

The Scientific Evidence For Materialism About Pains

Draft 2-4-13

ABSTRACT (added 5-5-14)

This paper argues in unprecedented empirical and philosophical detail that, given only what science has discovered about pain, we should prefer the materialist hypothesis that pains are purely material over the dualist hypothesis that they are immaterial. The empirical findings cited provide strong evidence for the thesis of empirical supervenience: that to every sort of introspectible change over time in pains, or variation among pains at a time, there corresponds *in fact* a certain sort of simultaneous neural change over time, or variation at a time. The empirical supervenience of pain on the neural is shown in turn to favor the hypothesis that pains are, in a sense that is made precise, purely material.

I

Philosophical discussions of the mind-body problem have often taken pain as their leading example of a phenomenally conscious mental state (see, e.g., Kripke 1980). In this paper, I discuss the implications for the mind-body problem of what science has to say about pain—and I aspire to do so in a way that is accessible to interested non-philosophers.

Science has clearly taught us much about the *etiology* of pain. It has taught us that there are several different types of specialized nociceptive neurons, distinct from tactile sensors and proprioceptors, that are sensitive to noxious stimuli of different kinds, e.g., to thermal, mechanical, or chemical stimuli (Price 1999, 76-79). It has taught us that bundles of these neurons run first to the spine, where they synapse with neurons that then run, along several distinct pathways, to various regions of the brain, some of which project further neurons to the cortex (Price 1999, 98-107). It has taught us that the presentation of a noxious stimulus reliably activates various distinct cortical areas—primary somatosensory cortex, secondary somatosensory cortex and its vicinity in the parietal operculum, insular cortex, anterior cingulate cortex, and prefrontal cortex (see, e.g., McMahon and Koltzenburg, 128).

Less clear but still—I claim—true is that the science of pain bears on the *nature* of pain, i.e., on the question whether pain sensations are—in some sense that needs to be made precise—immaterial rather than purely material states. My main conclusion in this paper is that what science has discovered about pain favors the hypothesis that pains are purely material over the rival hypothesis that they are immaterial; that is, given *only* the evidence that science has discovered, we should prefer the materialist hypothesis over the dualist (i.e., immaterialist) one.¹

¹ I do not argue against idealism in this paper, though it is, of course, an alternative to both materialism and dualism.

This conclusion is a modest one. It is consistent with the claim that, given *all* relevant rational considerations, we should *not* prefer the materialist hypothesis over the dualist one. For it may be that what science has discovered about pain does not exhaust the relevant rational considerations. It may be, as many philosophers think, that materialism can be refuted a priori, or by appeal to what is conceivable, or by appeal to what introspection tells us about the nature of pain. And it may be that such non-scientific considerations support dualism about pain strongly enough to outweigh the evidence against it provided by scientific discoveries. I strongly doubt that in fact they do, but it is beyond the scope of this paper to explain why.²

Does anyone actually *deny* that what science has discovered about pain favors the materialist hypothesis about pains over the dualist one? Some dualists may deny it. At any rate, I do not recall ever having read a dualist acknowledge that science at least makes it *appear* that pain is purely material; and some dualists resort rather readily to sociological explanation of the popularity of materialism among philosophers and scientists, as if the existence of apparent evidence for materialism could not explain it. Some pain scientists do not deny but still hesitate to affirm the main conclusion of this paper. For they characterize the relationship between pains and neural states in strikingly imprecise and non-committal terms, as if reluctant to assert definite materialist theses. Consider, for a representative example, a recent journal article that speaks of “brain areas *involved in* pain processing”, and “the neural *basis of* pain processing” (Schweinhart and Bushnell

² Criticism of arguments for dualism may be found in (Hill 1991, 19-43; Hill 2009, 100-127; Papineau 2002, 47-95; and Melnyk 2001).

2010, 3788; italics added); the article makes no attempt to sharpen the meanings of the italicized expressions. The main conclusion of this paper will presumably meet no resistance from materialist philosophers, but the way I reach it is controversial. The materialist philosophers Christopher Hill and Brian McLaughlin have argued for a conclusion similar to mine about sensations of all kinds (Hill 1991; McLaughlin 2010). But the way in which scientific findings favor materialism over dualism is not the same on their account as on mine—and my way has a few advantages, as we shall see.

It will throw the key issue raised by this paper into sharp relief to consider at length the epistemic position of *Aliens*. Aliens are non-human cognitive neuroscientists who want to know what sensations of (human) pain are, and who have access to everything that (human) science has discovered about pain, as well as to the testimony of pain-feeling humans who report on their own pains and what they are like. What distinguishes Aliens from human scientists is that they are so different from humans physiologically and phylogenetically (let's say they evolved from a distinct origin of life) that they hold—rightly—that nothing they know about their *own* mental lives through introspection (assuming they know anything at all about their own mental lives through introspection) has any evidential bearing on the nature of the internal states that the English-speaking objects of their inquiry call “pains”. Consequently, anything that Aliens come to think about what (human) pain sensations are must be evidenced *solely* by what (human) science has

discovered about pain, including what humans report about the occurrence and nature of their own pains.

Give these restrictions, what should Aliens conclude about the nature of human pains? Perhaps nothing; perhaps the science of pain has merely discovered information about the causal chain that typically begins with a noxious stimulus and ends with a sensation of pain, while leaving it open what the final link in this chain—pain—actually is, e.g., whether it is a neural state of some kind or some sort of immaterial state. I claim, however, that Aliens should conclude that pains are purely material (in a sense to be explained) rather than immaterial. First I argue that the science of pain has discovered evidence favoring the hypothesis that pains are purely material over the rival hypothesis that they are immaterial. Then I argue that the science of pain has discovered no evidence favoring the dualist hypothesis over the materialist hypothesis. Hence the totality of evidence available to Aliens favors the materialist hypothesis. The first argument occupies section III and most of section IV; the second occupies the balance of section IV. Section V infers from the conclusion about Aliens some morals about human inquirers. Section II is devoted to an essential preliminary.

II

We need to be more precise about what materialism and dualism amount to.³ For present purposes, materialism is best understood as the view that every mental state that a human can be in is *purely material* in the sense that it meets one of the following two conditions:

(1) it is identical with—one and the same thing as—some or other uncontroversially material state that a human can be in (e.g., a neural state);

(2) it is identical with—one and the same thing as—some or other *higher-order* state that a human can be in, *and* every actual case of a particular human's being in that higher-order state is *realized* by the human's being in some or other uncontroversially material state (e.g., a neural state).⁴

The expressions “higher-order” and “realized” that appear in (2) are philosophical terms of art. I understand them as follows. A *higher-order* state of a thing is a special sort of state that a thing can be in—a state such that the thing's being in that state just is the thing's being in any (lower-order) state that meets a certain condition, e.g., that plays a particular causal or functional role in relation to other

³ Most of what I say in this section about the formulation of materialism and dualism is a simplified version of the position I defend at length elsewhere (Melnyk 2003, 1-70). The position is, of course, controversial.

⁴ In claiming that a mental state is purely material if and only if it meets one of these two conditions, I set the bar for pain's being purely material higher than do Kenneth Sufka and Michael Lynch, who claim that pain is purely material if it “naturally supervenes on a distinct neural subsystem” (Sufka and Lynch 2000, 311).

(lower-order) states.⁵ And a thing's being in a particular higher-order state is *realized* by the thing's being in a particular (lower-order) state iff the (lower-order) state is one of those that meet the relevant condition for the higher-order state in question. If there is more than one (lower-order) state that can meet the relevant condition, the higher-order state in question is said to be *multiply realizable*.

Why is meeting condition (2) logically sufficient for materialism? If a thing happens to be in one of the (lower-order) material states that meet the relevant condition for a certain higher-order state, then the thing absolutely must be in that higher-order state.⁶ And if the high-order state is identical with—one and the same thing as—a certain mental state, then the thing absolutely must be in that mental state. It follows that, if the thing is in the (lower-order) material state in question, it absolutely must be in the mental state in question—which is a way of saying that there is no more to the thing's being in the mental state (on this occasion) than its being in that (lower-order) material state. Or, more colorfully, if God put you into the material state, he wouldn't need to do anything else to bring it about that you were in the mental state.

⁵ Two notes on terminology. (1) What I here call a “higher-order” state is usually called a “functional” state in the philosophical literature. (2) In this section, when I speak of a “mental state”, I mean something that multiple persons, or a single person at multiple times, can be *in*, i.e., what the literature calls a “mental state-type”. But there is also a usage of “mental states” in which a person's mental states are not the states (i.e., state-types) that she is in, but rather those particular states of affairs each of which consists of her being in some or other mental state (i.e., state-type); such particular states of affairs are called “mental state-tokens” in the literature.

⁶ Given the actual laws of nature. I shall take this qualification as read from now on.

A simple illustration of these very abstract ideas is provided by the state (that a gun can be in) of being *loaded*. A gun's being loaded is plausibly regarded as a higher-order state of the gun, specifically, as being one and the same as the gun's being in some or other state such that, if the gun's trigger is pulled, the gun rapidly emits a projectile. A particular gun's being loaded on a particular occasion is realized by its having a certain complex constitution and construction. But because different guns are made of different materials, and constructed on different principles, being loaded is a multiply realizable higher-order state. In consequence, being loaded is not identical with—one and the same thing as—any kind of (first-order) material state. But materialism about guns is still true, because every particular gun's being loaded is in fact realized by some particular material state of the gun.

In exactly the same way, pain is purely material if it turns out to be identical with a specific neural state, e.g., with the state of undergoing such-and-such activity in so-and-so parts of the pain-feeler's primary somatosensory cortex. But it's *also* purely material if, even though it isn't identical with a particular neural state, it turns out to be one and the same as a specific higher-order state, *and* every pain-feeler's being in that higher-order state turns out to be realized by the pain-feeler's being in some or other material state. The material state might be a human neural state, or a neural state but one that's different from any *human* neural state that realizes pain, or an electronic state of some microprocessor in a pain-feeling robot, or something else again. What might the higher-order state be? A natural—though

imprecise—hypothesis would be this: to be an organism that is in pain = to be a system containing a subsystem which (i) has the function of detecting imminent or actual damage to the system and of getting the system to respond appropriately, where appropriate response includes preventing the damage from occurring or making it stop and letting the system recover, and which (ii) is activated in such-and-such ways. But I only mention this hypothesis for the sake of giving a relatively concrete illustration.⁷

Two final points about the formulation of materialism. First, there is a difference between (i) saying that every mental state is identical with some or other material state, or materially-realized higher-order state, and (ii) actually specifying, for each mental state, *which* material state, or materially-realized higher-order state, that mental state is. The formulation of materialism given above does the former, but it does not aspire to do the latter. The second point is epistemological. That a certain mental state turns out to be identical with a certain material state, or materially-realizable higher-order state, is not something that we should expect to be discoverable by any a priori means, e.g., by reflecting on the meanings of the words (concepts) we use to talk about (think about) mental states. Rather, such identity claims must be discovered empirically, as was the identity of water with H₂O or of genes with segments of the DNA molecule: hypotheses that this kind of thing is one and the same as that kind of thing must be proposed and then tested against the evidence. A corollary is that materialism itself—which asserts the

⁷ Colin Klein has proposed an interestingly detailed and more plausible hypothesis (Klein 2007).

holding of an identity claim for every mental state—has the status of an empirical hypothesis, albeit one of great generality. It is analogous to the uncontroversially scientific hypothesis that every kind of atom is identical with some or other kind of microphysical structure composed of protons, electrons, and (in nearly all cases) neutrons.

Dualism can now be formulated as the view that every mental state—or perhaps every mental state of a certain sort—is *immaterial* in the sense that it meets *neither* of the two conditions above: it is neither identical (1) with any uncontroversially material state nor (2) with any higher-order state every particular instance of which is realized by some or other uncontroversially material state.

III

The science of pain has discovered certain remarkable correspondences between (i) changes in pains over time (or variation among pains at a time), as revealed by the introspective reports of experimental subjects, and (ii) changes over time (or variations at a time) in the subjects' simultaneous neural states, as revealed by various imaging techniques. In this section, I describe five such correspondences in some detail and then argue that they support quite a strong claim to the effect that pain depends on the neural. Part of the reason for going into detail is to show that the evidence I allege for materialism results from genuine empirical discovery, and not just the influence of materialist presuppositions; part is to show, by means

of extensive quotation, that I am not interpreting the science tendentiously; and part is to show how surprisingly fine-grained the neural dependence of pain is.

First, however, a brief preliminary. Pain researchers Price, Barrell, and Rainville (2002) report that

Psychophysical observers [i.e., the subjects in pain experiments] can be trained to detect very small differences in sensory qualities and intensities and to differentially judge magnitudes of different dimensions or qualities of their experience... (600)

One such difference that observers can be trained to detect is—perhaps surprisingly—between the *intensity* and the *unpleasantness* of the pains they are undergoing. Anecdotal evidence for this distinction is the case of a man, reported by Ploner, Freund, and Schnitzler (1999), who had a cortical lesion that altered his capacity to feel pain in the left hand when it was subjected to noxious thermal stimuli:

For left hand, up to an intensity of 600 mJ, no pain sensation could be elicited. However, at intensities of 350 mJ and more, the patient spontaneously described a ‘clearly unpleasant’ intensity dependent feeling emerging from an ill-localized and extended area ‘somewhere between fingertips and

shoulder', that he wanted to avoid. The fully cooperative and eloquent patient was completely unable to further describe quality, localization and intensity of the perceived stimulus. Suggestions from a given word list containing 'warm', 'hot', 'cold', 'touch', 'burning', 'pinprick-like', 'slight pain', 'moderate pain' and 'intense pain' were denied... (213)

The subject of the experiment is apparently reporting a sensation with some of the *affective* features of pain but without its usual *sensory* features. More systematic evidence is provided by two studies, Rainville *et al.* (1997) and Hofbauer *et al.* (2001), in which hypnotic suggestion was used to alter the reported *unpleasantness* of experimental subjects' pains without at the same time altering the pains' reported *intensity*—and conversely to alter the reported *intensity* of the pains without at the same time altering their reported *unpleasantness*.

Now for the first correspondence. Bushnell *et al.* (1999) report finding “that [somatosensory cortex region] S1 activation is modulated by cognitive manipulations that alter perceived pain intensity” (7709). Similarly, Hofbauer *et al.* (2001) report that

[a]s shown in previous brain imaging studies, noxious thermal stimuli presented during the alert and hypnosis-control conditions reliably activated contralateral structures, including primary somatosensory cortex (S1), secondary somatosensory cortex (S2), anterior cingulate cortex, and insular

cortex. *Hypnotic modulation of the intensity of the pain sensation led to significant changes in pain-evoked activity within S1... .* (402; my emphasis)

Corresponding to differences in the *felt intensity of different pains*, then, are differences in the simultaneous level of pain-evoked activity in S1.

A second correspondence concerns change in the felt intensity of a *single* pain over time. Porro *et al.* (1998) provide the following abstract of their study of “the time course of perceived pain intensity and the activity of discrete cortical populations during noxious somatic stimulation lasting several minutes” (3316):

We used a high-resolution functional magnetic resonance imaging (fMRI) technique in healthy right-handed volunteers to *demonstrate cortical areas displaying changes of activity significantly related to the time profile of the perceived intensity of experimental somatic pain over the course of several minutes*. Twenty-four subjects (ascorbic acid group) received a subcutaneous injection of a dilute ascorbic acid solution into the dorsum of one foot, inducing prolonged burning pain (peak pain intensity on a 0–100 scale: 48 ± 3 , mean \pm SE; duration: 11.9 ± 0.8 min). fMRI data sets were continuously acquired for ~ 20 min, beginning 5 min before and lasting 15 min after the onset of stimulation, from two sagittal planes on the medial hemispheric wall contralateral to the stimulated site, including the cingulate cortex and the putative foot representation area of the primary

somatosensory cortex (SI). *Neural clusters whose fMRI signal time courses were positively or negatively correlated ($P < 0.0005$) with the individual pain intensity curve were identified by cross-correlation statistics in all 24 volunteers. The spatial extent of the identified clusters was linearly related ($P < 0.0001$) to peak pain intensity.* Regional analyses showed that positively correlated clusters were present in the majority of subjects in SI, cingulate, motor, and premotor cortex. Negative correlations were found predominantly in medial parietal, perigenual cingulate, and medial prefrontal regions. To test whether these neural changes were due to aspecific arousal or emotional reactions, related either to anticipation or presence of pain, fMRI experiments were performed with the same protocol in two additional groups of volunteers, subjected either to subcutaneous saline injection (saline: $n = 16$), inducing mild short-lasting pain (peak pain intensity 23 ± 4 ; duration 2.8 ± 0.6 min) or to nonnoxious mechanical stimulation of the skin (controls: $n = 16$) at the same body site. Subjects did not know in advance which stimulus would occur. The spatial extent of neural clusters whose signal time courses were positively or negatively correlated with the mean pain intensity curve of subjects injected with ascorbic acid was significantly larger ($P < 0.001$) in the ascorbic acid group than both saline and controls, suggesting that the observed responses were specifically related to pain intensity and duration. These findings reveal distributed cortical systems, including parietal areas as well as cingulate and frontal regions, involved in dynamic encoding of pain intensity over time, a process of great biological

and clinical relevance. (3312; my emphasis)

Corresponding to changes in the felt intensity over time of a single pain, then, are simultaneous changes in the intensity of neural activity in certain regions (S1, cingulate, motor, and premotor cortex) of a pain-feeler's brain.

A third correspondence concerns the introspectible difference between the intensity and the unpleasantness of a single pain at a given time. As we saw, in two studies hypnotic suggestion was used to alter the reported unpleasantness of experimental subjects' pains without at the same time altering their reported intensity—and *vice versa*. Even more fascinating is what was revealed by these studies to happen in the subjects' brains as the reported unpleasantness and the reported intensity of their pains were modified independently of one another (Rainville *et al.* 1997; Hofbauer *et al.* 2001). In the first study, changes in the reported *unpleasantness* of pain (with no change in the reported *intensity* of pain) were correlated with changes in the level of activation in the anterior cingulate cortex, though there was no change in the level of activation in S1. In the second study, changes in the reported *intensity* of pain (with no change in the reported *unpleasantness* of pain) were correlated with changes in the level of activation in S1, though there was no change in the level of activation in the anterior cingulate cortex activation. As Hofbauer *et al.* put it,

This double dissociation of cortical modulation indicates a relative specialization of the sensory and the classical limbic cortical areas in the processing of the sensory and affective dimensions of pain. (402)

So, corresponding to changes in the introspectible sensory and affective dimensions of a single pain are, respectively, simultaneous changes in neural activity in two distinct regions of the brain, with changes in the affective dimension of a pain apparently varying with levels of activation in “the classical limbic cortical areas”.

A fourth correspondence concerns the phenomenon known as *first pain* and *second pain*, which can be characterized as follows (Ploner *et al.* 2002):

... single painful stimuli yield two successive and qualitatively distinct sensations referred to as first and second pain sensation [citation omitted]. First pain is brief, pricking, and well localized, whereas second pain is longer-lasting, burning, and less well localized. (12444)

A single painful stimulus yields two successive sensations because of the different arrival times of impulses from two different kinds of nociceptive nerve fibers—A δ fibers and C fibers—which differ in their myelination and hence conductance. But what happens when the impulses arrive at the cortex? Ploner *et al.* report as follows:

We...used magnetoencephalography to record and directly compare first and second pain-related cortical responses to cutaneous laser stimuli in humans. [...] Cortical activity was located in primary (S1) and secondary (S2) somatosensory cortices and anterior cingulate cortex. Time courses of activations disclosed that *first pain was particularly related to activation of S1 whereas second pain was closely related to anterior cingulate cortex activation*. Both sensations were associated with S2 activation. (12444; my emphasis)

In the early time window, the time courses of activations show significant activation of S1, bilateral S2, and ACC reflecting A δ fiber-mediated and first pain-related activation of these areas. In the late time window bilateral S2 and ACC show strong activations, whereas no significant activation is seen in S1 indicating C fiber-mediated and second pain-related activation of bilateral S2 and ACC but not of S1. (12446)

Corresponding to the introspectible differences between first pain and second pain, then, is a difference in the locations of simultaneous cortical activity.

A fifth correspondence concerns the *felt locations* that pains typically have. Various studies (e.g., Andersson *et al.* 1997; Bingel *et al.* 2004) have shown that S1 exhibits somatotopic organization: roughly, neighboring neurons in S1 can be traced back, via synaptic connections, to neighboring nociceptive neurons in the body. But although these studies aimed to reveal a correspondence between the *actual*

locations of noxious stimuli and activity in specific regions of S1, rather than between the *felt* locations of the pains caused by the stimuli and activity in specific regions of S1, they nonetheless support the existence of the latter correspondence. For in both studies the subjects did in fact report on their pains, and had the felt locations they reported not coincided with the actual locations of the noxious stimuli, the researchers would surely have noted the fact, which they did not do. I conclude that corresponding to the different felt locations of pains is simultaneous neural activity in different regions of S1.

Each of these five correspondences is an instance in which corresponding to a certain sort of introspectible change over time in pains, or variation among pains at a time, there is a certain sort of simultaneous neural change over time, or neural variation at a time. Moreover, despite extensive studies of the kinds cited above, no sort of introspectible change over time in pains, or variation among pains at a time, has been discovered to which there *fails* to correspond some sort of simultaneous change in, or variation among, neural states, even though discovering such failures of correspondence is readily conceivable and lies within our current observational abilities. But observed positive instances of a universal generalization (with no observed negative instances) provide inductive evidence that the universal generalization is true.⁸ Therefore, the five correspondences noted above provide inductive evidence for the conclusion that to *every* sort of introspectible change over

⁸ This inductive principle fails in certain well-known pathological cases that give rise to paradoxes of confirmation such as Nelson Goodman's "grue" paradox and Hempel's paradox of the ravens. But there is no reason to think that the present case is pathological.

time in pains, or variation among pains at a time, there corresponds a certain sort of simultaneous neural change over time, or variation at a time. Let us call this conclusion the *empirical supervenience claim*, since it says that *in fact* no sort of introspectible change or variation in pains occurs without a certain sort of simultaneous change or variation in neural state.

The empirical supervenience claim belongs to a family of empirical claims to the effect that the mental depends on the neural. It is worth comparing the strength of the empirical supervenience claim with that of other members of the family. It is obviously much stronger than the claim that, if one is ever in pain, then one has a properly functioning brain, for the same is true of a properly functioning circulatory system. It is stronger too than the claim that, if one is in pain (no matter of what sort), then one is in so-and-so neural state, for the latter claim leaves open the possibility that pains with *different* introspectible characters do *not* require being in *different* kinds of neural state. The empirical supervenience claim excludes this possibility, since it says that change or variation in the introspectible character of pain *never* occurs without a certain sort of neural change or variation.

On the other hand, the empirical supervenience claim is weaker than at least one of its kin. Christopher Hill and Brian McLaughlin's argument for materialism about sensations rests on a certain empirical claim that they call the "correlation thesis" (Hill 1991; McLaughlin 2010). McLaughlin formulates the thesis as follows:

For any type of state of phenomenal consciousness C there is a type of physical state P such that it is true and counterfactual supporting that a being is in C if and only if the being is in P (McLaughlin 2010, 267)

Even if the scope of the correlation thesis is limited to conscious states involved in pain, it goes well beyond what the empirical supervenience claim affirms, since it claims that for each such state there is a physical state that is not just necessary but also sufficient for the conscious state. The empirical supervenience claim affirms only that a certain sort of simultaneous neural change (or variation) is necessary for each sort of change (or variation) in pain. A corollary of this difference in logical strength is that it is much easier to come up with evidence for the empirical supervenience claim than for the Hill-McLaughlin correlation thesis. Indeed, evidencing the correlation thesis would seem to require discovering the so-called “neural correlates” of at least some conscious mental states. That doing so has proved to be problematic is presumably why neither Hill nor McLaughlin actually asserts the correlation thesis; they present their argument for materialism as *conditional* on our discovering in the future that the correlation thesis is true.

IV

What, then, is the evidential significance for Aliens of the empirical supervenience claim? The first point to note is that materialism—as formulated in section II—is consistent with and indeed predicts the empirical supervenience

claim. It is worth spelling out why. If materialism is true, then there are just two sensible possibilities regarding pain:

(1) Pain is one and the same as a certain complex neural state. The different kinds of pain (e.g., pain in the right foot, intensely unpleasant pain in the right foot, intensely unpleasant pain in the right foot of duration 18 seconds) are one and the same as certain more specific neural states.

(2) Pain is one and the same as a certain higher-order state, such that every particular pain-feeler's being in that higher-order state is realized by the pain-feeler's being in some or other neural state. The different kinds of pain are one and the same as certain more specific kinds of higher-order state, such that every particular pain-feeler's being in a higher-order state of any of these more specific kinds is realized by the pain-feeler's being in some or other more specific neural state.

Suppose that the first possibility is actual, and that kinds of pain just *are* kinds of neural state. Then obviously there must be a certain sort of neural change or variation corresponding to, and simultaneous with, each sort of change over time in pains (or variation at a time among pains). Nothing at all can possibly change or vary without itself changing or varying.

Suppose now that the second possibility is actual. The empirical supervenience claim must be true in this case too, though seeing why is harder. Consider Jan, whom we may safely assume to be a biologically normal human whose extra-cranial neuronal wiring doesn't spontaneously change in any significant way over the short term; the data we're trying to accommodate concern precisely such humans.⁹ Suppose that Jan is in pain of kind P at time t, and her being in pain of kind P at t is *realized* by her being in complex neural state N at t. By the definition of "realized" in section II, P must be identical with a certain higher-order state H, and N must meet the relevant condition for H; that is, N must meet condition C, where to be in H = to be in one of the states that meet C. Since Jan is in N, and since N meets C, Jan must be in H. Further, since H = P, she must be in P. So, since Jan is in N, she *must* be in P. But now imagine that Jan *ceases* to be in P; perhaps she enters a pain state of a different kind, or perhaps she stops being in pain of any kind. How can this happen? The physico-chemical laws governing neural states cannot change, and H cannot cease to be identical with P, since that would require that H cease to be identical with itself. So the only way in which Jan can cease to be in P is for her to undergo a change of neural state, i.e., to cease to be in N. The upshot, then, is that, if Jan ceases to be in P, then she undergoes some simultaneous change of neural state. And this upshot may be generalized: given possibility (2), *every* sort of introspectible diachronic change in pains is accompanied by a certain sort of simultaneous diachronic neural change. Finally, analogous reasoning shows that

⁹ The assumption means that we can disregard the possibility that Jan's state of pain can change simply because of a change in circumstances *external* to her brain; such a possibility exists if pain turns out to be an essentially *representational* state whose content is determined in part by circumstances external to the subject's brain.

possibility (2) also entails that the other half of the empirical supervenience claim is true, i.e., that every sort of introspectible variation among pains *at* a time is accompanied by a certain sort of simultaneous neural variation. The key point is that, if Jen's twin brother, Jon, were also in N, then he would have to be in P too; so if he is not also in P, he cannot be in N.

What about dualism? Consider, first, a version of dualism that regards pain as a state of the immaterial mind, that treats the immaterial mind as receiving sensory input from the brain and emitting motor instructions for the brain to execute, but that treats the brain as unnecessary for mentality *except* insofar as it is needed to send sensory input to the mind and to execute the mind's motor instructions. On this version of dualism, sensory states, including pain, can change without there being any corresponding simultaneous neural changes. Such a view is falsified by the empirical supervenience claim.

But there are, of course, dualist views which allow and indeed predict the empirical supervenience claim. One such dualist view regards pain as a state of the immaterial mind, but a state that the mind is *caused* to enter by a certain *simultaneous* neural state of the subject. Another such dualist view treats pain as an *immaterial* state that the subject's *brain* is caused to enter by a certain simultaneous *neural* state; it therefore assumes that the brain can instantiate immaterial properties.¹⁰ A third and a fourth dualist view can be formed by replacing the

¹⁰ It seems to be David Chalmers' positive view (Chalmers 1996).

appeal to simultaneous neural-to-mental *causation* in the two preceding views with an appeal to non-causal neural-to-mental *laws of association*.

Since there are versions of dualism that, like materialism, entail and are consistent with the empirical supervenience claim, it is tempting to conclude that the empirical supervenience claim cannot possibly favor materialism over the relevant versions of dualism. And this conclusion is correct if a naïve Popperian falsificationism is true according to which the *only* regulative role for evidence is to contradict hypotheses; for the empirical supervenience claim does *not* contradict the relevant versions of dualism. Indeed, I conjecture that the widespread endorsement of naïve Popperian falsificationism by scientists is what explains why, as noted in section I, pain scientists are reluctant to treat their findings as favoring materialism over dualism.¹¹ The conclusion that the empirical supervenience claim cannot possibly favor materialism over the relevant versions of dualism is *also* correct if an extreme form of empiricism is true according to which the empirical accuracy of competing hypotheses is the *only* feature relevant to their relative evidential status; for materialism and the relevant versions of dualism are *equally* accurate empirically. And I conjecture that the possibly tacit assumption of this form of empiricism is what explains why, as noted in section I, few dualists seem willing to acknowledge that science at least makes it appear that pain is purely material.

¹¹ One need only read the first chapter of half a dozen college textbooks to see the pervasive influence of naïve Popperian falsificationism in science.

But naïve Popperian falsificationism and the extreme form of empiricism in question are both open to serious objections (Newton-Smith 1981, Ch. 3; Laudan 1995). One especially important objection is that both views lead to an unacceptable skepticism. The crux is that, for pretty much *any* hypothesis that we presently favor, and *any* evidence that the hypothesis entails and is consistent with, we can concoct a *rival* hypothesis that entails and is consistent with the very same evidence. So, for example, the hypothesis that the universe is only ten minutes old can be formulated to be consistent with all the evidence usually taken to support the conventional view; likewise, of course, for creationism. And if either naïve Popperian falsificationism or the extreme form of empiricism in question is true, it follows that we have no evidential grounds for preferring the original hypothesis over the concocted one. For, by assumption, the concocted hypothesis is not contradicted by any of the evidence, and it is just as accurate empirically as the original hypothesis.

To avoid such skepticism, we need to allow that factors other than empirical accuracy can be relevant to the comparative evidential merits of competing hypotheses, such factors as parsimony and fit with background knowledge (these factors are often called “super-empirical criteria”). We can do so in a way adequate for present purposes by adopting the following principle of evidence:

Evidence E favors hypothesis H1 over hypothesis H2 if

- each of H1 and H2 entails and is consistent with E;

- H1 fares better than H2 on at least one super-empirical criterion;
- and
- H2 does not fare better than H1 on *any* super-empirical criterion.

This principle plausibly implies that the standard evidence favors the conventional view of the age of the universe over the ten-minute hypothesis, since the latter is so spectacularly unparsimonious, and fits so poorly with background knowledge, in comparison with the former. The principle could be embedded in different overall accounts of evidence and theory-choice, e.g., in a Bayesian account in which the initial prior probability of a hypothesis is assessed by reference to super-empirical criteria, or in Philip Kitcher's eliminativist account in which that hypothesis is preferred which is consistent with all the evidence and which achieves this consistency at the lowest cost in terms of the super-empirical criteria (Kitcher 1993, 237ff.).

I will, of course, use the principle to argue that the empirical supervenience claim favors materialism about pains over the relevant versions of dualism about pains. Arguing in this way enables me to sidestep an objection that Jaegwon Kim has made to the Hill-McLaughlin argument for materialism about sensations (Kim 2005, Ch. 5). The Hill-McLaughlin argument uses the principle of inference to the best explanation, whereby the *explanatory power* of a hypothesis counts in favor of its truth: they argue that their correlation thesis (see above) is better explained on the hypothesis that sensations are purely material than on any rival hypothesis

saying that they are immaterial. Kim's objection is that the materialist hypothesis is an *identity* claim (which, as we saw in section II, it is), and that identity claims never have genuine explanatory power. Using the principle above, however, I need not assume that identity claims have explanatory power.

Here is why the empirical supervenience claim favors materialism about pains over the relevant versions of dualism about pains. Both materialism and the relevant versions of dualism entail and are logically consistent with the empirical supervenience claim, as we have seen. But materialism fares better than the relevant versions of dualism on the super-empirical criteria of (i) parsimony and (ii) fit with background knowledge,¹² while the relevant versions of dualism do not fare better than materialism on any such criterion. With regard to parsimony, it's true that both materialism and these versions of dualism are committed to the neural states corresponding to the various kinds of pain *and* to the various kinds of pain themselves; but what this commitment comes to is not the same in each case.

According to any dualist view, even after God has created Jen in neural state N, and instituted the right physico-chemical laws, he has to do further work to put Jen into the state of pain P to which N corresponds; P is no ontological free lunch. According to materialism, however, once God has created Jen in neural state N and instituted the right physico-chemical laws, he has nothing further to do, for whether it turns out that $P = N$ or that $P = H$, if Jen is in N and the physico-chemical laws hold, she

¹² And perhaps also on the criterion of *explanatory power*. The versions of dualism in question certainly *entail* the empirical supervenience claim, but to entail something is not necessarily to *explain* it. Perhaps an empirical generalization is not explained, or explained only poorly, by saying merely that it holds as a matter of law.

must be in P. So the versions of dualism in question are less parsimonious than materialism, because they treat mental states as fundamental, non-neural states, rather than as real states that are, however, nothing over and above neural states.

The versions of dualism in question are less parsimonious than materialism in a second way too. As we have just seen, materialism can reductively explain the empirical supervenience thesis *without* taking it to reflect the holding of a myriad of irreducible neural-to-mental laws over and above the standard physico-chemical laws that ultimately govern neural states. However, the versions of dualism in question take mental states to be immaterial states, and so, since no physico-chemical laws govern immaterial states, they cannot reductively explain the empirical supervenience thesis, and therefore *must* take it to reflect the holding of a myriad of irreducible neural-to-mental laws. Like materialism, therefore, the versions of dualism in question are committed to the holding of the standard physico-chemical laws that ultimately govern neural states; but, unlike materialism, they are *also* committed to the myriad of irreducible neural-to-mental laws. That is what makes them less parsimonious than materialism.

Turning to fit with background knowledge (of Aliens), consider the thesis that all states *other* than mental states are ultimately realized by physical states, i.e., states describable in the proprietary vocabulary of physics.¹³ Such a claim is a commonplace among scientists, and in fact accepted by nearly all contemporary

¹³ To mental states should be added any sociological states partially constituted by mental states.

dualists—unsurprisingly, in light of the strong evidence for it.¹⁴ Now materialism about the mind coheres well with this thesis, since, if mental states are identical with, or realized by, neural states, then mental states turn out to be just like all other kinds of state in being ultimately realized by physical states. By contrast, however, all forms of dualism cohere poorly with the thesis, since, if mental states were immaterial, then they could not turn out to be ultimately realized by physical states; mental states would be the sole exceptions to an otherwise exceptionless generalization.

Let us now ask whether the relevant versions of dualism fare better than materialism on any of the super-empirical criteria. Apparently they do not. There is no reason to think it is part of the background knowledge of Aliens that all or most sensations *other* than pain are immaterial. And Aliens are so different from humans that they hold—rightly—that everything they know through introspection about their *own* mental lives is irrelevant to the nature of human mental states.

My argument that the empirical supervenience claim favors materialism about pains over the relevant versions of dualism about pains has assumed the legitimacy of appealing to super-empirical criteria in assessing the comparative evidential merits of rival hypotheses. And for some readers, despite my earlier remarks, this assumption will be unacceptable. I ask them to reconsider. Since the assumption raises fundamental issues about the justification of induction, I cannot

¹⁴ See Melnyk (2003, Ch. 6) for some of this evidence.

adequately defend it here; but I should add two brief remarks.¹⁵ First, the appeal to super-empirical criteria to discriminate between empirically equivalent hypotheses is pervasive in both science and everyday life. So we face a trilemma. We must do one of the following:

- (i) accept the appeal as legitimate;
- (ii) candidly acknowledge that we aren't really warranted in thinking that, for example, the universe is more than ten minutes old; or
- (iii) provide an alternative account of empirical evidence that vindicates our rejection of the ten-minute hypothesis but without appealing to super-empirical criteria.

I say that option (i) looks pretty good when compared to options (ii) and (iii); (ii) is preposterous, and no one knows how to do (iii).

Second, it is true that the appeal to, say, parsimony can be part of a reliable mode of inductive inference only if the world itself is parsimonious to some degree, i.e., correctly describable by theories that are parsimonious. It is true too that there is no a priori deductive guarantee that the world is parsimonious, and that any attempted inductive proof that it is would be circular. But so what? The points I have just conceded precisely mimic Hume's argument for skepticism about enumerative induction; the role played here by the parsimony of the world is played

¹⁵ For a fuller defense, see Melnyk (2003, 244-251).

in Hume's argument by the resemblance of the future to the past. But although philosophers disagree as to *where* precisely Hume's argument for skepticism about enumerative induction goes wrong, they nearly all agree that it must go wrong somewhere, and they do *not* recommend the abandonment of enumerative induction.¹⁶ I suggest that by parity of reasoning we should take exactly the same attitude as this toward skepticism about appeals to parsimony. Skeptics about appeals to parsimony almost invariably assume that such appeals are problematic in a way in which enumerative induction is not; but I see no basis for that assumption.

My conclusion thus far, then, is that, for Aliens, the empirical supervenience claim favors materialism about pains over dualism about pains. Let us now ask whether science has discovered anything about pains that points the other way, i.e., that favors dualism over materialism. The answer is that it has not. There are two kinds of evidence that, if discovered, would favor dualism over materialism, but neither kind—to the best of my knowledge—has actually been discovered. First, the discovery of changes over time in pains, or variations among pains at a time, that do *not* correspond to simultaneous changes in, or variations among, neural states would be irresistible evidence for dualism; but, as we noted in section III, no such changes or variations have been discovered. Second, a pain that is, or is realized by, a neural state must owe all its causal powers to that neural state. So there would be irresistible evidence for dualism if some of the known behavioral or neural effects of pains were found to be such that the neural states that were the best candidates for

¹⁶ I endorse James Van Cleve's line on Hume's inductive skepticism (Van Cleve 1984).

being identical with, or for realizing, pains were incapable in principle of causing the effects. The discovery of such effects would show that the physical was not *causally closed*, i.e., that there were physical effects for which there did not exist a sufficient physical cause. But I am unaware of any claims that such effects of pain have been discovered empirically.

V

The conclusion of sections III and IV is that, for Aliens, what science has discovered about pain favors materialism over dualism about pains. And this entails that, *for us*, what science has discovered about pain favors materialism over dualism about pains. For we are neither inferior nor superior to Aliens in our ability to appreciate the evidential force of what science has discovered about pain. It may be true that introspection can provide us humans (but not Aliens) with reason to think that pains are immaterial; I take no stand on the matter here. But even if it is true, and introspective considerations favor dualism over materialism, it does not contradict the claim that what science has discovered about pain favors materialism over dualism. Indeed, there is no contradiction even if introspective considerations favor dualism over materialism so strongly that, all things considered, we ought to prefer dualism. As noted in section I, the main conclusion of this paper is a modest one.

It might be objected that, if introspective considerations could favor dualism over materialism in the sense of raising the probability of dualism about pains to *one*, then a scientific discovery could not lower dualism's probability at all, and so presumably could not favor materialism over dualism.¹⁷ I reply that introspection could never generate such considerations, since any case for dualism based directly or indirectly on introspection, even if it raised the probability of dualism very high, could never raise it to one; in the real world of non-ideal reasoners, there is always the possibility that any such case might be defective in some hitherto unrecognized way. Proper defense of this fallibilist view would take us deep into epistemology, so I will say no more. But if it is correct, then the conclusion of sections III and IV does indeed entail that, for us, what science has discovered about pain favors materialism over dualism about pains.

The conclusion of this paper is modest in a second way too. As we saw in section II, there are, at the highest level of abstraction, just two ways in which pain could turn out to be purely material: it could turn out to be a certain kind of complex neural state or it could turn out to be a certain kind of neurally-realized higher-order state. But the conclusion of this paper is neutral between the two options; arguing for it has not required choosing between them. This is noteworthy, because, though choosing between them has traditionally been a major issue in the philosophy of mind, we may well be unable to do so on the basis of empirical

¹⁷ If the probability of dualism on introspective considerations is one, then the new prior probability of dualism, formed by conditionalization on this posterior probability, is also one; and, according to the Bayesian account of evidence, no evidence can lower the probability of a hypothesis whose prior probability is one.

evidence. Crucial evidence for the second option over the first would be the discovery that pain is multiply realized. But if pain is uniformly (i.e., non-multiply) realized in humans, which seems plausible, evidence of multiple realization would have to come from non-human animals: it would have to be evidence that certain animals (i) are in pain, but (ii) are not in any (human-type) neural state plausibly identifiable with pain. But what evidence could show that they are in pain? They obviously cannot tell us, and any behavior analogous to human pain-behavior might well be deemed inconclusive. That they were in a (human-type) neural state plausibly identifiable with pain would certainly be evidence that they are in pain, but would defeat the attempt to demonstrate multiple realization. That they were in a *higher-order* state plausibly identifiable with pain would also be evidence that they are in pain, but the plausibility of the identification would be challenged as question-begging by the advocates of the neural-state identity theory.

The conclusion of this paper is modest, finally, in that it carries no commitment to any particular specification, for each state of pain, of *which* neural state, or neurally-realized higher-order state, that state of pain is. One might have assumed the opposite, i.e., that, by the reasoning described in section IV, features of pain are (or are realized by) the neural states which correspond to them, e.g., that pain intensity = activity in such-and-such regions of S1, S2, anterior cingulate cortex, and insular cortex. But this assumption would be incorrect. Suppose that a certain sort of change over time in pains corresponds to a certain sort of simultaneous neural change over time. To account for this correspondence by hypothesizing that

pains are purely material states entails that the changing feature of pain in question is, or is realized by, some neural state that *overlaps* the simultaneous neural state; but no stronger claim is entailed.

It is noteworthy that one can provide evidence that pains are purely material without being committed to specific hypotheses as to which kinds of pain or features of pain are (or are realized by) which neural states.¹⁸ At least one other argument that pains are purely material does not have this feature. If the Hill-McLaughlin argument were applied to the particular case of pain, it would draw an inference to the best explanation from the fact that people are in pain when, and only when, they are in so-and-so neural state. The conclusion inferred would be the specific hypothesis that pain = so-and-so neural state.

It is a good thing, at least for materialists, that one can provide evidence that pains are purely material without being committed to specific hypotheses as to which kinds of pain or features of pain are (or are realized by) which neural states. For even if we knew necessary and sufficient neural conditions for every aspect of our state of mind when in pain, and even if we accepted that pains are purely material, it would still not be clear which neural states to treat as identical with (or as realizing) pain. The reason is that, when we are in pain, we are typically in a variety of distinguishable mental states, and it's not clear which of these mental states are parts, and which just concomitants, of being in pain. For example, when

¹⁸ If I understand them correctly, a similar claim is made (on different grounds) by Thomas Polger and Kenneth Sufka (Polger and Sufka 2005, 344).

experiencing the pain caused by a sharp object's penetrating the skin, is the pressure one may feel at the point of contact a part of the pain? What about the sudden anxiety that one might feel? There is no reason to think—and every reason to doubt—that either the everyday term “pain” or the ordinary concept of pain is precise enough to yield definite answers to all questions of this sort. But there is no threat to the thesis that pains are purely material so long as the thesis turns out to be true on every precisification of “pain”.¹⁹

REFERENCES

Andersson, Jesper L.R., Anders Lilja, Per Hartvig, Bengt Långström, Torsten Gordh, Hermann Handwerker, Erik Torebjörk, “Somatotopic organization along the central sulcus, for pain localization in humans, as revealed by positron emission tomography”, *Exp. Brain Res.*, 117, 1997, 192-199.

Bingel, U., J. Lorenz, V. Glauche, R. Knab, J. Gläscher, C. Weiller, and C. Büchela, “Somatotopic organization of human somatosensory cortices for pain: a single trial fMRI study”, *NeuroImage*, 23, 2004, 224–232.

Bushnell, M.C., G. H. Duncan, R. K. Hofbauer, B. Ha, J.-I. Chen, and B. Carrier, “Pain perception: Is there a role for primary somatosensory cortex?”, *PNAS*, 96.14, 1999, 7705-7709.

¹⁹ I am much indebted to an anonymous reader for very helpful comments that led to significant revisions of an earlier draft of this paper.

Chalmers, David, *The Conscious Mind: In Search of a Fundamental Theory*, New York: Oxford University Press, 1996.

Hill, Christopher, *Sensations: A Defense Of Type Materialism*, New York, Cambridge University Press, 1991.

Hill, Christopher, *Consciousness*, New York, Cambridge University Press, 2009.

Hofbauer, Robert K., Pierre Rainville, Gary H. Duncan, and M. Catherine Bushnell, "Cortical Representation of the Sensory Dimension of Pain", *Journal of Neurophysiology*, 86.1, 2001, 402-411.

Kim, Jaegwon, *Physicalism, Or Something Near Enough*, Princeton, Princeton University Press, 2005.

Kitcher, Philip, *The Advancement Of Science*, New York, Oxford University Press, 1993.

Klein, Colin, "An Imperative Theory of Pain", *The Journal of Philosophy*, 104.10, 2007, 517-532.

Kripke, Saul, *Naming and Necessity*, Oxford, Blackwell Publishing, 1980.

Laudan, L., "Damn the Consequences!", *Proceedings and Addresses of the American Philosophical Association*, 69(2), 1995, 27-34.

McLaughlin, Brian P., "Consciousness, Type Physicalism, And Inference To The Best Explanation", *Philosophical Issues*, 20, Philosophy of Mind, 2010, 266-304.

McMahon, Stephen B. and Martin Koltzenburg (eds.) *Wall and Melzack's Textbook of Pain*, 5th ed., Elsevier/Churchill Livingstone, Philadelphia, 2006.

Melnyk, A., "Physicalism Unfalsified: Chalmers' Inconclusive Conceivability Argument", in Carl Gillett and Barry Loewer (eds.) *Physicalism and Its Discontents*, New York, Cambridge University Press, 2001.

Melnyk, A., *A Physicalist Manifesto: Thoroughly Modern Materialism*, New York, Cambridge University Press, 2003.

Newton-Smith, W., *The Rationality of Science*, London, Routledge, 1981.

Papineau, David., *Thinking About Consciousness*, Oxford, Clarendon Press, 2002.

Ploner, M., H. -J. Freund, and A. Schnitzler, "Pain affect without pain sensation in a patient with a postcentral lesion", *Pain*, 81, Issues 1-2, 1999, 211-214.

Ploner, Markus, Joachim Gross, Lars Timmermann, Alfons Schnitzler, "Cortical representation of first and second pain sensation in humans", *PNAS*, 99.19, 2002, 12444-12448.

Polger, Thomas W. and Kenneth J. Sufka, "Closing the Gap on Pain: Mechanism, Theory, and Fit", in *Pain: New Essays On Its Nature And The Methodology Of Its Study*, M. Aydede (ed.), Cambridge, MA, The MIT Press, 2005, 325-350.

Porro, Carlo A., Valentina Cettolo, Maria Pia Francescato and Patrizia Baraldi, "Temporal and Intensity Coding of Pain in Human Cortex", *J. Neurophysiol.*, 80, 1998, 3312-3320.

Price, Donald D., *Psychological Mechanisms of Pain and Analgesia*, Seattle, IASP Press, 1999.

Price, Donald D., James J. Barrell, and Pierre Rainville, "Integrating experiential-phenomenological methods and neuroscience to study neural mechanisms of pain and consciousness", *Consciousness and Cognition* 11, 2002, 593-608.

Rainville, P., G. H. Duncan, D.D. Price, B. Carrier, & M.C. Bushnell, "Pain affect encoded in human anterior cingulate but not somatosensory cortex", *Science*, 277, 1997, 968-971.

Schweinhart, P, Bushnell, MC, "Pain imaging in health and disease—how far have we come?", *J Clin Invest.*, 120.11, 2010, 3788-3797.

Sufka, Kenneth J. & Michael P. Lynch, "Sensations and pain processes", *Philosophical Psychology*, 13.3, 2000, 299-311.

Van Cleve, James., "Reliability, Justification, And The Problem Of Induction", *Midwest Studies in Philosophy*, 9, 1984, 555-568.

Andrew Melnyk

Department of Philosophy

University of Missouri

USA

melnyka@missouri.edu