

Personal-Level Representation

Uriah Kriegel Arizona

Forthcoming in *Protosociology: Consciousness and Subjectivity*

Abstract :: The current orthodoxy on mental representation can be characterized in terms of three central ideas. The first is ontological, the second semantic, and the third methodological. After elucidating those, I argue that the emerging picture of mental representation is satisfactory only as an account of mental representation at the sub-personal level. It is unsatisfactory, in a principled way, as an account of mental representation at the personal level.

Introduction

The current orthodoxy on mental representation can be characterized in terms of three central ideas. The first is ontological, the second semantic, and the third methodological. The ontological tenet is that mental representation is a two-place relation holding between a representing state and a represented entity (object, event, state of affairs). The semantic tenet is that the relation in question is probably information-theoretic at heart, perhaps augmented teleologically, functionally, or teleo-functionally to cope with difficult cases. The methodological tenet is that mental representations are posited solely on third-person explanatory grounds.

In this paper, I argue that this picture of mental representation is satisfactory only as an account of mental representation at the *sub-personal* level. It is unsatisfactory, in a principled way, as an account of mental representation at the *personal* level.

A suitable conception of personal-level representation will comprise an alternative triad. Ontologically, personal-level representation is not a *two-place* but *three-place* relation: not the relation *x represents y*, but rather the relation *x represents y to z*. Semantically, a theory of personal-level representation will have to account not only for the *representation-of* component of personal-level representation (the relation between *x* and *y*), but also for its *representation-to* component (the relation between *x* and *z*), and the prospects for doing so in information-theoretic terms are unpromising. Methodologically, personal-level representations are posited not only on *third-person* grounds, but also on *first-person* ones: we admit their existence not only because they perform some explanatory function, but also because we encounter them in our personal experience.

I start, in §1, with an exposition of the mental representation orthodoxy. In §2, I introduce the personal/sub-personal distinction and discuss some options for drawing it with some precision. In §3, I argue for the poverty of the mental representation orthodoxy as applied to personal-level representation. I conclude, in §4, with some remarks on the significance for wider issues in the philosophy of mind.

1. Extant Theories of Mental Representation

Perhaps the most comprehensively articulated theory of mental representation is Dretske's (1981, 1986, 1988) "teleo-informational semantics." I start with an exposition of the theory, then use it to illustrate the aforementioned tenets of orthodoxy.

According to Dretske, a subject harbors a mental representation of, say, the fact that it is raining just in case she is in a brain state whose function is to carry information about the fact that it is raining. To understand what this means, we need to understand (i) what *information* is, (ii) what *carrying* information is, and (iii) what having the *function* of carrying information is.

(i). Every event or fact in the world *generates* a certain amount of information. This information consists, in some sense, in the set of all possibilities the event or fact rules out.¹ Thus, when it is raining, this fact rules out the possibilities that it is dry, that it is snowing, etc. We may think of the information generated by an event or fact in terms

of the narrowing of the set of all possible worlds to the set of all possible worlds consistent with the occurrence of the event or fact in question.

(ii). Some events and facts in the world depend on others, in that they cannot normally take place unless those other events or facts do. Moreover, sometimes this dependence is not accidental, but is due to the laws of nature. That is, some events and facts are *nominally dependent* upon other events and facts. Thus, it is a law of nature (a zoological law, as it happens) that certain kinds of snail do not normally venture out into the open unless it is raining (or has recently been raining). So the fact that Gail the snail ventured out on Sunday at noon is nominally dependent upon the fact that it was raining (or had recently been raining) on Sunday at noon.

This relation of nomic dependence is the key to information carriage. When fact *A* nominally depends upon fact *B*, *A* is said to *carry information* about *B*. More specifically, *A* carries the information generated by *B*. The fact that Gail the snail ventured out on Sunday at noon carries the information (generated by the fact) that it was raining (or had recently been raining) on Sunday at noon.

Of particular interest to us are certain events in, and facts regarding, the brains of sentient organisms. There are laws of nature – neurobiological laws, this time – that dictate which neurophysiological events take place in a brain under which conditions. Thus, it may be a matter of the laws of nature that a type of neurophysiological event *N* in Jim's brain cannot normally take place unless it is raining. If so, *N* is nominally dependent on its raining, and therefore carries the information that it is raining. Thus, if a token *N*-state took place in Jim's brain on Sunday at noon, it would carry the information that it was raining on Sunday at noon. The important point is that the present framework allows us to see how a brain state might be said, in rigorous information-theoretic terms, to carry information.

(iii). Some neurophysiological events in the brain are recruited to play a specific functional role in the cerebral economy of the organism, and some are so recruited in virtue of carrying the information they do. Thus, *N* may be recruited to play the kind of functional role involved in causing Jim to pick up an umbrella or close the window. When a neurophysiological event in the brain is recruited in this way, thanks to the information it carries, it is said to have the *function* of carrying that information.

On Dretske's proposal, a neurophysiological event in the brain that has the function of carrying information about the fact that it is raining *represents* that fact. More generally, whenever a brain event *A* has the function of carrying information about *B*, *A* constitutes a *mental representation* of *B*. *B* then constitutes the *representational content* of *A*. The fact that it is raining is the representational content of the mental state constituted (or realized) by the neurophysiological event nomically dependent thereupon.

Note that this account of mental representation employs physical ingredients exclusively. *A* is a physical event in, or fact regarding, the subject's brain. *B* is an event or fact in the subject's physical environment. The relation between *A* and *B* is (in an appropriate sense) a physical relation. Hence the physicalistic promise of teleo-informational semantics: the promise to account for mental representation in purely physicalistic terms.

It is one thing to offer an analysis of mental representation, quite another to show that there actually are items in the world that satisfy the analysans. A crucial part of Dretske's overall theory is the attempt to show that mental representations, as conceived in teleo-informational semantics, must exist, because they play an indispensable explanatory role. The main idea is that Dretskean representations earn their explanatory keep as structuring causes of behavior. To understand what this means, we must understand what is meant by (a) "behavior" and (b) "structuring cause."

(a). According to Dretske, a behavior, or a piece of behaving, is a bodily motion produced by a suitable internal state of the moving organism. Behavior is thus to be distinguished from sheer bodily movement. The same hand motion can serve (i) to waive hello or (ii) to chase a fly. The bodily motions are type-identical, but the behaviors type-different. The difference is due to the internal state that caused the relevant bodily motion. More radically, the same leg motion can be (i) the kicking of a chair or (ii) the result of a dog running into the leg. The former is a behavior, the latter is not. The difference is that the former was caused by an internal state (of the right sort), whereas the latter was not. The point is that what makes a given behavior the behavior it is, and behavior at all, is that its constituent bodily motion is caused by the right sort of internal state.

(b). A structuring cause is a specific kind of cause of causal connection (a specific kind of second-order cause, if you will). Pressing 8 on my remote control causes my television to switch to Channel 8. Why? Because the TV guy hooked it up this way. The TV guy's actions caused pressing 8 to cause the TV's switching to Channel 8. They structured the causal connection between pressing 8 and switching to Channel 8. In this sense, they constitute the connection's structural cause. Someone could do a number on my remote control, hooking it up to the microwave instead. She could cause pressing 8 to cause the microwave to heat for 8 minutes, rather than cause the TV to switch to Channel 8. If she did, her actions would constitute the structuring cause of a causal connection between pressing 8 and heating for 8 minutes.

Such structuring causes of causal connections are to be distinguished from *triggering causes* of causal connections. If my friend yells at me that the halftime must be over and the game is back on Channel 8, whereupon I press 8 and the television switches to Channel 8, we might say that my friend's actions caused my actions to cause the television's switching to Channel 8. So it is true both of my friend's actions and the TV guy's actions that they caused the cause of the TV's switching to Channel 8. But the senses in which this is true are very different. My friend caused the causing of the effect by bringing about the cause of the effect, whereas the TV guy caused the causing of the effect by bringing about the relation between that cause and effect. Dretske flags that distinction by calling my friend's action the "triggering cause," and the TV guy's the "structuring cause," of the pressing's causing of the switching. The idea is that my friend's actions did not touch the causal relation between my pressing 8 and the TV's switching to 8 – they did not *structure* the causal connection. They merely *triggered* a preexisting causal connection.

Dretske's claim is that mental representations are the structuring causes of behaviors (and are such in virtue their representational content).² When I waive hello, my hand moves a certain way, M , and the cause of its movement is a certain internal state, S , say my mental representation of an acquaintance across the street. Recall that an internal state becomes a representation only when it is recruited to play a certain functional role in virtue of carrying the information it does. Part of S 's functional role is to cause M (in the right circumstances), and S was recruited to play this aspect of its functional role in virtue

carrying the information that an acquaintance is across the street. So the fact that *S* causes *M* is caused by the fact that *S* carries the information that an acquaintance is across the street. The fact that *S* carries the information it does causes *S* to cause *M*. Specifically, it is the *structuring* cause of *S*'s causing *M*. Since *S*'s causing *M* constitutes the hello-waiving behavior, the fact that *S* carries the information it does can be said to be the structuring cause of that behavior.

It is by being the structuring causes of behavior that representational facts – facts regarding a state's representational content – earn their explanatory keep. Mental representations need to be posited because something must do the job of explaining the causal connection between internal states and bodily motions, and facts about what internal states represent are best positioned to do so. Moreover, they do so in continuity with folk psychology, which cites beliefs, desires, perceptions, and emotions – all contentful states – as the causes of behavior.

Dretske's account of mental representation and its explanatory *raison d'être* is not without its critics.³ Some have attempted to modify or replace Dretske's conception of information carriage (e.g., Fodor 1990, 1994), some his conception of function (e.g., Millikan 1984, 1990). More have offered better, certainly simpler, accounts of the way mental representations earn their explanatory keep (Burge 1989). But the criticisms have often been of a technical order, in the sense that they did not undermine the proposition that something along the teleo-informational lines – what Fodor (1994) conveniently labels “The One True Informational Semantics” – combined with an appropriate account of explanatory role, could provide a fully satisfactory theory of mental representation.⁴

Teleo-informational semantics is a particular account of the relationship that holds between *A* and *B* when *A* represents *B*. It specifies what relation *A* has to bear to *something* in order to qualify as a mental representation. What it says is this: for any *x*, *x* qualifies as a mental representation iff there is a *y*, such that *x* bears the teleo-informational relation (the relation of having-the-function-of-carrying-information-about) to *y*.⁵

In this respect, teleo-informational semantics exemplifies a more general approach to mental representation. On this general approach, mental representation is a two-place relation between a representing entity and a represented entity. More precisely:

- (2P) x is a mental representation iff there is a y , such that x bears the appropriate relation R to y .

On this scheme, R is clearly a two-place relation. The relevant relation of representation is the two-place relation “ x represents y ” or “ x is a representation of y .”

It might be objected to this characterization that it does not take into account the fact that teleo-informational semantics, and other theories of the same mold, are supposed to apply to cases of hallucinatory misrepresentation as well.⁶ In such cases, there is no y such that x stands in R to y , and yet x qualifies as a mental representation.

This consideration is well placed, but it will not affect the argument of this paper. So we can simply bypass it – in one of the two ways. The first is to forego the *necessary-condition* part of (2P), focusing on the fact that bearing the right two-place relation is certainly a *sufficient* condition for qualifying as mental representation:

- (2Pa) x is a mental representation if there is a y , such that x bears the appropriate relation R to y .

The second is to restrict (2P) to *veridical* mental representations, thus eliminating hallucinatory cases from the scope of the claim:

- (2Pb) x is a veridical mental representation iff there is a y , such that x bears the appropriate relation R to y .

(2Pa) deletes “only if” from (2P), making it a claim about sufficient conditions only.

(2Pb) adds “veridical” to (2P), making it a claim about veridical mental representations exclusively. For the sake of convenience, we may also formulate the thesis that incorporates both amendments:

- (2Pc) x is a veridical mental representation if there is a y , such that x bears the appropriate relation R to y .

(2Pc) is an unquestionable commitment of extant theories of mental representation. And we can take commitment to (2Pc) as a litmus test for conforming to the ontological tenet of what I have called the mental representation orthodoxy. Dretske's theory thus exemplifies this ontological tenet. This is in addition to exemplifying – nay, being the chief exemplar of – the semantic tenet.

The theory also exemplifies the methodological tenet. To justify the postulation of mental representations, it seeks theoretical grounds on which to do so. That is, it attempts to identify a certain role for them to fulfill in a full objective theory of the world – a role nothing else could fulfill (at least not as well). This attempt might be cast in Quinean light, as the claim that the best theory of the world would (will) quantify over mental representations, therefore we should (will) be ontologically committed to them.⁷

Once such a third-person approach is taken to mental representations, it is not only their *existence* that is dictated by their explanatory role, it is also their *nature*. Once mental representations are treated as explanatory posits, which *properties* we are justified in ascribing to them is also determined by what explanatory role they have to perform. If the performance of that role does not require the assumption that some mental representation instantiates property *F*, then we ought not to construe that representation as *F*. Thus considerations of theoretical serviceability determine not only whether there are mental representations *at all*, but also what *kinds* of mental representation there are.

I have used Dretske's teleo-informational theory to illustrate the ontological, semantic, and methodological tenets of the mental representation orthodoxy. I have chosen Dretske's because it is, in my opinion, the most elegant theory of mental representation we have, and the most plausible version of the orthodoxy. But these opinions are inessential to the point of this paper. The illustration would have gone roughly the same had I employed Millikan's (1984) teleosemantics, Harman's (1987) long-armed conceptual role semantics, Fodor's (1990) version of covariational semantics, Papineau's (1993) success semantics, or any other theory in that genre.⁸ All share the same methodological and ontological assumptions that I want to question, and differ only to a limited degree along the semantic dimension.

I should note that, in characterizing the orthodoxy in terms of the three above tenets, I do not pretend that we can formulate a thesis – the conjunction of the three tenets – and find that the thesis is true on every extant theory of mental representation. Rather, what I have called the mental representation orthodoxy is supposed to capture a certain spirit, a spirit loosely identifiable in terms of this cluster of three ideas.⁹

2. The Personal and the Sub-Personal

One of the distinctions that have been most instrumental in the development and flourishing of the cognitive sciences is between cognition at the personal level and cognition at the sub-personal level. The distinction has been effectively operative in psychological research for about a century, but was drawn explicitly for the first time in Dennett 1969. In this section, I consider a number of possible ways of drawing the distinction with greater precision, including my favored one.

The distinction is of course theoretical, and the terms in which it is drawn are technical terms. So the exercise cannot be one of offering an accurate conceptual analysis of the terms. It is also not an exercise in Dennett exegesis, since it is common that a philosopher grasp for an important distinction that another philosopher ends up articulating better.¹⁰ The exercise is thus to articulate correctly the distinction which Dennett was grasping for. We get an informal handle on what that distinction is supposed to be by considering some paradigmatic examples (that is, ostensibly).

When I calculate the tip I want to leave for a waiter, I deliberately and self-consciously go through a certain process. The process is a cognitive process, and its product is a cognitive state. As such, both process and product are proper objects of the cognitive sciences. Processes and states of this type take place at the personal level: they are personal-level processes and states.

It was an insight of major foundational significance that the proper objects of cognitive science extend beyond such deliberate and self-conscious processes and states taking place at the personal level. When I see an ashtray and mindlessly move it aside, there are a host of processes and states taking place in me, and/or in my visual and

decision-making systems, that deserve to be called cognitive. These processes and states take place at a sub-personal level: they are sub-personal processes and states.

One reason such processes and states deserve to be called cognitive is that they can causally interact with deliberate and self-conscious processes at the personal level. In particular, many sub-personal processes have personal-level products. Thus, when I see the ashtray, a sub-personal process takes place in my visual system, but the process' product is a personal-level visual experience.

Another reason sub-personal processes and states deserve the title of 'cognition' is that they can resemble in essential respects to personal-level processes and states. Thus, a personal-level process and a sub-personal can be computationally very similar. Suppose I decide to leave the waiter a tip of 15-20% on a \$68 bill. My challenge is to calculate what would fall within that range. The very same values could constrain a sub-personal process in my visual system. Perhaps the ashtray subtends a 68 degree angle and I need to move it 15-20 inches leftward. The personal process of calculation and the sub-personal process of computation might proceed identically as far as the applied algorithm is concerned. It is just that, in the former case, it is I, the person, who applies it, whereas in the latter case, it is my visual system, a non-person, that applies it.

These two considerations – and there are probably others – make it unwise, indeed arbitrary, to study cognition by investigating personal-level processes and states exclusively. Furthermore, once sub-personal cognitive processes and states are admitted, conceptually, it quickly becomes clear that, empirically, the personal-level states and processes are only the tip of the cognition iceberg. The bulk of our cognitive life occurs at the sub-personal level. Coming to terms with this fact is perhaps the most important precondition for the kind of fruitful approach to the study of cognition that modern cognitive science has been able to devise.

In the interest of foundational clarity, however, it is also important to have a more precise formulation of the personal/sub-personal distinction. What is the one characteristic we find in all personal-level states and processes and no sub-personal states and processes? Or is the distinction not clearly drawn by appeal to a single characteristic, and is instead better understood in cluster terms?

In my informal presentation of the distinction, I characterized the personal-level process of calculation as one performed “deliberately and self-consciously.” Let us examine the “deliberately” part of this. When I calculate a tip, I am doing something. But when I see an ashtray, I do not do anything. Seeing an ashtray is not something that I do, it is something that happens to me. One suggestion might be that personal-level states and processes are things that we do, whereas sub-personal states and processes are things that happen to us. Call this the *agency approach* to the personal/sub-personal distinction.

One problem with this suggestion is that it does not naturally apply to *states*, only to *processes*. States are never things that we do. Entering a state is something that we might do, but entering a state is a process, not a state. It might be suggested that states be construed as personal when the process of entering them is personal, and sub-personal when that process is sub-personal. But this way of drawing the distinction returns the wrong results. My visual experience of the ashtray is a personal-level state, but would be classified as sub-personal on this way of drawing the distinction, since the process of entering it is sub-personal.

One thought might be that what matters about the fact that the calculation is something that I do is that it is *I* who is doing the doing, not that *doing* is what is done. That is, the important contrast is not between something that I do and something that happens to me, but rather between something that I, the whole system, do and something that some subsystem in me does. The suggestion here is that a state or process is personal if it is a state or process of the overall system; it is sub-personal if it is a state or process of a subsystem. This suggestion has the advantage of elegantly mapping the personal/sub-personal onto the systemic/sub-systemic. It also applies equally well to states as to processes. Call this the *systemic approach* to the distinction.¹¹

A problem with the systemic approach is that the system/subsystem distinction is relative, whereas the distinction personal/sub-personal is not. Arguably, every system but one is also a subsystem, and every subsystem but a handful is also a system. More weakly, many systems are subsystems and vice versa. Thus, I am not only a system, I am also a subsystem of my department. Conversely, my visual system is, well, a system – not only a subsystem of me. Yet we do not wish to say that my tip calculation is a personal

process of me but a sub-personal process of the university, nor that the computation in my visual system is a sub-personal process of me but a personal process of the system.

The suggestion might be modified in a way that overcomes this problems, by drawing the distinction as follows: a state or process is personal if it is a state or process of a system *that constitutes a person*; it is sub-personal if it is a state or process of a system that is a (proper) subsystem of a system that constitutes a person. This appears to have been Dennett's (1969) own take on the distinction.

There are two problems with the modified suggestion, however. First, it draws the personal/sub-personal distinction by explicit appeal to the person/non-person distinction. While this may not render the suggestion entirely vacuous, it does leave it call into question its explicatory cash value. Secondly, and to my mind more importantly, the suggestion still returns wrong results. We can readily envisage a global sub-personal computational process in my brain involving, say, sensorimotor parallel processing across the entire system. Such sub-personal global processes are easy to make sense of, but are impossible on the suggestion under consideration: in virtue of being global, they take place in a system that constitutes a person, and are therefore personal by the suggestion's lights.

In search of a safer suggestion, we might ask ourselves what it is that is missing from the just described sub-personal global process that makes it sub-personal. Intuitively, the fact that I have no idea it is taking place seems relevant. I am completely oblivious to the occurrence and unfolding of this process. We may even coherently suppose that the process is in principle inaccessible to me. This might suggest that the distinguishing feature of personal states and processes is that their subject is in some way, at least minimally, aware of them; sub-personal states and processes are those of which the subject is entirely unaware. On this suggestion, the key to the personal/sub-personal distinction is *awareness*. This would be to focus on the "self-consciously" part of the informal characterization of personal-level states and processes as occurring "deliberately and self-consciously." Call this the *awareness approach* to the personal/sub-personal distinction.

For the awareness approach to be at all workable, however, we must assume that awareness comes in degrees. You can be *very* aware of the fact that you are jealous of

your wife's new friends or only *dimly* aware of that fact. The degree of awareness might be determined by, or at least correlate with, attention: you can be very attentively aware of your jealousy or relatively inattentively aware of it. The awareness approach would be extremely implausible if we required of personal-level states and processes a high degree of awareness.¹² However, on the assumption that awareness comes in degrees, we might require of personal-level states and processes only the most minimal degree of awareness – any awareness greater than zero, so to speak. The suggestion then becomes much more plausible. When I am engrossed in the calculation of the tip, I am not being particularly attentive to the process of calculation. I am mostly aware of the numbers I am manipulating. Yet it would be a mistake to say that I am entirely unaware of the activity of calculating itself. I am aware of it, albeit only dimly and inattentively. It is not as though I have no idea that I am engaged in calculating, and the news that I am would surprise me as much as anyone.¹³

In addition, for the approach to be plausible the awareness of the personal state must be roughly simultaneous with that state. Future awareness of a mental state through memory clearly would not make the state a personal-level one.

The suggestion under consideration, then, is that a state or process is personal if its subject is simultaneously at least *minimally aware* of it; it is sub-personal if its subject is *completely unaware* of it. This suggestion returns the right results in all the cases we have considered thus far. The tip calculation is personal, and indeed I am minimally aware of it. My visual system's computation, leading to my seeing of the ashtray, is sub-personal, and indeed I am completely unaware of it (nor is my visual system aware of it, since my visual system is not the kind of thing that can be aware of anything). The seeing of the ashtray itself is personal, and indeed I am aware of it. The global parallel processing is sub-personal, and indeed I am unaware of it (as, again, is the global cognitive system that constitutes me).

The present suggestion is not without its problems, however. For one thing, there seem to be certain cases of absent-minded perception (as in the long-distance truck driver's perceptions of the road) that are certainly personal but of which the subject is, arguably, completely unaware. For another, the notion of awareness is not one of the

clearest in the philosophy of mind. The somewhat robust notion being availed of here, with its gradient structure and its minimal cases, may be thought particularly worrisome.

We might make some progress by turning the second problem against the first. Is it not possible that, although not remotely attentive to her perceptual states, the long-distance truck driver is nonetheless very dimly aware of them? We might hold, for instance, that the driver is minimally aware of such absent-minded perceptions, but they leave no mark on working memory (for whatever reason). Such claims are hard to assess, and this may augment our suspicion that the operative notion of awareness is obscure. Let me raise one brief consideration, though. Suppose that, through a cosmic incident, the driver becomes instantaneously color-blind during her trip. Plausibly, she would at that very instant notice the change in her perceptual experiences. Also plausibly, though much more controversially, she could not notice the change at that instant had she not been at least minimally aware of her perceptual experiences all along. It would follow that the driver was aware of her perceptual experience all along, albeit minimally.¹⁴

Another objection is that the awareness approach returns the wrong results for *standing states*. Jane has the standing belief that there are more than four countries in the world. But until we draw her attention to it, she is completely unaware of it. Yet it is Jane, the person, who believes that there are more than four countries in the world, not any subsystem within her. Thus Jane's standing belief is in some sense a personal-level state of which she is completely unaware.

Personally, I am not so impressed by this objection, because I am independently inclined to reject the existence of standing states. Consider the view that there are no *dispositional beliefs*, only *dispositions to believe*. The idea is that there is no need to posit such dispositional beliefs, since dispositions to believe would be sufficient to do any explanatory work we might want dispositional beliefs to do – and do so more economically (see Manfredi 1993). On this view, Jane has the disposition to believe that there are more than four countries in the world, but it does not follow that there is a dispositional belief to that effect that she has. By extension, we might claim that persons do not have standing states, only dispositions to enter occurrent states. If there are no standing states, *a fortiori* there are no personal-level ones.

There are other responses to the objection under consideration. One is to admit standing beliefs but deny that they are personal-level states.¹⁵ Another is to claim that although Jane has no *occurrent* awareness of her standing belief, she does have *standing* awareness of it.¹⁶ I will not elaborate on these, since as I take the very belief in standing states to be theoretically unmotivated.¹⁷

Yet another way of drawing the personal/sub-personal distinction might be in terms of *consciousness*. On this suggestion, conscious states and processes are personal, unconscious ones are sub-personal. Call this the *consciousness approach* to the distinction.

This approach faces a dilemma, however. There are two main ways to construe the notion of consciousness. One is in terms of sensuous qualities, another is in terms of awareness. If we adopt the first option, the approach returns the wrong results; if we adopt the second, it collapses into the awareness approach.

On the first construal, a mental state's being conscious is a matter of its exhibiting a qualitative character or sensuous quality, such as redness or bitterness.¹⁸ On this construe, it is simply false that all personal-level states and processes are conscious. Thus, there need not be a sensuous quality implicated in calculating a tip.¹⁹ More generally, while many *states* are conscious in this sense, it is difficult to see how *processes* could be said to be. Furthermore, even among states, there may well be personal-level states that are unconscious in the relevant sense. Thus, the occurrent personal-level thought that $2+2=4$ need not exhibit any qualitative character.²⁰

On the second construal of consciousness, conscious states are those that satisfy what Rosenthal (1993) calls "the transitivity principle": conscious states are states we are aware of. In this sense, the claim that all conscious states are personal and all unconscious ones are sub-personal is much more plausible, and does not obviously return wrong results. But it is not interestingly different, perhaps not different at all, from the awareness approach already discussed.

Thus we can, in fact, take this brief discussion of the consciousness approach to the personal/sub-personal distinction to provide further support for the awareness approach.

There is also the possibility that the notions of the personal and the sub-personal are cluster notions, in which all or most of the features considered above play some role or another. Thus, we might hold that prototypical personal-level states and processes are agentive states/processes of a global cognitive system that constitutes a person and of which the person is aware or conscious; and that non-prototypical personal-level states and processes are those that resemble the prototypical ones to a sufficient degree. Call this the *cluster approach* to the personal/sub-personal distinction.

This suggestion strikes me as ill-motivated in light of our discussion, however. A cluster view is called for when a number of competing suggestions seem to capture the spirit of what one is after, but happen to return wrong results at the margins. The problems we have raised for most of the above suggestions go to the heart of their plausibility as capturers of the target distinction's spirit. Moreover, the suggestions in question returned the wrong results not merely at the margins, in odd and artificially concocted cases. They returned the wrong results on straightforward core cases.

For these reasons, the awareness approach seems to me preferable to the cluster approach. Although admittedly imperfect, due mainly to an element of obscurity associated with the notion of awareness, the awareness approach appears to be the most plausible of the suggestions we have examined, in that the only case where it might be suspected of returning the wrong result (that of standing states) is a case that may not even be real. Principled considerations seem to support rather than undermine the awareness approach. In what follows, I will proceed with it as my working assumption about the personal/sub-personal divide. We may state it as follows: a mental state S (or process P) of a subject Z is a personal-level state (or process) just in case Z is at least simultaneously minimally aware of S (or P); S (or P) is a sub-personal state (or process) just in case Z is completely unaware of S (or P).

3. Mental Representation at the Personal Level

One way in which cognitive science has greatly benefited from the personal/sub-personal distinction is through the realization that even paradigmatically personal-level states,

such as visual experiences, have sub-personal counterparts that can be just as cognitively potent, so to speak, and sometimes even more so.

Examples are by now legion, and they come from studies of subliminal perception, visual masking and priming, brain lesion, and more. A striking example emanates from the widely accepted hypothesis that the visual system has two computationally independent pathways, the dorsal stream and the ventral stream, both of which construct concurrent representations of the visible environment (Milner and Goodale 1995; see Figure 1 overleaf). The interesting fact is that the representations produced in the dorsal stream are always sub-personal. The subject is never aware of them, and in fact *cannot become* aware of them. At the same time, they turn out to be more cognitively potent than their personal-level counterparts in the ventral stream, at least in this respect: online, on-the-fly visually guided action turns out to be determined by dorsal stream (sub-personal) representations, not ventral stream (potentially personal) ones.

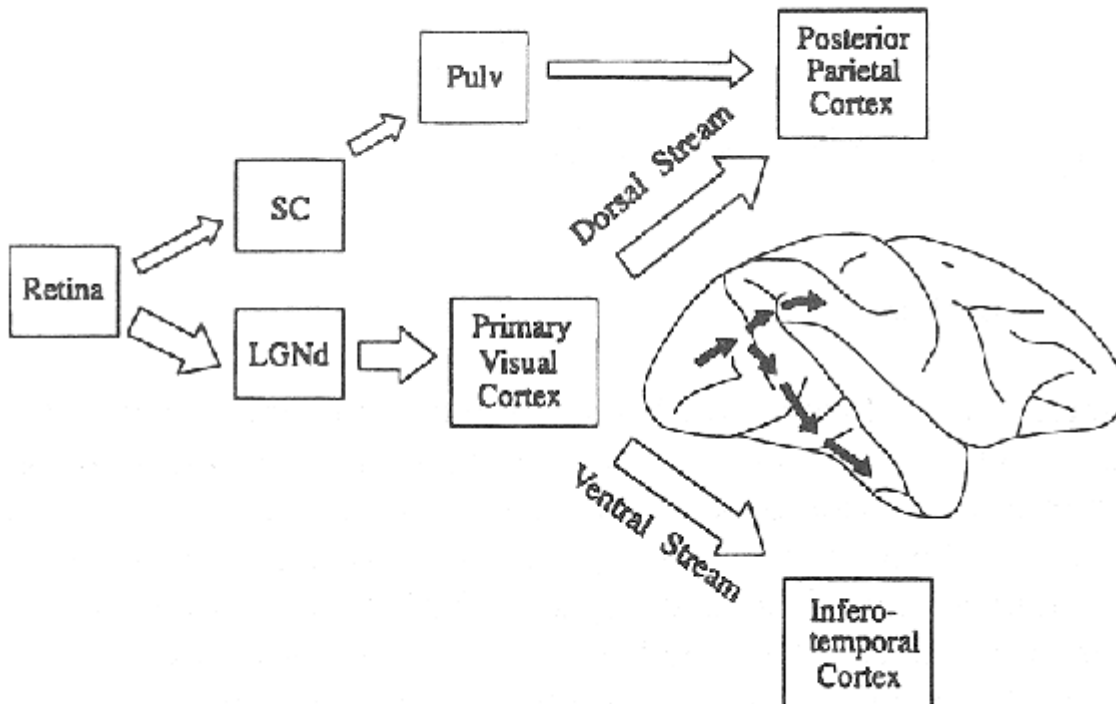


Figure 1. The Ventral and Dorsal Streams in the Visual Cortex

This can be seen in cases of Gestalt-effect-induced visual illusions that do affect our personal-level visual experiences but do not infect the sub-personal representations that guide our actions on the fly. Thus, in the Titchener illusion, a circle appears bigger when surrounded by smaller circles than when surrounded by bigger ones (see Figure 2 overleaf). A normal subject presented with two equisized circles differently surrounded will perceive one to be bigger than the other. More cautiously, we should say that she will perceive one to be bigger than the other *at the personal level*. At the sub-personal level, however, the perceptual system represents the circles to be equisized. For as it turns out, when subjects are asked to reach for a coin, they do *not* adjust their grip aperture (measured by the distance between the thumb and the index finger) in response to surrounding coins. That is, their grip apertures are the same for differently surrounded equisized coins. This suggests that, at some point in the visual processing, some representations of the coins' size are formed that are not infected by the Titchener illusion, and moreover, that it is these representations that guide the reaching and grasping behavior. The going hypothesis is that they are dorsal-stream visual representations (Ibid.). These are not representations we are aware of at the personal level, however. They are sub-personal representations.

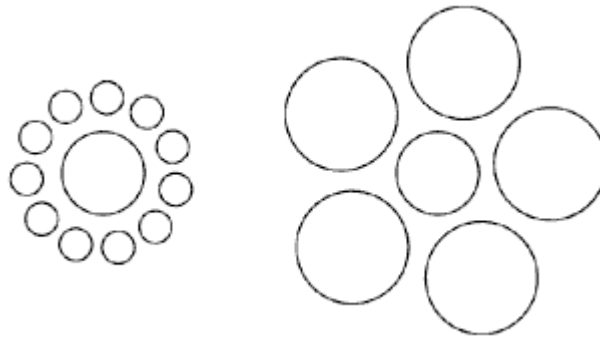


Figure 2. Titchener's Illusion

These findings force us to posit cognitively potent sub-personal representations. Such representations often have causal effects on personal-level representations. A

particularly dramatic illustration is the phenomenon of primesight, in which a blindsight patient who “blindstares” at a color stimulus long enough experiences after-images of complementary color (Weiskrantz et al. 2002, 2003). But there are innumerable others.

Findings of this sort are relatively recent, and their conceptualization is tremendously helped by the recognition of the personal/sub-personal distinction. Similar findings were thought immensely puzzling a century ago, when subliminal effects were first registered by Sidis (1898), Dunlap (1900), and others. For example, Dunlap managed to produce the Müller-Lyer effect at the personal level using masked angular arrowheads that were *not* visible at the personal level. Subjects presented with two lines of equal length, but accompanied by different shadowed arrowheads, reported that one line appeared longer than the other, despite not reporting seeing the arrowheads and indeed reporting not seeing the arrowheads (Dunlap 1900; see Figure 3 overleaf). The natural treatment of such a situation is to suppose that a sub-personal representation of the arrowheads was nonetheless formed, and interacted with the personal-level representations to generate the illusory effect at the personal level. That is, we have both visual representations of which we are aware and ones of which we are unaware, and the two interact. This is straightforward material today, but was considered a deep puzzle in Dunlap’s days.

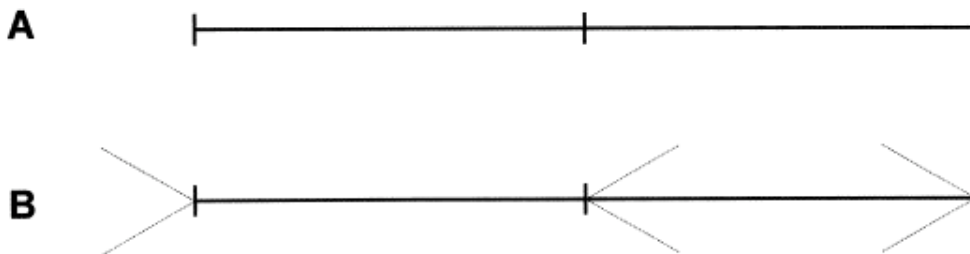


Figure 3. Stimuli used by Dunlap (1900: 436)

The point is that numerous experimental findings, rigorously obtained and diligently replicated, are bound to leave us bemused unless we posit sub-personal visual representations, which moreover cognitively interact with personal-level ones and are anyway independently cognitively potent vis-à-vis behavior. We may also assume that

many sub-personal visual representations overlap in content with personal-level representations, but that sometimes slight differences in content lurk among otherwise similar representations (as is the case with dorsal- and ventral-stream representations of Titchener coins). Finally, it would not be overly imprudent to suppose that personal-level visual representations are only the tip of the vision iceberg, and that the great majority of visual representations are sub-personal. (Thus, the great majority of representations posited in Marr's (1982) seminal computational model of vision – including his crucial “2.5D sketches” – are of the sub-personal variety.) To repeat, such suppositions have been vital to the inception and flourishing of cognitive science.

However, the felicitous introduction of sub-personal representations, and the center stage they have received in virtue of forming the massive bulk of the cognition iceberg, may have led to an undue neglect of personal-level representations. My contention is that the special properties of personal-level representations have been by and large either denied or ignored by philosophers and theoretical psychologists working on the foundations of cognitive science. And this in turn has led to theorization about mental representations as though either they were all sub-personal or they were essentially the same as sub-personal representations. Yet the most cursory examination of personal-level representations suggests that they may have certain characteristics that are both different and important.

3.1. *The Ontological Tenet*

Start with the ontology of representation. I will now argue that although personal-level representations may well be involve a two-place relation x represents y , and so orthodoxy is well-positioned to account for their ontology, personal-level representations involve the more complex three-place relation x represents y to z , where z ranges over persons.²¹ I start with some observations of how natural and fruitful it is to conceive of personal-level representation in this way, then turn to a more formal argument.

A personal-level representation is generally a representation of something *to someone* – to a *person*, that is to say. As I have my experience my visual experience of the laptop before me, I am in an internal state that represents the laptop *to me*. As you read these words, you are in an internal state that represents these words *to you*.²²

Now, when I look at a duck-rabbit figure, the figure represents to me both a duck and a rabbit. Suppose, however, that Jane has suffered a duck-related childhood trauma that causes her to repress all incoming duck-related information. It is natural to say that the same figure represents a rabbit to Jane, but does not represent a duck to her. Plausibly, this is because while she has sub-personal representations of the duck, she has no personal-level ones. Presumably, to know that there is something in need of repression, Jane's cognitive system must at some point recognize that a duck is presented. It must identify the threat. So at some processing stage, Jane must host a sub-personal representation of the duck. The representation is sub-personal because Jane is completely unaware of it. But although it represents the duck, this representation does not represent the duck *to Jane*. At the same time, some other representation – a personal-level one – does represent the rabbit *to Jane*.

Similarly, suppose John suffered a brain lesion, in the wake of which he is unable to recognize ducks (at the personal level). To him too, the figure represents a rabbit but not a duck. As it turns out, however, John's brain lesion incapacitated only the ventral stream in his visual cortex, sparing the dorsal stream. As a result, John is fully capable of hosting dorsal-stream representations of ducks, in spite of his inability to have ventral-stream ones.²³ These represent ducks, but not *to John*. As it happens, this type of scenario is not mere fancy. There are prosopagnostics (subjects who cannot recognize faces) who show signs of *covert* face recognition (revealed in priming tasks). This suggests that although at the personal level they do not have representations of face identity, they do have some such sub-personal representations. Bauer (1984) has hypothesized, on the basis of neuroanatomical and neurocognitive studies of one of his patients, that sometimes this is due to the fact that the dorsal stream was not affected by the brain lesion that put the ventral stream out of commission.

Upon reflection, it is plausible to say that even in the mundane case where *I* have a visual experience of the duck-rabbit figure, I cannot see it as a duck and as a rabbit *simultaneously*. Rather, the figure switches back and forth from duck to rabbit. It seems natural to describe this by saying that I undergo a succession of perceptual states that at first represent a duck to me, then represent a rabbit to me, then again represent a duck to me, and so on.

These facts – about Jane’s, John’s, and my personal-level representations as we look at the duck-rabbit figure – are described very naturally in terms of three-place representations. I am not ruling out that they might also be describable in terms of two-place representations. But such description is likely to be more cumbersome, incomplete, and less natural. In any case, these remarks are intended as illustration, not argumentation. They illustrate the notion of three-place representation, and how natural it is to use it in describing personal-level representations.

Note well: I am not denying that personal-level representations instantiate the two-place relation that extant theories of mental representation attempt to account for. I am merely insisting that, *in addition*, they also instantiate a three-place relation.

I now wish to make further claim, however. It is that unless a representation instantiates this three-place relation, it does not qualify as personal. That is, it is a necessary condition on personal-level representations that they represent *to the person* whose representations they are. The argument turns on the previous section’s awareness approach to the personal/sub-personal distinction.

Recall that the distinguishing mark of personal-level states is that their subject is aware of them. This awareness of a representational state might be thought to secure the state’s representing what it does to the subject who is aware of it. To a first approximation, then, the argument might be formulated as follows:

For any person P , representational state S , and content C , such that (i) P has S and (ii) S represents C ,²⁴

- 1) S is a personal-level representation only if P is aware of S ;
- 2) If P is aware of S , then S represents C to P ; therefore,
- 3) S is a personal-level representation only if S represents C to P .

The problem in this formulation is with Premise 2. If P is aware of S , but unaware that S represents C , then P ’s awareness of S would do nothing to secure S ’s representation of C to P .

This obstacle can be overcome, however, if we claim that personal-level representations must be transparent or diaphanous in the sense in which perceptual

experiences, as well as beliefs, are often – and in my view, justifiably – said to be.²⁵ In the relevant sense, for a representation to be transparent is for it to be such that one can only be aware of it by being aware of its content. This principle would justify the proposition that *P* cannot be aware of *S* unless *P* is aware of *S*'s representing *C*, or of *C*'s being the content of *S*. This in turn is equivalent to the proposition that, if *P* is aware of *S*, then *P* is aware of *S*'s representing *C*. This proposition plugs the whole in the argument and enables the following reformulation:

For any person *P*, representational state *S*, and content *C*, such that (i) *P* has *S* and (ii) *S* represents *C*,

- 1) *S* is a personal-level representation only if *P* is aware of *S*;
- 2) If *P* is aware of *S*, then *P* is aware of *S*'s representing *C*; therefore,
- 3) *S* is a personal-level representation only if *P* is aware of *S*'s representing *C*;
- 4) If *P* is aware of *S*'s representing *C*, then *S* represents *C* to *P*; therefore,
- 5) *S* is a personal-level representation only if *S* represents *C* to *P*.

Premise 2 here is guaranteed by the transparency of personal-level representations.

Premise 1 is the result of the previous section's discussion. And Premise 4 sounds more or less like a conceptual truth.

The conclusion of this argument undercuts the ontological tenet of the mental representation orthodoxy. Recall that the tenet was formulated, minimally, as follows:

(2Pc) *x* is a veridical mental representation if there is a *y*, such that *x* bears the appropriate relation *R* to *y*.

If this is supposed to apply to personal-level representations as much as sub-personal representations, then we might introduce that fact explicitly:

(2Pd) *x* is a veridical personal-level mental representation if (i) there is a *y*, such that (ii) *x* bears the appropriate relation *R* to *y*.

(2Pd) clashes with the conclusion (5) of the above argument. That conclusion entails the following:

- (3P) x is a veridical personal-level mental representation only if (i) there is a y , and (α) there is a z , such that (ii) x bears the appropriate relation R_1 to y , and (β) x bears the appropriate relation R_2 to z .

(2Pd) claims that clauses (i) and (ii) are a sufficient condition for x 's being a (veridical) personal-level mental representation. But (3P) denies that: it claims that (α) and (β) are also necessary, so that any combination of conditions that did not include (α) and (β) would be insufficient.²⁶

I have argued that personal-level representation must instantiate the three-place relation x represents y to z . If so, a personal-level representation, such as my visual experience of the laptop, instantiates two *two*-place representation relations. My experience bears the *representation-of* relation to the laptop and the *representation-to* relation to me. A full theory of personal-level representation would comprise two chapters: an account of representation-of and an account of representation-to.²⁷

Discussion of the representation-to component of representation is all but non-existent in the literature surrounding the mental representation orthodoxy. In one place, Dretske (1988: 55) considers summarily the claim that all representation is representation-to, which he dismisses by citing so-called natural signs. The rings on a tree represent the tree's age, and they do so whether or not anyone *takes* them to represent the tree's age (Stampe 1977). Such natural signs represent what they do regardless of what they are *taken* to represent, indeed of *whether* they are taken to represent. The tree rings would represent the tree's age even if nobody took them to represent the tree's age or everybody took them to represent the tree's height.

Dretske's point is well taken, but the claim I am pressing here is not meant to apply to all representations, only to personal-level ones. Rings on a tree represent, but not *to the tree*. They do not constitute personal-level representations. For personal-level representation, instantiation of the three-place representation relation is necessary.

3.2. *The Semantic Tenet*

With the exception of Ruth Millikan, whose views will be discussed toward the end of this subsection, I am not aware of any other comment by a proponent of extant theories of mental representation that addresses, or even acknowledges, the three-place representation relation. The representation-to component of personal-level representation has thus been a constant blindspot in those theories. To the extent that they do not address it at all, they cannot possibly constitute adequate accounts of it, and therefore of personal-level representation.

It is a separate question, however, whether the orthodoxy has the *resources* to account for representation-to. Even if no explicit attempt has been made to account for representation-to in informational, teleological, and/or functional terms, it may still be thought that it should be fairly easy to do so once one tries. In this subsection, I will examine some tempting avenues and indicate initial difficulties. I do not argue that extant theories categorically lack the resources to account for representation-to. I only wish to argue that it is not entirely obvious that they do not: it is not straightforward how the informational and teleological tools could be beneficially applied to the notion of representation-to. (In any case, it is quite a telling fact that no serious attempt has been made to account for representation-to. It betrays the way in which extant theories have clung tightly to the methodological and ontological tenets of orthodoxy.)

The proponent of the orthodoxy might attempt to deflate the challenge by arguing that representation-to amounts to nothing more than *ownership*. For a state or feature to represent something to someone is simply for that state or feature to represent what it does and be a state or feature *of* the relevant “someone.” On this account, “*x* represents *y* to *z*” can be analyzed as follows: (i) *x* represents *y* and (ii) *x* occurs in *z* (or: *x* “belongs to” *z* in the relevant sense).

However, this deflationary account is highly implausible. The tree rings (i) represent the tree’s age and (ii) are a feature of the tree, but they do not represent the tree’s age *to the tree*. So conditions (i) and (ii) can be fulfilled even when representation-to does not occur.

A modified deflationary account might restrict the analysans to persons, so that “*x* represents *y* to *z*” is analyzed as (i) *x* represents *y*, (ii) *x* occurs in *z*, and (iii) *z* is a person.

Nothing represents anything to trees, but internal states of persons can represent to them just in virtue of being theirs.

This analysis is still implausible. States of a person's skin can represent some environmental feature without representing it *to* the person. Thus, Jim's goosebumps represent that the ambient temperature is below 6°C. The state of being goosbumped thus (i) represents the ambient temperature and (ii) occurs in Jim, and (iii) Jim is a person. Yet it does not represent the temperature *to* Jim.²⁸

Even if we restrict ourselves to *mental* internal states, persons have unconscious states that represent but not *to* them. Blain the blindsight patient has an unconscious perceptual representation of an orange in his left visual field. Blain's perception (i) represents the orange and (ii) occurs in Blain, and (iii) Blain is a person. Yet clearly the state does not represent the orange *to* Blain.²⁹

The deflationary route is thus unpromising. A proponent of the mental representation orthodoxy would do better to try to accommodate the phenomenon rather than deflate it, using her informational, teleological, and functional tools.

One thought might be to employ some sort of second-order information carriage. The idea would be to account for representation-to in terms of second-order representation-of (which would be understood in terms of a second-order state that carries information, and has the function of carrying information, in the exact same way that first-order representations-of do). On this proposal, x represents y to z just in case (i) x represents y and (ii) z has a representation of x 's representation of y . More explicitly, we might say that x represents y to z iff (i) x represents y and (ii) z has an internal state w , such that w represents x 's representation of y .

The problem this informational account of representation-to faces is best brought out by a dilemma. Is w (the representation of x 's representation of y) a personal-level representation or a sub-personal representation? Neither answer is satisfactory. The first horn, where w is a personal-level representation, is doubly problematic. First, it launches a regress of personal-level representations. Secondly, it appeals to personal-level representation in accounting for the same, and is to that extent vacuous. The second horn is also problematic, though perhaps less flagrantly so. The main challenge it faces is to account for the appearance that all the components of personal-level representation are at

the personal level. Thus it seems that we are minimally aware, rather than completely unaware, of the fact that personal-level representations represent *to the person* who has them. It is possible that this challenge (and others that may arise) can be met by this horn of the informational approach. But more work must be done to show that this is so.

Another resource in extant theories' toolbox may be teleo-functional. I am thinking in particular of the sort of "consumer semantics" developed initially by Millikan (1984, 1989) and further by Carruthers (1996, 2000). Millikan (1989: 284) writes that a representation "must be one that functions as a sign or representation for the system itself," and proceeds to offer a functional account of this feature of representation. The suggestion is that a key aspect of a representation's functional role is its availability for consumption by certain downstream modules. Perhaps this could be cashed out in terms of availability to an executive control module. Perhaps it would be a matter of *global* availability to a number of high-profile, high-sophistication modules. But in any case, some such special and/or widespread consumability might account for representation-to. One might thus suggest that an internal state represents *to* the subject when it is readily available for use *by* the subject, and that personal-level representations enjoy heightened or sharpened availability.

This strikes me as the most promising avenue for extant theories to explore. There are nonetheless serious problems with the idea. For one thing, we can readily imagine a state that is available for consumption by any and every module in the subject's cognitive system, while remaining entirely outside the subject's awareness. This would be a sub-personal state that satisfied the consumer semanticist's requirements. It follows that these requirements do not specify a sufficient condition for being a personal-level representation.³⁰

Furthermore, having a certain functional role is a *dispositional* property, but being a representation-to seems to be an *occurrent*, hence non-dispositional, property. Thus, for a state to be *available* to certain modules, it need not actually be *availed of* by them. But my visual experience's representing of the laptop *to me* is not just a matter of things *possibly* happening. Something does actually happen when my experience represents the laptop *to me*.³¹

A consumerist account of representation-to might nonetheless be worth exploring. I cannot say with confidence that, everything said and done, it would fail (or succeed!) to offer a reasonable treatment of the notion. For that matter, the second-order information strategy might be worth exploring as well. Nonetheless it is an illuminating fact about extant theories that the explorations in question remain to be pursued, and have not been attempted with sustained seriousness. That fact by itself casts the current orthodoxy as inadequate and incomplete.

Observe that I have taken pains not to characterize the extant theories as naturalist. This is because although they are, it is not their naturalism that I want to impugn.³² None of the above undermines naturalism *per se*. As a naturalist myself, I certainly hope that personal-level representations prove amenable to naturalist treatment. But deflating the phenomena in order to make them thus amenable could only postpone that achievement.³³ Casting personal-level representations as essentially the same as sub-personal ones, and ignoring their representation-to component, strikes me as an instance of such undue deflation. The consequence is that adequate resources for accounting for representation-to have never been seriously developed within the mental representation orthodoxy.

3.3. *The Methodological Tenet*

In the nature of things, sub-personal representations can be posited only on purely theoretical grounds. We have no reason to believe in their existence apart from the sense their existence would make of observable behavior, that is, apart from their explanatory payoff. Accordingly, the only properties we can ascribe to them are those properties they would need to possess in order to carry out their explanatory job. Thus both their *existence* and their *nature* are fully determined on third-person grounds.

Not so with personal-level representations. Their existence can be recognized independently of any theoretical need to posit them. We encounter them in our personal experience and are presented with them all day long.³⁴ No theorizing need be implicated in our assumption that they exist. The reason I believe that I have a representation of my laptop right now is *not* that believing so would make sense of my hands' motion. It is simply that I am aware of this representation in real time.

So, whereas my sub-personal (dorsal-stream) representation of a Titchener coin is one I believe in only on the strength of third-person evidence, my personal-level (ventral-stream) representation of the same coin is one I believe in simply because I am aware of it – I am presented with it in my experience, whether I like it or not. This awareness is not infallible, and I may not have a representation I seem to myself to be having. But the seeming does provide the initial (defeasible) evidence on the basis of which I believe that I have a visual experience of the coin, and moreover, that the experience as *contentful*, that is, is an experience *of a coin* (or at least *of something*).³⁵ The seeming is thus a sort of first-person evidence of having personal-level representations, which is categorically different from the third-person evidence we have of having sub-personal representations.³⁶

Consider Strawson's (1994) weather-watchers, a hypothetical species of creatures who cannot move any part of their bodies, but can nonetheless sense changes in the weather and take great interest in them. They are excited when unexpected temperatures arrive to their locale. They agitate over the prediction of tomorrow's weather and try hard to remember that of days past, so that they may recognize patterns of warming and cooling, of which they are delighted to construct elementary mathematical models. These creatures have a mental life that is not only narrowly cognitive, but involves also emotion, memory, sensation, inductive reasoning, and more. In the conduct of this rich mental life, they deploy a panoply of mental representations (including, very possibly, sub-personal ones). However, while a weather-watcher has no evidence on which to ascribe *sub-personal* representations to her conspecifics or even herself, each weather-watcher has all the evidence she needs to ascribe *personal-level* representations to herself. Naturally, the evidence in question is first-person evidence.³⁷

The claim I am making is straightforward: we do not posit personal-level representations *solely* for theoretical reasons. If anything, we would need theoretical reasons to *unposit*, if you please, personal-level representations. After all, the initial appearances are that they do exist. These appearances are not incontrovertible and self-awareness is notoriously fallible. All the same, the initial appearance that a personal-level representation exists does provide initial, *prima facie*, and *de tanto* evidence that it does.

The evidence is defeasible, but no less evidence for that. Until defeated, such evidence would remain the last word on the subject.

There is a crucial disanalogy here with sub-personal representations. With respect to the latter, the initial appearances are that they do *not* exist – which is why theoretical considerations are needed to posit them. In any event, the evidential force of initial appearances is the same in the personal and sub-personal cases, but the appearance itself is different in each case: in the personal case, the initial appearance is of existence; in the sub-personal case, it is of non-existence.

I cannot stress enough that none of this presupposes heady Cartesian assumptions about infallibility, or for that matter the slightest special authority or privileged access. The only presuppositions operative in these considerations are, first, that there is such a thing as first-person evidence, and second, that it is distinct from third-person evidence. These presuppositions are minimal and hardly deniable. Right now I am visualizing a flying kangaroo. Having disclosed that this is so, both you and I know that a flying kangaroo is what I am right now visualizing. Yet it is clear that there is a fundamental difference in *how* we know this, and on what evidence we have come to know this. The way I know it is first-person and my evidence is first-person. The way you know it is third-person and the evidence you go on is third-person. None of this is particularly exotic, nor harmlessly deniable.³⁸

Because personal-level representations are not posited solely on explanatory grounds, the properties we attribute to them are not restricted to those that are necessary for them to perform their designated explanatory role. Since their *raison d'être* is not purely explanatory, we do not attribute to them only their explanatorily indispensable properties, but also the properties they present themselves to have when we encounter them in our experience. More loosely, theorizing about the nature of personal-level representation is simply not as hermetically constrained by explanatory considerations as theorizing about sub-personal representation is.

At the risk of seeming to overlabor the point, let me stress again that this is not to say that every fleeting impression we might have of a property of a personal-level representation provides equal warrant, let alone absolute or incontrovertible warrant, for ascription of the property in question. In fact, we could even insist that it is *never* justified

to ascribe to a personal-level representation a property that is *explanatorily detrimental*, that is, a property that undermines the representation's explanatory efficacy – even if the ascription of such a property is strongly suggested by first-person appearances. This would not only deny special authority to the first-person perspective, it would actually grant such authority to the third-person perspective. But this still allows that, with a personal-level representation, property ascription could be warranted even when it is not third-personally mandated, namely if it is first-personally encouraged. The result would be the following methodological principle:

- (M) For some representation R and property F , if R seems to be F (first-personally), and ascription of F to R is not explanatorily detrimental (third-personally), then one is justified in ascribing F to R .

Thus whereas with sub-personal representations, any ascription of a property that is not explanatorily useful is entirely groundless, in the case of personal-level representations it may not be.³⁹

Principle (M) allows us to see why a first-person-friendly approach to mental representation can be responsibly adopted, without falling into familiar Cartesian pitfalls. The approach would merely insist that, if I have a strong and stable first-person impression that my personal-level visual representations of kind K instantiate property F , then in some circumstances I might be justified in saying that they do – even if saying so contributes nothing to their explanatory dexterity. This is not the case with sub-personal representations. To them it is warranted to ascribe only those properties that contribute toward the satisfactory performance of the explanatory job they were posited to perform.

(M) is a relatively conservative principle we could adopt by way of insisting that the theory of personal-level representation is not hermetically constrained by third-person explanatory considerations, but is open to first-person, non-explanatory considerations. It is conservative in that it gives special authority to the third-person perspective: it dictates that we ascribe to mental representations, third-personally, any explanatorily useful property, and first-personally, any explanatorily non-detrimental property they seem to have. Other reasonable methodological principles might give a greater place to first-

person considerations. For example, we could treat the theory of personal-level representation as seeking a *reflective equilibrium* with regard to the full scope of third- and first-personal evidence.

(M*) For some representation R and property F , if the balance of first-person *and* third-person evidence suggests that R seems to be F , then one is justified in ascribing F to R .

How exactly (M*) plays out depends on how we cash out “balance.” But (M*) differs from (M) in that the third-person considerations do not automatically trump the first-person ones. For example, when the first-person seeming is very powerful and the explanatory detriment very limited, according to (M*) it might be reasonable to let the first-person consideration trump the third-person one.

Whichever principle we adopt, the upshot is that the theory of personal-level representation is not hermetically constrained by third-person explanatory considerations, but is open to first-person, non-explanatory considerations. If this is right, then the methodological tenet of the mental representation orthodoxy does not apply to personal-level representations. While there are good reasons to posit sub-personal representations that are nothing but explanatory role-occupants or placeholders, and while the latter may constitute the massive bulk of the cognition iceberg, there are nonetheless personal-level representations as well, and those are not mere explanatory posits. Whether or not they explain anything (doubtless some do, some don’t), belief in their existence is forced on us by their presence in our ordinary experience, and beliefs about their nature are *not* forced on us by explanatory considerations only.⁴⁰

Extant theories of mental representation do *not* take into account the first-person considerations, but are instead hermetically constrained by the third-person explanatory considerations. In that respect, they treat the theory of mental representation as the theory of *sub-personal* mental representation ought to be treated. Consequently, they may offer an adequate account of sub-personal representation, but are unlikely to offer an adequate account of personal-level representation.

4. Conclusion: Representation and Intentionality

I conclude that personal-level representation defies all three tenets of the mental representation orthodoxy. Contrary to the ontological tenet, personal-level representation is not (merely) a two-place relation holding between a representor and a represented, but (also) a three-place relation between a representor, a represented, and a represented-to. Contrary to the semantic tenet, it is not obvious that all components of this three-place relation could be equally well handled by a suitably augmented information-theoretic apparatus, and if they could be, that they should. Contrary to the methodological tenet, personal-level representations are not posited on purely theoretical, third-personal grounds, but also on experiential, first-personal grounds.

In some sense, the above critique, even if accepted wholesale, would not expose a widespread lacuna in the orthodoxy. After all, I have granted that the orthodoxy's treatment of sub-personal representation has the adequate resources and is on the right track, and moreover, that personal-level representations only constitute the tip of the cognitive iceberg. At most, then, there is a small minority of mental representations that defy treatment by the orthodoxy. Why should this matter?

The first reason it should matter is obvious and uninteresting. A full theory of mental representation ought to apply universally. If the orthodoxy does not, and does not have the resources to, this is something worth worrying about eventually.

A second reason may be more interesting. It is that personal-level representations are after all *paradigmatic* cases of mental representation. It should be slightly disconcerting if our theory of mental representation lacked the resources to account for paradigmatic cases of mental representation.

It would be *seriously* disconcerting if our theory lacked the resources to account for *any* paradigmatic mental representation. This would be the case if *only* personal-level representations were paradigmatic (so that no sub-personal ones were). This strikes me as not at all implausible, inasmuch as our pre-theoretic conception of mental representation is grounded in our ongoing stream of conscious occurrent thoughts and experiences and their representational character.⁴¹

A parallel point applies to the philosophical notion of intentionality. It is plausible to suppose that this notion derives from our personal-level representations, not from sub-personal ones. This might justify a preliminary suspicion that, to the extent that we seek in the theory of mental representation a solution to the traditional philosophical problem of intentionality, the solution is more likely to emanate from the theory of personal-level representation than from that of sub-personal representation.⁴²

I suspect there may be many more implications of clearly distinguishing personal-level representation. Thus, someone might hold, for whatever reason, that while content externalism is correct with regard to sub-personal representations, it is false of personal-level ones. Someone else might claim that Stich-style eliminativism and Dennett-style instrumentalism (Stich 1978, 1983, Dennett 1987), however well motivated with respect to sub-personal representations, defy credulity when it comes to personal-level representations. And so and so forth.

I am not going to pursue all these implications here. For that matter, the above remarks are already rather sketchy. I do not wish to establish or assert here any particular thesis regarding these wider issues. I am only concerned to underline the significance of the previous section's critique. If that critique of the mental representation orthodoxy is at all on the right track, then there may well be important consequences for some of the deepest, most "perennial-looking" philosophical issues surrounding mental representation.⁴³

References

- Adams, F. 1991. "Causal Contents." In McLaughlin (ed.), *Dretske and His Critics*. Oxford: Blackwell.
- Block, N. 1986. "Advertisement for a Semantics for Psychology." *Midwest Studies in Philosophy* 10: 615-677.
- Brandom, R. 1994. "Reasoning and Representing." In *Philosophy of Mind*, ed. M. Michael and J. O'Leary-Hawthorne. Dordrecht: Kluwer Academic Publishers.
- Burge, T. 1989. "Individuation and Causation in Psychology." *Pacific Philosophical Quarterly* 70: 303-322.

- Carruthers, P. 1996. *Language, Thought, and Consciousness*. Cambridge and New York: Cambridge UP.
- Carruthers, P. 2000. *Phenomenal Consciousness*. Cambridge: Cambridge UP.
- Chisholm, R. 1957. *Perceiving: A Philosophical Study*. Ithaca NY: Cornell UP.
- Cummins, R. 1989. *Mental Representation*. Cambridge MA: MIT Press.
- Dennet, D. C. *The Intentional Stance*. Cambridge MA: MIT Press.
- Dretske, F. I. 1981. *Knowledge and the Flow of Information*. Oxford: Clarendon.
- Dretske, F. I. 1986. "Misrepresentation." In R. Bogdan (ed.), *Belief: Form, Content, and Function*. Oxford: Oxford UP.
- Dretske, F. I. 1988. *Explaining Behavior*. Cambridge MA: MIT Press.
- Dunlap, K. 1900. "The Effect of Imperceptible Shadows on the Judgment of Distance." *Psychological Review* 7: 435-453.
- Field, H. 1977. "Logic, Meaning, and Conceptual Role." *Journal of Philosophy* 74: 379-409.
- Fodor, J. A. 1975. *The Language of Thought*. Cambridge MA: Harvard UP.
- Fodor, J. A. 1990. *A Theory of Content and Other Essays*. Cambridge MA: MIT Press.
- Fodor, J. A. 1994. *The Elm and the Tree*. Cambridge MA: MIT Press.
- Georgalis, N. 2006. *The Primacy of the Subjective*. Cambridge MA: MIT Press.
- Harman, G. 1987. "(Nonsolipsistic) Conceptual Role Semantics." In E. LePore (ed.), *New Directions in Semantics*. London: Academic Press.
- Harman, G. 1990. "The Intrinsic Quality of Experience." *Philosophical Perspectives* 4: 31-52.
- Horgan, T. and J. Tienson 2002. "The Intentionality of Phenomenology and the Phenomenology of Intentionality." In D. J. Chalmers (ed.), *Philosophy of Mind: Classical and Contemporary Readings*. Oxford and New York: Oxford UP.
- Horst, S. 1996. *Symbols, Computation and Intentionality*. Berkeley CA: University of California Press.
- Kim, J. 1991. "Dretske on How Reasons Explain Behavior." In McLaughlin (ed.), *Dretske and His Critics*. Oxford: Blackwell.
- Kriegel, U. 2002. "Phenomenal Content." *Erkenntnis* 57: 175-198.
- Kriegel, U. 2003. "Is Intentionality Dependent upon Consciousness?" *Philosophical Studies* 116: 271-307.
- Kriegel, U. 2004a. "Consciousness and Self-Consciousness." *Monist* 87: 185-209.
- Kriegel, U. 2004b. "The Functional Role of Consciousness: A Phenomenological Approach." *Phenomenology and the Cognitive Sciences* 3: 171-193.
- Kriegel, U. 2009. *Subjective Consciousness: A Self-Representational Theory*. Oxford: Oxford UP.
- Kriegel, U. 2010. "Cognitive Phenomenology as the Basis of Unconscious Content." In T. Bayne and M. Montague (eds.), *Cognitive Phenomenology*. Oxford and New York: Oxford UP.
- Loar, B. 1981. *Mind and Meaning*. London: Cambridge UP.

- Loar, B. 2002. "Phenomenal Intentionality as the Basis of Mental Content." In M. Hahn and B. T. Ramberg (eds.), *Reflections and Replies: Essays on the Philosophy of Tyler Burge*. Cambridge MA: MIT Press.
- McGinn, C. 1988. "Consciousness and Content." In *Proceedings of the British Academy* 76: 219-239. Reprinted in N. J. Block, O. Flanagan, and G. Güzeldere (eds.), *The Nature of Consciousness: Philosophical Debates*. Cambridge MA: MIT Press, 1997.
- Marr, D. 1982. *Vision*. San Francisco: WH Freeman Publishers.
- Millikan, R. G. 1984. *Language, Thought, and Other Biological Categories*. Cambridge MA: MIT Press.
- Millikan, R. G. 1989. "Biosemantics." *Journal of Philosophy* 86: 281-297.
- Millikan, R. G. 1990. "Seismograph Readings for *Explaining Behavior*." *Philosophy and Phenomenological Research* 50: 819-839.
- Milner, A. D., and M. A. Goodale 1995. *The Visual Brain in Action*. Oxford: Oxford University Press.
- Papineau, D. 1993. *Philosophical Naturalism*. Oxford: Blackwell.
- Rosenthal, D. M. 1993. "Thinking that One Thinks." In M. Davies and G. W. Humphreys (eds.), *Consciousness: Psychological and Philosophical Essays*. Oxford: Blackwell.
- Searle, J. R. 1991. "Consciousness, Unconsciousness, and Intentionality." *Philosophical Issues* 1: 45-66.
- Searle, J. R. 1992. *The Rediscovery of Mind*. Cambridge MA: MIT Press.
- Sidis, B. 1898. *The Psychology of Suggestion*. New York: D. Appleton and Company.
- Stampe D. 1977. "Towards a Causal Theory of Linguistic Representation." *Midwest Studies in Philosophy* 2: 42-63.
- Stich, S. P. 1978. "Autonomous Psychology and the Belief-Desire Thesis." *The Monist* 61: 573-591.
- Stich, S. P. 1983. *From Folk Psychology to Cognitive Science*. Cambridge MA: MIT Press.
- Strawson, G. 1994. *Mental Reality*. Cambridge MA: MIT Press
- Strawson, G. 2004. "Real Intentionality." *Phenomenology and the Cognitive Sciences* 3: 287-313.
- Weiskrantz, L., A. Cowey, and I. Hodinott-Hill 2002. "Prime-sight in a blindsight subject." *Nature Neuroscience* 5: 101-102.
- Weiskrantz, L., A. Rao, I. Hodinott-Hill, and A. Cowey 2003. "Brain potentials associated with conscious aftereffects induced by unseen stimuli in a blindsight subject." *Proceedings of the National Academy of Sciences USA* 100: 10500-10505.

¹ It is not particularly important for our present purposes to get clear on what events and facts are and what the difference between them is, but see Bennett 1988 for a comprehensive account (especially Chapter 2 on the difference between events and facts).

² It is important that the representations explain what they do in virtue of their representational content. As Dretske notes, an opera singer can cause a champagne glass to shatter, but because of the high pitch of the singing, not because of the content of the song. If all a representation's effect were of this sort, none would be effects that it has *because* it is a representation.

³ See Millikan 1990, Kim 1991, and Adams 1991 for a good sample.

⁴ It is usually further supposed that the relevant theory would provide a solution to the problem of intentionality – that is, that it would constitute a correct account of intentionality.

⁵ I speak of relations in this context even though it is well known that representation, or at least intentionality if it is a relation at all, would have to be a very unusual relation, in that it can hold even when not all of its relata exist. Some have, quite justifiably, refused on these grounds to call intentionality, or representation, a relation, claiming that a relation cannot be instantiated when the relata do not exist. But this would only require me to rephrase the way I express the point in the text. We could introduce the notion of a “relation*,” which is something just like a relation except that it does not require the existence of the “relata*.” More simply, we can focus on veridical representation, which is certainly a relation. This issue will be taken up momentarily in the text.

⁶ I use “hallucinatory misrepresentation” to denote misrepresentation of an entity that does not even *exist*, not just one that has different properties than the one it is represented to have. The term “targetless misrepresentation” is also sometimes used to denote the same.

⁷ The methodological assumption is often explicitly embraced, perhaps most notably by Fodor (1975) and Cummins (1989).

⁸ The reader will have noticed that I did not include in this list short-armed versions of conceptual or functional role semantics, such as Field's (1977), Loar's (1981), Block's (1986), and Brandom's (1994). This is mostly because the phrase “functional role semantics” is something an oxymoron: functional role is a matter of relations *among* mental states, whereas semantics are concerned with relations to (typically) *extra-mental* entities. Once this point is ignored, certain perversions are introduced into the theory of mental representation that should be immaterial to the main argument of this paper.

⁹ The mental representation orthodoxy, although hugely dominant in the literature, has not gone entirely unchallenged. Various parts of it have been questioned, or straightforwardly rejected, by such writers as McGinn (1988), Searle (1991, 1992), Strawson (1994, 2005), Horst (1996), Horgan and Tienson (2002), Loar (2002), Georgalis (2006), and Kriegel (2003, 2010). My own critique will make contact with theirs (especially McGinn's and Georgalis') at different points and in various ways that I will indicate in footnote in the right places. But it is also importantly different from theirs in several key respects, most manifestly in focusing on the personal/sub-personal.

¹⁰ Nor is Dennett particularly clear on how to draw the distinction more precisely.

¹¹ An immediate objection might target the construal of persons as systems. One *has*, not *is*, a cognitive system. It is quite implausible to identify me with my global cognitive system. But instead of assuming, problematically, that there is an identity relation between persons and global cognitive systems, the proponent of the systemic approach can assume, more plausibly, that there is a *constitution relation* between them. Thus, I am not a global cognitive system, but I may well be constituted by one.

¹² A very high degree of awareness might be quite rare in our mental life and characterizes only the more reflective or introspective episodes thereof, where we explicitly turn our attention inward and examine and scrutinize our internal states and processes.

¹³ I develop in greater detail the gradient model of awareness, and its relation to attention, in Kriegel 2004a, 2004b.

¹⁴ This little argument would require much defense of its premises before it could be treated as a major reason to believe that absent-minded perceptions are accompanied by awareness of them after all. I do not offer it here as such a major reason – merely as an initial consideration that might be pursued on a different occasion.

¹⁵ It might be argued that such states are precisely the sort cognitive science has started positing with the expansion of “the cognitive” beyond the occurrent happenings of which we are aware throughout our waking life. The fact that it is Jane, the person, who has the belief justifies treating the belief as a personal-level one only in a relatively unimportant sense of “personal-level.” The sense is unimportant because it is not the one that was crucial to the development of cognitive science.

¹⁶ This response requires a notion of standing awareness to match that of standing belief. The notion of standing awareness is not unfamiliar, however. We can imagine Tim saying to Tom “You know, 328.57 is greater than 174.16,” and Tom responding “Yes, Tim, I’m aware of that.” Tom is here giving voice to his standing awareness that $328.57 > 174.16$. And just as Tom has a standing awareness of this fact, Jane has a standing awareness of her *belief* that there are more than four countries in the world. (It would be a bit forced to say that standing awareness comes in degree, as occurrent awareness does. But the suggestion need not commit to degrees of standing awareness. Instead, it may claim that personal-level states and processes are such that their subject is minimally aware of them, whether occurrently or “standingly.”)

¹⁷ There is an intuitive case for standing states, inasmuch as it is intuitive to say that a person believes that $328.57 > 174.16$ and wants to remain employed. But it may well be that the terms “belief” and “want” as they are used in folk psychology are ambiguous as between denoting psychological states and denoting dispositions to enter psychological states.

¹⁸ I am using these terms here as referring to qualities of internal states, not as referring to qualities of external objects (e.g., red apples and bitter beer).

¹⁹ Perhaps there could be, but the important fact is that there need not be. That in itself shows that there is one possible personal-level process that is unconscious on this construe.

²⁰ Again, it might, but it need not (see previous endnote).

²¹ Here and throughout the paper, I do not reserve the term “person” to humans. Thus I certainly want to say that chimpanzees have personal-level representations. I use the term “person” in the way it is sometimes used in debates in applied ethics, whether some animals, some human fetuses, some future machines, etc. are persons. The assumption in these debates is that there is a pertinent notion of personhood that does not apply only to humans (nor to all humans) *by definition* (though it may do so contingently).

²² I develop this point in greater detail in Kriegel 2003.

²³ Since dorsal representations occur only in the context of action-guidance, we would have to envisage a situation in which John must *do something* with a picture of a rabbit.

²⁴ I speak here of representations *representing* their content. This is because I follow extant theories in construing the content of representation as what is being represented rather than what does the representing. At least this is the case with veridical representations.

²⁵ Since Harman's (1990) defense of the transparency thesis for perceptual experiences, the thesis has become extremely popular. I defend it in Kriegel 2002 and 2009.

²⁶ The thought, then, is something like this. There are traces in my brain that represent an external condition C in the same impersonal way that traces in the snow represent a thief's escape path even when nobody is aware of their existence. But such traces in my brain form a sub-personal representation of C precisely because I am unaware of them. In order for the representation of C to be personal, I would have to become aware of it, and once I do, the representation would represent C *to me*.

²⁷ These remarks are also in line with some of Colin McGinn's and Nicholas Georgalis' work on the difference between conscious and non-conscious representation (see McGinn 1988, Georgalis 2006). Their views can be harmlessly transplanted, perhaps with added plausibility, to the difference between personal-level and sub-personal representation. McGinn's claim, for instance, would be that personal-level representation is two-faced, whereas sub-personal representation is single-faced. Personal-level representations are Janus-headed, or two-faced, in that they involve the representational content's "presence to the subject." Compare a veridical personal-level and a veridical sub-personal (e.g., subliminal) perceptual representation of a red object (or surface or volume). The sub-personal representation involves only a relation between the representational state and the red object. The personal-level representation involves this relation as well, but it also involves, on top of that, a relation between the representational state and the subject. In such a personal-level representation of a red object, the relation between the representing state and the represented object is what McGinn calls the *outward-looking* face of the representation's content, while the relation between the representing state and the represented-to subject is its *inward-looking* face. The sub-personal representation has the very same outward-looking face, but lacks the inward-looking face. The latter is unique to personal-level representation.

²⁸ At least this is so if we stipulate that, throughout his life, Jim remains unaware that his goosebumps constitute the natural sign they do.

²⁹ The same holds for tacit beliefs. Tacitus believed that 1374.67 is greater than 873.92. His belief (i) represented the fact that $1374.67 > 873.92$ and (ii) occurred in Tacitus, and (iii) Tacitus was a conscious and sentient creature. Yet the belief did not represent the fact that $1374.67 > 873.92$ *to* Tacitus in any non-technical sense. At least this is so if we stipulate that, throughout his life, Tacitus never became aware of this tacit belief. Searle (1992) claims that tacit beliefs could not be anything but brute neurophysiological states. If so, Tacitus' tacit belief is no different from his goosebumps. Unconscious brain states and skin states are exactly the same in the respect.

³⁰ In other words, a consumer-semantic account of representation-to returns wrong results.

³¹ Two points about this paragraph. First, I am working here with an intuitive, unregimented notion of disposition, because delving into the metaphysics of dispositions, their manifestations, and their categorical bases would take us too far afield. Second, it might be objected that if my reasoning was correct, it would afflict functional role accounts of representation-of. My response is that indeed I think (and always have) that it does. Functional role semanticists assume that a representation acquires its representational content in virtue of its functional role. But this quite obviously puts the cart before the horse. More plausibly, a representation acquires the functional role it does precisely because of the representational content it has.

³² I am referring here to *ontological* naturalism – roughly, the view that there are no supernatural entities or phenomena – not the sort of *epistemological* naturalism based on the idea of methodological continuity between philosophy and the sciences. What I have said may well be consistent with such epistemological naturalism as well (depending on just how it is formulated, it may or may not be consistent with my rejection of the methodological tenet of orthodoxy), but in any case it is not in the epistemological sense that extant theories are referred to as naturalist.

³³ Nonetheless, it may well be that the adherence to naturalism of extant theories of representation is part of what has led them astray (if I am right that they were led astray), in that it may have encouraged deflating the phenomena to ensure that they do not elude the naturalist framework.

³⁴ There are important questions regarding what it means to encounter a representation, or be presented with it, in one's experience. Although I will make passing remarks on this issue, I will leave it largely open-ended, mostly because I do not wish to commit to anything that might becloud the overall point. The overall point of my discussion in this section should be neutral on just what it means to encounter a representation in one's experience, and I want it to remain such.

³⁵ This is in line with the transparency of phenomenal experience mentioned in the discussion of the ontological tenet.

³⁶ Moreover, a medieval peasant would have all the reason I have to believe that she has a mental representation when she is aware of having one. This, despite the fact that she may have no theoretical beliefs about an alleged sub-personal mental life.

³⁷ This is not quite right, actually. There may be instances of *mental behavior* (such as calculating or deliberating) whose explanation calls for the positing of sub-personal mechanisms employing sub-personal representations. Since the weather-watchers could "observe" some of their personal-level mental behavior, they might have grounds on which to posit the sub-personal representations needed to explain them. This makes the weather-watchers an imperfect illustration of the point that grounds for admitting personal-level representations are categorically different from grounds for admitting sub-personal representations. Perhaps the following modification would do, however. Let weather-epiwatchers be creatures whose mental states are virtually epiphenomenal, except that they do yield states of introspective awareness of them (in case the relation between a mental state and its introspecting is causal). Weather-epiwatchers cannot engage in mental behavior any more than in bodily behavior, since their mental states do not cause anything. Therefore, they have no grounds on which to posit sub-personal representations. Yet they have all the grounds they need for admitting their own personal-level representations.

³⁸ Although the issues here are complicated, it is possible that the presuppositions are even more innocuous than that, and shares much with ordinary perceptual experience. When I look out the window, I see a tree. I believe that the tree exists. Importantly, I do not believe that the tree exists because the supposition that it does is explanatorily useful to me. I believe that it exists because I *see* it. Theoretical and explanatory considerations may well convince me that it does not in fact exist. But until they do, I believe that the tree exists and believe so simply because the tree is present to me. It is possible that nothing more mysterious is involved in our awareness of personal-level representations. It is not because the supposition that my representation of my laptop exists explains something otherwise inexplicable that I believe the representation exists. I believe that it exists because I am *aware* of it – in the same way I believe that the tree exists because I *see* it. (Some might insist that I posit the tree as an explanation of my tree experiences. The tree is posited as the thing that is causally responsible for my having the experiences I do. This strikes me as false to the facts, but even if it were conceded, a similar story could be told about my awareness of my laptop experience. The laptop representation would be conceived as a posit necessary to account for my awareness of the laptop representation. It is still the case that this sort of evidentiary relation is different from others, and characterizes first-person evidence exclusively.)

³⁹ This principle could also be thought of as applying with equal force to sub-personal representations, but in the nature of things, sub-personal representations never seem first-personally to have any properties.

⁴⁰ In rejecting the methodological tenet of orthodoxy, I join most notably Nicholas Georgalis (2006), who has argued thoroughly and repeatedly against it.

⁴¹ It may be objected that we do not *have* a pre-theoretic conception of mental representation. But this seems to me to be false. We do not pre-theoretically use the *term* "mental representation" (or at least do not

do so often), but we do have a *concept* of mental representation. (And even the term is not as theoretical-sounding as, say, “lepton.”)

⁴² The problem of intentionality, as developed originally in the writings of Brentano, Frege, and their contemporaries, and later consolidated and elucidated by Chisholm (1957), is in effect the problem of *intensionality* (“intensionality-with-an-s”). The idea is that an activity is intentional if it is denoted by an intensional verb, and a verb ϕ is intensional just in case the following hold: (i) from the fact that $a \phi$'s b it does not follow that there is an x , such that $a \phi$'s x ; (ii) from the facts that $a \phi$'s b and $b=c$ it does not follow that $a \phi$'s c . In other words, both existential generalization and substitution of co-referential terms do not necessarily preserve truth.

Extant theories of mental representation have long struggled to account for intensionality so construed. The question whether they succeed, or have the resources to eventually succeed, is one of the biggest issues in contemporary philosophy of mind, and cannot be taken up here. To some, the attempt by extant theories to strike the water of intensionality out of the rock of teleo-informational relations has appeared hopeless. Others think that the problem is of a technical rather than principled order and intensionality will eventually succumb to teleo-informational domestication. Certainly some progress has been made over the past three decades in addressing the challenge, though perhaps more so regarding (i) than (ii).

As I said, this issue cannot be seriously discussed here. But the suggestion might be made that we have sought an account of intensionality in the wrong place. It is possible to maintain that sub-personal representations simply do not *feature* intensionality, or at least do so only partially. For instance, it might be held that all our tacit beliefs about the morning star are equally about the evening star – in contravention of (ii).⁴² In other words, the transitive verb “tacitly believes” might not be (fully) intensional. Likewise, it is plausible to say that our dorsal-stream representations of the duck in the duck-rabbit figure are equally representations of the rabbit. If so, the theory of sub-personal representation is not *supposed* to account for (ii). Against the background of wholesale acceptance of the previous section’s critique, this would relieve extant theories of the duty to account for intensionality (or part of it), and reassigns that duty to the as-yet-to-be-pursued theory of personal-level representation.

If all this is correct, then the dominant approach to the problem of intentionality in current philosophy of mind has “rested on a mistake,” as they say. The mistake is to seek an answer to the puzzle of intentionality in the theory of entities that do not bear (some of) the marks of intentionality. More generally, the theory of mental representation, in the style we have become accustomed to, is good philosophy of psychology (as a branch of philosophy of science), and has been impressively successful in elucidating the nature of the scientific posits of psychology and the cognitive sciences. But it is not the right kind of theory to approach the puzzle of intentionality with. This point, or a similar one, is belabored by Horst (1996).

⁴³ This paper was written during a research fellowship at the University of Sydney, to which I am greatly indebted. For comments on a previous draft of the paper, I would like to thank Nicholas Georgalis, George Graham, David Rosenthal, and especially David Chalmers. The paper was presented at the Institut Jean Nicod, and materials from it also at the Australian National University, California State University at Los Angeles, and Macquarie University. I am grateful to audiences at all these places, especially Mark Ballagher, Tim Bayne, Dick Carter, David Chalmers, Stephen Davis, Jerome Dokic, Pierre Jacob, Kristie Miller, Elizabeth Pacherie, David Pitt, and Barry Smith.