Naturalism Meets the Personal Level: How Mixed Modelling Flattens the Mind

Robert D. Rupert

I. Introduction and Overview

It is a philosophical truth (almost) universally recognized that a proper understanding of the human mind – and of its relation to cognitive science and neuroscience – rests on a distinction between the personal and subpersonal levels. Appeals to this distinction veritably permeate contemporary work in philosophy of mind and philosophy of cognitive science; they appear frequently in the broader philosophical literature as well, perhaps most notably to epistemology.[1]

Widespread commitment to the importance the personal-subpersonal distinction and to its applicability to humans raises a host of puzzling questions, however. Is the distinction meant to mark a metaphysical divide, a robust boundary between two (relatively) autonomous dimensions or domains of reality? Or, does it signal only a difference between two, noncompeting styles of explanation, both of which play a legitimate role in accounting for the same or closely related phenomena?[2] Questions also arise about authors' purposes when they put the distinction to work, as well as the conditions that must be met in order for those purposes to be well served. Does the

[1] The distinction is introduced in Dennett (1969). Here is a selective sampling of more recent work in which the distinction is endorsed, appealed to, or presupposed: Hurley (1998), Rowlands (2009, 2010), Schroeter (2008), McDowell (1985, 1994), Fridland (2021), Shea (2013, 2018), Miracchi (2017), Levy (2016), Debus (2017), Davies (2000*a*, 2000*b*), Hornsby (1997, 2000), Boghossian (2008), de Vignemont (2014), Lyons (2016), Ismael (2014), Fodor (1987), Cain (1999), Robins (2017), Frankish (2009, 2016), Clark (2015, 2019), Davies and Egan (2013), Holton (2016), Spaulding (2010), Schwitzgebel (2012), Buckner (2018), Wilkinson (2015), Block (2017), Gładziejewski and Miłkowski (2017), Ramsey (2016), Mallon (2016), Mole (2011), Gallagher (2008), and Alsmith and de Vignemont (2012). In some cases, terminology differs without marking any significant difference in content. Burge (2010), for example, shows a marked preference for 'individual level' and 'subindividual level'; but more than once (*ibid.*, 93, 369 n3), he explicitly uses the language of the 'person-level' and the 'subpersonal', and treats this talk as equivalent to his preferred language of individual and subindividual levels (a similar point applies to Huebner 2008, 101).

[2] McDowell (1994) holds that both explanatory schemes are legitimate and not in competition with each other: "This [personal-level talk] is a different metaphor of telling, not in competition with the 'sub-personal' one" (197; McDowell surrounds 'sub-personal' in scare-quotes not because he is skeptical about the idea of a subpersonal level in humans, but rather because he makes a point about the human case by analogy to the case of a frog, to which, on McDowell's view, 'person' has no literal application).

distinction effectively insulate pretheoretical (or otherwise nonscientific) reflection on the mind from empirical research, in cognitive neuroscience, for instance, or in computational cognitive psychology (Davies 2000*a*, 45, Hornsby 2000, n29)? A naturalistically minded philosopher might insist that, if the distinction is to be appealed to effectively in philosophical contexts, the distinction must play a productive role in the empirical enterprise, regardless of whether the distinction is construed metaphysically or as marking a difference only between styles of explanation. Would such insistence be reasonable?

The absence of clear, well-justified answers to many of these questions has, in this author's estimation, impeded the effort to understand human thought and cognitive processing and their relation to behavior. The difficulties so engendered make especial mischief when philosophers with naturalistic leanings turn to topics of longstanding disciplinary interest, such as the nature of the self, mental content, consciousness, and first-person authority; but appeals to the personal-subpersonal distinction also muddy the waters when matters turn topical or applied, toward questions about, for instance, implicit bias or addiction.

In this essay, I argue that progress, at least for the naturalist, lies along the eliminativist path: I contend that there is no personal level – no distinct level, layer, or domain of reality at which the human person or particularly person-related properties reside, and no autonomous (or relatively autonomous) personal-level causal-explanatory vocabulary that earns its empirical keep. Accordingly, I recommend that philosophers abandon the distinction between the personal and subpersonal levels in their discussion of the human mind and related matters, at least if such discussion is meant to respect naturalism of even a moderate sort.

In a more positive vein, I plump for an alternative vision of the human mental architecture, one consistent with elimination of the personal level. On this view, our best

accounts of intelligent behavior are opportunistic and often mixed (Wilson 2010). Individual models of the processes that produce intelligent behavior draw on a variety of resources, including properties, states, and processes commonly associated with the personal level as well as those commonly associated with the subpersonal level. Components of this admixture co-contribute in varying proportions – sometimes interactively, sometimes relatively independently, sometimes modulating the contribution of others – within the context of a mental (or cognitive) architecture that is not segregated into levels.

Trends in cognitive-scientific modeling provide both grist for the essay's critical mill and support for its positive proposal – the picture of a mind "flattened from above." The success of mixed models of intelligent behavior does not depend upon the assignment of various contributors to disjoint groups, as denizens of distinct levels or elements of distinct explanatory vocabularies: differentially placing so-called personal and subpersonal contributors at distinct levels plays no substantive causal-explanatory or theoretical role in the modeling of intelligent human behavior – and thus the critical conclusion. At the same time, extant modeling practices reveal the emerging shape of a mental and cognitive architecture *sans* distinct personal and subpersonal levels.

The remainder emphasizes a methodological moral. Although there may be coherent ways in which to spell out the personal-subpersonal distinction – one can, after all, draw as many distinctions as one likes – we should not take the domain of states and processes associated with the personal level to be sufficiently metaphysically or epistemically isolated as to merit talk of an autonomous (or largely autonomous) area of enquiry, and, in particular, not one about which we can acquire substantive knowledge by conceptual analysis or philosophical reflection alone, uninformed by empirical research into the variety and mixture of causes of intelligent behavior.

There are no sound, principled grounds for bracketing matters to do with so-called subpersonal states and processes when enquiring about the mind or explaining human behavior (or action).[3]

To be clear, eliminativism about the personal *level* does not entail eliminativism about the self, consciousness, or folk psychological states; the flattened view (as I will often write, for ease of exposition) is consistent with these more widely discussed forms of eliminativism but is also consistent with their rejection. It is not unlikely that, in the end, our best models of the processes that produce intelligent behavior will include conscious states, beliefs, and desires;[4] it is simply that those models will not mark such states as appearing at a metaphysically distinctive level of

[3] Rey (2001), Bermúdez (2000), and Drayson (2012, 2014) also challenge prevailing views about the personal-subpersonal distinction. For seeds of the positive view offered here, see Hursthouse (1991) and Gendler (2008), in particular, Hursthouse's rejection of a standard belief-desire-based formula for the explanation of some actions and her acknowledging of contributions that, at least by some standards, are subpersonal and Gendler's provocative description of beliefs and aliefs as contributing "alongside" one another (*ibid*., 635–637); *cf.* the perspective of Schwitzgebel who both is a proponent of a distinct personal level (2010, 2012) and seems sympathetic to a mixed-model view of behavior-production (2010, 540–41).

The Dennett of *Consciousness Explained* (1991) most directly inspires the work done here. Dennett might seem thusly miscast, however, given that he himself introduced the personal-subpersonal distinction (1969, 90) and given the extent to which he seemed committed to its application to humans. Note, though, that Dennett saw personal-level vocabulary as distinctive only of a style of explanation – epistemically autonomous and indispensable, but not indicative of a genuine metaphysical level; in fact, he treated personal-level terms as nonreferring (*ibid*., 96). (For further discussion of the nuances of Dennett's overall view, see Hornsby 2000, Roth 2015, and Wilkinson 2015). Holding that a certain form of talk is merely a style of explanation does not, however, insulate that form of talk from evaluation; I shall argue that, as styles of explanation go, appealing to or observing a personal-subpersonal distinction does not earn its keep.

[4] Much depends on one's view of the semantics of such terms as 'belief', for example, on the extent to which one is willing to countenance revision to our conception of folk psychological states without shifts in the reference of the apposite terms (Stich 1996). If, for instance, it is a necessary condition on beliefs that they appear at a distinctive personal level, then eliminativism about beliefs follows from eliminativism about the personal level. And, if one takes personal-level states to be the distinctive of the mental, then, if successful, my arguments for the flattened view imply that even a relatively modest form of naturalism precludes the existence of minds (*cf.* McDowell 1985, 330). A more likely moral of the story is that minds are not what, for instance, McDowell thinks they are. One can conceive of minds in various ways, and I see nothing privileged about McDowell's conception, particularly if one is committed to figuring out what these minds here – the ones humans have – are, as opposed simply to exploring one's prescientific conception of what minds are. Similar remarks apply to belief. According to McDowell, the idea of a belief is unintelligible unless we have in mind an ideal according to which a given belief is situated in a deductive network of an infinity of rationally connected states (1985, section 3). It would seem impossible, though, for a human to have such a network definitely in mind (as opposed to representing a summary description of such a network), and, regardless, a naturalist might reasonably doubt that states fitting the description of McDowell's ideal play a causal-explanatory role in the modeling of human behavior. Thus, one might rightly doubt the utility of McDowell's conception of beliefs and be willing, instead, to consider alternatives (such as states that have many other properties associated with beliefs and appear in mixed models, regardless of their not being situated in a deductive network of an infinity of rationally connected states).

nature (or as part of an epistemically autonomous, naturalistically successful level of explanation). There is no such thing as the personal *level*, on this view, although the states traditionally associated with the personal level – biographical memories, for instance – may well exist and contribute to the production of intelligent human behavior.

What, the reader might wonder, becomes of the terminology 'personal' and 'subpersonal' upon adoption of the flattened view? On one gloss, what have commonly been characterized as personal-level states and processes are, in fact, more subpersonal states and processes; for, what were commonly thought of as personal-level states are, with regard to their role in the causal-explanatory enterprise and their epistemic status, of a piece with those that have typically been labeled 'subpersonal'. On this view, *all* mental and cognitive states and processes turn out to be subpersonal.[5] One might, however, just as well take the flattened view to entail that neither a personal nor a subpersonal level exists and thus, presumably, that philosophers should abandon all use of 'personal' and 'subpersonal' (as applied to humans – take this as read hereafter). After all, upon rejection of the personal level, no contrast remains for the terminology to mark, nothing for the so-called subpersonal to stand below. I favor the latter way of understanding the flattened view's implications, partly on grounds of a general skepticism about levels. Nevertheless, the substance of the view matters more than the resolution of this terminological question.

At various turns in what follows, I appeal to trends in cognitive science's efforts to account for intelligent human behavior, broadly construed,[6] and to trends in philosophy of

[5] To be clear, I do not contend that there are no distinctions between the various states and processes represented by our best models of intelligent behavior. I contend only that the differences do not support a distinction between levels. The causal properties of gases differ from those of solids, but that hardly entails the existence of distinct solid and gaseous levels of nature.

[6] I presuppose only a weak notion of intelligent behavior, treating intelligent human behavior, first and foremost, as deviation from the null hypothesis – in particular, deviations that are of a striking and distinctive enough sort as to be grouped together for the purposes of cognitive-scientific enquiry, but the grouping of which is tentative and subject to revision as the causal-explanatory project proceeds, as is typical in the sciences.

science as it pertains to levels. Philosophers who embrace a thoroughgoing anti-naturalism might thus react to the present work with detached indifference or a simple shrug. What do they care about, for instance, trends in cognitive-scientific modeling? It is fair enough to note this dialectical limitation. One rarely encounters such thoroughgoing anti-naturalism, however, in contemporary philosophy of mind and philosophy of cognitive science. A substantive question to be addressed below, then, is just how weak one's form of naturalism must be in order to preserve a proprietary personal level. I argue that the only form of naturalism likely to vindicate a personal level abandons methodological naturalism entirely; it consists only in a metaphysical naturalism so lacking in teeth that it vindicates virtually any distinction humans care to draw. It seems to follow, then, that many of the philosophers whose work is cited in note 1 occupy an unstable position, by helping themselves to a distinction that their naturalism is not likely to legitimate – for they would not be willing to limit their naturalistic commitment to the nearly vacuous form of metaphysical naturalism in question.

II. The Distinction and the Dialectic

A. The distinction

As it is typically understood, the subpersonal level is the target of scientific research into the subconscious causes of intelligent behavior – as such research is pursued, for example, in cognitive neuroscience, perceptual psychology, cognitive psychology, and psycholinguistics, and particularly in so far as such fields deliver mathematically formulated process-models of the subconscious sequences of states that produce behavior. The precise boundary of the subpersonal level is contested; at what might be considered the lower end, most removed from a folk conception of the mind, perhaps the subpersonal level comprises gastrointestinal representations (Block 2017, 8). Acknowledging room for disagreement about details, the general idea seems to

be that an account of a process (and its component states) that produces intelligent behavior is subpersonal if and only if *(a)* subjects do not have direct or introspective access to it, *(b)* it does not figure into pre-scientific, folk-psychological explanations of human behavior, and *(c)* it is not properly attributable to an organism referred to using a proper name or personal pronoun;[7] perhaps it must also be, in a broad sense, mechanistic.

To illustrate, Zoe Drayson (2014, 342) reviews two of the best-known cases of so-called subpersonal theorizing in the cognitive sciences, drawn from the study of language processing and the study of early visual processing. The nuances of language processing, as they are captured by models of, say, the deployment of an axiomatic grammar, clearly lie beyond the scrutiny of introspection and folk-psychological ken.[8] Human subjects also clearly lack introspective access to the processes involved in the kind of visual processing famously investigated by David Marr (1982) and his colleagues. For example, it is news to common sense and entirely unavailable to introspection that the process begins with sensitivity to zero-crossings in the intensity of light falling on the retina, as detected by the relevant retinal and immediately post-retinal cellular arrays.

What of the personal level? How is it typically characterized, in contrast to the subpersonal? What follows is an attempt at a relatively comprehensive list of properties and state-types supposed to make their home at the personal level. Items on this list might be better

[7] Condition *(c)* is the least central of the three, for reasons briefly commented on below.

[8] Whether such deployment is properly attributable to persons is a less straightforward matter. Subconscious linguistic processing can be attributed to the person in the way in which, from the perspective of surface grammar, any form of so-called subpersonal processing can be attributed to a person – "John is such that his cognitive system encodes and applies an axiomatic grammar sensitive to 'wh'-traces," for instance. Thus, the work of condition *(c)* rests on intuitions about what can *properly* be attributed to the person, which seems to be a nebulous affair, often degenerating into a question-begging insistence on the correctness of one's intuitions about which attributions make sense or are nondeviant. To my mind, we would do best first to determine on independent grounds whether a personal level exists, and then let the chips fall where they may with regard to the proper attribution of what have been thought to be personal and subpersonal properties.

thought of as widely endorsed indicators of a state's (or process's) being at the personal level than as individually necessary and jointly sufficient conditions.

1. personal-level states are conscious (Shea 2013, 1065; McDowell 1994, 196, 203; Burge 2010, 93; Davies 2000*a*, 90; Frankish 2009, 91)[9] and, as such, are the appropriate objects of introspection and introspection-based verbal report;

2. personal-level states are subjective; they present or manifest a first-person perspective (Metzinger 2003, 131) or are states of which the subject experiences ownership;

3. personal-level states are states of whole organisms, rather than of organisms' parts (McDowell 1994, 193, Lyons 2016, 249, Frankish 2009, 90);

[9] Some authors are less committal on this point, holding only that all conscious states are at the personal level. Burge, for example, takes personal-level states commonly to be conscious; but because he takes some subconscious states to be properly attributable to the person, he allows the possibility of unconscious personal-level states (2010, 368, 374, 470 n66), as does Kriegel (2012, 91) at least if we construe conscious states as just those possessing sensuous qualities.

4. personal-level states are states of whole persons, as such, not merely whole organisms (Hornsby 1997, 161*ff.*); [10, 11]

5. personal-level states are properly attributable to persons (Burge 2010, 95, 292 n1, 369; Lyons 2016, 249–250; Frankish 2009, 90);

[10] This seems to be the characteristic at issue when authors attempt to convey the distinctive nature of the personal level by the use of italicized pronouns: "*I* reason about what to believe, not a part of my brain" (Boghossian 2008, 116; for further examples, see McDowell 1994, 193–194, and Rowlands 2010, 141–142, 145, 151, 215–216). That being said, it is sometimes difficult to tell when an author has in mind the organism-oriented reading of 'whole person' or 'whole animal' (e.g., Burge 2010, 103, 189, 272, 333, 373) and when it is a matter of something more distinctively person-related. McDowell (1994, 196) conceives of the personal level (or its nonhuman equivalent) as somehow constituted by a creature's competently inhabiting its environment. It is worth remarking, though, that an interest in the competent inhabiting of an environment – the sort of thing one might study in ecology or ethology (Gallistel et al. 1991) – does not alone yield a personal level or its animal equivalent. It is of decided interest that humans and frogs can function in their environments in ways that strike observers as skilled and successful (say, in the near optimal use, when foraging, of biological resources available); but it is another matter entirely whether our best science of that behavior establishes, presupposes, or posits an autonomous or quasi-autonomous *level* the nature of which supports *a prioristic*, commonsense-based, or introspection-based enquiry into the causes of that behavior.

[11] Many proponents of the personal level are more interested in entry 4 than entry 3 (Hornsby 2000, 9); but the distinction is not always clearly marked (see Rowlands 1997, 288, for passages that either run the two features together or indicate a resistance to a choice between the two). Thus, it is worth making a pair of remarks about 3's relation to 4. First, proponents of a personal level who emphasize 4 seem committed to the claim that, even if the person is something more than or distinct from the organism, personal-level properties cannot be properties of parts of the organism, and thus if a property is obviously a property of a proper part of the organism, it is not a personal-level property, whereas if a property seems to be a whole-organism property, it might or might not be a property of the person. Second, entry 3 cannot reasonably be asked to carry much weight by the proponents of a robust personal level. Cognitive scientists surely do count and track organisms. It is fundamental to the methodology of "running subjects" that experimenters accurately count the distinct bodies participating in the experimental procedure in question (the methodology section reports that "n = …") and that experimenters index various data points (collected in repeated trials, for instance) to the specific organisms that produced them. This practice goes no distance toward showing that, say, experimental psychology or cognitive neuroscience presupposes a personal level of the relevant sort, that is, about which we can know substantive, domain-specific truths *a priori* or by common sense or philosophical reflection alone; from the standpoint of experimental methodology, the cause of behavioral output in any such experiment could just as well as have been a process occurring only in a proper part of the organism as it was a process of the whole organism; and besides, neither possibility would imply the existence of a distinctive level of the epistemic standing in question. Thus, the gross methodology of experimental psychology (i.e., of counting subjects by counting distinct organisms and indexing data points to them) provides no support for 3 (or 4) as the articulation of a property distinctive of a proprietary personal level. This observation suggests that something has gone wrong when Bermúdez (2000, 65) categorizes reaction times as personal-level behaviors, whether he has the organismic reading of 'personal' in mind or has in mind the more fulsome reading associated with entry 4. Such behavior is, in the first instance, neither personal nor subpersonal. It is a measured effect, collected in a specified context, caused by some parts or other of the, or possibly the whole, organism.

6. personal-level states are the states recognized by folk psychology, such states as beliefs, desires, hopes, and fears (and perhaps more – Bermúdez 2000, 65, Frankish 2009, 90–91); the questions addressed at the personal level are ones "we can know the answers to by interacting, as commonsense psychological subjects, with others" (Hornsby 1997, 167);

7. personal-level states rationalize action; they ground intentional explanations of human behavior as, for instance, components of belief-desire pairs the contents of which allow us to make sense of human actions (Shea 2013, 1065; Boghossian 2008, 119; Bermúdez 2000, 69–70; Frankish 2009, 90); understanding such an explanation makes "a person intelligible as rationally motivated" (Hornsby 1997, 161; also see Hornsby 2000, 15).

8. personal-level states appear in a space of reasons, as opposed to a space of mere causes, and, as such, have normative properties or enter into normatively governed relations; they populate the domain of "explanations in which things are made intelligible by being revealed to be, or to approximate to being, as they rationally ought to be" (McDowell 1985/1998, 328);

9. personal-level states are states of rationally coherent agents (Lyons 2016, 251);

10. personal-level states and processes include all of thought, properly understood, particularly states related by genuine inference (Davies 2000*a*);

11. personal-level states are such that attributions or descriptions of them are intensional (Davies 2000*b*, 88);

12. personal-level states have a kind of content that is partly constitutive of conscious experience – so-called phenomenal content (Thompson 2007, 157) – not possessed by subpersonal states.

Although the preceding list may seem unmanageable, the entries form natural, though not necessarily exclusive, clusters. Entries 1, 2, and 12 pertain especially to consciousness. Entries 3, 4, and 5 relate at least partly to questions of scale. Entries 6 and 7 are drawn from folk psychological explanation and the commonsense concept of a person (as, perhaps, refined by philosophical reflection). Entries 7, 8, 9, and 10 pertain to reason, rational thought, and the associated notion of normativity. Entries 11 and 12 focus specifically on content, homing in on what is meant to be a specific kind of content appearing only at the personal level. To be sure, there is something artificial about this clustering exercise, and it threatens to obscure relations among the various characteristics. For instance, if folk-psychological explanation presupposes that a person's actions are to be understood by their relation to the content of the mental states producing them, and if the relations between the contents of those states is essentially normative, there is a natural connection between entries to do with folk psychological states and explanations, on the one hand, and those concerning rationality and normativity, on the other.

Construction of the preceding list serves two primary purposes. First, it conveys a deeper sense of the supposed nature of the personal level. Second, given the depth and the complexity of the list and the interrelations among its items, the list illustrates how much seems to ride on the development of a naturalistically defensible conception of the personal level and, accordingly, how much philosophical work might need to be re-considered, if no such level exists.[12]

[12] To be clear, elimination of the personal level does not necessarily eliminate the properties previously thought to be distinctive of the personal level. Such properties could, instead, be instantiated by some states contributing to the

B. The dialectical use of the distinction

Whether naturalism vindicates a distinction between the personal and subpersonal levels depends partly on the content of such a distinction; for this reason, it has been of some use to list properties associated with two supposed levels. Additionally, though, one should wonder about the grounds of the inference from two contrasting lists of properties to two distinct levels. Statues and train cars have many contrasting properties, but that hardly entails the existence of two distinct levels of reality, a statue level and a train-car level.

Thus, progress on the present issues might seem to require a theory of levels – of their individuation criteria, for instance, and of what determines the ordering of levels, beyond mere lists of contrasting properties.[13] The variety of concepts of levels that one finds in the literature immediately complicates matters, however (Craver 2007, Chapter 5). Consideration of that variety threatens to push the discussion too far afield. Instead, for present purposes, I propose to identify functionally the kind of level in play, by the way it is appealed to dialectically. Of central importance is the treatment of the personal level as sufficiently autonomous (or isolated, or segregated), relative to supposedly subpersonal matters, that the latter can be treated as irrelevant to the prosecution of questions about the former, which questions can be answered productively by focusing exclusively on the supposed personal level. To be clear, the sort of autonomy in question is not typically meant to be absolute; given the widespread acceptance of some brand of physicalism, materialism, or ontological naturalism, it is no surprise that discussions of the personal level frequently acknowledge constraints from the subpersonal level

---

production of intelligent behavior, though not in a way that entails the existence of a proprietary level at which those states appear – perhaps because those states play a role in our best models of the production of behavior alongside a hodge-podge of other co-contributing states, many of which do not have the properties supposedly distinctive of personal-level states and without those models building in or appealing to a distinction between levels.

[13] For a classic discussion, see Oppenheim and Putnam (1958). For an influential alternative, see Wimsatt (1992).

– constraints that become relevant in, for example, cases of breakdown or malfunction. Nevertheless, the personal level is typically thought to be sufficiently autonomous that investigation of many of the deepest and most important questions about the mind can proceed in relative isolation, independent of consideration of the subpersonal level, often in a way that explicitly brackets facts about the subpersonal level that might (as the thinking goes) otherwise erroneously appear to be germane. The present subsection illustrates and challenges this isolationist (or quasi-isolationist – take this as read hereafter) tendency.

At its core, the isolationist strategy involves the shifting of argumentative ground, in the following manner: "One might have thought that consideration R pertains to debates about question Q, but R concerns the subpersonal level and Q concerns personal-level matters; thus R does not pertain to Q and can be ignored." Consider an example, taken from debates in philosophy of perception. Should we, in our account of visual perception, put perception first (in order of logical or causal priority) or experience first (Miracchi 2017)? On the experience-first view, one kind of perceptual state – experience – is common to veridical visual perception, illusion, and hallucination. In contrast to the experience-first view, Miracchi advocates for a perception-first view, according to which "successful perceptual engagement with the world, perceiving things as they are, is metaphysically and explanatorily prior to all other perceptual states" (Miracchi 2017, 629). In order to make a case for her view, which builds on the foundational idea of perceptual competence, Miracchi must clear away (or at least present an appealing alternative to) the leading, naturalistically grounded competitor: representationalism. The representationalist holds that common to, for example, the veridical perception of a waterfall, an illusory seeing of a waterfall, and a hallucination as of a waterfall is an experience with the same representational content, something like *waterfall there*. In rejecting

representationalism (and the experience-first view more generally), Miracchi wants "to make it clear how there might be a view that explains the commonalities between perceiving things as they are and failing to do so without positing a more fundamental mental state in common to the two cases" (*ibid*., 633).

In the course of developing her competence-oriented, perception-first view, Miracchi invokes what are characterized as subpersonal bases for competences. The question arises, then, whether Miracchi offers a genuine alternative to the representationalist approach. After all, the obvious candidates for competence-bases include just the sort of states invoked by representationalists. Miracchi responds, "…the charge that I am surreptitiously appealing to mental representational states is unfounded. I am appealing to the same independently specifiable features of a subject's sub-personal cognitive system, but proposing a different connection between them and mental events" (*ibid*., 657). Miracchi's defense here requires rejecting what she takes to be a common representationalist gloss of vision science: that visual experiences, considered as mental states, are to be (at least token) identified with subpersonal computational states that possess representational content, that is, with the kind of subpersonal state that populates models of visual processing found in Marr's work (Marr 1982). Miracchi protests that, although the basis of a perceptual competence might be common to seeing veridically, illusory seeing, and hallucinating, such bases are merely subpersonal; she does not identify such bases with experiential states, and thus she has not surreptitiously appealed to any common *mental* component shared by all three kinds of state in question. *Sans* the sort of identification in question, the shared states pertain only to the subpersonal level, and thus, their being shared across cases, if they are, is not relevant to the debate in question, which is a debate about personal-level phenomena. Subpersonal states can be appealed to for many purposes – as

the bases of competences, for instance – but, in the absence of cross-level identification, they are not relevant to the central question, *i.e.*, whether there is a personal-level state common to veridical perception, illusion, and hallucination.

Miracchi takes it that there is an important personal-level matter the nature of which is such that claims about the subpersonal level are irrelevant to it; showing that the common state-type in question (involving competence bases) is subpersonal thereby shows that facts about such states have no bearing on matters at hand. This illustrates the dialectical role often played by an appeal to the personal-subpersonal distinction. To be fair, though, we might ask whether considerations distinctive of the subject matter support isolationism in this case. It would appear not, and the reasons why not are instructive.

On Miracchi's view, subpersonal states can drive our engagement with the environment (*ibid.*, 676), but she also includes personal-level *explanans* in her picture. It is difficult, however, to locate cognitive-scientific reasons to think the models that best account for visually guided engagement with the world – models of data collected during such engagement, for instance – contain *explanans* residing at two distinct levels. Miracchi offers examples, if only in passing, of social sciences in which, she claims, personal-level states or capacities serve as *explanans* (*ibid.*, 668). Much more would need to be said, though, to establish that such sciences deal in a personal-subpersonal distinction of a sort that licenses the dialectical move just criticized, as opposed simply to counting organisms and modeling their behavior in ways that do not appeal to a distinctively personal level. (Compare the remarks in notes 10 and 11 above and note 19 below.) It is one thing to observe that psychologists recognize human organisms and that economists talk about rational agents; it is another to show that we best explain the practice of successful experimental psychology or successful micro-economic modeling by invoking a

robust personal-subpersonal distinction, one that supports isolationist methodology in its approach to the supposed personal level.[14]

In contrast to the suggestion that personal-level states and capacities serve as *explanans*, Miracchi sometimes suggests that personal-level perceptual states are *explananda*. She appeals to dynamical systems theory to motivate this thought (*ibid.*, 668–669), comparing (personal-level) perceptual states to such phenomena as convection rolls. This seems to confuse the issue, though. Convection rolls are observable, measurable phenomena to be accounted for or modeled, and in that sense, are *explananda*. Similarly, organisms produce behavior and engage with their environments, which is where we find the opportunity for measurement and the collection of data, and thus a parallel, in perceptual science, to convection rolls. From a naturalistic perspective, distinctively personal-level perceptual events are, in contrast, *explanans*, if they appear anywhere in our best models at all. This reveals a striking disanalogy between cases of convection rolls and personal-level perceptual states. In the case of convection rolls, there are two dimensions to the discussion: the measurable rolls, as *explananda*, and the dynamical-systems-based models of the appearance of the rolls (models that, e.g., identify the temperature-dependent movement of the parts of the fluid in question) as *explanans*. But, in the case of perception, Miracchi's introduces three terms: the behavior (measured by reaction times and the like), subpersonal states or processes, and the supposed personal-level ones. The inclusions of personal-level perceptual states as *explananda* is thus unmotivated. The proposed analogy to

[14] Compare Fodor's dim view of the personal-subpersonal distinction: "There is, obviously, a horribly difficult problem about what determines what a person (as distinct from his body, or parts of his body) did. Many philosophers care terrifically about drawing this distinction ... [B]ut whatever relevance the distinction between states of the organism and states of its nervous system may have for some purposes, there is no reason to suppose that it is relevant to the purposes of cognitive psychology" (1975, 52) How can this be, given that Fodor also takes folk psychology to have significant empirical power and takes computational psychology to vindicate folk psychology as an empirical theory (1987, Chapter 1). The implication would seem to be that the personal-subpersonal distinction, if it is part of folk psychological thinking, is not a part of folk psychology that is vindicated naturalistically, by computational psychology.

convection rolls – which establishes only the importance of an observable, measurable phenomenon and a model of it – does not support the idea that personal-level perceptual states are *explananda* to be modeled or accounted for.

C. Relation between the two levels

This subsection canvasses and evaluates possible relations between the personal and subpersonal levels, particularly as these might establish a legitimate role for a personal level in the causal-explanatory enterprise.

1. *Personal-level capacities as* explananda

One attempt at naturalistic vindication – suggested above in the discussion of Miracchi's work – holds that personal-level capacities or abilities are the given *explananda* of cognitive science. On this view, the job of cognitive science is to discover the processes and mechanisms that implement or enable the appearance of personal-level capacities or abilities (Hornsby 2000, 16, Bermúdez 2000, Burge 2010, Drayson 2012).[15] On this view, introspection, common sense, conceptual analysis, or *a priori* reflection establishes the existence of certain human capacities (for perception, for instance – Burge 2010, 369–370, 536). Cognitive science becomes, as one

[15] Davies, too, holds a view in this vicinity, though he does not emphasize capacities and abilities. He argues (2000, section 4) that a reflective understanding of the personal level – as involving rational, rule-based, conceptual thought – imposes constraints on subpersonal models, in particular, that they have an articulated structure of the sort associated with a Fodorian language of thought (Fodor 1975). To the extent that this is an exercise in explanation – that the experience of structured thought, at the personal level, is best explained by structured processing at the subpersonal level – Davies treats personal-level phenomena as *explananda* of cognitive science.

Note that what might appear to be personal-level capacities come to the fore when we view cognitive science through the lens of homuncular functionalism or functional decomposition. Drayson (2012) makes this case by appeal to the work of a host of influential philosophers of mind working in the 1960s, 1970s, and 1980s – including Dennett, Lycan, Cummins, and Fodor. Whether this approach has the virtues that proponents of the personal level value – in particular, whether it offers nonscientifically established *explananda* for cognitive science – depends on the extent to which one takes the capacities in question to be given *a priori* or by introspection, common sense, or conceptual analysis, in contrast to being, say, things the positing of which is itself grounded in empirical data.

might say, an engineering exercise; its job is to provide a casual-mechanical account of the basis or implementation of those independently established capacities.

As philosophy of science, this picture seems fundamentally misguided. Cognitive science is a science, responsive primarily to data. As Newell, Shaw, and Simon put it, at the dawn of cognitive science, "What questions should a theory of problem solving answer? First, it should predict the performance of a problem solver handling specified tasks" (1958, 151).[16] First and foremost, cognitive science develops causal (or otherwise explanatory) models of behavioral data, as constrained nowadays by such physiologically oriented sources as imaging data, lesion-studies, and interventions that rely on, for instance, transcranial magnetic stimulation. Painting with a broad brush, in such models, relations between quantities instantiated in model elements (such as cells) explain instances of human behavior (the data to be modeled); they do so on the assumption that kinds of entities and quantities appearing in a model correspond to entities and properties in the human that produce the behavioral tokens in question, in the standard causal way (that is, in a horizontal way – for the horizontal-vertical distinction, see Bermúdez 2000, Drayson 2012, Colombo 2013). The purpose of the enterprise is to model the data, which data are not, in the first instance, capacities of individuals, but rather a set of token results or measurement outputs, collected or produced under specified circumstances.

Of course, data do not wear their best model(s) on their face. One should remain open to the possibility that the best model of a given set of data – or, more likely, a collection of sets of data – contains structure that reifies, or is best interpreted as reifying, personal-level capacities. But, that way of understanding the cognitive-scientific enterprise turns the picture under consideration on its head. In that case, personal-level capacities would not be *explananda*; rather,

[16] For a more contemporary discussion, entirely unremarkable in its adoption of the same sort of view, see Botvinick and Cohen (2014, for instance at 1255).

they would enter the scene only as part of the *explanans*; they would have been introduced – explicitly or implicitly – into models of the data in just the way any theoretical entity, force, or relation might be introduced as part of the *explanans*, so as to do causal-explanatory work. That being said, so far as I can tell, reified personal-level capacities are not even *explanans* in the current state of cognitive science.

This is not to deny that talk of human capacities sometimes plays a role in cognitive science. Such talk can help to generate new ideas for experimental designs and can serve as a catalyst for the generation of hypotheses, and it can provide a way to summarize patterns in behavioral data. The serving of such purposes does not, however, confer model-disconfirming power upon independently arrived at claims about personal-level capacities. That is the bottom line. Such power is reserved for data on performance (and relations between competing models of them – here the theoretical virtues may well come into play in the complex ways in which they do in the empirical sciences). I find no evidence that the cognitive scientific community succeeds in its work by choosing otherwise worse models of the data over better models of the data simply or largely because the former better explain (supposedly) independently given facts about personal-level cognitive capacities. In this sense – that is, in the sense in which *explananda* constrain or set adequacy conditions for model selection – personal-level capacities are not the *explananda* of cognitive science.

Consider, for example, one of the most influential Bayesians research programs, that of Tom Griffiths and colleagues (e.g., Griffiths et al. 2010). As they describe their work, *a priori* proofs concerning optimal behavior provide a route to the discovery of the actual mechanisms that produce often messy, non-optimal human behavior. It is, on their view, a more efficient way to guide a search for the accurate mechanistic models of the production of human behavior, a

method more efficient (they are betting) than a bottom-up approach that begins with data about internal processes – for example, neural imaging data – and then attempts to use that data to build models of performance (*ibid.* 358, 362–363). Marr's arguments for the importance of a computational level – which might be identified with a personal level – rest on similar grounds of epistemic efficiency (Marr 1982, 27). If we are to organize the data effectively, in a way that facilitates the construction of successful process models, it has to be given conceptual shape. According to Griffiths and Marr, the most efficient (possibly the only manageable) way to do so is to think in terms of what would be useful or even optimal for the organism to compute. But, the point of such a strategy is not to reify the Bayesian or computational level – to make it an *a priori* truth about human cognition, which is then set up as *explanandum*. Rather, it is a methodological tool for discovering the best models of the data, which likely do not exemplify Bayesian optimality or compute exactly what it would make sense for the organism to compute.

It is also worth noting the many forms of human behavior that seem not obviously subject to fruitful Bayesian methods (of searching the space of possible models more efficiently by making optimality assumptions): priming, depression, mind-wandering, the adjusting of one's motor behavior to that of one's interlocutor, monothematic delusions, development of a certain self-narrative in response to one's sociological surroundings, addiction, the ruffling of the hair of one's child (on through the list of Hursthouse's "arational" actions) and so on – a list inspired by phenomena famously put to a different purpose by Churchland (1981), of dreaming, creativity, and mental illness. This is not to say that Bayesian assumptions have not been applied in attempts to model some of these phenomena, but the jury is out in most cases, and, in the typical case, non-Bayesian empirical assumptions must be added to an otherwise Bayesian model (e.g., Davies and Egan 2013, on monothematic delusions) to get it to fit the data. In other words, the

model consists in more than Bayesian optimality calculations. In many cases, it takes incredible ingenuity to frame these phenomena as exercises of (approximate) Bayesian rationality; it is not plausibly part of our concept of a person or available to introspection. Thus, even if there is some sense in which Bayesian optimality assumptions determine a level of analysis, it is not the sort of level that philosophical proponents of a personal level need, one that delivers the sort of isolationism that allows the setting aside of so-called subpersonal considerations. The best models discovered by the use of Bayesian optimality thinking, as models of human data, contain all manner of subpersonal elements – parameter settings, noise terms, nonoptimal strategies for search, and so on.

To muddy the waters, psychologists sometimes apply the language of levels to matters concerning scale or something akin to a distinction between phenomenal matters and models of the processes that produce phenomena, in ways that do not deliver isolationist methodology to proponents of a personal level. For instance, De Houwer, Gawronski, and Barnes-Holmes (2013) articulate what they describe as two levels of theorizing in social psychology, which they call the 'functional' and the 'cognitive' levels. On their view, social psychology often proceeds by discovering empirical dependencies – or functional relations – between stimuli and responses, as, for instance, in affective priming paradigms: if a subject has been exposed to a positively valenced stimulus prior to viewing a further stimulus item of ambiguous valence, the subject is more likely than they are at baseline to select a positive interpretation of the second, ambiguous stimulus. Such functional-level dependencies are then accounted for by process models – models of the neural or computational processes the playing out of which explain the functional dependencies in question. These models appear at what De Houwer et al. call the 'cognitive level'. This use of 'levels' does not yield a personal level of the relevant sort, however, one that

allows subpersonal science to be readily dismissed during enquiry into personal-level matters. For, on the view of De Houwer et al., it is a central task of a science of the mind to integrate the functional data and the mechanistic cognitive models of it. In fact, what they call a 'level' is something like a subset of the data, one involving correlations between responses to various stimuli (including patterns of response to stimuli that are contingent on past exposure to other kinds of stimuli). To the extent that there is a notion of level in play, it may be a matter of scale, the presumption being that the scale of organismic response is larger than the scale of cognitive process models; the organism is larger than entities (neurons, data structures, and so on) appearing in process models.

2. *Cognitive science as the attempt to make the personal level intelligible*

Much of the enthusiasm for the personal level rests on a commitment to the autonomy of personal-level explanation (Hornsby 2017, 176) and to a conception of humans that, in some way, places them, because of their agency, outside of the orthodox scientific picture of causal relations between states or properties in nature.

Proponents of this kind of view sometimes seem to want to insulate *a priori* or commonsense claims about human action from the sciences, claiming that nothing cognitive science delivers could disprove or even count strongly against what we know independently about the personal level. On a straightforward reading of such claims, they express a squarely anti-naturalist view, one beyond the scope of the present essay; for, they place mental processing outside the realm of integrated naturalistic investigation. At the same time, however, authors who eschew a role for science in our attempt to understand the personal level sometimes are at pains to establish their naturalistic credentials. Such conciliatory protestations take at least three forms:

(i) that all of the entities that possess personal-level states are themselves physical entities; (ii) that the subpersonal level (or other physical facts) *enable* the appearance of personal-level phenomena; or (iii) that study of the subpersonal level can make personal-level phenomena intelligible. As attempts at naturalistic vindication, the general idea seems to be that if the personal level is at least tied down by, given rise to by, or systematically related to the natural world, proponents of a personal level can reasonably claim to have fit the personal level into the natural order.

In what follows, I focus on (iii). For, by itself, (i) does almost nothing to secure a naturalistic approach. After all, someone who believes in supernatural magical powers might claim that the entities possessing such magical powers are physical entities; that would hardly secure the naturalistic credentials of a theory of magical powers. And (ii), either on its own or conjoined with (i), is too vague. Enabling must be fleshed out; if enabling conditions are simply consistent causal contributors or background conditions – such as the appropriate amount of air pressure – the enabling-relation does not deliver epistemically insulated levels. The move to (iii) provides a way to flesh out the claim of enabling and thus, potentially, to justify a substantive claim about levels.

What is it, though, for events or processes at one level to make intelligible those at another level? An extant strategy focuses on structural or dynamical aspects of personal-level phenomena, which could then be partly accounted for by isomorphic structure or dynamical mirroring at the subpersonal level. Assume that rational inference is a personal-level phenomenon (Davies 2000*a*). A central feature of rational inference is that it respects the syntactically characterized transitions distinctive of valid argumentation (as captured by, for example, such formal rules of inference as *modus ponens*). Imagine now that cognitive scientists

have identified neural activity that exhibits activation patterns isomorphic to the syntactically characterized transitions of interest, activity that occurs when human subjects are drawing inferences (McDowell 1994, 198–199) and that involves neural activity independently thought to contribute to attentive rule-following (e.g., the activation of frontoparietal loops – see, for example, Goldman-Rakic 1987). Proponents of a personal level typically hold that this neural activity is not genuine inference, because they hold that inference requires an agent who performs the inference (and who understands the content of the relevant inferentially connected statements). Nevertheless, they might claim that this neural discovery renders inferential acts partly intelligible by revealing a suitable tool, available to the agent, for doing at least some of what the agent does when drawing an inference. And as regards content, elements of a subpersonal structure might bear a stable relation to objects or features beyond the organism. Consider McDowell: "[C]an we understand how possession of a visual system that works like this makes it possible for a creature to become informed of just *these* features of its environment?" (McDowell 1994, 196; see also Burge 2010, 447, on the systematic relation between the subpersonal, nonperceptual tracking of bodies in the visual system and the capacity for the tracking of objects in genuine perception, which is, by Burge's lights, an individual- or personal-level affair). The idea would seem to be that the obtaining of a subpersonal state differentially sensitive to such-and-such feature in the environment renders it intelligible that the subject can enter a personal-level perceptual state representing just that feature.

This attempt at naturalistic vindication must overcome at least four significant difficulties. First, it seems to presuppose that human capacities are given *explananda*, that we have independent access to the structure of personal-level states or activities, which cognitive

science might then help to explain, by making that structure intelligible. That account of the cognitive-scientific project was roundly criticized in the immediately preceding subsection.

Second, although we may take ourselves to have independent access to personal-level phenomena, for all we know, the best models of the human behavior in question – of, for instance, our saying with conviction that there are personal-level phenomena with such-and-such features – do not codify, appeal to, or presuppose a distinction between two levels of contributors to the production of that behavior. Dennett (1991, ch. 10, section 4) has described consciousness as a "user illusion," belief in which is a side effect of the architecture of the human cognitive system that does not contain consciousness, at least not of the sort commonly believed in by Cartesians and so-called qualia freaks (Jackson 1982, Chalmers 1996). The negative thesis of the present work might be fruitfully understood in a parallel way, as the claim that the personal level is a user illusion. A naturalist seems hostage to empirical fortune in this regard – hostage to a proper psychology of philosophy, as it might be put – and fortune does not appear to favor an outcome friendly to the proponents of the personal level. From the standpoint of cognitive science, our claims about the personal level and about the causes of our making those claims do not appear to be produced by activity at a distinctively personal level or by a combination of states or processes some of which appear at a distinctively personal level sufficiently autonomous to perform the epistemic work it is typically recruited to do.

Third, the view on offer faces a dilemma of motivation. Either supposedly personal-level *explananda* consist in more than the structural or relational properties accounted for by subpersonal-level states and processes, or they do not. If they consist in more, what reason do we have to think such "more" exists, beyond intuitions related to such things as folk psychological practice, which intuitions themselves stand in need of psychological explanation? If they do not

consist in something more, then the claim to a personal level looks like a claim to duplicative structure.[17] There appears to be a trade-off: the more intelligibility one's models produce – by finding processes at the subpersonal level that correspond to the supposed personal-level ones – the less reason there is accept the existence of distinctive personal-level phenomena there to be made intelligible.

Lastly, I wonder about the nature of the supposed explanatory relation between the personal and the subpersonal levels. If the two levels are genuinely distinct, how does structural matching *explain* anything? It may be of interest that similar structure appears in two distinct domains of reality, but why should the presence of a certain structure in one domain of reality explain the presence of that structure in another? The concern seems especially pointed if the personal-subpersonal distinction is not meant to be simply a matter of scale, but captures some deep difference in the nature of the two domains in question or the forms of explanation apposite to them (consider again the contrast between entries 3 and 4 on the earlier list of supposed personal-level properties). It is as if the proponent of substance dualism were to claim to make the functioning of mental substance intelligible by citing the existence of corresponding structure in the physical material of the body. Without some clear and systematic account of the relation between mental substance and the physical body, we are in no position to judge whether the

[17] Consider the way these concerns dovetail with the literature on consciousness and the so-called hard problem (Chalmers 1996). Begin with the easy problems of consciousness, which we solve by modeling the relevant data. For the sake of concreteness, assume that such a solution takes the form of Global Workspace Theory. In this case, there is no reason to treat conscious states as states at the personal level. Treated as an element in a model, the global workspace, from which information is globally broadcast, is not at a different level from the input and output systems – the contributing and consumer systems – as represented in the global workspace model. The distinction between the global workspace and input and output systems concerns the richness of information flow and the communication channels among various states and processes; it has nothing to do with levels. There seems to be no scientific reason to reduplicate structure, by positing personal-level conscious states all the same – saying that any state in the global workspace, although at the subpersonal level, corresponds to a personal-level conscious state – unless one has independent reason to think there is an additional component or aspect to such posited personal-level states. In the literature on the hard problem, those who advocate for this additional component do not claim to have discovered it in the scientific work, but rather motivate it *a priori*, by the consideration of zombies and so forth. And here, of course, the distinction between naturalism and nonnaturalism is on full display.

presence of a certain neural structure makes the presence of a given mental processes (with whatever structure) more intelligible. In the case of supposed personal-level facts, the only promising naturalistic way to ground the structural-explanatory inferences in question is to follow the lead of adjacent sciences, to ape the relation between, say, chemistry and biology (or, more specifically, say, biochemistry and molecular biology). But, to do so undermines the isolationist tendency that runs so strongly through the philosophical literature. It is not credible for a molecular biologist to dismiss otherwise apparently relevant biochemical facts by claiming to work at a different, epistemically autonomous level from biochemistry; such a move would be legitimate only if it could be shown that the research question at hand, together with the pertinent empirical details, renders the specific biochemical facts at issue empirically irrelevant. Citing a difference in levels or difference in research area or scientific domain does not suffice.

Again we encounter a problematic trade-off: the more intelligibility the proponent of the personal-level tries to derive from the subpersonal, the closer together (in the sense of nearness found in adjacent or overlapping empirical sciences) become the personal and the subpersonal; but the closer together they are, the less plausible it is to maintain the autonomy of the personal level, that is, the less plausible it is to make the isolationist's dialectical move. The more that intelligibility requires in the way of systematic explanatory relations between the subpersonal and personal levels, the more likely it is that commitments *vis-à-vis* the personal level become hostage to empirical results, laying claims about the personal level bare to refutation or radical revision. The proponent of intelligibility seems faced with an unappealing (by her lights) trade-off: to the extent that she can secure intelligibility-making facts at the subpersonal level, she therefore risks rendering the supposedly fixed, personal-level facts subject to revision (or even elimination). If, in contrast, one holds that the personal-level capacities are only partially

explained by subpersonal factors and that full explanation requires the addition of something distinctively personal (e.g., robust normativity), the proponent of the personal level must motivate that addition (and its levels-creating status, in particular) by showing that it does work in the causal-explanatory enterprise, at the risk of abandoning naturalism. And such motivation seems lacking.

3. *A more liberal naturalism?*

Perhaps I have been stacking the deck against naturalistically inclined supporters of a personal level, by presupposing an excessively demanding conception of naturalism. Many domains of human interest – ethics, aesthetics, epistemology – provide a home to properties that do not seem to play a causal-explanatory role in successful scientific research programs. If a conception of naturalism recommends that we banish the relevant properties and kinds – right and wrong, the beautiful and the gaudy, the justified and the ill-founded – from ontologically serious discourse, then so much the worse for that conception of naturalism (Kim 1988). Defenders of the personal level might take the personal-subpersonal distinction (as applied to humans) to be of this ilk. On such a view, we should embrace a more liberal conception of naturalism, according to which, in order that a property be naturalistically legitimate, it is required only that the property have some physical (or perhaps, more broadly, scientific) basis or other.

How might this more liberal conception of naturalism vindicate the personal level? Assume, for the sake of argument, that one can successfully contrast personal and subpersonal states using only legitimate scientific vocabulary and concepts (ones that, in some significant range of contexts, indicate kinds, properties, or quantities that play an established causal-explanatory role or, as in the case of logical or mathematical vocabulary, serve as indispensable

resources in the causal-explanatory enterprise). Perhaps, for instance, personal-level states are those and only those that produce verbal report; to gesture toward concreteness, imagine that such states are highly correlated, in humans, with a specific complex of activity in Broca's area, superior temporal sulcus, and motor cortex, which areas do not contribute – at least not in the same way – to the physical states highly correlated with the various subpersonal processes. In that case, liberal naturalism would vindicate the personal-subpersonal distinction, because the states so distinguished correlate with genuinely physical states (or otherwise scientifically legitimate states – one can embrace both methodological naturalism and a scientific anti-reductionism [see Fodor 1974]), physical states that differ from each other in a way that can be marked using established scientific vocabulary. There is a legitimate personal-subpersonal distinction simply because there is a scientifically grounded distinction between the two collections of neural states. Even if the two collections of scientifically legitimate states are relative hodge-podges, the mere fact that the two hodge-podge classes are disjoint and the token states can be picked out using naturalistically legitimate language vindicates the personal-subpersonal distinction.[18]

The form of naturalism under consideration sets the bar for naturalistic legitimacy too low. As the story goes – and as things could well have gone – humans have sometimes consistently distinguished between (supposed) witches and nonwitches. And physical bases underlie the relatively consistent drawing of that distinction: appearance-related traits (e.g., having a crooked nose, warts, and ragged clothing) and behavioral traits (e.g., living in relative isolation, not appearing in church during Sunday services) that can be picked out, at least in

[18] The idea might be fleshed out in more complex form in order to allow for multiple realizability, but doing so serves no purpose in the current dialectical context. What is important is only a dependence on the physical (or legitimately scientific) – that is, that token physical (or otherwise legitimately scientific) states determine the presence of the (supposedly) personal and subpersonal states.

token instances, using scientifically respectable language. (This is at least as plausible as that supposedly personal and subpersonal states can be so specified.) But, any naturalist worth their salt should want to hold that witches do not exist, largely because the physical bases on which the distinction is drawn do not correlate with other supposed traits of witches – that is, the physical bases of categorization do no causal-explanatory work *vis-à-vis* what are thought to be the interesting properties in play (e.g., having the power to curse others). The witch-nonwitch distinction is the tiny tip of the iceberg. Such examples can be multiplied *ad nauseum*. Any distinction humans might draw even somewhat reliably has a physical basis; otherwise, humans could not draw it reliably, given that humans are physical creatures whose responses are, or depend on, physical processing. Moreover, bracketing human limitations, the physical world can be gerrymandered into an infinite number of contrasting classes of states. This would seem to be a *reductio* of liberal naturalism. The core insight of naturalism is that we best understand (or arrive at truths about) our universe by using scientific methods. In contrast, the perusing of an outrageously enormous list of abundant properties (Lewis 1983) the entries on which are produced by the gerrymandering of physical states or property instances promises no significant insight into the workings or nature of the universe.

Notice, too, the contrast between, on the one hand, using physicalistically or naturalistically acceptable resources to characterize two sets of properties and, on the other hand, those two sets of properties' being at different levels and where, additionally, the levels in question are of a sort that would support the isolationist's dialectical move. It might be that states typically classified as personal level are correlated with states of high activity in, for instance, pre-frontal cortex, whereas states typically classified as subpersonal correlate with activity in, for example, lateral geniculate nucleus (or LGN) in the case of early visual processing. This alone

would provide no support for a claim about a difference in levels. Activity in prefrontal cortex does not appear at a different level from activity in the LGN. So, even if there is a scientifically legitimate description of the difference between the two sets of states – those typically considered personal and those typically considered subpersonal – and even if these scientific descriptions identify differences that are relevant from a causal-explanatory perspective (one kind of state controls slow, deliberate, verbal report, while none of the others do, let us imagine), that fact does not establish (or even tend to suggest) that the two sets of properties or states appear at different levels, in the relevant sense. In order for proponents of the personal-subpersonal distinction to put liberal naturalism to productive use, they must plausibly locate a physical (or otherwise naturalistic) basis that captures the contrast between cases in which we judge "*x* is at the same level as *y*" and cases in which we judge "*x* is at a different level from *y*" (where *x* and *y* could be entities, properties, state-types, etc.). If judgements about sameness and difference of level are not coherent (Eronen 2015), liberal naturalism offers no hope for vindication of the personal-subpersonal distinction. And, even if such a physical demarcation can be found, it requires a further step – one proponents of the personal-subpersonal have not made – to show that the physically marked distinction between what we recognize as different levels does causal-explanatory work, in general, and specifically in the case of the personal-subpersonal distinction (and, to wit, in a way that supports the isolationist dialectical strategy).

What makes naturalism plausible is its wild success as a set of attitudes and methodologies (Papineau 2001). To the extent that we have good reason to accept a metaphysical naturalism – according to which there is no supernatural stuff, no "minded nature" – it is because the assumption of metaphysical naturalism has contributed to the enormous success of scientific methods. Thus, metaphysical naturalism derives its primary epistemic credentials from the

success of methodological naturalism. Liberal naturalism, however, seems to rest on an interpretation of metaphysical naturalism that is at odds with methodological naturalism. Liberal naturalism takes for granted metaphysical naturalism's prohibition on the belief in entirely supernatural properties, states, or entities (those not dependent on the physical realm or on otherwise scientifically legitimate properties, states, or entities). Liberal naturalism adds to this assumption a form of permission: the "right" to countenance any property, state, or entity that depends metaphysically on physical (or otherwise scientifically legitimate) states, properties, or entities. This creates an irresolvable tension. Liberal naturalism both wants to rely on the success of methodological naturalism (specifically on the contribution of metaphysical naturalism to that success) and, at the same time, flout methodological naturalism, by giving permission to countenance a wide range of properties, states, and entities that play no role in the causal-explanatory enterprise underwritten by methodological naturalism. Methodological naturalism has succeeded wildly not by countenancing just any property, kind, relation, etc. that has a physical or otherwise scientifically legitimate basis, but rather by countenancing only the properties and kinds that play a causal-explanatory role. Thus, the liberal naturalism on the table forfeits the rational justification rightly associated with naturalism, by countenancing kinds and properties grounded only in a metaphysical naturalism divorced from methodological naturalism, and thus unmoored from the rational justification for metaphysical naturalism.

Philosophers sometimes write as if philosophical naturalism commits one only to consistency with established scientific theories or results. I am not inclined to frame the matter in that way. But if it is framed in that way, my point become this: science trades in properties that play a causal-explanatory role in the scientific picture of the universe; if science contains a closure clause ("and nothing that is not included in a final science"), then proponents of the

personal level seem, from our present standpoint, to be committed to something inconsistent with science. Thus, liberal naturalism falls flat even on the minimal view that philosophical naturalism requires only that one not flout science or commit to propositions inconsistent with our best science.

In the end, then, liberal naturalism offers only the weakest and unsatisfying naturalistic vindication of the personal-subpersonal distinction. If liberal naturalism is the strongest form of naturalism consistent with commitment to a distinct personal level, that should be a cause of genuine concern for philosophers of mind and cognitive science who are inclined toward naturalism and who deploy the personal-subpersonal distinction in philosophical contexts.

III. The mind as flattened from above

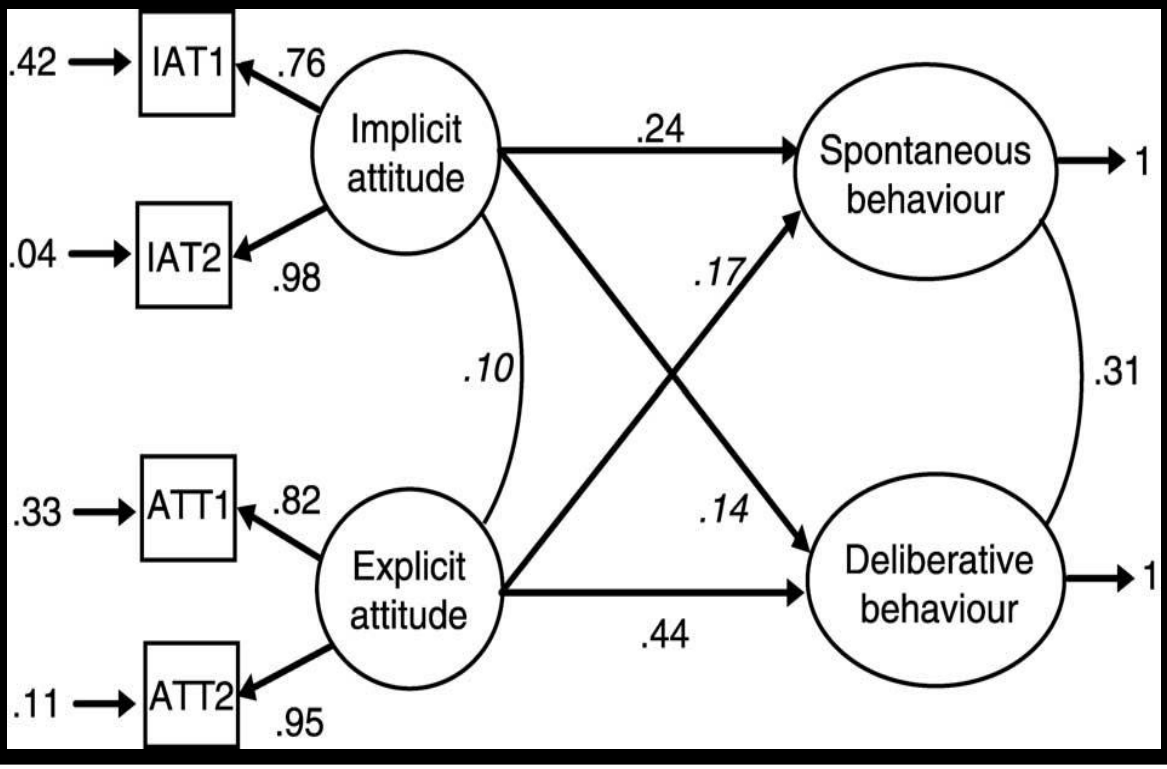
What is the alternative to the prevailing layered model of human mind and cognition? When one examines models produced by contemporary cognitive scientists, one frequently encounters what I have been referring to as mixed models, in which various states, properties, and processes from different scales co-contribute to the production of intelligent behavior, doing so in various proportions and in more or less interactive ways (e.g., the strength of one contributing factor amplifying the strength of the contribution of another factor). Such factors often include, within a single model, what would typically be considered personal-level states or processes as well as so-called subpersonal ones, where no distinction between the two kinds of model elements, as appearing at different levels, plays a role in accounting for the data. Moreover, the extent and nature of such co-contribution calls into question the thought that there is something distinctive about the so-called personal-level contributors; for, even if, in some cases, their distinctive effects can be isolated and known, the overall picture of the mental architecture is one of co-contribution, interactive contribution, and mutual modulation. Success in sometimes identifying

the distinctive contribution of so-called personal-level states provides no more reason to think that there exists a distinctive *level* – the personal level – than does our success in sometimes isolating the effects of pressure on the weather provide reason to think that there is an epistemically autonomous "pressure" level in the natural world.

My arguments for this flattened, co-contribution model, and the attendant elimination of a distinctively personal level, appeal to methodological naturalism together with prevailing trends in cognitive modeling. Take the data that cognitive science is out to explain – data of the concrete sort, such as differences in reaction times, error rates, or sorting patterns. Cognitive scientists typically account for such data by the construction of models. Since a picture is worth a thousand (or more) words, consider the following figure, taken from a paper by Marco Perugini

(2005).

Figure 2. Standardized parameters for the structural equation model testing for the double dissociation pattern (Study 2 n = 109) (from Marco Perugini, "Predictive models of implicit and explicit attitudes," *British Journal of Social Psychology*, 2005)



If the reader extracts only one message from this figure, let it be the following. There is nothing in Perugini's analysis to suggest that a difference in levels separates explicit attitudes – those correlated with states supposedly at the personal level – from implicit attitudes – commonly associated with the subpersonal level. Both kinds of state are part of a network of causes and effects that has no apparent ontological layering built into it (or epistemically privileged subspaces of the sort required by the proponents of a personal level). Even if some philosophers would like to insert ontological layers into the picture, the figure itself does not contain them, and an assumption of ontological layers does no work in the modeling depicted by the figure; it

is nowhere to be found in the statistical modeling of the relations between the various forms of behavior and contributors to them.

It may be worth filling in some details. In the paper in question, Perugini reports the results of two experiments, both of which contribute to the now vast body of research on implicit attitudes. The topics of interest in the experiments are attitudes toward smoking and attitudes toward junk food, the basic question being whether subjects' explicitly avowed attitudes or rather their implicit attitudes toward smoking or junk food govern certain forms of behavior. First, Perugini uses four measures – including an Implicit Association Test and a questionnaire – to probe the presence of implicit and explicit attitudes; these appear in the figure as IAT1, IAT2, ATT1, and ATT2. In the experimental trials, Perugini collects data on both deliberative behavior and more time-pressured, spontaneous responses. He models the resulting data using three different approaches – additive, double-dissociative, and interactive. An additive approach represents each of the kinds of attitude, implicit and explicit, as making a linearly scaling contribution to the production of each kind of behavior, deliberative and spontaneous. A double-dissociative model takes explicit attitudes to produce deliberative behavior and implicit attitudes to produce spontaneous behavior. And, the multiplicative approach takes the two kinds of attitudes to interact (in the statistical sense) in a non-linear way to produce at least some forms of behavior.

Figure 1 depicts the results of the analysis of Perugini's snacks-and-junk-food experiment, which are fit best by a double-dissociative model, the only of the three approaches amenable to an interpretation that jibes with the drawing of a personal-subpersonal distinction.[19]

[19] This conciliatory remark must be hedged, for, while the double-dissociative model may be consistent with the existence of a personal level, it does not alone provide reason to think that the subject can tell *a priori* or by introspection what contribution explicit states make to the production of a given form of deliberative behavior; nor does it entail that the subject can tell in which cases a double-dissociative model governs the production of behavior

This illustrates the point that, in some cases, the modeling of experimental results can be made to fit with a distinction between levels, even though the model itself is neutral on the matter.[20] The smoking-related data is, however, best fit by a multiplicative model; when trying to predict whether a subject smokes, how much difference is made by a difference in the subject's implicit attitude toward smoking depends on the strength and valence of the subject's explicit attitude toward smoking (as these are measured, recall, by independent tests). For technical reasons, Perugini applies a different form of statistical analysis to his smoking-related data – that is, different from the structural-equation modeling depicted in the figure above representing the snacks-related data. But, a distinction between levels plays no role in that statistical analysis, and Perugini makes no effort to layer his presentation of that data, separating streams into levels. So, the flattened architecture one sees in the reproduced figure can fairly be taken to indicate Perugini's general conception of the architecture as it pertains to the roles of implicit and explicit attitudes. Thus, contributions of two streams of processing – whether additive, multiplicative, or even sometimes dissociative – does nothing to support the idea of two different levels in the

---

and when, instead, a more interactive model governs behavior. Thus, the fact that a double-dissociation model fits some sets of data does not support the isolationist strategy embraced by proponents of a robust personal level. A version of this concern applies with even greater force to the proposal that the additive model lends itself to an application of the personal-subpersonal distinction. For, even if empirical evidence seems to show that explicit states make a distinctive (not interactive, not modulating) contribution to the production of a particular form of behavior – alongside implicit states, which make their own distinctive and separable contribution to the production of that form of behavior – that provides no reason to think the subject can identify *a priori* or by introspection which aspects of that form of behavior are caused by which kind of state or how much of the variance in a particular form of behavior each kind of state accounts for. Neither does it provide reason to think the subject can tell in which cases a double-dissociative model governs the production of behavior and when, instead, a more interactive model governs behavior.

[20] The fact that there are some cases of the dissociative sort does not alone entail (or highly probabilify) the existence of distinct levels of reality (nor does it naturalistically validate two disjoint styles of explanation). The effects of auditory stimulus can sometimes be dissociated from the effects of visual stimulus. That does not establish that audition and vision are at two different levels of reality (or are subject to two epistemically disjoint, or quasi-disjoint, explanatory styles or vocabularies).

cognitive system. The fact that additive and multiplicative models are sometimes the best fit provides further reason to doubt the existence of two distinct cognitive levels.[21]

Perugini's approach to modeling speaks against the view that explicit attitudes appear at a distinctive ontological or epistemic level, the personal level. Instead, there are explicit attitudes and implicit attitudes (perhaps with various sub-kinds in each category), any of which can contribute to the production of either kind of behavior: deliberative or spontaneous. And, each of the two kinds of attitudes can, and sometimes do, modulate the contribution of the other with

[21] See Gawronski and Bodenhausen (2014, for example at 188–189) for a different kind of approach to implicit attitudes research that also supports skepticism concerning any parallel between implicit and explicit attitudes, on the one hand, and the personal-subpersonal levels, on the other.

Note that some authors (e.g., Lyons, 2016, 252) claim that implicit attitudes are at the personal level. I find the view puzzling, perhaps because I find the entire idea of the personal level puzzling, but also because implicit attitudes seem to fail the salient tests for being personal-level states, except possibly the most nebulous one ("being properly attributable to the person"). After all, part of the puzzle of implicit attitudes is that they drive behavior without being consciously accessible, without being part of our introspection-based self-conception, without their effects being under intentional control, and without the behaviors they produce being best accounted for by rationalizing explanations.

Some empirically oriented philosophers of mind endorse an explicitly layered picture according to which personal-level states are realized by subpersonal-level states and processes. Inspired by work in contemporary cognitive science, Peter Carruthers (2009) and Keith Frankish (2009, 2016) both suggest a realization-based approach to the personal level, at least to the extent that one takes the so-called System 2 to be the natural home of personal-level states and processes. For example, Carruthers says, "I shall outline an architecture for System 2 that sees it as realized in cycles of operation of System 1 (rather than existing alongside the latter)" (*ibid.*, 112). For his part, Frankish (2009) explicitly identifies System 2 with the personal level and claims that it is realized by System 1; in later work (2016, 35), Frankish hedges this claim, though on questionable grounds, which, to my mind, illustrate one of the challenges that results from trying to enforce a personal-subpersonal distinction in the human case. Frankish (2016) holds that (a) System 1 realizes System 2, (b) implicit attitudes appear in System 1, (c) implicit attitudes can be attributed to the whole person, and (d) System 2 is the personal level. This quartet of propositions would appear to be inconsistent: implicit attitudes cannot both be attributed to the whole person, and thus appear at the personal level, but also appear at the level of processes (System 1) that realize personal-level processes, at least not if we treat realization as it normally is treated, as an inter-level relation.

The flattened approach dissolves the tension that Frankish's view puts on display. States associated with System-2-style processing (slow, deliberate, conscious) appear in models that also include states associated with System-1-style processing (fast, automatic, subconscious), both kinds of processes accounting in differing proportions for variance in the data, in ways that involve differing amounts of interaction and mutual modulation, depending on the case, but with no distinction between levels.

Ron Mallon also represents an interesting case (Mallon 2016). He argues for what he pitches as a personal-level account of stereotype threat, but the overall picture he presents combines the effects of states and processes that would normally be placed at the personal level with states and processes that would normally be placed at the subpersonal level, without arguing that treating these two contributors as being at different levels produces a causal-explanatory payoff, that is, without showing that the *levels* distinction makes any causal-explanatory difference. It is important to recognize the possibility that two distinct states can, respectively, pass extant tests for being at different levels, whilst also making distinct causal contributions to an outcome, without their supposed difference *in levels* playing any role in explaining the outcome.

regard to either kind of behavior. This last point is of special relevance, for it suggests that, even if we limit ourselves to thinking of the personal level as a level of description, a self-standing vocabulary of the sort Dennett originally had in mind, contemporary cognitive science issues a skeptical verdict. Mixed modeling cuts against the idea that that the personal-level vocabulary fruitfully isolates a domain that is best approached using only the language of the personal level. And, so far as I can tell, Perguini's approach is entirely unremarkable in the cognitive sciences, in its lack of deference to and its lack of causal explanatory appeal to a personal-subpersonal distinction.

Here is a way of interpreting the lesson of such models as Perugini's. They show that there is no intentional level distinct from the level of representations and mechanisms (broadly construed). States that have been thought of as (relatively) robustly intentional – not merely representational in the way a bit of feature-sensitive cortex might be – are, in fact, part of the collection of explanatory resources from which cognitive scientific modelers draw materials, and this collection includes various other sorts of resource (including bits of feature-sensitive cortex that represent in what is normally thought to be a more minimal way), which are freely combined to account for the data. This belies the claim that two kinds of resource exist at different levels – at least if levels are supposed to be metaphysically distinct domains of nature or epistemically autonomous (or quasi-autonomous) explanatory vocabularies (to the correct application of which we have access *a priori* or via conceptual analysis, introspection, or common sense).

This flattened approach to cognitive or mental architecture is a natural outgrowth of recent developments in philosophy of science, in particular, growing skepticism about the usefulness of the notion of levels. Some such skepticism rests on the observation that successful scientific work often produces so-called mixed models, including models composed of elements

appearing at different scales (Kim 2002, Thalos 2006, Wilson 2010, Potochnik and McGill 2012). In some cases, the differences in scale are enormous: the geographical distribution of a given animal species might be accounted for by a model that includes everything from prevalence of certain pathogens, to the distribution of species of tree, to variations in weather patterns. Consider the following observation, made Bill Bechtel:

> …a model for the intracellular oscillator is supplemented by terms and equations capturing a hypothesized pattern of interaction of neurons. There is an intuitive sense in which interactions between neurons is at a higher level of organization than operations within individual neurons, so that if the interaction is modulating the activity of individual neurons, one is dealing with top-down effects of a higher-level system on its components. But interestingly a differentiation of levels did not figure in these accounts. Rather, the modelers simply extended the scale of their model by adding equations to capture effects on a neuron's behavior from operations outside it. (Bechtel 2016, 725)

In other words, larger-scale effects appear alongside smaller-scale effects within a single model of intracellular oscillation, without any presupposition of autonomous levels (and certainly not ones to which we have epistemic access of the sort required by proponents of the personal level).

Some efforts have been made to save the notion of a level, particularly within the context of "new mechanist" philosophy of science; but, to the extent that such efforts succeed, they do so in a context-specific way that does not deliver a picture of any use to proponents of the personal level (Craver 2014, 2015, Gillett 2021, DiFrisco 2021; *cf.* Eronen 2013, 2015), for such a

context-specific distinction between levels does not, by itself, justify dismissing questions about lower-level entities or processes when asking questions about higher-level ones (in a compositional hierarchy of mechanisms) or provide epistemic autonomy in the consideration of questions at the supposedly higher level.

IV. Objections and replies

A. Indispensability. The first objection to be considered rests on an indispensability claim and an associated claim of self-defeat. According to the critic, we can interpret and draw inferences from empirical work in the cognitive sciences only by presupposing a distinction between the personal and subpersonal levels. The scientific investigation of human behavior begins with an understanding of humans as persons, as rational subjects, as beings with a subjective point of view. We cannot make sense of psychological or cognitive-scientific investigation absent an interpretive framework that commits one to beings of that sort. Thus, an argument for eliminativism about the personal level cannot appeal to trends in cognitive-scientific results; to do so would be self-defeating, for to conclude that there is no personal level invalidates the very research programs meant to support the eliminativist conclusion empirically.

These are sweeping and imprecise claims. Surely there is *some* sense in which experimenters and experimental subjects think of each other in, loosely speaking, everyday, folk-psychological ways. It is quite a distance, however, from that observation to the claim that their interaction presupposes a distinction between levels, of the sort that marks out a deep ontological or epistemic divide and the facts about one of which can be known *a priori* or by introspection or common sense. For example, experimenters might think of subjects as likely to produce certain forms of behavior subsequent to engaging in certain kinds of reasoning processes; but, once that behavior has been tested for, the modeling of the processes that produced behavioral results –

whatever they turn out to be – might not reify the scientist's process-related intuitions that motivated the experimental design. Moreover, even at the early stages of intuition-driven experimental design, the idea of a distinct, autonomous (or quasi-autonomous) level have played no role.

Can the concern at issue be pressed in a more pointed way? Here is one possibility: Experimental work in the sciences of the mind typically requires communication between experimenter and subject, which presupposes that experimenters and subjects interact as persons, consciously appreciating the content of what is said or written, tracking entailment relations among sentences and understanding the various implications of the instructions given.

I grant that, at a first pass, it seems apt to describe the subject as understanding the experimenter's instructions, as grasping or comprehending them. But, that first pass includes no rigorous account of what understanding, grasping, or comprehending amounts to. I contend that the process of comprehending instructions is itself best modeled as involving the co-contribution of "mixed" states and processes – that is, by models that appeal to states typically associated with the personal level operating alongside states typically associated with the subpersonal level, with no levels-implicating distinction built into those models or presupposed by their success. Thus, the point is well taken that subjects must comprehend experimenters' instructions, but I reject the claim that comprehension presupposes a distinct (robust, semi-autonomous) personal level. It may be said that comprehending is something *persons* do. Perhaps, but let us not allow italics to do so much philosophical work. Clearly there is a sense in which it is something organisms do, and maybe it is something persons, in a minimal sense, do; but it is a further question whether comprehending is done by a person who resides at, or instantiates properties that appear at, a distinct personal level – one that brings with it the metaphysical and epistemic

commitments embraced by proponents of the personal level. Our best answers to that question depend at least partly on the fruits of the causal-explanatory enterprise. I have argued that the causal-explanatory enterprise is not likely to yield models of comprehension-related behavior, or any other kind of behavior, that build in or appeal to a robust distinction between the personal and subpersonal levels – that is, models the success of which depends on the marking of various states as being at different levels (of the relevant sort).

Bear in mind, too, that even if some experimenters themselves are committed to the existence of a personal level, it is a further question whether that commitment explains the success of their research programs. A gap yawns between "the experimenter describes their research program as having property *P*" and "the research program has *P* and having *P* explains its success." Navigating these matters requires especial vigilance given how often scientists and philosophers use the same terms, but with importantly different associated meanings. Many a psychologist has used *a priori* theory to mean something like "what we thought prior to running our experiments," which is not what mainstream analytic epistemologists mean to express when they refer to a claim or theory as *a priori* (as known on the basis of a purely rational grasp of meanings, or the like). Similar remarks at least sometimes apply to the use of 'person', 'normative', 'rational', and a host of other terms in this vicinity.

B. Folk psychology as simplified model or as model fragment.

Consider the familiar point that folk psychological prediction and explanation frequently succeed (Fodor 1987, Chapter 1). Why, a critic might wonder, is that observation alone not enough to support the positing of a personal level?

A number of points militate against this defense of the personal level. First, bear in mind the forceful case Paul Churchland has made against the success of folk psychology, emphasizing its limitations and failures (Churchland 1981). I am not inclined to take sides on this issue. Folk psychology is successful in some ways, not in others. In the present context, it is more helpful to ask (a) whether folk psychology is committed to the existence of a personal *level* and (b) if it is so committed, whether that commitment accounts for the success of folk psychology. I am not sure whether folk psychology is committed to a personal level (partly because I doubt that there is a single, well-behaved beast that is folk psychology). Regardless, when we turn to (b), the answer seems clearly to be no. The deployment of a model of folk psychological states may well help me to coordinate interactions with my fellows, so that they and I, for instance, successfully meet for dinner on a given night. But, that success depends not at all on whether, for instance, the motivational states I attribute to my fellows appear at a different level from the subconscious edge-detection mechanisms they deploy as they visually locate the door of the restaurant.

Perhaps this objection could be made more forceful if were translated into the talk of models. Might an explicit appeal to the autonomy and utility of folk-psychological models ground a more apt objection to the flattened approach? Consider a remark Ron Giere makes in connection with distributed cognition in scientific enquiry: “The question is whether the various models generated by the principles of the theory fit actual situations of interest well enough for the intended purposes. On this measure, folk psychology does remarkably well” (2006, 717). Science is driven by modeling, and models come in various forms and include varying degrees of detail, even within the same domain and aimed at roughly the same phenomenon. Which model is applied in a given context, relative to a given target system, is often a matter of adequacy for purpose. Folk psychological models are adequate for many purposes, Giere claims, and thus, one

might think, their posits should be treated as we would the posits of any successful models – that is, taken ontologically seriously and, in Giere's case, attributed to groups, if doing so is adequate to certain purposes of sufficient importance.

Giere's general position seems sensible, though one might reasonably hope for a more detailed account. Are not some models more successful than others? Should not our ontological attitude (or attitude toward distinct explanatory vocabularies) take differential levels of success into consideration, including different extents of application? A model with only limited range of application, adequate for only a small number of purposes surely will be given less ontological deference than a model that, relative to the same general domain, covers a much broader range of systems, is adequate for a much wider range of purposes, and is adequate for more demanding or exacting purposes (accounting for more variance, for instance, in the cases where both models are adequate for some purposes of interest).

The more germane question concerns an inference to a folk-psychological *level*, an inference that Giere himself does not make. Such an inference would seem reckless. Relative to a range of target systems (radioactive processes, for example), some models that are sometimes useful contain more detail; others that are sometimes useful contain less; some models represent certain forces or entities, while others advert to the contributions of a different collection of forces and entities. The idea that each of these models establishes an autonomous (or semi-autonomous) level in nature (never mind one about which we gain knowledge *a priori* or by introspection or commonsense reflection) is entirely implausible, even in a case in which one model is more abstract or coarse-grained than the other, as folk psychology is often thought to be relative to the typical cognitive-scientific process-model. The property of being a table is more

abstract than the property of being an end-table. But that hardly entails that there are two different levels of reality or two different styles of explanation applicable to the sitting room.

More to the point, imagine we have in hand two models of some behavioral data. One contains only elements recognized by folk psychology. The second model includes those elements (or elements much like them) as well as what are typically considered subpersonal elements. The second model outperforms the first in most contexts, even though the first is adequate for certain purposes. The obtaining of this situation gives us no reason to think that the first models limns a distinct level of reality or has its own integrity as a distinctive explanatory framework. After all, when we want more fine-grained explanations or predictions and thus work with the second model, we get those enhanced or more fine-grained results by mixing together elements normally thought to be personal and elements normally thought to be subpersonal; we do not achieve such results by appealing to a levels-based distinction between those elements. Certain states control verbal report; they produce the report *P* in circumstances in which we are interested to see whether subjects or organisms will say that *P*. And the fact that subjects say *P* may be correlated with other quantities in which we are interested. But that's just a point about the causal powers of the 'P'-producing state, not a matter of its being on a different level from states typically thought to be subpersonal.

Let us imagine that elements in a more complex and detailed model cluster into two sets, each of which can, in some circumstances, be applied independently to the relevant target systems to generate reasonably accurate predictions or can account for data in a reasonably satisfactory manner. (And imagine one of these clusters consists of elements recognizable from folk psychology.) That is hardly reason to think that each cluster corresponds to a different level of reality or what is an epistemically autonomous domain. Factors related to unemployment rates

may partly account for high rates of inflation. Factors related to pandemic-induced supply-chain glitches may partly account for high rates of inflation. We should not conclude from such observations that there is an unemployment level of reality and a distinct supply-chain level of reality. Of course, one can talk that way, if one wishes, but to call these both 'levels' is to strip levels talk of any import. Such a notion of levels certainly cannot do the sort of dialectical work the distinction between the supposed personal and subpersonal levels is meant to do in the philosophical literature.

C. The unaccounted for explanandum: action

Discussion of the immediately preceding objection might seem insensitive to one of the most oft-made points in this vicinity (e.g., Hornsby, 2000, 19): that, when we appeal to folk psychological states and processes, we target distinctive *explananda*, actions, which can *only* be accounted for by rationalizing explanations. Take a now classic example, that of raising one's arm at an intersection while driving (Davidson 1963). Precisely the same physical raising of the arm that might, in one context, be a waving (e.g., to a friend) might, in a different context, be a signaling (of, for instance, an impending right turn). To make sense of the behavior in either case – to see the *explanandum* for what it is – one must see it as the expression of the intention of a person, either to greet a friend or to signal a turn. According to my imagined critic, this gap, between action and mere behavior, and the sort of explanation distinctively suited to the former, entails the existence of a personal level in the relevant sense. There is a distinctive domain, of actions, populated by events that are explained by a self-contained vocabulary – the only vocabulary suited to the task of accounting for action.

Perhaps, but there are alternative, relatively straightforward ways to understand the difference between signaling, waving, and, say, twitching. In the case of signaling, the states that contribute to the production of the arm-raising include (subpersonal) representations of extant traffic-related events, of rules of driving, and the goal-state that represents the creating, in other drivers, of representations of the future movement of the automobile of the organism doing the signaling, all among other states. In the case of waving, the states that contribute to the production of the arm-raising include (subpersonal) representations of a particular conspecific and a goal state that represents the activation, in that conspecific, of a representation of the organism raising its arm, among other states. A mere arm-raising, such as a twitching, is an arm-raising that occurs without the contribution of such constellations of states – one that is best modeled as the activation only of motor commands, perhaps as the result of implanted electrodes or some such. None of this requires the introduction of levels; rather, we are helping ourselves to various selections from a grab-bag of causal-explanatory resources, as is suited to each case.

Wherever there is action of the supposedly distinctive sort, there is behavior. And in most cases, one best explains the behavior partly by appealing to states normally associated with the personal level and partly by appealing to states normally associated with the subpersonal level. Perhaps we tend to categorize the behavior differently partly because there are different states among the causes (though I would contend that it is also because behavior with those sorts of difference in cause tend to occur in different contexts, in which, accordingly, the behavior in question has different effects – someone shouts ‘hello’ or someone else depresses their brake pedal – which partly accounts for differences in the categorization of the behavior in question). And, among the states that contribute to the behavior, it may be of special interest that certain ones appear (representations of a friend, for instance). But, none of this creates a new levels-

birthing *explanandum* – action – in addition to behavior. There is one kind of movement, with three possible explanations, as the case may be. There is nothing about any of the three explanations that requires the introduction of a new category of *explanandum* that can only be accounted for by the introduction of states appearing at a proprietary and autonomous level. There are simply three (or more) models, most likely mixed.

In the bigger picture, this insistence on a missing *explanandum* should be viewed with skepticism. It would appear that we are being told by proponents of a theory, call it T1, that there is an *explanandum* left unexplained by competitors to T1 (and for the explanation of which we must introduce levels of a certain sort). And when asked why we should accept the existence of such this *explanandum*, the response seems to be little more than that because T1 entails its existence.

VI. Concluding Remarks

The primary goals of this essay have been (1) to show that, given prevailing cognitive-scientific practices, substantive naturalism cuts strongly against commitment to a personal level of the sort commonly adverted to in the philosophical literature, and (2) to argue for a flattened view of the human mental and cognitive architecture. I prosecuted the former goal partly by appeal to empirical work, which work also informs the prosecution of the latter. A distinction between personal and subpersonal levels does not appear in successful models of the data of cognitive science; rather, many of our best models are "mixed," with states typically placed at the personal level co-contributing (sometimes in interactive, mutually modulating ways) alongside states typically placed at the subpersonal level to the production of the behavior of interest. Moreover, the collection of behavioral data of interest presupposes nothing more than the importance of organisms that can be reidentified and to which various data points can be indexed.

Of course, the term ‘level’ has been used in many ways, and thus it is essential that we keep in mind what is at stake epistemically and methodologically. The import of the preceding arguments is that there is no epistemically autonomous or semi-autonomous domain of the personal level and no level independent knowledge about which sets the explanatory target for cognitive science. If there is a robust, naturalistically legitimate personal level, the facts about it will emerge from successful modeling of data; such supposed facts are not given to us *a priori* or otherwise nonscientifically, as fixed *explananda* or data for cognitive science to account for or explain.

The preceding paragraphs make bold claims, which many readers may resist. Such resistance may ultimately carry the day, though I see little reason for optimism in this regard. The most reasonable response to the extant challenge is, rather, to embrace a mixed approach to modeling human behavior, to embrace an architecture flattened from above, and to accept the epistemic limitations – and opportunities – that follow. To be clear, this does not amount to eliminativism of the states and processes normally placed at the personal level. Whether eliminativism looms may be a function of one’s semantics, whether one is willing to accept revision in one’s conception of the states and processes normally placed at the personal level. Naturalism, it is worth noting, and a naturalistic attitude toward scientific enquiry itself, would recommend flexibility in this regard (Wilson 2006), which renders commitment to a flattened architecture significantly less likely to bring eliminativism about mental states in its train.

Works Cited

Alsmith, Adrian John Tetteh, and Frédérique de Vignemont. 2012. “Embodying the Mind and Representing the Body.” *Review of Philosophy and Psychology* 3: 1–13.

Bechtel, William. 2016. “Mechanists Must be Holists Too! Perspectives from Circadian Biology.” *Journal of the History of Biology* 49: 705–731.

Bermúdez, José Luis. 2000. “Personal and Sub-personal: A Difference without a Distinction.” *Philosophical Explorations* 3, 1: 63–82.

Block, Ned. 2017. “Unconscious Perception within Conscious Perception.” *Neuroscience of Consciousness*, 3, 1: 7–9.

Boghossian, Paul. 2008. “Epistemic Rules.” *Journal of Philosophy* 105, 9: 472–500.

Botvinick, Matthew B, and Jonathan D. Cohen. 2014. “The Computational and Neural Basis of Cognitive Control: Charted Territory and New Frontiers.” *Cognitive Science* 38: 1249–1285.

Buckner, Cameron. 2018. “Empiricism without Magic: Transformational Abstraction in Deep Convolutional Neural Networks.” *Synthese* 195: 5339–5372.

Burge, Tyler. 2010. *Origins of Objectivity*. Oxford: Oxford University Press.

Cain, Mark J. 1999. “Fodor’s Attempt to Naturalize Mental Content.” *Philosophical Quarterly* 49, 197: 520–526.

Carruthers, Peter. 2009. “An Architecture for Dual Reasoning.” In *In two minds: Dual Processes and Beyond*, edited by J. S. B. T. Evans & K. Frankish (Oxford: Oxford University Press), pp. 109–127).

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

Churchland, Paul M. 1981. “Eliminative Materialism and the Propositional Attitudes.” *Journal of Philosophy* 78: 67–90.

Clark, Andy. 2015. *Surfing Uncertainty: Prediction, Action, and the Embodied* Mind. Oxford:

Oxford University Press.

Clark, Andy. 2019. “Replies to Critics: In Search of the Embodied, Extended, Enactive, Predictive (EEE-P) Mind.” In *Andy Clark and His Critics*, edited by Matteo Colombo, Elizabeth Irvine, and Mog Stapleton (Oxford: Oxford University Press), pp. 266–302.

Colombo, Matteo. 2013. “Constitutive Relevance and the Personal/Subpersonal Distinction.” *Philosophical Psychology* 26, 4: 547–570.

Craver, Carl F. 2007. *Explaining the Brain*. Oxford: Oxford University Press.

Craver, Carl F. 2014. “The Ontic Account of Scientific Explanation.” In *Explanation in the Special Sciences*, edited by Marie I. Kaiser, Oliver R. Scholz, Daniel Plenge, and Andreas Hüttemann, Synthese Library, vol. 367 (Dordrecht: Springer), pp. 27–52.

Craver, Carl F. 2015. “Levels.” In *Open MIND: 8(T)*, edited by T. Metzinger & J. M. Windt (Frankfurt am Main: MIND Group), pp. 1–26.

Davidson, Donald. 1963. “Actions, Reasons, and Causes.” *Journal of Philosophy* 60, 23: 685–700.

Davies, Martin. 2000*a*. “Interaction without Reduction: The Relationship between Personal and Sub-personal Levels of Description.” *Mind & Society* 1: 87–105.

Davies, Martin. 2000*b*. “Persons and Their Underpinnings.” Philosophical Explorations 3, no. 1: 43–62.

Davies, Martin, and Andy Egan, A. 2013. “Delusion: Cognitive Approaches – Bayesian Inference and Compartmentalization.” In *Oxford Handbook of Philosophy and Psychiatry*, edited by K. W. M. Fulford, M. Davies, R. G. T. Gipps, G. Graham, J. Z. Sadler, G. Stanghellini, and T. Thornton (Oxford: Oxford University Press), pp. 689–728.

De Houwer, Jan, Bertram Gawronski, and Dermot Barnes-Holmes. 2013. “A Functional-

Cognitive Framework for Attitude Research." *European Review of Social Psychology* 24, no. 1: 252–87.

De Vignemont, Frédérique. 2014. "A Multimodal Conception of Bodily Awareness." *Mind* 123, 492: 989–1020.

Debus, Dorothea. 2017. "Memory Causation." In *Routledge Handbook of the Philosophy of Memory*, edited by Sven Bernecker and Kourken Michaelian (Abingdon, UK: Routledge), pp. 63–75.

Dennett, Daniel C. 1969. *Content and Consciousness*. Abingdon, UK: Routledge.

Dennett, Daniel C. 1991. *Consciousness Explained*. Boston, MA: Little, Brown and Company.

DiFrisco, James. 2021. "Integrating Composition and Process in Levels of Developmental Evolution." In *Levels of Organization in the Biological Sciences*, edited by D. S. Brooks, J. DiFrisco, and W. C. Wimsatt (Cambridge, MA: MIT Press), pp. 111–134.

Drayson, Zoe. 2012. "The Uses and Abuses of the Personal/Subpersonal Distinction." *Philosophical Perspectives* 26, no. 1: 1–18.

Drayson, Zoe. 2014. "The Personal/Subpersonal Distinction." *Philosophy Compass* 9, no. 5: 338–46.

Eronen, Markus I. 2013. "No Levels, No Problems: Downward Causation in Neuroscience." *Philosophy of Science* 80: 1042–1052.

Eronen, Markus. 2015. "Levels of Organization: A Deflationary Account." *Biology & Philosophy* 30: 39–58.

Fodor, Jerry A. 1974. "Special Sciences (Or: The Disunity of Science as a Working Hypothesis)" *Synthese* 28, no. 2: 97–115.

Fodor, Jerry A. 1975. *The Language of Thought*. Cambridge, MA: Harvard University Press.

Fodor, Jerry A. 1987. *Psychosemantics: The Problem of Meaning in the Philosophy of Mind.* Cambridge: MIT Press.

Frankish, Keith. 2009. "Systems and Levels: Dual-system Theories and the Personal-Subpersonal Distinction." In *In Two Minds: Dual Processes and Beyond*, edited by J. St. B. T. Evans and K. Frankish (Oxford: Oxford University Press), pp. 89–107.

Frankish, Keith. 2016. "Playing Double: Implicit Bias, Dual Levels, and Self-Control." In *Implicit Bias and Philosophy, Volume 1: Metaphysics and Epistemology*, edited by Michael Brownstein and Jennifer Saul (Oxford: Oxford University Press), pp. 23–46.

Fridland, Ellen. 2021. "Intention at the Interface." *Review of Philosophy and Psychology* 12: 481–505.

Gallagher, Shaun. 2008. "Are Minimal Representations Still Representations?" *International Journal of Philosophical Studies* 16: 351–369.

Gallistel, C. R., Ann L. Brown, Susan Carey, Rochel Gelman, Frank C. Keil. 1991. "Lessons from Animal Learning for the Study of Cognitive Development." In *The Epigenesis of Mind: Essays on Biology and Cognition*, edited by Susan Carey and Rochel Gelman (New York: Psychology Press), pp 3–36.

Gawronski, Bertram, and Galen V. Bodenhausen. 2014. "The Associative–Propositional Evaluation Model: Operating Principles and Operating Conditions of Evaluation." In J. W. Sherman, B. Gawronski, and Y. Trope (eds.), *Dual-Process Theories of the Social Mind* (New York: Guilford Press), pp. 188–203.

Gendler, Tamar. 2008. "Alief and Belief." *Journal of Philosophy* 105, no. 10: 634–663.

Giere, Ronald N. 2006. "The Role of Agency in Distributed Cognitive Systems." *Philosophy of Science* 73: 710–719.

Gillett, Carl. 2021. “Using Compositional Explanations to Understand Compositional Levels: An Integrative Account.” In *Levels of Organization in the Biological Sciences*, edited by D. S. Brooks, J. DiFrisco, and W. C. Wimsatt (Cambridge, MA: MIT Press), pp. 233–259.

Gładziejewski, Pawel, and Marcin Miłkowski. 2017. “Structural Representations: Causally Relevant and Different from Detectors.” *Biology & Philosophy* 32: 337–355.

Goldman-Rakic, Patricia. 1987. “Circuitry of Primate Prefrontal Cortex and Regulation of Behavior by Representational Memory.” In *Handbook of Physiology*, vol. 5, edited by F. Plum and V. Mountcastle (Bethesda, MD: American Physiological Society), pp. 373–417.

Griffiths, Thomas L., Nick Chater, Charles Kemp, Amy Perfors, and Joshua B. Tenenbaum. 2010. “Probabilistic Models of Cognition: Exploring Representations and Inductive Biases.” *Trends in Cognitive Sciences* 14: 357–364

Holton, Richard. 2016. “Review of Andy Clark, *Surfing Uncertainty*.” *Times Literary Supplement* 7 October 2016.

Hornsby, Jennifer. 1997. *Simple-Mindedness: In Defense of Naïve Naturalism in Philosophy of Mind*. Cambridge, MA: Harvard University Press.

Hornsby, Jennifer. 2000. “Personal and Sub-Personal: A Defence of Dennett’s Early Distinction.” *Philosophical Explorations* 3 no. 1: 6–24.

Heubner, Bryce. 2008. “Do You See What We See? An Investigation of an Argument against Collective Representation.” *Philosophical Psychology* 21, 1: 91–112.

Hurley, Susan. 1998. “Vehicles, Contents, Conceptual Structure, and Externalism.” *Analysis* 58, no. 1: 1–6.

Hursthouse, Rosalind. 1991. “Arational Actions.” *Journal of Philosophy* 88, 2: 57–68.

Ismael, Jenann T. 2014. “On Being Someone.” In *Surrounding Free Will: Philosophy, Psychology, Neuroscience,* edited by A. R. Mele (Oxford: Oxford University Press), pp. 274–297.

Jackson, Frank. 1982. “Epiphenomenal Qualia.” *Philosophical Quarterly* 32, 127: 127–136.

Kim, Jaegwon. 1988. “What Is ‘Naturalized Epistemology’?” *Philosophical Perspectives*, Vol. 2, *Epistemology,* edited by J. E. Tomberlin (Atascadero, CA: Ridgeview), pp. 381–405.

Kim, Jaegwon. 2002. “The Layered Model: Metaphysical Considerations.” *Philosophical Explorations* 5 (1): 2–20.

Kriegel, Uriah. 2012. “Personal-Level Representation.” In *Consciousness and Subjectivity,* edited by G. Preyer and S. Miguens (Frankfurt: Ontos) pp. 77–114.

Levy, Neil. 2016. “ ‘My Name Is Joe and I’m an Alcoholic’: Addiction, Self-knowledge and the Dangers of Rationalism.” *Mind & Language* 31, no. 3: 265–76.

Lewis, David. 1983. “New Work for a Theory of Universals. *Australasian Journal of Philosophy* 61, 4: 343–377.

Lyons, Jack. 2016. “Unconscious Evidence.” *Philosophical Issues,* 26, *Knowledge and Mind*: 243–262.

Mallon, Ron. 2016. “Stereotype Threat and Persons.” In *Implicit Bias and Philosophy, Volume 1: Metaphysics and Epistemology,* edited by M. Brownstein and J. Saul (Oxford: Oxford University Press), pp. 130–154.

Marr, David. 1982. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. New York: W.H. Freeman and Company.

McDowell, John. 1985. “Functionalism and Anomalous Monism.” In *Actions and Events:*

Perspectives on the Philosophy of Donald Davidson, edited by Ernest Lepore and Brian McLaughlin, 387–98. Blackwell: Oxford. Reprinted in John McDowell, *Mind, Value, and Reality* (Cambridge, MA: Harvard University Press, 1998), pp. 325–340; page references are to the reprinted version.

McDowell, John. 1994. "The Content of Perceptual Experience." *Philosophical Quarterly* 44, no. 175: 190–205.

Metzinger, Thomas. 2003. *Being No One: The Self-Model Theory of Subjectivity*. Cambridge, MA: MIT Press.

Miracchi, Lisa. 2017. "Perception First." *Journal of Philosophy* 114, 12: 629–77.

Mole, Christopher. 2011. *Attention Is Cognitive Unison: An Essay in Philosophical Psychology*. Oxford: Oxford University Press.

Newell, Allen, John C. Shaw, and Herbert A. Simon. 1958. "Elements of a Theory of Human Problem Solving." *Psychological Review* 65, 3: 151–166.

Oppenheim, Paul, and Hilary Putnam. 1958. "Unity of Science as a Working Hypothesis." *Minnesota Studies in the Philosophy of Science*, vol. 2 (Minneapolis: University of Minnesota Press), pp. 3–36.

Papineau, David. 2001. "The Rise of Physicalism." In *Physicalism and Its Discontents*, edited by C. Gillett and B Loewer (Cambridge: Cambridge University Press), pp. 3–36.

Perugini, Marco. 2005. "Predictive Models of Implicit and Explicit Attitudes." *British Journal of Social Psychology* 44: 29–45.

Potochnik, Angela, and Brian McGill. 2012. "The Limitations of Hierarchical Organization." *Philosophy of Science*, 79: 120–140.

Ramsey, William. 2016. "Untangling Two Questions about Mental Representation." *New Ideas*

in Psychology 40: 3–12.

Rey, Georges. 2001. “Physicalism and Psychology: A Plea for a Substantive Philosophy of Mind.” In *Physicalism and Its Discontents*, edited by Carl Gillett and Barry Loewer, 99–128. Cambridge: Cambridge University Press.

Robins, Sarah. K. 2017. “Memory Traces.” In *Routledge Handbook of the Philosophy of Memory*, edited by Sven Bernecker and Kourken Michaelian, 76–87. Abingdon, UK: Routledge.

Roth, Martin. 2015. “I Am Large, I Contain Multitudes: The Personal, the Sub-personal, and the Extended.” In *Content and Consciousness Revisited, Studies in Brain and Mind* 7, edited by C. Muñoz-Suárez and F. De Brigard (New York: Springer), pp. 129–142.

Rowlands, Mark. 1997. “Teleological Semantics.” *Mind* 106, no. 422: 279–303.

Rowlands, Mark. 2009. “Extended Cognition and the Mark of the Cognitive.” *Philosophical Psychology* 22, no. 1: 1–19.

Rowlands, Mark. 2010. *The New Science of the Mind: From Extended Mind to Embodied Phenomenology*. Cambridge, MA: MIT Press.

Schroeter, Laura. 2008. “Why Be an Anti-individualist?” *Philosophy and Phenomenological Research* 77, no. 1: 105–41.

Schwitzgebel, Eric. 2010. “Acting Contrary to Our Professed Beliefs or the Gulf between Occurrent Judgment and Dispositional Belief.” *Pacific Philosophical Quarterly* 91, 4: 531–553.

Schwitzgebel, Eric. 2012. “Introspection, What?.” In *Introspection and Consciousness*, edited by D. Smithies and D. Stoljar (Oxford: Oxford University Press), pp. 29–48.

Shea, Nicholas. 2013. “Neural Mechanisms of Decision-Making and the Personal Level.” In

Oxford Handbook of Philosophy and Psychiatry, edited by K.W.M. Fulford, Martin Davies, Richard Gipps, George Graham, John Z. Sadler, Giovanni Stanghellini, and Tim Thornton, 1063–82. Oxford: Oxford University Press.

Shea, Nicholas. 2018. *Representation in Cognitive Science*. Oxford: Oxford University Press.

Spaulding, Shannon. 2010. "Embodied Cognition and Mindreading." *Mind & Language* 25, 1: 119–140.

Stich, Stephen P. 1996. *Deconstructing the Mind*. Oxford: Oxford University Press.

Thalos, Mariam. 2006. "Nonreductive Physics." *Synthese* 149: 133–178.

Thompson, Evan. 2007. "Look Again: Phenomenology and Mental Imagery." *Phenomenology and the Cognitive Sciences* 6: 137–170.

Wilkinson, Sam. 2015. "Dennett's Personal/Subpersonal Distinction in the Light of Cognitive Neuropsychiatry." In *Content and Consciousness Revisited, Studies in Brain and Mind* 7, edited by C. Muñoz-Suárez and F. De Brigard (New York: Springer), pp. 111–127.

Wilson, Mark. 2006. *Wandering Significance*. Oxford: Oxford University Press.

Wilson, Mark. 2010. "Mixed-Level Explanation." *Philosophy of Science* 77: 933–946.

Wimsatt, William C. 1992. "The Ontology of Complex Systems: Levels of Organization, Perspectives, and Causal Thickets." *Canadian Journal of Philosophy*, Supplementary Volume 20: 207–274.