g., chimps, dolphins, elephants, dogs, etc.) also do. The alternative—that *Homo sapiens sapiens* is literally the *only* (and, presumably, first) species on the planet to attain phenomenology—seems as improbable as solipsism; indeed, the proposal seems grounded in a sort of species-level solipsism. Common as such extreme anthropocentrism has historically been (and, regrettably, may still be), this grand solipsism of mankind simply must be found untenable in our post-Darwinian age. And fortunately, most reasonable people today would readily concede that at least *some* animals have genuine subjective experience(s)—but the closer one looks, as Chalmers notes, the harder it becomes to draw clear lines. If "higher mammals" are conscious (as they certainly seem to be), why

[60] Eddington 1928, 258–260; quoted by Strawson in Freeman 2006, 10–11.
[61] Chalmers 1996, 154.

not extend the benefit of the doubt to all mammals, or also to birds, reptiles, amphibians, and fish? In an oddly poignant passage, Chalmers asks us first to consider the mouse:

> Mice may not have much of a sense of self, and may not be given to introspection, but it seems entirely plausible that there is *something* it is like to be a mouse…. The natural hypothesis is that corresponding to the mouse's "perceptual manifold," which we know they have, there is a "phenomenal manifold." …There does not seem to be much reason to suppose that phenomenology should wink out while a reasonably complex perceptual psychology persists. If it does, then either there is a radical discontinuity from complex experiences to none at all, or somewhere along the line phenomenology begins to fall out of synchrony with perception, so that for a while, there is a relatively rich perceptual manifold accompanied by a much more impoverished phenomenal manifold. The first hypothesis seems unlikely, and the second suggests that the intermediate systems would have inner lives strangely dissociated from their cognitive capacities. The alternative is surely at least as plausible. Presumably it is much less interesting to be a fish than to be a human, with a simpler phenomenology corresponding to its simpler psychology, but it seems reasonable enough that there is *something* there.[62]

Chalmers is, in essence, running a *reductio* against the all-too-common implicit assumption that certain forms of life just aren't "complex enough" to "need" consciousness, as well as against the cluster of implicit assumptions *behind* that implicit assumption. These latter, namely, are (1) that consciousness emerged relatively late in the evolutionary game, (2) that it did so *in response to* an increased physical complexity, (3) that at a certain point, the proper functioning of this increased physical complexity somehow suddenly *required* phenomenal conscious experience, (4) that previous, lower levels of physical complexity possess(ed) *no* such "phenomenal requirement," and (5) that because they possess(ed) no phenomenal requirement, these less-evolved organisms possess(ed) *no phenomenology whatsoever*. Brought into plain sight, each assumption is clearly tendentious on its own—and as Chalmers ably demonstrates, together they imply situations wherein bizarrely asymmetric relationships hold between perception and

[62] Chalmers 1996, 294–295. Chalmers is by no means the first to articulate this approach to mindedness; in many ways, in fact, it underscores much modern ethical philosophy regarding non-human (e.g., animal) beings. In discussions of environmental ethics, the obligatory quotation comes from the venerable Jeremy Bentham: "The question is not, 'Can they reason?' nor, 'Can they talk?' but, 'Can they suffer?'" [Bentham, Jeremy. *An Introduction to the Principles of Morals and Legislation*. Oxford: Clarendon Press, 1907. Retrieved November 10, 2011 from http://www.econlib.org/library/Bentham/bnthPML0.html.]

phenomenology.[63]

Now perhaps, having granted the consciousness of mice and fish, one might still wish to reserve consciousness for vertebrate animals—but even that timeworn distinction seems rather arbitrary in light, e.g., of studies suggesting that the invertebrate octopus is among the handful of animals on earth capable of tool use.[64] From all external appearances, the octopus seems to be "more conscious" than any vertebrate fish. Yet if either an octopus or a fish could be conscious, why not, perhaps, a spider?[65] Why should we assume there is literally *nothing it is like* for the spider to spin its web or stalk its prey? Just as long as we are careful to distinguish basic phenomenal consciousness from more advanced features like self-consciousness, Chalmers sees no reason not to continue on down the branches of the great tree of life, perhaps even to its pre-biotic roots:

> As we move along the scale from fish and slugs through simple neural networks all the way to thermostats, where should consciousness wink out? The phenomenology of fish and slugs will likely not be primitive but relatively complex, reflecting the various distinctions they can make. Before phenomenology winks out altogether, we presumably will get to some sort of maximally simple phenomenology. It seems to me that the most natural place for this to occur is in a system with a corresponding simple "perceptual psychology," such as a thermostat. The thermostat seems to realize the sort of information processing in a fish or a slug stripped down to its simplest form, so perhaps it might also have the corresponding sort of phenomenology in its most stripped-down form. It makes one or two relevant distinctions on which action depends; to me, at least, it does not seem unreasonable that there might be associated distinctions in experience.[66]

---

[63] For that matter, one might further respond to the "insufficiently complex to *need* consciousness" skeptic by asking whether it isn't at least as strange that something as complex as a human brain could *maintain* consciousness! That is, if it is plausible that a simple system be *too simple to need* consciousness, why isn't it just as plausible that a complex system be *too complex to maintain* consciousness? The desideratum here, clearly, is some account of the relation between physical complexity and phenomenology; this is a topic to which we shall return in a later section concerning the so-called *combination problem*.

[64] Julian K. Finn, Tom Tregenza, and Mark D. Norman, "Defensive Tool Use in a Coconut-Carrying Octopus," *Current Biology* 19, issue 23 (Dec. 15, 2009): 1069–1070. See also Cell Press, "Coconut-Carrying Octopus: Tool Use in an Invertebrate," *Science Daily* (Dec. 15, 2009), www.sciencedaily.com/releases/2009/12/091214121953.htm.

[65] Note that even in the case of the octopus, intelligence is being treated as an indicator for consciousness; this seems ill-advised in light of Chalmers's "consciousness detector" comments and Bentham's lately mentioned concerns. Perhaps, as the late David Foster Wallace suggested in a provocative and well-researched article, we would do just as well to "Consider the Lobster." Or the spider…

[66] Chalmers 1996, 295. It is perhaps worth noting here, in the context of Chalmers's slug, that even reductive functionalist Fred Dretske allows that his own proposed "recipe for thought…. might, after all, be realized in a snail" (see Fred Dretske, "A Recipe for Thought," in Chalmers 2002, 497).

And "if there is experience associated with thermostats," Chalmers muses, "there is probably experience *everywhere*: wherever there is a causal interaction, there is information, and wherever there is information, there is experience."[67] In Chalmers's hands, then, the Darwinian argument brings us from the acknowledgement of non-human phenomenology to the implication of an *informational panpsychism*—with, perhaps, some bullets bitten along the way. But before examining in greater detail the notorious case of the conscious thermostat, a brief detour is in order.

As he breezes from "relatively complex" slug consciousness through to truly "primitive" thermostat consciousness, Chalmers skips over a possibility that might provisionally be labeled *biological panpsychism*. In its weaker form, biological panpsychism merely(!) asserts that all living things are conscious; in its stronger form, it also asserts that *only* living things are (i.e., can be) conscious, thus ruling out the possibility of so-called Strong Artificial Intelligence (to say nothing of conscious thermostats!). Neither form of the hypothesis has, to my knowledge, been directly addressed in the literature on panpsychism,[68] but one can find it articulated elsewhere. For example, the late biologist Lynn Margulis and science writer Dorion Sagan (son of Margulis and astronomer Carl Sagan) have made the following claim:

> Not just animals are conscious, but every organic being, every autopoietic cell is conscious. In the simplest sense, consciousness is an awareness of the outside world. And this world need not be the world outside one's mammalian fur. It may also be the world outside one's cell membrane. Certainly some level of awareness, of responsiveness owing to that awareness, is implied in all autopoietic systems.[69]

---

[67] Chalmers 1996, 297.

[68] Technically, any theory of mind that restricts experience to certain systems takes a bit of the "pan-" out of "panpsychism" (a point, as will be seen, that Strawson presses against Chalmers's arguments for informational pan*proto*psychism), but "biological panpsychism" still strikes me as the most intelligible label for the position(s) to be outlined here.

[69] Lynn Margulis and Dorion Sagan, *What is Life?* (New York: Simon & Schuster, 1995), 122; quoted in Evan Thompson, *Mind in Life: Biology, Phenomenology, and the Sciences of Mind* (Cambridge, MA: Belknap/Harvard University Press, 2007), 161. "Autopoesis" is a term introduced by Thompson's late mentor, Francisco Varela, in collaboration with Humberto Maturana:

> An autopoietic machine is a machine organized (defined as a unity) as a network of processes of production (transformation and destruction) of components which: (i) through their interactions and transformations continuously regenerate and realize the network of processes (relations) that produced them; and (ii) constitute it (the machine) as a concrete unity in space in which they (the components) exist by specifying the topological domain of its realization as such

Philosopher and cognitive-science researcher Evan Thompson calls into question Margulis and Sagan's assertion, finding it "unlikely that minimal autopoietic selfhood involves phenomenal selfhood or subjectivity."[70] It is, Thompson reminds us, "important to situate consciousness in relation to dynamic, unconscious processes of life regulation. This effort becomes difficult, perhaps impossible, if one projects consciousness down to the cellular level."[71] A later, more speculative section of the present essay will consider some intriguing empirical evidence against Thompson's latter claim, though it can also be questioned here at a conceptual level by running a variant of the Darwinian argument, not from humans down to slugs, but from brains down to neurons. That is, given that there is consciousness in the brain, why assume there is *no consciousness whatsoever* in the brain's component neurons?[72] And if there *is* consciousness in neurons, why not in (all) other single-celled organisms as well?

Regardless, Thompson's own thesis is that "where there is life there is mind, and mind in its most articulated forms *belongs* to life,"[73] and this doesn't seem far from the strong form of biological panpsychism. It does, though, suggest a somewhat nuanced notion of "life." Consider the hypothesis that so-called Strong Artificial Intelligence (or Strong AI) is possible: that human intelligence might be replicated or even exceeded by computers.[74] It would seem that for Thompson, if the development of Strong AI (an artificial mind in "its most articulated form") is possible, it will necessarily require and/or entail the development of *artificial life*, a development he does entertain as possible. Now, whether one could properly call artificial life "biological" is no doubt open to question, but in some sense, to be a living thing—even an *artificial* living thing!—is to be "biological," or at least "artificially biological" (the only apparent alternative to conceding this broadened

---

a network. [Humberto Maturana and Francisco Varela, *Autopoiesis and Cognition: The Realization of the Living*, eds. Robert S. Cohen and Marx W. Wartofsky (Dordecht: D. Reidel Publishing Co., 1980; 1st edition 1973), 78.]

[70] Thompson 2007, 162.

[71] Ibid.

[72] Though Chalmers is certainly not a biological panpsychist, his argument against "fading qualia" can be seen as pressing this point from a different angle, considering the effects of progressively replacing each neuron in the brain with a functionally identical microchip; see Chalmers 1996, 253–263.

[73] Thompson 2007, ix; emphasis mine.

[74] Though the focus of Strong AI is technically on artificial *intelligence* rather than artificial *experience*, the operative notion is, after all, that there is nothing the human brain can do that a machine, in principle, could not. A human brain can have conscious experiences; ergo, the proponent of Strong AI will want to say that a machine can also have conscious experiences.

sense of "biological" would be to push for some sort of genuine vitalism). But Strong AI aside, insofar as Thompson claims that mind "belongs to life," it could be said that he has, ironically, articulated a stronger form of biological panpsychism than Margulis and Sagan, whose assertions he finds far-fetched!

For Chalmers, by contrast, biology (however defined) is irrelevant; at least, "it is hard to see why that should make a principled difference."[75] Rather, he argues that even thermostats—which none would claim are forms of "life" in any sense—may possess rudimentary phenomenal consciousness. What matters is information processing: the thermostat's detection of a change in the environment and its corresponding response (e.g., turning the heater on). Chalmers grounds all this in the "double-aspect principle," the essential ingredient in his signature "naturalistic dualism": observing that there seem to be two aspects of the brain (i.e., physical and phenomenal), and that the physical aspect of the brain appears to be a vastly complex information-processing system, Chalmers extrapolates the hypothesis that *every* information-processing system possesses a phenomenal aspect corresponding in "richness" to the complexity of its physical organization (as he drolly concedes, this means "it will not be very interesting to be a thermostat," and one imagines the same holds for individual neurons[76]). In such "unconstrained" form, the double-aspect principle is, Chalmers acknowledges, a pathway to full-blown panpsychism. And whether or not he's inclined to follow that pathway on any given day, Chalmers's overarching naturalistic dualism does seem to entail that experience is, in some sense, a fundamental property of the entire universe. All the same, he has certain reservations about panpsychism that are, in fact, also rooted in his informational approach:

> I would not quite say that a rock *has experiences*… [or] *is conscious*…. A rock, unlike a thermostat, is not picked out as an information-processing system. It is simply picked out as an object…. It may be better to say that a rock contains systems that are conscious: presumably there are many such subsystems, none of whose experiences count canonically as the rock's (any more than my experiences count as my office's). For the thermostat, by contrast, there is a canonical associated information space, so it seems more reasonable to talk of

[75] Chalmers 1996, 296.
[76] Chalmers 1996, 293.

the thermostat's canonical experiences.[77]

Even a panpsychism grounded in the "unconstrained" double-aspect principle, then, will involve *certain* constraints on what constitutes a "canonical" conscious entity—as opposed, say, to a mere lump of latently experiential matter (this point will be touched on again later in this paper, not only as concerns the common "conscious rock" objection but, more importantly, as concerns the contrast between Strawson's full-bore panpsychism and Chalmers's panprotopsychist model). For various reasons, Chalmers himself doesn't prefer the term "panpsychism"—not only because he rejects the possibility of conscious rocks *qua* rocks, but because he sets the bar for mindedness (i.e., "psyche") higher than for consciousness *simpliciter* (i.e., experience). As he explains, "having experiences may fall well short of what we usually think of as having a mind, although it may qualify as mind in its simplest form."[78] Neither, however, does Chalmers call himself a "panexperientialist," and he allows that with his "caveats noted, it is probably fair to say that the view is a variety of panpsychism."[79]

Nevertheless, Chalmers is quick to point out that "panpsychism is not at the metaphysical foundation of my view: what is rather at the foundation is naturalistic dualism with psychophysical laws." In other words, Chalmers, like Strawson, has reached his panpsychist implications logically, not presumptively. Unlike Strawson, however, Chalmers holds out the possibility of "less extreme" readings of Russellian neutral monism on which either the intrinsic properties of matter—Armstrong's "stuffing"—are "protophenomenal" (i.e., panprotopsychism), or on which "some are neither phenomenal nor protophenomenal" (i.e., micropsychism).[80] As we shall presently see, this disagreement between Strawson and Chalmers is surprisingly nontrivial.

4. The Chalmers/Strawson Divide: Panpsychism versus Panprotopsychism

What Chalmers calls the type-F monist view could, depending on one's preference and/or degree of specificity, also be called neutral monism, Russellian monism, Spinozan

[77] Chalmers 1996, 297–298.
[78] Chalmers 1996, 299.
[79] Ibid.
[80] Chalmers 1996, 155.

monism, dual-aspect monism, or even property dualism (while Chalmers also entertains a property-dualist view of type-D dualism, the paradigmatic type-D view remains "Cartesian" substance dualism[81]). While Chalmers remains, aforementioned aesthetic considerations notwithstanding, ostensibly agnostic between type-D dualism and type-F monism, we shall assume for the purposes of this discussion that his preferred, naturalistically dualistic, nonreductively functionalist view is best understood as a type-F monist position. Certainly, the mere fact that Chalmers calls his basic position "naturalistic dualism" carries an implication against *substance* dualism, whether we construe "natural" as meaning "non-supernatural" or merely "physical" *sensu* Strawson. And in one of his rare attempts at Dennettian sloganeering, Chalmers has proffered this unambiguously monistic morsel: "Experience is information from the inside; physics is information from the outside."[82]

Even granting that Chalmers's is a type-F position, though, we find some ambiguity in his categorization of the view, wherein he speaks of "(proto)phenomenal"[83] qualities and properties, more or less leaving it to the reader to decide whether and how to read the parenthetical "proto." One could either characterize the type-F view, he says, "as a sort of panpsychism, with phenomenal properties ubiquitous at the fundamental level," or one could "give the view in its most general form the name *panprotopsychism*, with either protophenomenal or phenomenal properties underlying all of physical reality."[84] Chalmers obviously recognizes a difference between panpsychism and panprotopsychism, but he doesn't seem to find it a difference that *makes* a difference, inasmuch as he subsumes both views under type-F monism (and both, ultimately, under "the name *panprotopsychism*"). That said, Chalmers does tend to associate Russellian neutral monism with the "less extreme" panprotopsychist and micropsychist positions,[85]

---

[81] Whether Chalmers's proposed type-D *non-substance-monistic property dualism* represents a coherent position is a discussion for another day, but *prima facie*, there appears to be deep instability in such a view. Nevertheless, Chalmers has, over the years, only grown more open-minded toward the type-D view, after "setting [it] aside in *The Conscious Mind* for reasons of compatibility with physics… [that were] much too glib" (Chalmers 2010, 128, f. 26).
[82] Chalmers 1996, 305.
[83] Chalmers 2010, 134.
[84] Ibid.
[85] Cf. Chalmers 1996, 154 ("Perhaps, as Russell suggested, *at least some* of the intrinsic properties of the physical are themselves a variety of phenomenal property"; emphasis mine) and 155 ("less extreme cases[s

and generally conveys a sense of being more comfortable with these than with panpsychism proper.[86]

In all of this, Chalmers differs markedly from the more strident Strawson, who in "Realistic Monism" will countenance nothing less than what Chalmers deems "extreme" panpsychism. As we have seen, Strawson dismisses micropsychism: no less incongruous than "the idea that some but not all physical ultimates are spatio-temporal," he remarks, micropsychism's "radical heterogeneity at the very bottom of things" makes it "hard to see why this view would not count as a form of dualism." Applying Chalmers's categories, it seems fair to say that Strawson sees micropsychism as collapsing into a type-D view involving interaction between bits of (proto?)phenomenal matter and bits of wholly non-experiential "zombie" matter. In this sense, according to Strawson, either Russellian neutral monism is really a type-D view, in which case Chalmers must be mistaken in treating it as the paradigmatic type-F view, or Chalmers must be mistaken in associating micropsychism with Russell.

But Chalmers himself doesn't spend much time on micropsychism (never, so far as I am aware, identifying it by name when he does discuss it). The true Chalmers/Strawson Divide, so to speak, rests on the "proto" issue. While Chalmers, as mentioned, seems more confident in panprotopsychism than in panpsychism proper, Strawson doubts whether panprotopsychism, with its purported protophenomenal properties, is even coherent as an alternative to brute emergence. In turn, much of Chalmers's resistance to panpsychism à la Strawson is grounded in his concerns over the so-called *combination problem*: how can trillions of individual subatomic experience(r)s add up to a single human experience(r)? Together, the horns of emergence and combination present an especially troublesome dilemma for pan(proto)psychism, so it will be worthwhile to explore the differing approaches taken by leading contemporary proponents of the position. The remainder of this section will consider the challenges Strawson has put to Chalmers's panprotopsychist view(s); Chalmers's resistance to full-bore panpsychism

---

are] perhaps best regarded as a version of Russell's neutral monism"), as well as Chalmers 2010, 134 ("In its protophenomenal form, the [type-F] view can be seen as a sort of neutral monism").

[86] See Chalmers's stated qualms about the label "panpsychism," noted at the end of the previous section and to be discussed in greater detail in the following section.

will in turn be considered in the following section, along with other common concerns about the view.

Regarding, then, the putative "protophenomenal" properties at play in Chalmers-style panprotopsychism, Strawson believes some clarification is in order. His concern is that if protophenomenal properties are "ultimately non-experiential in themselves," then panprotopsychism "doesn't escape the problem, it merely changes the terms":

> 'Proto-experiential' [read: 'protophenomenal'] now means 'intrinsically suited to constituting certain sorts of experiential phenomena in certain circumstances', and clearly—necessarily—for X to be intrinsically suited to or for constituting Y in certain circumstances is for there to be something about X's nature in virtue of which X is so suited. If there is no such in-virtue-of-ness, no such intrinsic suitability, then any supposed emergence is left brute, in which case it is not emergence at all, it is magic, and everything is permitted.

Nor, Strawson insists, can we avoid the problem by substituting "*giving rise to* Y" or "*producing* Y" for "constituting Y":

> The idea will be that X remains *in itself* wholly and utterly non-experiential, but *gives rise to* something wholly ontologically distinct from itself, i.e. Y. But real physicalists can't make this substitution. For everything real and concrete is physical, on their view, and experiential phenomena are real and concrete, on their view, and none of them will I think want to throw away the conservation principles and say that brand new physical stuff (mass/energy) is produced or given rise to when experiences are emergent from the non-experiential.… That is magic again, and I am assured that nothing like this happens with liquidity…[87]

Like *just must* emerge from like, on Strawson's view, or we're back to the "magical" doctrine of brute emergence. But given Chalmers's own deep discomfort with brute emergence, it's safe to say this isn't the account of protophenomenology he would prefer to give—not that he's entirely clear on the matter. At one point, he defines the protophenomenal as "properties that collectively constitute phenomenal properties when organized in the appropriate way,"[88] which might seem to play right into Strawson's original critique (compare Chalmers's phrasing here with Strawson's "intrinsically suited to constituting certain sorts of experiential phenomena in certain circumstances"). Yet for

[87] Strawson in Freeman 2006, 23–24.
[88] Chalmers 2010, 151. The chapter containing this passage, "The Two-Dimensional Argument Against Materialism," was originally published as an article of the same name. Chalmers has added an afterword ("Other Anti-Materialist Arguments") to the latest edition; see Chalmers 2010, 141–206.

the critique to have force, Chalmers would need to be denying the "in-virtue-of-ness" and "intrinsic suitability" that Strawson insists be present for the kosher emergence of the phenomenal from the protophenomenal. Chalmers doesn't appear to be making such a denial; that is, he doesn't posit the emergence of like from unlike. Indeed, Chalmers seems to recognize as well as anyone (e.g. Strawson and McGinn, among others) that for theoretical purposes, fully *non*-phenomenal protophenomenal properties might just as well be old-fashioned non-phenomenal *physical* properties. To avoid brute emergence (not to mention a collapse from type-F monism into reductive materialism), protophenomenal properties cannot be *non*-phenomenal, nor even *pre*-phenomenal—they must, Chalmers and Strawson would seem to agree, be phenomenal *in some way*.

But in what way? Is Chalmers positing the emergence of like from *kinda*-like? In a passage aimed more or less directly at Chalmers,[89] Strawson puts the point starkly:

> If you take the word 'proto-experiential' to mean 'not actually experiential, but just what is needed for experience', then the gap is unbridged. If you take it to mean 'already intrinsically (occurently) experiential, although very different, qualitatively, from the experience whose realizing ground we are supposing it be', you have conceded the fundamental point [i.e., proper panpsychism].[90]

In other words, Strawson thinks any non-brute, "intrinsically experiential" theory of panprotopsychism simply collapses into proper panpsychism. On this account, it wouldn't even be right to call such panprotopsychism a *form* of panpsychism (recall, by contrast, that Chalmers sees panprotopsychism, not panpsychism, as the "most general form" of type-F monism). Rather, for Strawson, any legitimate type-F position[91] *will just be panpsychism*, albeit by another, perhaps more "politically correct" name. And any illegitimate type-F position will be just that: illegitimate and non–gap-bridging.

Thus far, Chalmers has made little public comment on Strawson's general case for panpsychism, much less any direct response to these specific protophenomenal concerns. However, Chalmers does have more nuanced explanations of panprotopsychism on offer,

---

[89] We are urged, in a footnote, to compare Strawson's critique of "proto-experiential" with the "protophenomenal" speculation in Chalmers 1996; see Strawson in Freeman 2006, 22f37.
[90] Strawson in Freeman 2006, 21–22.
[91] Strawson openly accedes to Chalmers's taxonomy, writing that "Chalmers correctly classified me as holding a version of the position he now calls 'Type-F monism'" (Galen Strawson, "Panpsychism? Reply to Commentators with a Celebration of Descartes," in Freeman 2006, 186).

with which Strawson is almost certainly familiar but has himself not addressed in detail. In any case, Chalmers's panprotopsychist story runs as follows: "it could be that consciousness is not itself fundamental but is necessitated by some more primitive fundamental feature *X* that is not itself necessitated by physics. In this case, we might call *X* a *protophenomenal* property, and we can say that protophenomenal properties are fundamental."[92] At first blush, this response seems to slip right through Strawson's gauntlet: since the fundamental protophenomenal *X* is, like ordinary phenomenology, "not itself necessitated by physics" (non-physicSal, in Strawson's language), Chalmers might seem to (1) offer a kosher emergence of like from like that (2) avoids the "extreme" move of positing consciousness as fundamental by positing instead "some more primitive fundamental feature *X*." This move may seem especially compelling if we consider Chalmers's fundamental, non-physicSal, consciousness-necessitating *X* in the context of a passage quoted earlier from Strawson (where consciousness, of course, is *Y*):

> It seems plain that there must be a fundamental sense in which any emergent phenomenon, say Y, is wholly dependent on that which it emerges from, say X. It seems, in fact, that this must be true by definition of 'emergent'; for if there is not this total dependence then it will not be true after all, not true without qualification, to say that Y is emergent from X. For in this case at least some part or aspect of Y will have to hail from somewhere else and will therefore not be emergent from X. Plainly this is not how it is with liquidity.[93]

Nor with Chalmers's panprotopsychism, apparently: though the protophenomenal *X* may lack consciousness just as fundamental matter (or, depending how one runs the analogy, *the property of mass*) lacks liquidity, nothing truly *new* emerges here: everything in the phenomenal universe reduces to the non-physicSal, protophenomenal *X*. Thus we might seem to get the best of both worlds, preserving the elegant, holistic simplicity of panpsychism while respecting our strong intuition that the phenomenal difference between a human and an electron is one of *kind*, not merely one of degree. But the matter is not that simple, as Chalmers at least partially acknowledges:

> One might… object that we do not have any conception of what protophenomenal properties might be like or of how they could constitute

[92] Chalmers 2010, 125 (from "Consciousness and its Place in Nature," the source of the "type A–F" taxonomy to which Strawson accedes).
[93] Strawson in Freeman 2006, 14.

> phenomenal properties. This is true, but one could suggest that this [is] merely a product of our ignorance. In the case of familiar physical properties, there were principled reasons (based on the character of physical concepts) for denying a constitutive connection to phenomenal properties. Here, there are no such principled reasons. At most, there is ignorance of a connection. Of course, it would be very desirable to form a positive conception of protophenomenal properties. Perhaps we can do this indirectly by some sort of theoretical inference from the character of phenomenal properties to their underlying constituents, or perhaps knowledge of the nature of protophenomenal properties will remain beyond us. Either way, this is no reason to reject the truth of the view.[94]

Nevertheless, Chalmers does seem to overstate his case here in claiming that "at most, there is ignorance of a connection"—for not only is there ignorance of a *connection*, but by hypothesis, there is deep present (and perhaps permanent) ignorance "of what protophenomenal properties might be like" *whatsoever*! And there are other considerations in the vicinity. First, one could well ask whether the elegant simplicity of panpsychism is sundered by the introduction of "some more primitive fundamental feature *X*." After all, we *know* that consciousness exists; adding to our ontology an underlying *X* that is by hypothesis *unknowable* via either experience or scientific detection (and perhaps not even by "theoretical inference"!) seems unparsimonious, even unwarranted. Second, is there only one underlying protophenomenal property, or are there many? Third, one might wonder whether Chalmers's fundamental protophenomenal property/ies is/are supposed to be ubiquitous or only present in certain forms of matter; this latter might be called the question of "microprotopsychism" or "protomicropsychism" (one imagines Chalmers would lean toward ubiquity, given the putative role played by *X* in "constituting the physical domain"). But most crucially, one might question whether the emergence of like from like is truly kosher here, especially in light of comments later in Chalmers's discussion that "in its protophenomenal form, the view can be seen as a sort of neutral monism: there are underlying neutral properties X (the protophenomenal properties), such that X properties are simultaneously responsible for constituting the physical domain (by their relations) and the phenomenal domain (by their collective intrinsic nature)."[95] Consider the putative information we have about the protophenomenal and/or "underlying neutral" property/ies *X*:

[94] Chalmers 2010, 136.
[95] Chalmers 2010, 134.

(1) *X* is a "neutral," "primitive fundamental feature" (of reality/matter etc.).
(2) Consciousness is "necessitated" by the "collective intrinsic nature" of *X*.
(3) The physic(S)al domain is constituted by the "relations" of *X*.
(4) *X* "is not itself necessitated by physics."

It is premise (4) that is supposed to assure us safe passage from like to like. Assuming we are onboard with Chalmers's basic Hard Problem program, we grant the background premise that consciousness, like *X*, "is not itself necessitated by physics." And so it is, Chalmers asserts, that we can get like from like: since *X* is not necessitated by physics, there are no "principled reasons" deriving from "the character of physical concepts" that might otherwise lead us to deny "a constitutive connection" from protophenomenal to phenomenal properties. But merely *not being necessitated by physics* is hardly the distinctive feature of consciousness that someone like Strawson (or even, presumably, Chalmers) should wish to preserve in any story of emergence of like from like. After all, any number of hypothetical "features" are not themselves "necessitated by physics," including, *inter alia*, Cartesian soul-stuff.

Nor is *not being a certain way* a particularly strong form of similarity; I am not a table, and neither is the book in front of me, but the shared "non-tableness" of the book and me seems a trivial similarity. Nor, were it a person in front of me instead of a book, would the shared non-tableness of that person and me specify a particularly non-trivial similarity between us! Rather, it would specify no greater degree of similarity between me and the person than between me and the book: we would, all three of us, be equal in our non-tableness. Yet indubitably I share other, more pertinent similarities with the person than I do with the book. In the case of panprotopsychism, then, it seems arguable that the putative not-necessitated-by-physics-ness (henceforth, after Strawson, "non-physicSality") shared by the protophenomenal *X* and phenomenal consciousness just isn't a strong enough similarity to give us a meaningful story of emergence of like from like and a definitive ruling-out of brutality. Consider: that I emerged from the person standing in front of me is quite possible, presuming that person is my mother, while my emergence from the book in front of me is, of course, impossible (or would, in any case, be quite brute were it possible)—yet these two cases are *indistinguishable* based solely on the information that "something in front of me shares my non-tableness."

Now, this line of argumentation isn't quite fair; after all, the non-physicSality of consciousness is no minor detail, but is arguably the mystifying issue at the heart of the Hard Problem. An underlying fundamental property that was itself non-physicSal would, one must concede, be more significantly similar to consciousness (in respect of being non-physicSal) than I am similar to a book (in respect of not being a table). But while non-physicSality may be an important and mysterious aspect of consciousness, it is not nearly as *prima facie* crucial an aspect as *experientiality itself*. That is, it doesn't seem unfair to ask that a serious panprotopsychist theory of mind get us *Y* in virtue of *X* by way of *experientiality* rather than by way of *non-physicSality*. Indeed, the Hard Problem applies to the latter but not the former: just as we agonizingly wonder how experience could possibly emerge from the physicSal, we might agonizingly wonder in this case how experience could possibly emerge from the non-physicSal—yet the emergence of experience from the experiential, by contrast, presents no such problem (though, as will be discussed, it does present others). So even granting that experience is non-physicSal, it's not so much the *non-physicSality of experience* that needs to be accounted for, emergence-wise, as, for lack of a better phrase, the *experientiality of experience*. Unless some entailment from non-physicSality to experientiality can be shown or even proposed (analogous to the entailment, in the case of water, from physicSality to liquidity), it's not clear that Chalmers's protophenomenal story truly delivers us the *relevant* like from like. The desideratum here is presumably an account of emergence in virtue of some *positive* similarity between the protophenomenal and the phenomenal (e.g., experientiality); the "negative similarity" of non-physicSality just doesn't seem sufficient to get us from there to here—especially if Chalmers's protophenomenal *X* is not in itself experiential in any way (an issue on which Chalmers is, as noted, somewhat vague). Or, as Strawson has pointedly put it: "If one can have PROTO with no what-it's-likeness, then we're back with radical emergence."[96]

Here, then, is the crucial point of divergence between Chalmers and Strawson: the former can accept the possibility of experience emerging from a non-physicSal but non-experiential substrate; the latter simply cannot. This divide explains the at-first-surprising

[96] Galen Strawson, personal correspondence, April 29, 2010.

fact that Strawson, despite his paradigmatically type-F deployment of the Russell-Eddington line, distances himself from the "neutral monism" most commonly associated with Russell (and which, as noted, Chalmers considers equivalent to the "protophenomenal form" of type-F monism). As Strawson writes:

> The central idea of neutral monism is that there is a fundamental, correct way of conceiving things—let us say that it involves conceiving of them in terms of 'Z' properties—given which all concrete phenomena, experiential and non-experiential, are on a par in all being equally Z phenomena. This proposal, however, merely confirms the current position [i.e., Strawson's panpsychism]. For what we do, when we give a satisfactory account of how liquidity emerges from non-liquidity, is show that there aren't really any new properties involved at all. Carrying this over to the experiential case, we get the claim that what happens, when experientiality emerges from non-experientiality, is that there aren't really any new properties involved at all. This, however, means that *there were experiential properties all along*; which is, precisely, the present claim. One cannot oppose it by appealing to 'neutral monism' in any version that holds that really only the Z properties are ultimately real, if this involves the view that experiential and non-experiential properties are at bottom only appearances or seemings. Such a view is incoherent, because experience—appearance, if you like—cannot itself be only appearance, i.e. not really real, because there must be experience for there to be appearance.[97]

Strawson stresses that final point elsewhere as well: "Experiential phenomena… cannot be mere appearance, if only because all appearance depends on their existence."[98] Of course, Chalmers would be loath to claim that experiential phenomena are mere appearance—but if he's right that the "protophenomenal form" of type-F monism can "be seen as a sort of neutral monism" in which the underlying protophenomenal *X* is "neutral," not only does he, by Strawson's lights, commit himself to brute emergence, but also to the "incoherent" view that both "experiential and non-experiential properties are at bottom only appearances or seemings." And whether or not such a view is truly incoherent, Strawson seems on the mark as far as the tenets of proper neutral monism. Though the view is not without significant variations, Leopold Stubenberg asserts its fundamental premise thusly in his *Stanford Encyclopedia of Philosophy* entry: "What distinguishes neutral monism from its better known monistic rivals is the claim that the

[97] Strawson in Freeman 2006, 23. Emphasis mine.
[98] Strawson in Freeman 2006, 17.

intrinsic nature of ultimate reality is *neither* mental nor physical."[99] So if Chalmers's protophenomenal *X* is truly "neutral," then the intrinsic nature of ultimate reality is *not* mental after all; the Russell-Eddington line then no longer holds, and one of the major conceptual planks of panpsychism (and, arguably, of type-F monism in general) is lost. Indeed, as Stubenberg notes, "On the face of it, panpsychism and neutral monism are as different as could be. Neutral monism reduces mental and physical phenomena, panpsychism does not; neutral monism holds that the 'materia prima' is neutral, panpsychism does not. And neutral monism assumes that there are genuinely physical (i.e., nonmental) phenomena that need reducing, panpsychism does not."[100] One wonders, then, whether the view Chalmers presents even warrants the name "panprotopsychism"—not merely because of the conceptual gulf between "nonmental" and "proto-mental," but because the underlying *X* seems not merely protophenomenal (whatever, in the end, that means), but also *protophysic(S)al*. These factors make the term "panprotopsychism" doubly misleading: it both wrongly implies a form of panpsychism (from which it is "as different as could be") and represents, in name, only half the ontology it purports to account for. "Panprotoholism" might be more accurate, but "neutral monism" will do just as well.

As noted, however, idiosyncratic interpretations of neutral monism abound, and the "experience as mere appearance" version addressed above is perhaps not exactly what Chalmers has in mind. Though he characterizes panprotopsychism as "a sort of neutral monism," in the immediately prior passage he writes that the view "can be seen as a sort of dualism."[101] Specifically, Chalmers is referring to a position variously denoted as property dualism, dual-aspect theory, and even dual-aspect monism. While the association of monism (neutral or otherwise) with dualism may initially seem incoherent, Chalmers is not out on a limb here—or at least not alone on the limb. As Stubenberg notes:

> The view with which neutral monism is most often compared or identified is the dual aspect theory…. All versions of the theory appear to be committed to the

[99] Leopold Stubenberg, "Neutral Monism," in *The Stanford Encyclopedia of Philosophy (Spring 2010 Edition)*, ed. Edward N. Zalta, plato.stanford.edu/archives/spr2010/entries/neutral-monism/.
[100] Stubenberg 2010, section 9.2.
[101] Chalmers 2010, 134.

> view that there are certain substances—god or nature (Spinoza 1677), persons ([P. F.] Strawson 1959), body or brain (Thomas Nagel, 1986), information (a view explored by David Chalmers 1996)—that are intrinsically neither material nor mental. Nevertheless these substances can present themselves under the aspect of the mental and the aspect of the physical. And these aspects are distinct yet inseparable and basic in the sense of being irreducible to each other or to anything else.[102]

Such a view brings us closer to Galen Strawson's panpsychism, given the irreducibility (and, *pari passu*, reality, i.e., not-mere-appearance-ness) of the putative dual aspects. On this sort of conception, to borrow Stubenberg's description, "the allegedly neutral entities are not neutral between mind and matter but really a little bit of each; and… the activity that the neutral monists describe as 'reducing physical and mental phenomena to constructions of neutral entities' is really not quite that, but whatever it is that panpsychists do when they explain how it all hangs together."[103] But this is not *proper* neutral monism; Stubenberg quotes William James, a founding father of neutral monism, griping that the dual-aspect view "is a monism in name only":

> It poses an unknown reality, but tells us that this reality always presents itself under two "aspects," the conscious aspect and the material aspect, and these two sides remain as irreducible as extension and thought, the fundamental attributes of Spinoza's God. In effect, contemporary monism is pure Spinozism…. [Neutral monism is] absolutely opposed to the so-called bilateral monism of the scientific positivist or Spinozist.[104]

Yet even Russell himself—whose view on neutral monism, under heavy influence from James, evolved from skepticism to acceptance to advocacy over many years—might be read as endorsing just such a "Spinozist" dual-aspect position. Consider Russell's claim that under the framework of neutral monism, "It becomes possible to regard both a mind and a piece of matter as logical constructions formed out of materials not differing vitally and sometimes actually identical."[105] Taken at face value, this view, as we have seen, qualifies as micropsychism, and if the "materials" in question are *always* "actually identical" (i.e., if micropsychism is rejected), a dual-aspect view is clearly implied.

---

[102] Stubenberg 2010, section 9.4.
[103] Stubenberg 2010, section 9.2.
[104] William James 1905, 109 and 117, quoted in Stubenberg 2010, section 9.4 (source unclear).
[105] Bertrand Russell, *My Philosophical Development* (Reprinted: New York: Simon and Schuster, 1975), 103; quoted in Stubenberg 2010, section 4.4.

Might such a view be capable of delivering a viable pan*proto*psychism with kosher emergence? Certainly, it would seem to permit the idiosyncratic type-identity claim Strawson admits to making ("I am happy to say… that experience is 'really just neurons firing'… But when I say these words I mean… that experiential phenomena 'just are' physical, so that there is a lot more to neurons than physics and neurophysiology record (or can record)"[106]). But here again, one would perhaps be surprised to find that Strawson disavows property dualism as unflinchingly as he does neutral monism:

> One needs to grasp fully the point that 'property dualism', applied to intrinsic, non-relational properties, is strictly incoherent (or just a way of saying that there are two very different kinds of properties) insofar as it purports to be genuinely distinct from substance dualism, because there is nothing more to a thing's being than its intrinsic, non-relational propertiedness."[107]

At this point, the deeper currents of Strawson's metaphysical thought threaten to sweep us away; must we simply just accept that "there is nothing more to a thing's being than its intrinsic, non-relational propertiedness"? In light of the Russell-Eddington line, this is actually a natural conclusion to draw: extrinsic, relational properties show us what a thing *does*, but intrinsic, non-relational properties determine what a thing *is*. In his published response to critics of "Realistic Monism" and at greater length in his 2009 book *Selves: An Essay in Revisionary Metaphysics*, Strawson provides ample argumentation for this claim, but to assess its merit here would take us significantly off course. For now, just two important points should be made regarding Strawson's take on property dualism. First, note that so long as property dualism is "just a way of saying that there are two very different kinds of properties" (i.e. the phenomenal and the physicSal), Strawson takes no conceptual umbrage; presumably, his only gripe would be that such property dualists (including Chalmers, with his "double-aspect principle") should just do the brave thing and admit that they are really panpsychists. Second, though we needn't here evaluate Strawson's claim that "there is nothing more to a thing's being than its intrinsic, non-relational propertiedness," it is worth noting that another major disagreement between Strawson and Chalmers hangs on precisely this claim. Specifically, Strawson rejects the logical possibility and even the coherent conceivability of Chalmers's infamous non-

[106] Strawson in Freeman 2006, 7.
[107] Strawson in Freeman 2006, 28.

conscious zombies.[108] The very idea of such "Australian zombies," Strawson sniffs, "rules out real physicalism,"[109] because after all, real physicalism is defined by the thesis that "experience is a real concrete phenomenon and every real concrete phenomenon is physical." Having thus defined the "intrinsic, non-relational property" of experientiality into the (really) physical, and furthermore maintaining that "there is nothing more to a thing's being than its intrinsic, non-relational propertiedness," clearly Strawson will want to say that any physical duplicate of a conscious entity must necessarily have the same intrinsic, non-relational propertiedness (i.e. consciousness) as the original entity—hence the impossibility of zombies. As for their coherent conceivability, Strawson opines:

> One reason why Chalmers accepts it may be that he subscribes to the idea that one can in counterfactual speculation suppose that one is talking about qualitatively the same physical objects when one varies the physical laws that govern them. It seems plain to me that the laws of physics are constitutive of the nature of the physical in such a way that one cannot do this.[110]

As it happens, Chalmers preemptively accounts for Strawson's line of objection in his original description of the type-F view; in fact, he outlines three basic positions that a type-F monist may take regarding "the zombie argument against materialism":

> Some type-F monists may hold that a complete physical description must be expanded to include an intrinsic description and may consequently deny that zombies are conceivable. (We only think we are conceiving of a physically identical system because we overlook intrinsic properties.) Others… [could] accept conceivability but deny possibility. On this interpretation, we misdescribe the conceived world as physically identical to our when in fact it is just structurally identical. Finally, a type-F monist might hold that physical concepts refer to dispositional properties, so that zombies are both conceivable and possible, and the intrinsic properties are not physical properties. The differences among these three attitudes seem to be ultimately terminological rather than substantive.[111]

Clearly, Strawson takes the first view while Chalmers (when he's wearing his type-F hat)

---

[108] There are other intriguing differences in intuition in this vicinity between Strawson and Chalmers; for example, Strawson denies the thermostat status as a conscious entity (i.e., as "any sort of locus of experience larger than fundamental panpsychism"; Strawson, personal correspondence, May 14, 2010). Strawson's reasoning here is not altogether clear to me, however, so I shall not pursue the point in any further detail.

[109] Strawson in Freeman 2006, 22 f.37.

[110] Strawson, "Panpsychism?" in Freeman 2006, 225 f.85.

[111] Chalmers 2010, 134–135.

takes the third. Whether Strawson could be persuaded to concede the viability of the second view is an open question, but what is certain is that the third view—on which "the intrinsic properties are not physical properties"—stands in starkest opposition to the expanded definition of "physical" that, for Strawson, *entails* panpsychism. It's also a bit puzzling how Chalmers could cash out a view on which intrinsic properties are *not* physical properties, given his fondness for the Russell-Eddington line regarding the intrinsic nature of the physical; the story would presumably entail that *the intrinsic nature of the physical is non-physical* (not just non-physicSal)—which only seems workable, as we've seen, as "panprotoholistic" neutral monism, not as genuine panpsychism.

One might therefore wonder whether Chalmers is too sanguine in finding this difference of views about zombies "ultimately terminological rather than substantive." Not only, as we have seen, does Chalmers's panprotopsychism fail to square with Strawson's panpsychism (unless it collapses into it), but even a full-blown panpsychist account from Chalmers would—by dint of denying the physicality of intrinsic properties, i.e. "real physicalism"—be *equally incompatible* with Strawson's proposal! But Chalmers may be unlikely to deliver a full-blown panpsychist account any time soon; while he has not offered direct criticism of Strawson's view, he does, as mentioned earlier, have certain generalized reservations regarding panpsychism—reservations shared by many who have considered panpsychism (Strawsonian or otherwise) over the years. Such concerns will be the focus of the following section.

5. Answering Objections to Panpsychism

Let us begin by taking a closer look at the reasons Chalmers gives for "not generally us[ing] the term" panpsychism to describe his own view:

> (1) because I think that having experiences may fall well short of what we usually think of as having a mind, although it may qualify as mind in its simplest form; (2) because protophenomenal properties may be even further away from the usual concept of "mind"; (3) because I do not think it is strictly accurate to say that rocks (for example) have experiences…. [(4)] Perhaps the central reason why the term is misleading, though, is that it suggests a view in which the experiences in simple systems such as atoms are fundamental, and in which complex experiences are somehow the sum of such simpler experiences. While this is one way things could go… complex experiences may be more

> autonomous…. In particular, the informational view suggests a picture on which complex experiences are determined more holistically than this.[112]

As previously discussed, Chalmers's first reservation turns on a literalist reading of "psyche" as "mind." This is a fair enough distinction to draw, but it doesn't really touch on the substance of the contemporary debate; by that standard, what Strawson and Skrbina offer wouldn't technically be panpsychism either, but rather panexperientialism (a term with which Chalmers seems no more comfortable, despite its greater technical accuracy[113]). The second reservation simply keeps open the option of protophenomenal properties, which we have already discussed at length. The third reservation is, again, one that would presumably put Strawson and Skrbina out of the running as genuine panpsychists, for neither would assent to the claim that it "is strictly accurate to say that rocks… have experiences" ("I don't believe this for a moment," writes Strawson, while Skrbina calls the idea "obviously ludicrous"[114]). This leaves standing only the unnumbered (by Chalmers) fourth reservation, but it is a doozy: the notorious *combination problem*, as elucidated by, e.g., James, Seager, and Stoljar.

The problem, in short, is this: how can a countless multitude of individually experiential ultimates add up to a single, unified human experience(r)? Leaving aside his reservations regarding panpsychism *per se*, Chalmers evidently considers the combination problem the only "principled problem in the vicinity" of a "distinctive type-F monism."[115] And certainly it is a principled problem for Strawson's panpsychism; as the latter freely admits near the end of "Realistic Monism," a developed version of his view "will need to address [the] well known objection to the idea that many subjects of experience can somehow constitute a single 'larger' subject of experience…. we will have to wonder how macroexperientiality arises from microexperientiality, where by microexperientiality I mean the experientiality of particles relative to which all evolved experientiality is macroexperientiality."[116] Chalmers likewise surmises that any attempt to answer the combination problem will require "a much better understanding of the compositional

[112] Chalmers 1996, 299.
[113] Strawson treats these equivalently in his response to critics in Freeman 2006.
[114] Strawson in Freeman 2006, 26; Skrbina 2005, 17.
[115] Chalmers 2010, 136.
[116] Strawson in Freeman 2006, 26.

principles of phenomenology: that is, the principles by which phenomenal properties can be composed or constituted from underlying phenomenal properties, or protophenomenal properties."[117]

Easier said than done, of course (even leaving aside Chalmers's problematic "protophenomenal properties"). The challenge laid down by the venerable William James in 1890 remains the obligatory combinatory *reductio* of panpsychism, so I shall quote it here once again:

> Take a sentence of a dozen words, and take twelve men and tell to each one word. Then stand the men in a row or jam them in a bunch, and let each think of his word as intently as he will; nowhere will there be a consciousness of the whole sentence.… Where the elemental units are supposed to be feelings, the case is in no wise altered. Take a hundred of them, shuffle them and pack them as close together as you can (whatever that might mean); still each remains the same feeling it always was, shut in its own skin, windowless, ignorant of what the other feelings are and mean. There would be a hundred-and-first feeling there, if, when a group or series of such feeling were set up, a consciousness *belonging to the group as such* should emerge. And this 101st feeling would be a totally new fact; the 100 original feelings might, by a curious physical law, be a signal for its *creation*, when they came together; but they would have no substantial identity with it, nor it with them, and one could never deduce the one from the others, or (in any intelligible sense) say that they *evolved* it.[118]

Is this an intractable problem for panpsychism? Many have assumed so, but it noteworthy that despite having authored this most pointed objection, James himself would actually embrace panpsychism in years to come. Skrbina finds in James's later lecture notes such statements as "Our only intelligible notion of an object in itself is that it should be an object for itself, and this lands us in panpsychism and a belief that our physical perceptions are effects on us of 'psychical' realities," along with outright advocacy of "a great empirical movement towards a pluralistic panpsychic view of the universe" (where "empirical," Skrbina notes, "refers to James' 'radical empiricism,' in which everything consists of pure experience").[119] How, then, did James resolve (or attempt to resolve) the combination problem he once found so problematic? Disappointingly, with a move few

[117] Chalmers 2010, 136.
[118] William James, *The Principles of Psychology*, vol. 1 (New York: Henry Holt and Co., 1890; reprint 1950, New York: Dover, 1950), 160; quoted in Stubenberg 2010, section 5.
[119] Skrbina 2005, 146.

contemporary philosophers would likely countenance; witness Skrbina's account of James's "final solution to the combination problem":

> Formerly [James] had argued that any collective experience had to be unlike the constituent experiences; they had to be "logically distinct." The result, logically speaking, was that combination was impossible. Now James realizes that this situation is "almost intolerable" because "it makes the universe discontinuous." Such logic forces one to conclude that the universe is a "contradiction incarnate." If analytic logic compels one to this view, "so much the worse for logic." For James, logic is an intellectual tool of the cynical, materialistic philosophers, and so he transcends it. He adds this: "Reality, life, experience, concreteness, immediacy, use what word you will, exceeds our logic, overflows and surrounds it." Thus, combination is possible after all, and in fact it maintains the continuity of mind throughout the universe.[120]

Whatever our sympathies regarding the analytic method or "cynical, materialistic philosophers," we should presumably expect more from a panpsychist account of combination than a glib "so much the worse for logic." But where, exactly, might we start? Strawson makes a perhaps unexpected suggestion:

> It is at this point… that the notion of emergence begins to recover some respectability in its application to the case of experience…. For we can take it that human or sea snail experientiality emerges from experientiality that is not of the human or sea snail type, just as the shape-size-mass-charge-etc. phenomenon of liquidity emerges from shape-size-mass-charge-etc. phenomena that do not involve liquidity.[121]

Strawson admits this doesn't get us very far and "has nothing to offer to scientific test,"[122] but the idea is that just as liquidity is a wholly physicSal property not possessed by any *individual* molecule of water, human consciousness ("macroexperientiality") could be a wholly non-physicSal property not possessed by any *individual* organic molecule. This isn't to say that those individual organic molecules would possess no non-physicSal (read: experiential) properties of their own. Indeed, they would possess "microexperientiality" (not merely Chalmers-style "protoexperientiality," of course!), but wouldn't possess the property of "macroexperiential" human consciousness any more than an isolated molecule of water at room temperature possesses the property of

[120] Skrbina 2005, 147–148.
[121] Strawson in Freeman 2006, 27.
[122] Strawson in Freeman 2006, 28.

liquidity.[123] Still, the analogy seems somewhat strained: put any number of water molecules together in a bucket at room temperature and liquidity will naturally emerge, yes—but put the same number of organic molecules (and some water) in a similar bucket and human consciousness will almost certainly *not* emerge.

This, I think, is the point at which Chalmers, wearing his proper (non-"proto") panpsychist hat, has much to offer Strawson's argument. After all, the obvious reason we don't get macroexperiential human consciousness from a bucket of microexperiential organic molecules is that macroexperientiality requires not merely the *addition* or *combination* of microexperientiality, but also some sort of *organization*. Chalmers's informational panpsychism thus seems like just the sort of thing Strawson's story needs to have any hope of explanatory power. Now, as it happens, Strawson resists Chalmers's conscious thermostat, the very poster child for informational panpsychism (as Strawson explains, "I don't think a thermostat constitutes any sort of subject of experience… any sort of locus of experience larger than fundamental [panpsychism]").[124] But Strawson's own preferred story involves, as it clearly must, "relatively unorganized matter" that, via evolutionary processes, "organized into increasingly complex forms," such that "just as there was spectacular enlargement and fine-tuning of non-experiential forms (the bodies of living things), so too there was spectacular enlargement and fine-tuning of experiential forms."[125] So whether or not Strawson sets the bar for macroexperientiality higher than the thermostat, at some point, something akin to Chalmers's informational/organizational approach needs to enter the story. Now, as we've seen, there's simply no room for that level of sophistication in the liquidity analogy—but perhaps we can find a better analogy.

Strawson himself provides a hint, in an especially bold bit of speculation: "The heart of experience, perhaps, is electromagnetism in some or all its forms; but electromagnetism

[123] That Strawson would treat the intrinsic, non-relational "stuffing for matter" (that which purportedly *just is* experience) as a *property* may seem quite odd until one has fully digested Strawson's revisionary metaphysics; recall: "there is nothing more to a thing's being than its intrinsic, non-relational propertiedness." One needs, he insists, "a vivid sense of the respect in which… every object is a process; one needs to abandon the idea that there is any sharp or categorial [sic] distinction between an object and its propertiedness." In this approach, Strawson sees himself as "following Nagarjuna, Nietzsche, James, Ramsey and many others"; see Strawson in Freeman 2006, 28.
[124] Strawson, personal correspondence, May 14, 2010. "Fundamental panpsychism" is presumably equivalent to "microexperientiality."
[125] Strawson in Freeman 2006, 27.

in all its forms is no doubt just one expression of some single force whose being is intrinsically experiential, whatever else it is or is not."[126] Now, given the strong correlation between electrical activity and consciousness in the human brain, Strawson might not be far off here—but whether or not experience is *literally* "electromagnetism from the inside" will be irrelevant to the present analogy. And the analogy is this: while electromagnetism is a fundamental and pervasive force in all matter, its most interesting effects occur when complex conditions obtain, e.g., in an electrical circuit. Electromagnetism is undoubtedly *present* in an electron or rock, but not the way it is in a thermostat, hair dryer, computer, or human brain; these latter constitute electromagnetic *systems* in a way the electron and rock do not. Likewise, we might posit, with experientiality.

This way of thinking about panpsychism seems to me a powerful way to defuse the combination problem. For consider the simple electric current supplied to your home, in which electrons are carried via a conducting wire from a generating plant to an electrical device such as your computer. Each individual electron is an elementary particle (that is to say that an electron is, so far as science is aware, not divisible into any smaller parts), and each carries a fixed charge of –1 (or, depending on one's level of precision, $-e$, where $e = 1.602\times10^{-19}$ coulombs). And that's all: from various manipulations of these electrons, each identically simple in its negative charge, emerges every type of electrical system yet conceived by mankind. Take away the flow of electrons to your home and you have a power outage; return it and you have the capacity, among other things, to refrigerate and microwave food, illuminate and climate-control your room, listen to your stereo, watch television, locate your precise coordinates on the planet, and download content from the Internet at rates nearing (or even exceeding) one terabyte per second. All of this multifarious technology—thermal, audiovisual, navigational, computational, etc.—depends, ultimately, on the uniform simple negative charge of all electrons. Yet while I am admittedly not an electrical engineer, I am confident that not a single electrical device currently in manufacture relies on *brute emergence* for its effect. That is, I am confident that in every case, a sensible story of the emergence of like from like can be

[126] Ibid. Strawson dryly adds that "I do not, however, foresee any kind of scientific research programme."

told, in which *something* in the nature of electrons (and the substrates through which they flow or with which they interact) accounts for the light coming from my lamp, the sound coming from my stereo, the calculations being performed by my computer, etc. Nor, I am confident, is it necessary that every electron participating in the function of a given electrical device "contain," in any way, the whole or even any portion of the aforementioned light, sound, calculations, etc. The electrons themselves contain none of these: they just have their uniform simple negative charge (as well, of course, as mass and spin, but neither do I suppose these properties carry the aforementioned light, sound, calculations, etc.). And so somehow—albeit in a way perhaps only intelligible to electrical engineers—it seems as though the flow of countless simple electrons, via proper *combination* and *organization* in one or more substrates, "gives rise to" discrete electrical devices with novel emergent properties entirely predictable (indeed, deliberately designed!) via intelligible (i.e., non-brute) laws of nature.

Where, then, is the combination problem for electrical devices? Do we remain skeptical that each individual negatively charged electron could somehow "add up" to a computational system capable of teraflop speeds (even if precious few of us actually understand the specifics of how such a system "emerges")? We do not, I trust. Partly this is because we defer to the expertise of electrical engineers when it comes to explanations (explanations we, if we are not electrical engineers, likely do not want to hear and would not understand anyway). Partly, though, it is because we see clear evidence around us every day that somehow all this electrical functionality *does* emerge from the flow of innumerable simple electrons through the sockets in our walls, and we (if we are not given to magical thinking) tend to assume that this emergence of discrete, novel properties from countless homogenous ultimates isn't just brute magic, and represents no contradiction in analytic logic. We understand, too, that even though electromagnetism is literally everywhere in the universe, not everything is an electromagnetic *device*. And it seems natural to us that the *fundamental electromagnetic force* underlying all electromagnetic devices needn't itself possess—nor possess in its "constituent" electrons—all the novel emergent properties to be found in electromagnetic devices.

The intended analogy to experientiality should by now be obvious. I must, of course,

concede that at present, we lack the equivalent of electrical engineers; when it comes to consciousness, no one can yet offer us an analog to the ubiquitous elementary school demonstration of a simple electrical circuit. If we are panpsychists (and even if we are not), we may count ourselves as "knowing" that such a demonstration is, in principle, possible, though likely not without presently unimaginable breakthroughs and/or paradigm shifts in science. Or it may be the case, as "mysterians" like McGinn suggest,[127] that we will simply *never* have the equivalent of electrical engineers (experiential engineers?) when it comes to consciousness; our youngsters will never build simple "experiential devices" for their science-fair projects. But whether or not the analogy I offer could be rendered into any sort of research program, the force of it against the combination problem remains untouched—for whether or not we could ever understand *how* the combination and organization of experiential ultimates (or, perhaps, some underlying "experiential force") yields a discrete experiential subject with novel emergent properties (i.e., sense modalities, emotions, etc.), by analogy to electrons, electromagnetism, and electromagnetic devices we can at least appreciate that the possibility of such an intelligible explanation is not, in principle, incoherent.

This analogy also has force against a series of related objections to panpsychism presented by McGinn in response to Strawson. The first is what McGinn calls the *derivation problem*: "how are higher-level experiences derived from lower-level ones?"[128] This objection is a sort of hybrid of the emergence and combination problems (a third horn of the panpsychist's dilemma, one might say), so it bears quoting at length:

> [In the spatial world] we can get a lot of different things by spatially arranging a smallish number of physical primitives. But there is no analogous notion of combination for qualia—there is no analogue for spatial arrangement (you can't put qualia end-to-end). We cannot therefore envisage a small number of experiential primitives yielding a rich variety of phenomenologies; we have to postulate richness all the way down, more or less. An easy way to see this is to note that you can't derive one sort of experience from another: you can't get

[127] McGinn, like most so-called mysterians, does not prefer the term, what with its mystical overtones. Chalmers classifies the mysterian view as one branch of type-C materialism; the other branch, one supposes, might be deemed "temporary mysterians" or, even more derisively, the "scientifically faithful." Chalmers sees both type-C views as inherently unstable, with the faith-in-science view collapsing into type-A Dennett-style eliminative materialism, while McGinn's mysterian view likely collapses into type-F monism (Chalmers 2010, 118–123, 136 f.33).
[128] McGinn in Freeman 2006, 96.

> pains from experiences of colour, or emotions from thoughts, or thoughts from acts of will. There are a large number of phenomenal primitives. Accordingly, we cannot formulate panpsychism in terms of a small number of phenomenal primitives—say, one for each type of elementary particle—and hope to derive the rest. We have to postulate richness at the basis. It would be impossible, say, to begin with simply an array of faint experiences of shades of grey and then hope to derive all of human phenomenology! For the same reason, we cannot suppose that the particles have an alien phenomenology perhaps more suitable to their limited and peculiar 'form of life'… because there is no coherent way to derive from such an alien form of experience the kinds of familiar experiences that we enjoy. To suppose otherwise is to fall victim to the kind of magical thinking that the brute emergentist indulges in; there can be no miraculous transformation of one type of experience into some other quite distinct type…. (and if anyone mentions synaesthesia at this point I will scream).[129]

McGinn insists we not retreat to "faint and blurry qualia"—some vague, infinitesimal analogs to our experience—for "even the faint and blurry is phenomenology too much for the humble electron."[130] Rather, "we can solve the emergence problem only if we credit the ultimates with a rich enough phenomenology to form an adequate basis for a full-bodied human mind, or else we have to suppose input from outside to pump up the volume (and hence relinquish [kosher] emergence)"; therefore, McGinn concludes his attempted *reductio*, "there is really no alternative but to accept that particles have minds in much the same way we (and other animals) do."[131] But the point to be put to McGinn is clear: the "humble electron" is somehow capable, physicSally, of yielding an astonishing array of modalities in physicSal devices—including, for example, the images, sounds, mechanical fan-spinning, and sundry calculations that "emerge" in/from/via my computer. One might even note, to McGinn's chagrin, a certain synesthetic aspect in the way both audio and visual data can be reduced to "modality-neutral" media files, but we needn't push the analogy too far. What is crucial here is that McGinn would presumably never posit that the richness of modalities clearly emergent from the combination and organization of electromagnetic ultimates into/from/via electromagnetic devices requires "richness all the way down"; it would be ludicrous to insist that every complex electromagnetic process involving a given electron must somehow be fully *present* in that electron. The sound and picture emerging from the electromagnetic device in front of me

[129] McGinn in Freeman 2006, 95–96.
[130] McGinn in Freeman 2006, 95.
[131] Ibid.

is, I take it, a naturally emergent property of that device by virtue of the combined and organized operation of electrons that do not, each one in itself individually, bear either the sound or the picture. Likewise, there seems no compelling reason to accept McGinn's insistence that panpsychism requires "full-bodied" minds all the way down. If *physicSally non-audiovisual* electrons can be harnessed to yield either an auditory or visual physicSal result—which we perceive and phenomenally experience as either a sound or an image—there seems no reason that *phenomenally non-audiovisual* electrons could not be harnessed to yield either an auditory or visual phenomenal result (i.e., our very experience of that sound or image).

Nor is it clear that it matters whether one can "put qualia end-to-end"; can one really put even the aforementioned functions of my computer "end-to-end," much less all permutations of electromagnetic activity "end-to-end"? We may grant McGinn's assertion that "you can't get pains from experiences of colour," but likewise you can't get word processing from artificial light—yet both, indubitably, are emergent properties of certain electromagnetic devices whose very *status* as functional electromagnetic devices depends upon the simple charge of electrons and, ultimately, the fundamental force of electromagnetism. You don't need to get word processing from artificial light to get both from the proper electromagnetic combinations and organizations, nor of course do you need word processing and artificial light to be "full-bodied" properties of any individual electron or of the electromagnetic force.

What the electromagnetic analogy leaves vague, I admit, are the criteria by which a given arrangement of fundamentally experiential matter (particles, say, with an "experiential charge" of 1, –1, or neutral) is properly to be understood as an "experiential device." We might not seem to have any example in the offing of what even a simple "experiential circuit" might look like (although I am tempted to side with Douglas Hofstadter in seeing the phenomenon of video feedback as a potential model, if not outright instantiation, of such a circuit[132]). But the point, again, is (1) that McGinn's talk of "phenomenal primitives" may well be misguided, and (2) that his insistence that "the honest

[132] See Douglas Hofstadter, *I Am a Strange Loop* (New York: Basic Books, 2007). For a striking demonstration of the effect, see www.youtube.com/watch?v=sfmAf7CMLp4. Consider also, in this context, Uriah Kriegel's "self-representational" theory of mind.

panpsychist"[133] must commit to "full-bodied" (i.e. human) phenomenology in the ultimates needn't be taken seriously. Indeed, the very notion of "phenomenal primitives" seems to stand in contrast to a certain reading of "full-bodied" phenomenology; one needn't invoke synesthesia to recognize the complexity of, e.g., taste phenomenology, which involves a combination not only of five (or more) senses of taste but the indispensible and fiendishly holistic olfactory sense(s), with texture and temperature making tactile contributions as well. The taste of a Granny Smith apple is not just the *taste* of a Granny Smith apple, but also its smell and, to at least some extent, its texture and temperature. McGinn's insistence on "phenomenological primitives" is intended to do justice to the many facets of human conscious experience, but it threatens to do the opposite by ignoring the crucially holistic aspect of phenomenology that is, after all, the very heart of the combination problem! Furthermore, we must ask: given that flavor, smell, and tactile phenomenology can and do *combine* in our experience to yield the "taste" of a Granny Smith apple, do we really find ourselves, as McGinn claims, entirely without a meaningful "notion of combination for qualia"? It seems to me that the combination of qualia, even and *especially* of different modalities, makes for the better part of the "richness" of conscious experience as we know it.[134]

Still, this leaves two important questions in the vicinity, beyond our acknowledged ignorance of any science of "experiential engineering." The first emerges from the holistic aspect of phenomenology discussed above, which again is the defining conundrum of the combination problem for panpsychists and "Dennettians" alike: how does the sense of being a *whole conscious entity*—rather than (1) of being a composite of (un)conscious subparts or (2) of not even "realizing" one is such a composite—emerge in us? Again, I can only appeal to some sort of "experiential engineering" on the part of Mother Nature; whatever conditions must be met to instantiate an "experiential circuit," they were apparently reached and subsequently manipulated to great effect. A circuit is, after all, a closed loop, and thus perhaps any "experiential circuit" would *just be* an "experiential device"—a.k.a. a conscious entity, along the lines of Chalmers's thermostat. Or perhaps the combination and organization of simple individual "experiential circuits"

[133] McGinn in Freeman 2006, 95.
[134] See "beef and wine" section of Skrbina 2009.

might be necessary to yield a true conscious entity. Either way, if there are experiential laws analogous to electromagnetic laws, and if these laws entail that at least one "experiential loop" must be closed to constitute a conscious entity, this would seem to account for the heart of the mind/body problem: our intuitive sense of "standing apart" from the "outside world" (which, assuming panpsychism, is both phenomenal and physical—and of which, assuming just science, we know ourselves to, despite appearances, in fact be nothing more than a part). If being an experiential device—a conscious entity—entails being a closed experiential circuit ("insulated," as it were, from the experiential "current" of others), then the infamous *problem of other minds* begins to look at least theoretically tractable. That having been said, I do not presume here to resolve the many epistemological and ontological riddles surrounding, e.g., indexicality and attitudes *de se*; if the analogy from electromagnetism merely persuades the reader to reconsider the combination problem's allegedly crippling threat to panpsychism, I will have achieved my ambitions.

But as mentioned, there is a second consideration left somewhat open in the analogy from electromagnetism, specifically the *phenomenal nature of the primitives*. Of course, for many, that this issue should even be considered is a self-evident *reductio* of panpsychism; the "incredulous stare" tends to be accompanied with a rhetorical question along the lines of, "Does that mean there is *something it is like* to be an electron?" The motivation behind this question—to portray panpsychism as essentially *crazy*—might be called the *argument from imagination*. That is, we are supposed to find it utterly unimaginable that there could be something it is like to be an electron. How could anything that small, that simple, that primitive, possibly constitute any sort of locus for phenomenal experience? Intuitively, the notion of conscious electrons just doesn't seem to make sense, and it seems a fair-minded panpsychist needs to acknowledge this *prima facie* difficulty. Chalmers does just that in his discussion of thermostat (as opposed to human) phenomenology:

> We should really expect something much simpler, for which there is no analog in our experience. We will likely be unable to sympathetically imagine these experiences any better than a blind person can imagine sight, or than a human can imagine what it is like to be a bat; but we can at least intellectually know

> something about their basic structure.[135]

Beyond this table-turning *embrace* of unimaginability, however, I believe there are valid moves open to the panpsychist when presented with the argument from imagination. First and foremost, the panpsychist should be sure the rules of the dialectical game are fair and consistent; specifically, the panpsychist might ask whether (or to what extent) the objector can imagine the uncontroversial *physicality* of an electron. I suspect that for most of us, the physical smallness of an electron is *every bit as unimaginable* as its proposed "phenomenological smallness." Speaking for myself, as I stare at the glass of water on my desk, I have serious doubts as to whether I can really properly imagine the scale of an individual water molecule—an object orders of magnitude larger than, say, an individual electron. It seems uncontroversial to say that none of us would even *suspect* the existence of physical entities as small as electrons had we not learned (directly or indirectly) of them through science. But because we've been told electrons exist—because some very smart people *discovered* that electrons exist—we blithely assume we can *imagine* the physicality of an electron. I'm not so sure we can; I'm not even sure physicists who "work with electrons," as it were, can *properly imagine* their physicality in the relevant sense. Try it: try imagining the smallest object you can possibly imagine, and I'll bet you won't even come close to properly imagining the true physical *infinitesimality* of an electron (or even an atom, or even a water molecule). Now, this is not to say that one cannot have a perfectly valid set of calculations demonstrating precisely how small an electron is—but just as there is a big difference between having a calculation that yields ∞ and actually *imagining infinity*, or between having a calculation that yields i or $-i$ and actually *imagining the square root of –1*, there is a big difference between having calculations that imply (or experiments that verify) the existence of electrons and actually *imagining an electron*. (Likewise, while most of us are roughly aware of calculations regarding the size of the universe, or the number of stars therein, I suspect precious few can actually, properly *imagine* how big and full of stars the universe *really is*, to say nothing of its non-Euclidean "finite but unbounded" multidimensional geometry.)

[135] Chalmers 1996, 294.

Note that so far, I've spoken only of imagining the electron as a determinate object, a "Newtonian" particle, which it most certainly is not; to properly imagine the physicality of an electron, we need not only to scale down our imagination by vast orders of magnitude (so vast as to themselves be arguably unimaginable), but must also wrap our minds around the reality of wave/particle duality. So again, try imagining the smallest object you can imagine—except that the "object" isn't really even an "object" so much as a *presence*, a wavelike smear of matter. If you (or anyone) should happen to possess such expansive powers of imagination, it should hardly be a problem to imagine human conscious experience "scaled down" by equivalent orders of magnitude—but more likely, you are like me, and simply *cannot think that small*. Stubenberg makes a similar point in his *SEP* entry on panpsychism: "After all, the effects of gravitation are invisible at the level of extremely small sizes and masses but this does not mean that gravitation is insignificant in the universe, nor that it is not a ubiquitous and fundamental feature of the world, of which every existing thing partakes."[136]

The upshot of this is that the argument from imagination appears to rely on a rather arbitrary double standard. It is only, I contend, because science *tells us* things as small as electrons physically do exist that we assume we *can* properly imagine the physicality of a electrons—but we cannot, not really, not *properly*. What force, then, is the argument from imagination supposed to have against panpsychism? "If you can't even properly imagine the *physicality* of an electron," the panpsychist might ask the skeptic, "why should your inability (or just plain unwillingness?) to properly imagine the *experientiality* of an electron matter a whit here?"

Still, we should prefer *some* sort of positive account of the phenomenology of the ultimates. Merely saying the ultimates are "potentially experiential," as panprotopsychism would imply, simply will not do, at least on a Strawsonian account—or, as McGinn acerbically notes:

> [P]lease don't say that the particles are only required to be *potentially* experience-endowed for panpsychism to be true, since this is common ground for *any* view of the relation between experience and the wider world—*of course*

[136] Stubenberg 2010, section 5.

> matter must have the potential to generate mind, since it patently does (unless we are radical dualists). The whole question is, in virtue of what sort of property—and the honest panpsychist at least has a nontrivial answer, viz. experiential properties. The potentiality move simply says that particles produce minds when combined into brains, and hence have that potential; but that is not a *theory* at all, just the datum we are trying to explain.[137]

What's more, McGinn insists, the retreat to "watered-down… faint and blurry qualia, along the lines possibly of those in the nascent mind of a foetus" is a "weaselly line."[138] Certainly, it would seem at odds with the arguments against fading, dancing, and absent qualia marshaled by Chalmers,[139] which not only support his own type-F "nonreductive functionalist" pan(proto)psychism but offer, we have suggested, a much-needed adjunct to Strawson's general panpsychist framework. Let us not, then, take the weaselly line. Now, for reasons already discussed, we needn't postulate full phenomenal richness all the way down; the ultimates needn't be credited with recognizable human consciousness. But instead of postulating that phenomenology "fades out" as one moves down the scale of complexity, we might consider that the simplest, most fundamental phenomenology is a sort of "white out." In other words, perhaps experientiality in its most basic form is modally, intentionally, and informationally *undifferentiated*: not watered-down at all, but rather a super-saturated ur-experientiality. Consider what it would be like to experience *everything at once*: hot and cold, rising and falling, quiet and loud, light and dark, wet and dry, happy and sad, etc., *ad infinitum*. Not, again, that we wish to claim the humble electron actually feels any of *those* specific things; the point, rather, is to get a sense for what "experiencing everything at once" actually implies: an utter *undifferentiation* that is, in a certain sense, the antithesis of consciousness as we know it. To experience every experiential state *E* simultaneously with every experiential state *not-E* would be to have *no useful phenomenal content* concerning either oneself or one's environment. It would be a senseless seething stasis of phenomenology—but phenomenology it would be!

If we think of our putative experiential ultimates this way, the natural emergence of our own modally, intentionally, and informationally *differentiated* consciousness, with all its dynamic richness, seems more sensible than if we posit watered-down qualia (or, per the

[137] McGinn in Freeman 2006, 95–96.
[138] McGinn in Freeman 2006, 95.
[139] See Chalmers 1996, chapter 7, 247–275.

anti-panpsychists, none at all) in the ultimates. That is, it might be the case that complex phenomenology (of, e.g., the human sort) requires not only experiential *addition* and *combination*, but also experiential *subtraction* and *differentiation*. Perhaps differentiated experiences—simple at first, but evolving in tandem with an organism's complexity—emerge from undifferentiated ones. McGinn would likely protest that it simply *can't* be the case that our sense modalities could be derived via differentiation from some pan-modal, synesthetic ur-experience; he seems quite resolute in his claim that "there are a large number of phenomenal primitives." But perhaps he's wrong: perhaps the apparent differences of kind are actually differences of degree. On such a view, there is only one relevant type of experience, and one needn't invoke synesthesia to see it. We don't find radical differences of kind in the brain's gray matter, nor in matter in general (for the most part)—so why posit that experiences alone come in such ontologically distinct flavors? Why *can't* there be ur-experientiality in the electron? Such an explanation is, of course, entirely speculative, and perhaps McGinn would still find it weaselly, but I think it shows we can posit a sort of phenomenal "richness" (overabundance, really) in the ultimates without needing to *anthropomorphize* electrons, quarks, etc.

Needless to say, even if this speculative account of ur-experientiality happened to be on target, many details would need to be worked out, not least the "laws of experiential differentiation." And while Strawson might happen to be right that experience is just electromagnetism from the inside, Chalmers is more inclined to think of consciousness as a fundamental force distinct from any of the known fundamental forces (though, of course, amenable to assimilation with them into a Grand Unified Theory). So the ur-experience could be that of an electron, but it might be the case that we're dealing with a different sort of "particle" here: a *qualion*, perhaps.[140] This suggestion, it must be said, brings us within spitting distance of type-D substance dualism, but certainly nothing supernatural or "homuncular" is being posited here. But again, this is all speculation—and again, there is no shame in just admitting the unimaginability (though not the inconceivability!) of, e.g., an electron's phenomenology. Though I confess I do find the extended analogy from electromagnetism compelling, I cannot disagree with Chalmers's

[140] Chalmers prefers "psychon" (see Chalmers 2010, 137 f.34), but given his qualms about the "psyche" part of "panpsychism," I'd think "qualion" would be the better choice.

assessment that "the combination problem is easily the most serious problem for the type-F monist view, and at this point, it is an open question whether or not it can be solved."[141] I also, however, agree with Chalmers that "this is an area that deserves much close attention,"[142] and it is to that end that I offer the above speculations, whatever use they may be.

Ultimately, as with all the most tantalizing theories in philosophy, it seems the world would look the same whether or not panpsychism (or panprotopsychism) were true. This leads to a final frequently-aired objection to the view: that it is not a verifiable and/or falsifiable hypothesis. I cannot refute this, but I'm not sure how much force it holds in a philosophical context. However, there is one scientifically credible way in which the presence of consciousness might at least be inferred: in the collapse of the quantum wave-function (i.e., the collapse of quantum superposition into determinate, "Newtonian" reality). As noted by Chalmers in his discussion of type-D dualism, while there do exist "alternative interpretations of quantum mechanics on which there are no collapses or on which measurement has no special role in collapse," the "standard formulation" holds that "collapses occur only… on measurement"—and while "there is no widely agreed definition of what a measurement is… there is one sort of event that everyone agrees is a measurement: observation by a conscious observer."[143] Chalmers has his own informed speculations about whether the causal role of conscious observation might be empirically verifiable, but I would prefer to direct the reader's attention to recent discoveries regarding photosynthesis made by scientists at the U.S. Department of Energy's Lawrence Berkeley National Laboratory. The evidence suggests that the staggeringly efficient energy conversion involved in photosynthesis is the result of the manipulation of quantum forces on the part of plants (and other plant-like photosynthetic organisms).[144] As the Berkeley press release states:

> Electronic spectroscopy measurements made on a femtosecond (millionths of a

[141] Chalmers 2010, 136.
[142] Ibid.
[143] Chalmers 2010, 127.
[144] Gregory S. Engel et al., "Evidence for Wavelike Energy Transfer Through Quantum Coherence in Photosynthetic Systems," *Nature* 446: 782–786, www.nature.com/nature/journal/v446/n7137/abs/nature05678.html.

> billionth of a second) time-scale showed… oscillations meeting and interfering constructively, forming wavelike motions of energy (superposition states) that can explore all potential energy pathways simultaneously and reversibly, meaning they can retreat from wrong pathways with no penalty.[145]

In short, as I understand it, the proteins involved in photosynthesis can "match" the wave-like superposition of incoming photons, such that wherever a photon ultimately collapses into a particle, a protein will *already be there*, so to speak. But one wonders about the nature of the collapse of these superpositional states: if indeed consciousness is essential to superpositional collapse, and if indeed photosynthesis involves the formation and collapse of superpositional states, then the natural conclusion would seem to be that photosynthesis is, in some fundamental sense, a *conscious process*. This doesn't get us all the way to panpsychism, nor even to biological panpsychism (lacking any implication that plants are conscious *qua* plants), but it should at least get us (re)thinking.[146]

6. Closing Remarks

My attempt in this essay has not been to answer every objection levied against panpsychism and other type-F monist views, nor even every objection to the specific supporting arguments of Strawson and Chalmers. Such an effort would be Herculean if not Sisyphean, given the remarkable amount of discussion these two philosophers have generated (some might say *provoked*). My primary objective here has been to give a general lay of the land: to draw out the substance of Strawson and Chalmers's respective pan(proto)psychist positions—which turn out to be strikingly different and even metaphysically incompatible—and to attempt to evaluate their relative merits. My own sense is that Strawson-style panpsychism is more viable than Chalmers-style panprotopsychism (or, for that matter, any form of biological panpsychism), although I believe Chalmers's informational and organizational focus will nevertheless be indispensible if we are to have any hope of developing a robust panpsychist theory. I do not, however, presume to pass judgment on whose story is ultimately true (and I hasten to remind the reader that Chalmers remains seemingly quite amenable to Strawsonian

[145] See "Quantum Secrets of Photosynthesis Revealed," *Research News: Berkeley Lab*, April 12, 2007, www.lbl.gov/Science-Articles/Archive/PBD-quantum-secrets.html.

[146] One might also note that if mere algae can actively manipulate quantum forces, there is no immediately obvious reason that a human brain could not, as some have suggested, be a quantum device as well [DEFOOTNOTE AND/OR MENTION PENROSE-HAMEROFF ORCH OR BY NAME?].

panpsychism anyway).

For those constitutionally disinclined toward panpsychism, this whole comparative analysis will surely seem pointless and fantastical. Although I have attempted to provide thorough and straight-shooting responses to the most common objections to panpsychism, I harbor no delusions about converting the incredulous. I would hope, though, that the philosophical legitimacy of panpsychism might be bolstered by "internal debate" of the sort I offer here. A hashing out of the long-promised details might indicate to skeptics that panpsychism needn't be some inchoate, borderline-religious just-so story of last resort: there *is* substance here, and a far greater ratio of method to madness than has often been assumed. And for those already sympathetic to panpsychism (or any type-F monist view), I hope the foregoing will serve not only as a reminder of the work left to be done in developing the theory, but as encouragement that real progress can, at least in principle, be made.