STILL IN PROGRESS: Please read it accordingly.
Comments / criticisms / corrections are most welcome!
Draft 3.5, July 2001 — (Draft 1: August 1999)

# Concepts, Introspection, and Phenomenal Consciousness: An Information-Theoretic Account


MURAT AYDEDE
University of Florida
Department of Philosophy
330 Griffin-Floyd Hall
P.O. Box 118545
Gainesville, FL 32611-8545
maydede@phil.ufl.edu

GÜVEN GÜZELDERE
Duke University
Department of Philosophy
201 West Duke Building
Box 90743
Durham, NC 27708
guven@duke.edu



**ABSTRACT**. This essay is a sustained information-theoretic attempt to bring new light on some of the perennial problems in the philosophy of mind surrounding phenomenal consciousness and introspection. Following Dretske (1981), we present and develop an informational psychosemantics as it applies to what we call *sensory concepts*, concepts that apply, roughly, to so-called secondary qualities of objects. We show that these concepts have a special informational character and semantic structure that closely tie them to the brain states realizing conscious qualitative experiences. We then develop an account of introspection utilizing this special nature of sensory concepts. The result is a principled naturalistic account of a class of concepts used in the introspection of experiences, which, following recent terminology, we call *phenomenal concepts*. Contrary to widespread opinion, we show that information theory contains all the resources to satisfy internalist intuitions about phenomenal consciousness, while not offending externalist ones. A consequence of this account is that it explains and predicts the so-called conceivability arguments against physicalism on the basis of the special nature of sensory and phenomenal concepts. Thus we not only show why physicalism is not threatened by such arguments, but also demonstrate its strength in virtue of its ability to predict and explain away such arguments in a principled way. However, we take the main contribution of this work not so much as a response to conceivability arguments as a substantive account of the interface between sensory and conceptual systems, and the mechanisms of introspection as based on the special nature of the information flow between them.

# Concepts, Introspection, and Phenomenal Consciousness: An Information-Theoretic Account

## 1. INTRODUCTION

Levine, in his recent book (2001), summarizes his lengthy discussion of the explanatory gap as follows:

> What emerges from our discussion is that the explanatory gap is intimately connected to the special nature of phenomenal concepts. E-type materialists try to save materialism from the conceivability argument by arguing that phenomenal concepts are special in some way. Well, I grant that, but then we have the problem of providing an explanation in physicalistic terms of that very specialness, and we don't seem to have one. If we could explain the explanatory gap, then either it would go away or we would just learn to live with it. But it seems we can't do that without a good account of phenomenal concepts, and that's something we don't have. We lack both an account of phenomenal properties and phenomenal concepts. (2001: 86)

In this essay, we try to provide both — and more. This work is a sustained information-theoretic attempt to bring new light on some of the perennial problems in the philosophy of mind surrounding phenomenal consciousness and introspection. Following Dretske (1981), we present and develop an informational psychosemantics as it applies to what we call *sensory concepts*, concepts that apply, roughly, to so-called secondary qualities of objects. We show that these concepts have a special informational character and semantic structure that tie them closely to the brain states realizing conscious experiences and their qualities. We then develop an account of introspection utilizing this special character of sensory concepts. The result is a naturalistic account of a class of concepts used in the introspection of experiences, which, following recent terminology, we call *phenomenal concepts*. Contrary to widespread opinion, we

show that informational psychosemantics contains the resources to satisfy, in a principled way, internalist intuitions about phenomenal consciousness, while not offending externalist ones.

Our account contributes to what appears to be a growing convergence of views, sometimes loosely grouped under the label of "perspectivalist materialism or physicalism",[1] which in recent years have been developed in response to so-called anti-physicalist conceivability arguments.[2] In its typical versions, perspectivalism is advanced in three stages. First, it diagnoses the puzzle involved in attempting to conceive the phenomenal in terms of the physical as a Frege puzzle — namely, as one arising from distinct but co-referential concepts. Second, it points out that the Frege case at hand is special in a way that marks it off from all other Frege cases we know of, and that this specialness needs accounting for. Third, it postulates a group of concepts, typically called "phenomenal concepts", whose nature is said to be perspectival, revealing what is so special about our epistemic access to the phenomenal qualities of our experiences.

We think that extant perspectivalist accounts are lacking in precisely those respects that are crucial for the plausibility of the physicalist project responding to anti-physicalist arguments. One such respect relates to a proper understanding of the link between sensory concepts (e.g., of red) and phenomenal concepts (e.g., of experiencing red). As far as we can tell, no perspectivalist account to date has properly explained the proper relationship between

[1] Churchland (1985, 1989), Loar (1990 / 1997, 1999), Lycan (1987, 1996), Papineau (1993), Pereboom (1994), Sturgeon (1994, 2000), Rey (1997), Hill (1997), Hill & McLaughlin (1999), Tye (1995, 1999, 2000), and Perry (2001). Rosenthal (2001) and Shoemaker (2001) also come close to this line.

[2] Descartes and Locke are the most prominent historical expounders of conceivability arguments. Recent versions can be found in Farrell (1950), Feigl (1967), Kripke (1970), Nagel (1974), Jackson (1982, 1986), Robinson (1982), Hart (1989), McGinn (1991), and Chalmers (1996). For an excellent critical discussion of such arguments, see Levine (1993, 1998, 2001) and Biro (1991). Levine does not endorse the metaphysical conclusion of these arguments; in fact, he argues against drawing such a conclusion. But his discussion of the explanatory gap between the phenomenal and the physical has contributed significantly to a more articulate development of conceivability arguments.

these two kinds of concepts — in fact, this distinction is sometimes obscured.[3] Not only that; no detailed perspectivalist account has been developed about how we acquire these phenomenal concepts, what it is that we apply these concepts to, and how. Most perspectivalists commit themselves to the claim that phenomenal concepts apply to brain states realizing conscious qualitative experiences. Again, we believe that very little has been said about why phenomenal concepts don't present to us physical properties of these brain states *as such*, which is at the source of the puzzle. As a result, the postulation of phenomenal concepts has seemed *ad hoc*, even though, from a technical point of view, it seems to block some forms of conceivability arguments (e.g., Jackson's Knowledge Argument). Our account offers a principled and integrated story about the special nature of phenomenal concepts, based on the special informational structure of sensory concepts, from which an information-theoretic account of introspection naturally falls out.

As will be apparent in the coming sections, our account is an unexpected synthesis of otherwise quite diverse views. For instance, we will see that a pure information-theoretic psychosemantics for sensory and phenomenal concepts has consequences that should satisfy internalists about phenomenal consciousness. Likewise, our account integrates intuitions that motivate both higher-order perception (e.g., Armstrong, Lycan) and higher-order thought (e.g., Rosenthal) accounts of introspection — even though it does not itself fall under these labels.

We conclude the essay by discussing the flaws of anti-physicalist conceivability arguments as revealed by our account. We ask the reader to bear with us until we come to this last section while we develop our account of sensory and phenomenal concepts. This will take some time. But we believe that in the absence of a substantive and detailed theory about how a psychosemantics of such concepts fits naturally into a general information-theoretic account, special pleading for these concepts in response to conceivability arguments will of necessity

[3] E.g., Churchland (1985, 1989). Tye (1999) mentions the distinction, but then he goes on to write in a way which suggests that he thinks the distinction is not important.

lack plausibility and appear *ad hoc*, and hence will not remove the mystery surrounding phenomenal consciousness.

## 2. THE ARCHITECTURE OF INFORMATION FLOW IN COGNITION

In a somewhat reconstructed form and sometimes using different terminology, we will present the information-theoretic account of concept formation (from sensory experiences) provided in Fred Dretske's seminal work, *Knowledge and the Flow of Information* (1981). And we will modify it as we go along.[4] This account postulates an *architectural* distinction between sensory systems and a central cognitive system controlling the behavior of the organism. The sensory system has the job of providing information about one's environment to the cognitive system, and normally affects behavior only indirectly, via intermediary cognitive structures. Sensory systems hook up with the environment via transducers whose job is to transform the particular forms of energy impinging on the peripheral sensory organs into forms usable by the internal perceptual systems. The output of transducers and much of the subsequent processing in the sensory and perceptual systems appear to be unconscious (but see below). The output of the sensory system is a sensory representation of (some aspects of) the distal layout that is made available to the cognitive system.

In this framework, the sensory representations are *conscious* only insofar as the information contained therein is *available* to the conceptual system, even if it is not necessarily put to use. Put differently, for the most part, the information contained in sensory representations is *consciously* available to the organism only insofar as the organism *can* conceptualize this information, i.e., only insofar as the information can be used in the acquisition and/or deployment of the relevant concepts. Hence, we will use "having an experience" (which is a conscious affair) and "tokening a sensory representation" interchangeably. We will come back to this point later on, but for the rest of this paper we will concern ourselves about sensory

[4] For the ease of our presentation, we won't make a special effort to tell when and where our story differs from Dretske's.

representations that are thus conscious in this way. Hence we will not discuss those modular (pre- or intra-perceptual) processes whose state-transitions and outputs are not consciously accessible —that is, which do not constitute direct inputs to the cognitive system.[5]

What are the functional determinants of this architectural distinction? We have already touched on one: sensory representations don't normally affect behavior directly. It is largely the central cognitive system which controls behavior. So a necessary condition of a cognitive structure's being conceptual as opposed to sensory is its executive connections to behavior. A representation is sensory, on the other hand, only if it makes information about one's environment (internal / bodily as well as external) available to the conceptual system for further processing, which normally also makes the representation (experience) conscious. Besides this, the most important characteristics underpinning the architectural distinction are to be found in the following distinctions.

## 2.1. Vertical vs. Horizontal

First, sensory experiences are supposed to track changes in the environment. In this they are (non-conceptual) representations whose primary job is to make available to the hosts temporally indexed information about the environment. The crucial point here is that sensory

[5] Availability to the central conceptual system is necessary for a sensory representation to be a conscious state. Although we will talk as if it were also sufficient in what follows, sufficiency is probably secured by detailing some constraints or further conditions on the role the information in the experience can play, after its delivery, within the conceptual system on its way to setting the behavioral parameters, especially about its availability to speech systems. Cases where the subliminal / unconscious perception of a stimulus affects behavior are probably cases where some information in the experience is made available to the conceptual system influencing behavior but nevertheless is not available to all central cognitive capacities — its processing within the conceptual system is limited. This seems to suggest that the central system may not be entirely non-modular in the Fodorian sense, and thus may be composed of loosely connected central faculties, or better, processing streams, and that there are conditions under which global availability may be hindered.

We are aware of the discussion surrounding accounts of state consciousness, and our stipulation here about sensory consiousness closely follows Dretske's position. But we think that at least as a necessary condition it should also be acceptable to those who defend a higher-order-representation (HOR — perception or thought) theories of state consciousness. As will become apparent later on, our view accommodates the central insights of HOR theories.

experiences normally carry information about features of the environment: they are *responses* to environmental events. As such their informational value is typically restricted within a time frame sufficient for the organism to act back on the environment on the basis of this information. In short, sensory representations are normally stimulus-driven (*a fortiori*, not directly voluntary). We will call this aspect of information processing, *vertical* information processing.

By contrast, central cognitive processes, such as thinking, reasoning, remembering (memory), imagining, and daydreaming, are normally *horizontal* forms of information processing. By this we mean that they *can*, and pretty frequently *do*, occur in the absence of a direct or immediate causal (vertical / informational) relation with the things being thought or reasoned about. This is perhaps the most important hallmark of human intentionality. In contrast to sensory systems, central cognitive systems harbor representational processes defined over concepts that are not directly prompted by what those concepts represent.[6]

Although all concepts can be informationally decoupled from their referents in horizontal processes, most of them *can* also be used vertically, so that their tokenings carry information about the (instantiation of the) property they denote. In this *extended* sense of a vertical process, experience is the necessary intermediary.[7]

---

[6] Horizontal uses of concepts may be either voluntary or involuntary. To give you the flavor, here is a hypothetical situation. You read in the newspaper that a friend you had not seen since elementary school days is the featured pianist in Boston Symphony Orchestra's New Year's concert. You start thinking about the days together when you used to take music lessons together. Then you reminisce about your quirky piano teacher and the jokes he used to tell you, which makes you think that you should perhaps get in touch with your friend… Just think about the range of things that such an event could prompt you to think about: at the moment of your thinking almost none of the objects of your thoughts has any direct or obvious causal / informational relation to your thinking. (Horizontal prompting of a thought may be an involuntary occurrence, but then you may voluntarily prompt thoughts in yourself.)

[7] Information-carrying tokenings of a concept can arise in two ways: as a direct and immediate response to its "proper" sensory base, or by a reliable inferential process from "improper" sensory bases, including speech perception. If we were to label all information-carrying tokenings as vertical, we would need to distinguish them from horizontal "non-serious" tokenings as in thinking, free association, imagining, etc. in which they can be correctly tokened without carrying information about their denotations or truth-conditions. These two kinds of vertical tokenings complicate the picture (as Dretkse insisted in personal communication). However, it is clear that those vertical tokenings based on inference from "improper" sensory

In brief, conceptual representations *are* the kind of cognitive structures that are capable of engaging in horizontal processing, whereas sensory representations are not.[8] So one important determinant of the architectural distinction is the distinction between vertical and horizontal information processing.

## 2.2. Sensation vs. Perception

There is a useful sense in which perception, unlike sensation or mere raw experience, is the vertical informational process whereby objects of sensation and their sensible qualities are discriminated and recognized, i.e., categorized or classified under concepts.[9] For most perceptual / observational concepts, this normally takes the form of recovering the information already (mostly) in the sensory array by computational processes that eventuate in the tokening of a concept applying to the object of perception. We see this process mainly as one of information extraction by digitalization / abstraction from a rich array of information present in analog form in the experience. The mechanism underlying the formation of primitive sensory

---

bases require that the semantics of these concepts be already in place. Then the question becomes whether their semantics is acquired from "proper" sensory bases or something theoretical or descriptive, like Mary's concept of red before her release. Here we restrict the information-carrying / vertical uses of *sensory concepts*, which we will define below, to those prompted by "proper" sensory representations.

[8] There is much evidence that imagination, which is a central / horizontal capacity, uses the same representations as those involved in some stage in sensory processes. But we think these representations are special sensory / perceptual concepts which turn sensations into perceptions as we characterize this central process below.

[9] The contrast between sensation and perception is sometimes described in terms of a distinction between *seeing* and *seeing as* (*hearing* and *hearing as*, etc.). Having normal vision, you can certainly see an aardvark, but having no idea of what aardvarks are (not having the concept of one), you cannot see it *as* an aardvark. In '*S* sees *x*', '*x*' occurs transparently (i.e., could be replaced by any co-referring expression without changing its truth value), but the occurrence of '*F*' in '*S* sees *x* as *F*' is opaque, reflecting the fact that the truth-value of the statement depends on whether *S* has the concept expressed by '*F*' and applies it to *x* as a consequence of standing to *x* in the *seeing* relation. However, we should note that we are not cognitivists about conscious experience in the way Dretske (1994) characterizes Dennett's position (for Dennetts response, see his (1994)). Although we think that concept possession is necessary for consciousness in the sense required by our characterization of state consciousness above, we grant that having sensory experiences, strictly speaking, does not require concepts, while perception does.

concepts and their vertical deployment is probably hard-wired in concept-using organisms like us.

So, according to this scheme, visual object recognition, for instance, however automatic it may be, is mostly a central process,[10] since it involves categorizing an object under a "visual concept". Although the process itself appears to be unconscious, many features of the output representation (like variation in light intensities over the visual field, and consequent to these, lines, edges, colors, distance, orientation, texture, relative position, etc.), apparently utilized in the extraction process, are also consciously (hence centrally / globally) available. So perception is a central process in our sense and should be treated as a species of conception. (More on this below.)

## 2.3. Analog vs. Digital

Another determinant of the architecture, most important for our purposes, is captured by a distinction between the ways in which information is coded in the representations. We follow Dretske's original characterization:

- The most specific information a signal, *r*, carries about a source, *s*, is the information *r* carries about *s* in *digital* form.
- If *r* carries more information about *s* [or, about *t* (≠*s*)] *in virtue of* carrying this digital information about *s*, then this extra information is said to be carried by *s* in *analog* form.
- Analog information is information *nested* (nomologically or analytically) in the information carried in digital form.

[10] Or perhaps, one of the vertical streams within the central system dedicated to extract information from the visual *sensations* — but central nonetheless in that most of the information in the stream is globally available, even though the *processing itself* is not voluntary or introspectable, and may even be open to top-down cognitive influences.

The cognitive value of a sensory representation lies largely in the information about the distal layout it carries in analog form. Its digital informational content is the most specific information it carries about this layout, which is very rich not only in detail but also in its amount. The conceptual system is mostly keyed to the information nested in this specific and rich information. The analogy here between pictures and sensory representations is helpful. If we take a color picture of a cubical object, the picture will carry very rich, detailed, and determinate information about the size, texture, and orientation of the object, as well as its position relative to other objects, the illumination conditions, its determinate shades of color and their brightness across its surface, and so forth. We can think of this very specific and detailed information as expressible by a very long conjunctive sentence. But nested in this information there will be less specific information implied by it, such as the information that the object is cubical, that it has six faces, that it has eight corners, that it is colored, etc. Normally we are interested in the analog information carried by the picture. I may be interested merely to know that the object depicted is cubical — discarding the more specific information about its determinate color, size, orientation, etc. Or, depending on the situation, I may be interested only in its size or color, discarding the rest, and so forth.

Similarly with sensory representations. The conceptual system mostly exploits the analog information provided by sensory representations. In fact, according to Dretske, part of what makes a cognitive structure a conceptual representation is the way it digitalizes the analog information contained in the sensory representations. That is to say, concepts are those representations (subject to the above architectural constraints) whose most specific informational content (i.e., their digital informational content) is acquired from information carried (mostly) in analog form by sensory representations. Concepts (except sensory ones) are designed to selectively respond to and utilize the analog information contained in sensory representations. So, for instance, even though we cannot *sensorially* represent a triangle without at the same time representing its determinate size, shape, orientation, etc. we can *conceptually* represent an object simply as a triangle without representing anything more specific or

determinate about it if we have the concept of a triangle. Concepts on this scheme are those structures that are *acquired* from sensory representations, mostly on the basis of the analog information they carry.

Dretske identifies the *semantic content* of a concept with the information it carries in digital form. The informational content of a concept, however, is not unique in the way the semantic content is supposed to be, since a vertical tokening of a concept will carry *all* the information nested in its digital informational content. So, for example, when you identify a geometrical shape as an isosceles triangle, your identification carries more information about the object nested in its being such a triangle, e.g., that it has three sides, that it has three corners, that it is a geometrical shape, that it has a surface area, etc. Thus, these pieces of information are carried in analog form.

## 2.4. Extractable vs. Non-Extractable Analog Information

Both sensory and conceptual representations carry information in analog and digital form. But they encode analog information in fundamentally different ways:

- Whereas there is *some* analog information sensory representations carry in *extractable* format, the (primitive) conceptual representations carry *all* their analog information (contained in their vertical tokenings) in *non-extractable* format.[11]

A sensory representation is physically realized in such a way that its complex structure allows the analog information contained in it to be extracted by the conceptual system operating on it. Of course, what information can be extracted from the sensory representation doesn't depend solely on its complex informational structure; it also depends on the capabilities and the

[11] When we talk about a concept's carrying information we have of course its vertical tokenings in mind. Most of the time we will omit this qualification in what follows since the context should make it clear what sort of concept tokening we have in mind.

sophistication of the conceptual system. But, subject to these constraints, it is necessary for conceptualization that the analog information in the sensory representation is carried in a way that is extractable, and not all information carried in analog form is.

To illustrate, consider the example Dretske uses (1981: 138–39). It *is* possible to carry *all* the information encoded by a picture of a scene with a simple / primitive signal, say a buzzer system. Suppose the buzzer is activated when and only when a camera attached to the buzzer detects the occurrence of a situation *exactly* like the one depicted in the picture. As Dretske notes, computer recognition programs that rely on whole-template matching procedures approximate this kind of transition from one form of coding to another. Both structures carry exactly the same information, both digital and analog. However, we will say that the buzzer's buzzing carries the analog information carried by the picture in a way that is not extractable, whereas the picture carries it in an extractable form. This distinction needs to be developed in more detail in terms of physical constraints on the structures realizing the representations,[12] but what is intuitively obvious — and all we need for present purposes — is that the representational format which allows for information extraction must consist in a physical / functional structure *complex* enough to be the only source for subsequent digitalizations / abstractions based on it.[13] The activation of the buzzer, though it carries all the information carried by the picture, is structured in such a way that does not allow for digitalization of the information it carries in analog form. Primitive conceptual representations are like the buzzer system: although their vertical tokenings carry analog information nested in their digital content, they are structured in such a way that they cannot serve as the (only) basis

[12] For an elaboration of this distinction in insightful ways, see Kulvicki (2001). He uses this distinction to give an original and general account of isomorphism.

[13] For the purposes of exposition, we are restricting the abstraction base to the one provided only by the sensory representation / picture. Of course, this need not be the case once a threshold is passed and a certain basic conceptual repertoire is in place. What can be learned from a signal depends not only on the signal itself but also on what is independently known about the source, and this requires inferential deployment of other concepts and collateral information. See below.

for digitalization / abstraction of this information. This is part of the reason why primitive concepts are sometimes characterized as *discrete* representational structures or symbols.

### 2.5. Acquisition vs. Deployment

Although the distinction between the *acquisition* and *deployment* of concepts is not a functional determinant of the informational architecture, it is important to keep in mind for clarificatory purposes. Both acquisition and deployment can be *vertical* and *horizontal* in some intuitively extended sense. So, for instance, we can acquire concepts by reading, or by being talked to, by looking at pictures, by engaging in inference to the best explanation, etc.[14] This would be horizontal acquisition of concepts. Also note that for the moment we are using "acquisition" in a way that is neutral between *triggering* and *learning*.

Now that we have the functional determinants of the architecture of information flow in cognition and the relevant distinctions,[15] we focus on the nature of the concepts this architecture gives rise to.

## 3. CONCEPTS AND THEIR SENSORY BASES

Among the concepts directly and immediately acquired from sensory experiences are what we will call *sensory concepts*. These form a special class of concepts that will be important for what follows. Intuitively and roughly put (to be qualified in a moment), sensory concepts are those concepts whose digital informational content is also part of the digital informational content of the sensory representations from which they are acquired, so that the abstraction / digitalization distance between the concepts and these experiences is maximally short.

[14] We will argue below that *sensory* concepts in the sense we will introduce cannot be acquired horizontally.

[15] For a fuller and more detailed account we refer the reader to Drestke's seminal work, (1981).

The digital informational content of sensory representations is rich along several dimensions. We can think of these dimensions as presenting determinables such that the resolution of our sensory experiences marks the limit of their most determinate values about which we can gather sensory information. To the extent that we can separate these dimensions, we can speak of that part of the total digital information content of an experience which belongs to one of these dimensions fixed by the modality of the experience. So, for instance, under conditions that are optimal for color vision, seeing a ripe tomato will involve a visual experience whose total digital content contains the most specific information about the color of the tomato: it will represent the tomato as having a determinate shade of red, say, $red_{16}$. This is part of the total digital content of the visual experience containing information about the color of the object seen. As mentioned above, we can conceive of this total digital content as being expressed by a very long conjunction detailing all the most specific information it carries. The particular shade of color a region in the visual field has, then, would be one of the conjuncts. Sensory concepts are those concepts that are closest (in terms of abstraction distance) to the digital informational content of experiences from which they are acquired along these different dimensions.

If the property of being $red_{16}$ is a disjunctive property whose disjuncts are particular spectral reflectances, then the information the sensory representation carries about $red_{16}$ is about this disjunctive property. Every disjunct would be a particular ratio fixed by the percentage of light that the surface of an object reflects at each of the three characteristic wavelengths determined by the response sensitivity of three retinal cone types.[16] But whatever feature of sensory representation is responsible for carrying this information, it carries it without revealing its complex and disjunctive character. For instance, this feature, by carrying information about a surface's being $red_{16}$, also carries the analog information that it has a spectral reflectance, or that it reflects light at different wavelengths. These are nested in the information that the surface is $red_{16}$. But these pieces of analog information cannot be recovered or extracted from

the signal, i.e., from whatever feature of the sensory representation carries the color information in question.

There is, however, still some abstraction / digitalization — some loss of information — in this process. This can be explained in terms of a distinction between concepts used in synchronic discriminatory tasks and concepts used in diachronic recognitional / identification tasks. In fact, we typically reserve the notion of a concept for those cognitive structures involved in the latter sort of task. Consider the tomato again. If the conditions are appropriate, it will be possible to discriminate slight variations in the shade of red across the surface of the tomato. But when the same shades of color are shown to us suitably diachronically we may not be able to discriminate among them: most of the time, the best we can do is to identify and co-classify them as, say, dark red. Both kinds of task involve discrimination and categorization (sorting) of different color stimuli, and so, in this minimal sense, require conceptual capacities. In what follows, however, when we talk about sensory concepts, we will have in mind the most specific concepts one can have as revealed by diachronic recognitional tasks, which involve memory. It is clear that the abstraction distance between sensory experiences and the sensory concepts conceived in this way is still maximally short, although there is still some distance — some information is lost. Notice that in the case of color concepts this distance can be explained entirely in terms of set-theoretic inclusion. When these concepts are vertically deployed, the information they carry is disjunctive: they say something like "it is either $red_1$, or $red_2$, or $red_3$, or … $red_n$", where *n* is finite and $red_i$ is the most determinate shade of red one's visual experiences can carry information about and thus be synchronically discriminated.

It is important to note that the disjuncts here are still colors — determinate shades of red. This is important because the abstraction process here is not based on information about the constituents of colors (whatever objective properties color experiences / concepts detect), which are not themselves colors. So, for example, if color vision detects sets of individual surface

[16] Although it is controversial, we will assume a primary quality view of secondary qualities like color à la Hilbert (1987). We don't think, however, that much hangs on this choice in what

spectral reflectances, color sensations don't represent them *by* representing their constituent properties, say, individual reflectances or whatever further properties constitute these reflectances. This is to say that color sensations represent colors as lacking internal constituent structure, or as we will say sometimes for convenience, as simple / atomic properties.[17,18]

Contrast this to the visual representation of shapes. Our visual system is such that we can't visually represent a geometrical figure (in such a way that we can then recognize it as what it is, say, as a square) without simultaneously representing the lines, angles, curves, edges, and corners that, in some intuitive sense, constitute the figure. These constituents, notice, are not more determinate instances of the same figure type, so that even the concept of a most

---

follows: any objective property would do.

[17] This is not exactly true, but the way in which it is not true won't be important for what follows. There are at least three further dimensions along which we gather information about colors: hue, brightness, and saturation. Furthermore, there are also phenomena like the one exhibited by so-called binary hues: they are represented in experience in a way that these hues seem to be composed of relevant unique hues (e.g., orange is represented as containing, in some sense, red and yellow, whereas red, like other unique hues, is not represented as being constituted by other hues). These complicate the claims made in the main text, but not in a way that alters the main point, which is that there is a limit to what determinables our sensory systems can discriminate; at bottom, some determinate values of these will have to be represented as simple / atomic. Whatever these determinables are, sensory concepts will be those based on these such that the abstraction distance will be maximally short. This is true for all sensory modalities. As we said, we will ignore this complication in what follows.

Also, we will talk for convenience as if there were no abstraction distance between the sensory representations of, say, red, which represent determinate shades of red, and the most specific concepts of shades of red we can diachronically discriminate. So we will say that this distance is maximally short.

However, we believe that there is a deep point to be made on the basis of these reflections: autonomous representational systems are nomologically bound to be hooked up to their environments in a way that at some level of abstraction they will always harbor sensory representations that will represent complex physical properties in their environment as simple or atomic, or rather as having no constituents. Furthermore, if an autonomous intentional organism has concepts at all or a conceptual system (as opposed to just sensory representations), however primitive or sophisticated, then necessarily it has *some* sensory concepts in our sense. It is these necessities which partly create the mystery around phenomenal consciousness. If this is so, it would be interesting to know what it is about our world that generates these necessities.

[18] The point about how our sensory representations represent certain determinables (as determinate simples) has been made by a number of philosophers before (Armstrong, Shoemaker, Lycan, as well as the British empiricists in general), but for a detailed elaboration of this idea in the context of discussing sensory / pictorial representations, see Kulvicki (2001). He tries to give a principled distinction between primary and secondary qualities in terms of how

determinate geometrical figure of that type will not, in our sense, have a maximally short abstraction distance between it and the sensory base it is directly acquired from — even though these sensory bases are the sole authoritative source of acquisition for such concepts. We will call such concepts *perceptual concepts*. The information necessary and sufficient for the correct application of these concepts, whose abstraction distance is nevertheless not maximally short (but shorter than what we will call observational concepts), is contained in the sensory base from which they are directly acquired. Typical perceptual concepts in the case of vision include concepts of spatial / temporal relations, geometrical figures, and shapes.

For the sake of completeness, we can distinguish sensory and perceptual concepts from *observational concepts* like the concept of an apple, a robin, a tree, a lake, and a truck. These concepts are also *typically* acquired from an appropriate sensory base, but need not be, and sometimes are not. However, the information contained in experiences required in the correct application of these concepts is necessarily more impoverished, in the sense that it typically underdetermines correct categorization. In other words, although the information about the denotations of these concepts can be perceptually available, its delivery requires that certain channel conditions *external* to the sensory systems be in place. The abstraction distance between these concepts and the sensory bases from which they may be acquired is considerably greater than in the case of sensory and perceptual concepts. What seems to mark the difference is that (most of) the sensory information used in the acquisition and deployment of observational concepts is typically only contingently related to the denotation of these concepts.

It is no accident that thought experiments involving spectrum inversion are carried out in terms of sensory bases of sensory concepts, where the property detected and denoted is represented as simple or atomic.[19] Although *we* cannot conceive of inversion with respect to the

---

they are represented in conscious sensation, i.e., in epistemic terms, in opposition to the more traditional way of drawing the distinction in metaphysical terms.

[19] Strictly speaking, we should rather say that the sensory representations of secondary qualities do not represent them as having a complex constituent structure. This is different from saying that they are represented as simple or atomic. But we will be relaxed about this in what follows. For the distinction, see Armstrong's discussion of the headless woman fallacy in his (1968).

properties denoted by perceptual concepts (e.g., of shapes) and their sensory bases, there is nothing preventing a differently organized cognitive system from performing this feat. We can imagine and even construct devices that "sensorially" detect geometrical shapes (quite abstract from our cognitive point of view) by outputting simple and primitive sensory representations. For instance, we can construct a detector that responds with a green light when it detects a square (*any* square) and with a red light when it detects a circle (*any* circle). Suppose that all the information it uses in making its responses is lost at the final stage of output. When this device, a 2D geometrical shape detector, lights up green, its relevant state carries information that something it is informationally connected to is square. If it lights up red, its state carries the information that something is circular. But even if the "sensory" outputs of the device carry these pieces of information, they are structured in such a way that there is no way to recover any information about the structural relationships holding among the internal constituents of these shapes. Of course, these "sensory" outputs also carry information about the constituent properties (necessarily so), but only in analog form which is not extractable. Nor is it possible to extract any topological information that obtains between these different shapes — if the device carries information about the shapes in a spatial expanse. For all the device "knows", whatever is being represented by these colored lights, it is simple and atomic. There are no computational/formal constraints stemming from the representations themselves that would make the thought experiment of an "inverted shape" unintelligible here. For all the device "knows", circles could look exactly the same to it as squares do now, and vice versa.

If this device is also equipped with a central conceptual system that can acquire concepts from such "sensory" representations, the concept of a circle the device directly acquires from its "experiences" will be a *sensory* (as opposed to *perceptual*) concept in our sense. *Our* concept of a circle is not sensory because the sensory representations from which it is acquired don't carry the information that something is a circle as part of its total digital informational content so that when our conceptual system digitalizes this piece of information there is always more specific information that is lost but nevertheless available to the central cognitive system for

conceptualization / digitalization;[20] furthermore it is this lost information that seems to be used in the acquisition and vertical deployment of the target concept. What prevents the abstraction distance from being maximally short here is the existence of more specific but used-and-then-discarded information that is nevertheless available to the conceptual system for digitalization (which, subject to some further conditions, makes this used-but-then-discarded information contained in the experience *consciously available*).

In contrast to our perceptual system, the architecture of the device is such that the abstraction distance between the "sensory" and "conceptual" representations of circles / squares is maximally short. Not surprisingly, we are not such machines. But it is important to keep in mind that there is no *logical* necessity in our having the perceptual and cognitive architecture that we do, including set of particular abstraction distances it gives rise to — although there are probably evolutionary / ecological reasons for this architecture.

Another way to see what makes sensory concepts so special is to understand the nature of the abstraction distance between them and their sensory bases. As we have said, this distance is maximally short (subject to the qualification we have just introduced), which is what marks these concepts off from the rest. Following Fodor (1990) and Margolis (1998), we will call the mechanisms that mediate the informational link between the vertical tokenings of a concept and the instantiations of the property it denotes *sustaining mechanisms*.[21] The intra-cranial portion of the sustaining mechanisms for sensory concepts is not cognitive: since there is (almost) no loss of information in the acquisition of color concepts, there is nothing to be made available to the central system for digitalization. Acquisition of sensory concepts is therefore brute and primitive: to acquire these concepts it is enough to occupy the relevant sensory states

[20] Lost in the sense that the tokening of the target concept, which is more abstract, does not carry it.

[21] We prefer this notion to Dretske's notion of channel conditions because it is more specific and suggests mechanisms internal to the agent, which is what we would like to emphasize here. Although cognitive factors (what is independently known about the information source) can be part of channel conditions, Dretske, with this notion, emphasizes those conditions external to the agent, or at least external to his mind.

for an organism equipped with an appropriate conceptual system — i.e., by an information pick-up system operating on the sensory representations. This is why the notion of *learning* is not appropriate for the acquisition of these concepts. Rather, the preferred term for this, for both empiricists and nativists, is "triggering". So one sense in which the abstraction distance is maximally short is that the process underlying the acquisition and vertical deployment of sensory concepts does not involve any loss of information which is nevertheless available to the conceptual system for further digitalization.

Contrast this to the intra-cranial sustaining mechanisms for other concepts, which are (partially but essentially) *cognitive*. The acquisition and deployment of perceptual concepts may be automatic and innate in some sense, but these still involve a digitalization process with considerable loss of information, information that is still available for digitalization. When we visually recognize shapes of objects or geometrical figures, most of the information about their spatially distributed and organized constituents (illumination gradients, edges, corners, curves, color, etc.) are still consciously available. It isn't that *we* consciously use this information somehow inferentially in the acquisition and deployment of such concepts — this is something our perceptual (as opposed to sensory) systems automatically do for us. But what is interesting is that even though this process may be automatic and unconscious, (most of) the information used in the process (which is then discarded) is available to *us*, to the central cognitive system, and thus conscious in just that sense. Because of the importance and centrality of perceptual concepts, their acquisition, although cognitive in our minimal sense, may still be innately determined — i.e., such concepts may be triggered rather than learned. We leave this issue open.

The notion of learning seems most appropriately applied to the acquisition of observational (and for that matter, theoretical) concepts. The sustaining mechanisms for those concepts are heavily cognitive, involving the use and loss of a great amount of information,

which is also normally consciously available.[22] Normally, the more cognitive the sustaining mechanisms of a concept, the greater the abstraction distance between it and the sensory bases it is acquired from.

## 4. WHAT MAKES SENSORY CONCEPTS SPECIAL

It is not accidental that the distinction we drew between sensory and perceptual concepts is roughly co-extensive with the distinction traditionally drawn between concepts of secondary and primary qualities, respectively.[23] Secondary qualities are those which are represented in our experiences in a primitive way: sensory representations carry information about them in a way that makes the information carried about their constituents analog but non-extractable (we need science to extract these!). Hence, sensory experiences carry the most specific information about these properties without revealing their internal structure. This is why the abstraction distance between the concepts of secondary qualities and their sensory bases is maximally short; equivalently, this is why the acquisition of these concepts is non-cognitive or brute.

Sensory concepts apply, in the first instance, to the objects of perception, to whatever it is that our sensory experiences represent.[24] This is so despite the fact that they are directly and immediately acquired from sensory representations. The flow of information required for their acquisition (and vertical deployment) necessitates the presence of sensory intermediaries that

---

[22] See Margolis (1998), and Laurence and Margolis (forthcoming) for a parallel account of concept learning that involves *cognitive* sustaining mechanisms, which is nevertheless not a hypothesis forming and testing model à la Fodor (1975, 1981). They show, within a similar framework, that many lexical concepts may be primitive despite being learned from experience; hence they deny that atomism implies a radical nativism of the sort endorsed by Fodor. This is good news for the Language of Thought Hypothesis (LOTH), and concept atomism because it frees them from one of their main burdens. Their work nicely supports and complements the Dretskean account given here.

[23] "Roughly" because we think that the match may not be perfect. If some spatiotemporal properties / relations (like being a point or an expanse), as we suspect, turn out to be primitively represented in our experiences, then they may turn out to be categorized as secondary, contrary to the tradition. Although we think that there are principled ways to avoid this consequence, still we are happily prepared to live with this consequence if it turns out we can't avoid it. See Kulvicki (2001) for further discussion.

carry information about the properties denoted by these concepts. Indeed, this is one of the main differences between sensory concepts and observational concepts.[25] There is an asymmetry in their acquisition: while sensory concepts are necessarily acquired from the experiences sensorially representing the properties they denote, observational concepts are different. Observational concepts can be and typically are acquired from experiences representing their denotations, but this is not necessary. We can acquire them "horizontally", i.e., by sensory means (speech perception, seeing pictures, reading books / newspapers, inference, etc.) that are only very indirectly related to, and hence don't carry information about, their denotations.

There is a deep reason for this asymmetry which we haven't touched on so far but will be very important for what follows: the information about the secondary qualities contained in experiences cannot be *completely digitalized* by the conceptual system, whereas the conceptual system can completely digitalize the information contained in experiences about the properties denoted by observational concepts.[26] The notion of "complete digitalization" is a technical term introduced by Dretske which indicates a necessary condition for a piece of information to count as the semantic content of a concept. Recall that the semantic content of a concept is the most specific information its vertical tokenings carry about the objects it applies to, which is equated with its digital informational content. But Dretske eventually refines this definition by requiring that the semantic content is that piece of information which is *completely digitalized*. Here is the definition (1981: 185):

---

[24] Exceptions to this claim are what Armstrong (1962, 1968) called "intransitive bodily sensations" like pains, itches, and tickles, which we will take up later on.

[25] We would like to put aside perceptual concepts for the moment. Their discussion requires empirical data we are presently ignorant about. Accordingly, their treatment requires an empirically informed solution to "Molyneux's problem", which we don't presently have.

[26] Again the verdict about perceptual concepts is not clear: they fall in between sensory and observational concepts. But our intuition is that they will be closer to sensory concepts in that they can't completely digitalize the information about their denotations. Dretske himself seems to think this is the case, see fn. 4 in (1981: 261), and below.

Structure *S* has the fact that *t* is *F* as its semantic content [i.e., *S* is the concept of an *F*] =df.

(a) *S* carries the information that *t* is *F* and

(b) *S* carries no other piece of information, *r* is *G*, which is such that the information that *t* is *F* is nested (nomically or analytically) in *r*'s being *G*.

Intuitively, the intention is to rule out those cases where concepts carry the most specific distal information about an object by carrying information about their proximal causes, in our case their sensory bases.[27] So, for instance, the concept ROBIN, when acquired from experiences that carry information about robins, should not carry information about the structure of sensory representations which give rise to ROBIN.[28] Since we are working in a naturalistic framework, if concepts carried information about sensory representations from which they are acquired, this information would be information about the instantiation of certain neurophysiological properties (or disjunctive sets of such properties) realizing these sensations. Hence our concepts would be selectively responding to such properties in the first place. And this would imply either that our concepts represent neurophysiological conditions, or that our sensory concepts have dual semantic content, and therefore are systematically ambiguous.

Interestingly, Dretske does not make a point about the empirical impossibility of complete digitalization; nor does he talk about the fact that, *de facto*, complete digitalization is

[27] There is a parallel condition intended to rule out such cases in Fodor's version of informational semantics (Fodor 1987, 1990): the asymmetric-dependency condition, which says that a cognitive structure (concept) *C* has the semantic content *X* in virtue of an informational law between them such that if there is another such law between *C* and anything *Y*, then this law asymmetrically depends on the existence of the law between *C* and *X*. In other words, breaking the law between *C* and *X* breaks the law between *C* and *Y* but not vice versa. Like Dretske's condition, this account, too, fails to assign the distal secondary property to a sensory concept as its semantic content, and ends up assigning, instead, proximal properties underlying the sensory representation of the distal property. For it is clear that breaking the law between the sensory representation of red and the concept RED will break the law between the property of redness and its concept, but not vice versa.

[28] We will use uppercase letters to name concepts and italics to name properties denoted by concepts. To many ears, talking about a property as the denotation of a concept may sound strained. Although nothing very important hangs on this, we not only find this usage convenient, but also are prepared to justify it on the basis of informational semantics, which freely traffics in property instantiations.

routinely violated in the case of sensory concepts.[29] If sensory representations of secondary qualities are realized by a homogeneous set of neurophysiological properties, or by a finite disjunction of such properties, then vertical tokenings of sensory concepts carry information about their distal causes (instantiations of secondary qualities) by carrying information about the instantiations of these proximal physical properties. Whether or not experiences of such qualities are physically realized in a homogeneous way is ultimately an empirical question, but we think that there is enormous empirical as well as a priori evidence that this is the case — certainly intrapersonally, and most probably interpersonally.[30] What is important for our purposes is the minimal claim that the neurophysiological realization bases of sensory representations of such qualities are not arbitrarily / indefinitely varied, but consist of a finite disjunctive set of physical properties, and are more or less homogeneous in just this sense. We think that this claim is true, but we stand ready to be corrected by future empirical evidence.

There are also overwhelmingly strong engineering reasons for this claim: whenever you make an architectural distinction between a sensory buffer and a conceptual system that extracts information about the distal layout from this buffer (and whose behavior is causally sensitive to what this buffer contains), there will be a need to correlate the information carried by concepts and the elements of the buffer in such a way that matches up with the distal layout. If the only way the conceptual system carries information about the distal properties is through a physically realized sensorium, then it had better be the case that the same elements of this sensorium carried the same information, at least in the case of secondary qualities where the abstraction distance is maximally short. Otherwise, the informational efforts of the conceptual system will be fooled. From an engineering perspective, it is unclear how such an architectural

[29] In fn. 4 to Chapter 8 of his (1981: 261), Dretske shows signs of being cognizant of the problem here. He draws a distinction between a *visual concept* of a robin and the concept of a robin *tout court*, and says that the former is not completely digitalized. But the point is not elaborated.

[30] Interpersonal cases pose special problems (of the sort Shoemaker 1981 / 97 highlights) that we will discuss elsewhere, in the context of spectrum inversion thought experiments against functionalism about qualia.

design can be constructed without making the realization bases of those sensory representations more or less physically homogeneous (not arbitrarily varied), at least within a single system.[31]

Notice that in the case of observational concepts there is no real problem about complete digitalization. There are indefinitely many ways robins, trucks, etc. can strike our sensory receptors, and thus many ways in which they can be represented in experience. In such cases, the standard information-theoretic remedy is to say that these concepts track their distal causes without tracking proximal sensory representations, since the alternative is to say that they track a massively (probably open-ended) disjunctive proximal property. We believe that the former is indeed more plausible than the latter. But if so, we can now see better why sensory representations carrying information about properties denoted by observational concepts are not necessary for acquiring such concepts. These concepts, though observational, are modality-neutral (amodal), and to that extent not perspectival. But that is not to say that their cognitive sustaining mechanisms don't involve sensory / perceptual channels and concepts; they do. It is to say, however, that the sustaining mechanisms involved provide information only (mostly) contingently related to the denotation of these concepts.

It is the failure of complete digitalization that makes sensory concepts special by giving them a *perspectival* and (*quasi-)indexical* character. Their acquisition, semantics, and vertical deployment are essentially host-unique in two senses:

(i) it matters essentially for whose cognitive system these cognitive structures function as concepts, and

[31] We in fact suspect that even a stronger claim is true: sensory concepts are those tokens of the structures realizing the sensory representations of secondary qualities (even perhaps primary ones), but used differently by somehow being recruited by the central system. For sensory concepts we are prepared to accept what we take to be the central claim of Barsalou (1999) and Prinz (forthcoming), namely that there is no fundamental distinction between percepts and concepts, and that concepts are percepts (only used and organized differently): hence, necessarily, sensory concepts are not amodal.

(ii) they track features of the environment (instantiations of secondary qualities however objectively understood) essentially by tracking something *about their host*, namely, the sensory experiences from which they are directly acquired.

In other words, these are concepts which a properly functioning conceptual system *cannot* normally acquire unless suitably hooked up to a properly functioning sensory / perceptual information delivery system *of the same host* that has *actually* delivered the necessary information, i.e., carried information about the properties denoted by these sensory concepts.[32] We also want to emphasize that their acquisition is *direct* and *immediate*, by which we mean this: their sustaining / acquisition mechanisms are not cognitive, but primitive / brute; that is, they don't involve the exploitation of consciously available information which is then discarded in the digitalization process. This is roughly to say that the abstraction distance involved in their acquisition is maximally short. In the context of our discussion above, this implies that no information about the internal constituents of the properties denoted by sensory concepts is available in an extractable format: they don't represent their denotations as having a complex internal structure.[33] All these points about sensory concepts will be crucially important later on, when we criticize conceivability arguments against physicalism.

---

[32] We want to emphasize the qualifier 'normally' here. Although we adopt an actualist psychosemantics in this paper in order not to further complicate exposition, strictly speaking, we would like to remain, in general, neutral between an actualist psychosemantics (à la Dretske) and a purely counterfactualist one (à la Fodor (1987) and (1994)). If we adopt a counterfactual account, then we can cast the point in the text in terms of a narrow content conceived as a partial function from contexts to semantic content, in which case actual delivery of information for genuine sensory concepts is not required, but what is required is a "sensory" state that would track in a given context what it would in the actual world, namely secondary qualities of objects conceived as objective properties. Systematic hallucinations of such a quality would then enable one to acquire a genuine sensory concept. See below.

[33] We remind the reader about the distinction between "they represent their denotations as simples / atoms" vs. "they don't represent their denotations as having a complex internal structure, as having constituents". We always mean the latter if and when we sometimes talk in the former way for convenience.

Also, it is clear that more needs to be said about the nature of the perspectivalness involved: in particular, we need to distinguish *pure indexicality* from what we might call *quasi-indexical predication*. Although sensory concepts in our sense are mental predicates, there is obviously something indexical about them: their semantics is hostage to where and when their

## 5. FIXING THE SEMANTIC CONTENT OF SENSORY CONCEPTS

What justifies the claim that, despite the failure of complete digitalization, the semantic content of a sensory concept, say RED, is the secondary quality, *redness*, possessed by the objects of the sensory experiences from which we directly acquire it?[34] Irrespective of what semantic content our theories assign to these concepts, there should be no doubt about what their semantic contents are: they are the qualities that our experiences represent the external objects as having. Our (exteroceptive) experiences place these qualities in the world of objects external to our bodies. So do our sensory concepts. Given this, the question before us is how to reconcile a Dretskean informational semantics with the failure of complete digitalization. For surely, even if we rightly want to be able to say that RED represents *redness* despite the failure of complete digitalization, what justifies rejecting the option, which seems to be a consequence of the theory, that the semantic content of RED is the *experience* of *redness*, i.e., E-red, realized by a certain set of neurophysiological properties?

Here is another way of putting the problem. Informational semantics starts with the information carried by a structure on its way to working out how to determine its semantic content (**SC**). We have seen that Dretske wants to assign the completely digitalized informational content of a concept (*C*) as its semantic content: in other words,

- the semantic content is supposed to be the most specific information carried by *C* about a source (*o*) such that there is no separate structure (*e*) such that *C* carries the most specific information about *o* by carrying the most specific information about *e*.

---

tokenings occur. Fully developing the notion of quasi-indexical predication will require a paper on its own, which we will take up elsewhere.

[34] We will use the abbreviation E-p to denote the experience of property *p*. So, our terminology includes, for example, RED (the concept), *red* or *redness* (the property), and E-red (the experience).

But this assigns E-red as the semantic content of RED — assuming, as we do, there is no further informational intermediary of the relevant sort between E-red and RED tokenings. The theory gives us the wrong result.

Let us say that the (most specific) informational content (**IC**) of RED, which interests us, can be given by an ordered pair:

**IC**(RED) = <*redness*, E-red>.

The structure of the information flow is such that RED carries information about *redness* by carrying information about E-red.[35] If we want to insist, as we should, that

**SC**(RED) = *redness*

despite this informational alignment, we have to modify the content-assigning mechanisms of a Dretskean informational semantics, and we have to do that in a principled way.

One option, which Dretske himself might be tempted to take, is to invoke teleology: the semantic content of RED is determined by whatever indicator function the tokenings of RED are supposed to serve. It might be plausibly claimed that it is the *redness* of external surfaces that RED has the function of carrying information about, and not E-red. Indeed, we think this claim is not difficult to justify on the basis of evolutionary considerations, by appealing to the idea of the adaptiveness of cognitive structures given our practical needs and interests in our environment. In fact, we believe it to be true. But it doesn't solve the problem. For, as Dretske (1986) himself is aware, if we try to determine semantic content in terms of indicator functions, the problem about the indeterminacy of semantic content has a tendency to translate into a problem about the indeterminacy of function. For then it is possible to argue in the following

way. The true function of RED is to indicate E-red, but since E-red is perfectly informationally correlated with *redness*, any need or interest that the organism might have related to *redness* will be satisfied by a structure whose job it is to indicate E-red. In other words, we can equally well claim that RED has the function of indicating *redness* in virtue of having the function to indicate E-red. After all, when, in abnormal circumstances, E-red fails to correlate with *redness*, RED's functioning is not to blame; it does its job just fine, it is the world that doesn't cooperate — or so the intuition goes. Dretske's solution to this problem in his (1986) is not applicable to sensory concepts since his proposal comes down to an acknowledgement that the problem of indeterminacy of function can be solved only for those concepts which can completely digitalize the most specific information they receive about the things they are supposed to denote, i.e., only for those concepts whose abstraction distance is large enough to allow them to be acquired from an indefinitely large set of proper sensory bases, and sensory concepts are not among these.

So, according to informational semantics, what anchors the semantic content of RED to *redness*, instead of E-red? A crucial part of the answer, we believe, can be gleaned by reflecting on the integration of the information coming from a variety of intra- and inter-perceptual sources. Consider the visual information that our cognitive system uses in the acquisition and deployment of observational concepts, like ROBIN, CAR, and TOMATO. We have said that there is no serious problem about the complete digitalization of information with respect to these concepts. So we can safely claim that these concepts apply to external objects: they are much further away, in terms of abstraction distance, from the sensory experiences that give rise to them. But the acquisition and vertical deployment of these concepts utilize lots of sensory and perceptual information that is also consciously available (this is what makes their sustaining mechanisms heavily cognitive), which is to say, apt for digitalization by the same central cognitive system. Between the sensory experiences and the vertical tokening of ROBIN,

[35] The order in the pair is meant to reflect this dependency relation. The relevant information is carried by the appropriate *tokenings* of RED and E-red about the *instantiations* of the property

there is a lot going on informationally. Even though this process may be inferential/computational, it seems mostly automatic and unconscious. But much of the information used in the process is consciously available, such as the determinate size, shape (even the particular lines, curves, edges, etc.), texture (even the smaller changes in light intensities reflected), orientation, distance, and color (its variations across the surface) of the robin that has occasioned the tokenings of ROBIN. When we say this information is lost or discarded, we don't mean that it is forever hidden from consciousness; rather, it is lost from the perspective of the tokening of ROBIN, which is to say that the tokening of ROBIN no longer carries information about these more specific values along dimensions just mentioned. It is the loss of this sort of more specific information that enables me to visually recognize this object as a robin. But this information is also the same information used and integrated in the recognitional process and is consciously available, the conceptualization of some of which has a shorter abstraction distance (and, in the case of color concepts deployed, a maximally short one). As the abstraction distance gets closer and closer to the sensory experiences, concepts start to lose their completely digitalized character. Now if the conceptual system uses and integrates information about external things (e.g., about their determinate color, variations in light intensities, edge here, curve there, etc.) delivered by sensory representations *on its way* to categorizing these things as external objects (as robins, cars, tomatoes), then the conceptualization of the former kind of information had better anchor their semantic content outside of the organism; otherwise the conceptual system will not present to me a coherent picture of my environment, which is bad for my survival chances. It makes no sense (theoretically or biologically) to anchor the semantic content of RED to E-red if I am able to vertically classify the thing in front of me *as a red tomato*, especially when E-red involved in the sensory intermediary *actually* delivers the information about the *redness* of the tomato to the central cognitive system. RED *actually* carries this information (by carrying information about

---

*redness*. Most of the time, we will forgo talking this way for convenience.

E-red), and, precisely because of this, enables me to recognize the object as a red tomato. Indeed, otherwise what would be the semantics of RED TOMATO? The mind boggles.

Consider the 2D geometrical figure detector introduced above. Logically speaking, we could be like such a device with respect to the recognition of middle-size objects such as tomatoes, robins, and cars. (But in that case we would not live long, needless to say.) For, if we were like that, the information leading up to our recognition of these objects would not be consciously available to us. This means that we could not acquire any concepts usable in the discrimination of more specific information about these objects — even though our pre-perceptual system might actually make use of this very information in the recognition process, denied to the central system for purposes of conceptualization.[36] If we were like this, we would have the same problem about how to anchor the semantic content of concepts like ROBIN, CAR, TOMATO — these would be sensory concepts which would not represent their systematic distal and proximal causes (about which they carry information) as having complex internal structure.

Briefly, the partial answer to our original question comes down to the need for *coherent* integration of information. It is the pressure exerted by our practical / utilitarian interests in having a *coherent* global representation of our external environment that forces the conceptual system to pick out the first element, *redness*, in **IC**(RED)=<*redness*, E-red> as the semantic content of RED. We have seen how the processes integrating various sorts of information in the acquisition and vertical deployment of observational concepts generate a need for coherence. We have also seen how adaptive / evolutionary forces determine where to put this coherence: on the global representation of a reality *external* to one's sensory and cognitive systems. To point

[36] Blindsight subjects *might* be precisely in this position: even though they may recognize and categorize some things in their blind field under forced-choice conditions, they can't voluntarily apply the concepts to objects in their blind field despite the fact that they may have (some form of) intact sensory representations. In other words, the problem might not be a problem of sensory representation, but rather it may be related to what information the conceptual system is allowed to vertically pick up from such a representation. So, for instance, it is allowed to pick up the information that there is a horizontal line before one's eyes, but no further vision-specific

this out is, of course, not to specify the mechanisms by which this is accomplished. But this latter task is rightly to be left to psychologists or neuroscientists.[37]

## 6. CONCEPTS OF BODILY SENSATIONS

This partial account predicts that the less need there is for coherent integration (because of less information, the scarcity of its sources, or a redirection of immediate interest due to the affective / hedonic value of the experienced stimuli, etc.), the less pressure there is to anchor the semantic content of a sensory concept to the outside. We indeed find the gradual change implied by this in all sensory modalities and submodalities. Vision is the paradigm source of information generating sensory concepts whose semantic content is unequivocally external to the subjects. Things get increasingly less clear as we look at other modalities (hearing, smell, taste, and touch) according to how close to the body the detected properties are, how rich the information provided by the experiences in these modalities is, how much information from other channels is used, how impoverished the quality space determined by sensory experiences is, and / or how many quality spaces each (sub)modality determines along its different dimensions. The limit in this direction are the interoceptive submodalities of touch producing so-called intransitive bodily sensations, such as pains, itches, tickles, and the like.

If we assume that these sensory experiences carry information about, and thus have come to represent, certain bodily conditions like tissue damage (or, some such bodily condition), then one would naturally expect the same, or at least a very similar, informational division of labor that we find in exteroception, say, vision. We would expect that the job of these bodily sensations is to deliver information about bodily conditions to the central cognitive

---

information can be directly conceptualized. And since the patient doesn't know where the information about the horizontal line is coming from, he is not volunteering this information.

[37] Actually, we can further speculate on this by pointing out that certain syntactic requirements of a *system* of mental representations within the central conceptual system may generate the coherence requirement naturally: for instance, certain syntactic positions in this system may require predication of concepts denoting distal objects. Thanks to Jonathan Weinberg for pointing out this line to us.

system for digitalization resulting in the acquisition of sensory concepts which apply in the first instance to aspects of these bodily conditions. But curiously enough, this is not what we find. The sensory concepts PAIN, ITCH, and TICKLE apply to token experiences, to bodily sensations — well… to pains, itches, and tickles — not to the bodily conditions these sensations represent. The result is a curious asymmetry between sensory concepts (like RED) and concepts of bodily sensations. Despite identical information flow, sensory concepts acquired directly and immediately from the relevant experiences apply to different states (see **FIGURE 1**).[38]

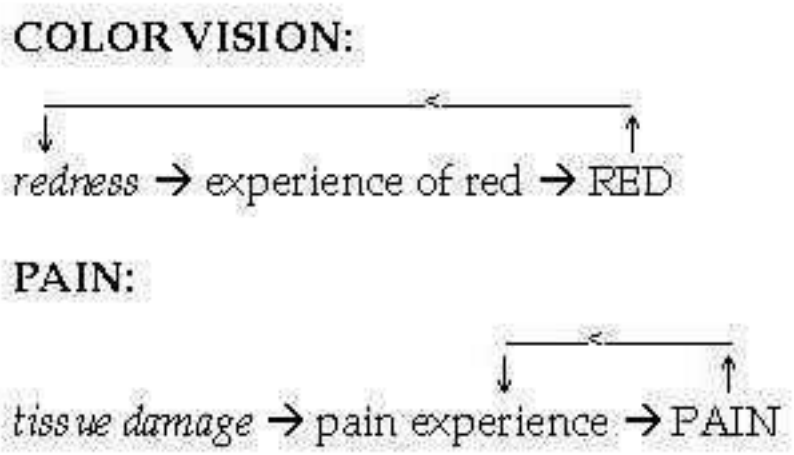


**FIGURE 1**: *Asymmetry in concept application despite identical information flow.*

Let us assume that the most specific informational content of PAIN, which interests us now, can be given by an ordered pair consisting of a certain sort of tissue damage and a sensory representation thereof, call it E-damage, more or less homogeneously realized by physical properties in the brain and delivering this information to the conceptual system. So,

[38] For an elaboration of this theme about pain and other bodily sensations in the context of an argument against pure representational theories of qualia, see Aydede (forthcoming). Note that even advocates of direct perceptual or representationalist theories of pain grant that our dominant / ordinary concept of pain (and of other bodily sensations) applies to token experiences. They insist, however, that these experiences represent certain bodily conditions, and that their qualitative content is exhausted by their representational content.

**IC**(PAIN) = <*tissue damage*, E-damage>.

PAIN carries information about tissue damage by carrying information about E-damage. So PAIN doesn't completely digitalize the most specific information it gets about the tissue damage. In this case, as a matter of fact, the semantic content of PAIN is anchored to E-damage:

**SC**(PAIN) = E-damage.

Why is this different from the case of RED?

If our partial answer to the parallel question above is right, then we can discern one part of the reason that explains the difference. It relates to the nature of the perceptual object, i.e., the object sensorially represented. Tissue damage is, of course, only one of many causes of pain experiences — we have used it as a stand-in for whatever it is that specific sorts of pains represent. These are mostly internal conditions of the body, normally not open to other sensory channels. So inter-modal sensory information for integration is either non-existent or extremely limited. Not only that; the quality space generated by noxious stimuli is quite impoverished compared to exteroception, especially vision. Although pain experiences sort out the noxious stimuli both temporally and according to a spatially articulated somatosensory map, there is not much information *integration* going on in a way so as to epistemically clue the conceptual system in on what it is that is being perceived. There is certainly information available in the pain experience to sort out different kinds of bodily disorders or damage. But again, this does not help to generate concepts whose abstraction distance is sufficient for complete digitalization. On the contrary, it appears to be the case that the quality space created by pain experiences gives rise to a corresponding set of sensory concepts whose abstraction distance is maximally short. There is certainly a lot of analog information about bodily conditions contained in pain experiences, but clearly it is in a format not extractable by the conceptual

system. Pain experiences don't represent bodily conditions they carry information about as composed of complex properties. They are also informationally impoverished, in that the conceptual system cannot digitalize concepts with sufficient abstraction distance from them.

As we have seen, however, the ability to generate concepts with sufficient abstraction distance such that complete digitalization can be obtained (which results in putting the semantic content of these more abstract concepts outside) was the key to the integrative processes which resulted in the need for a coherent picture of an external reality. For it is the more specific information about the qualities of the denotations of these more abstract concepts that is being used in their acquisition. So it is imperative for the purposes of (re)presenting a coherent reality that if this more specific information is also available for conceptualization it be attributed to the very same objects denoted by the more abstract concepts. We don't have sufficiently rich information to yield completely digitalized concepts in the case of bodily sensations: hence, the sensory concepts they give rise to apply to their proximal causes.

But there is a more important reason why E-damage is picked out as the semantic content of PAIN: whether or not E-damage represents tissue damage, and for that matter, misrepresents it, it hurts! If pain experiences hurt irrespective of whether they are veridical, then it is not surprising that our immediate epistemic and practical focus is directed, in the first place, onto the experience itself.[39] The affective / hedonic tone of the experience puts a heavy demand on the cognitive centers to urgently re-allocate cognitive and behavioral resources and response priorities for stopping the experience by doing whatever is necessary to get rid of its cause. Again, what we see here is the cognitive adaptations of the conceptual system to align its semantics with the pragmatic and evolutionary needs of the organism.

Note that there is some room — albeit a small amount — in the folk conception for thinking of pains as pure bodily conditions. There are situations where we find it natural to talk

[39] For an elaboration of this theme, see again Aydede (forthcoming). We should note, however, that to the extent to which the awareness of pain is also awareness of this affective (hurtful) aspect of pain, introspecting it will involve mechanisms partly different from the ones we will

in a way that there may be unfelt pains: when we do that, we are talking of pains as disordered states of our bodies gone unnoticed. We talk about our headaches lasting, say, during a heated discussion even though we have not felt them most of the time. When we talk about the very same pain coming back, we find it natural to conceive of it as if it had already been there, unnoticed, unfelt (in fact, some cognitive / behavioral therapies for chronic pain utilize this phenomenon). This is certainly not the dominant conception of pain. But that such use has kept a foothold in the folk usage is remarkable, since it is precisely what our account of sensory concepts predicts. The information about the bodily condition is there: the vertical tokening of PAIN does carry it. Our ambivalence about what to say in such rare cases when pressed is generated by this double informational content, making it somehow possible to alter the semantic focus (see below).

A proper understanding of bodily sensations and their conceptualization is crucially important for a proper account of the informational architecture of the cognitive mind and the special role sensory concepts play right at the interface between sensory and conceptual systems. Notice that "pain perception" is, technically speaking, a form of introspection — if introspection is the means by which we learn about our own mental states from a first-person perspective. Our first-person knowledge of our pains, itches, and tickles is knowledge of our experiences. To know we have them is to know we have experiences. And to come to know that is to engage in introspection.

A proper understanding of bodily sensations is crucial because here we see the basic mechanisms of introspective access to our own experiences in their barest form, being located at one extreme of the spectrum of sensory representations. It is by working from this extreme that we hope to develop an account of introspection of experiences in general.

---

offer below — insofar as this affective quality of pain is non-sensory or non-representational. For an extension of our present account to cover affect, see our (in prep.).

## 7. INTROSPECTION

Knowledge, including introspective knowledge, requires discriminative and recognitional capacities, which are conceptual capacities, as we explained previously. This meshes well with the case of sensory concepts of bodily sensations: when PAIN is vertically deployed, i.e., when it is used as a classificatory response to noxious stimuli and their effects on body in a way that carries information about them by carrying information about the corresponding experience, E-damage, the *semantic content* of the *de re* judgment made is that [the E-damage to which PAIN is informationally connected] is pain.[40] But this judgment is an introspective judgment, i.e., a judgment about a token experience, a mental event. Insofar as the concept of pain is, intuitively, a mental concept, the judgment classifies a certain neurophysiological event in the brain carrying information about tissue damage under a mental concept.

But what exactly is the informational value of this judgment? What information does it convey? We have been talking about sensory concepts carrying information about the sensory representations from which they are acquired. But even if the kind of weak type-type identity theory we assume for sensory experiences of secondary qualities is true, what generates the information? When there is information in a signal about a source, what makes this possible is the elimination of alternative possibilities at the source that could have occurred, and the nomological dependency of the signal on these. If we want to talk about sensory concepts carrying information about experiences, we have to treat experiences as information-generating sources *on their own* — even when much of the information thus generated at the sensory level nomologically depends on the elimination of possibilities at a source beyond them, i.e., in the world.

---

[40] Single square brackets will be used in what follows as a way of indicating that the expression occupying the linguistic position marked by them is to be read as occurring transparently. The judgments expressed by such brackets in the subject position will be *de re* judgments, expressing singular propositions. The term 'pain' in the expression of this judgment is meant to express the *sensory* concept PAIN. For even people who are congenitally insensitive to pain (sic! — but this is the technical term standardly used) can have a concept of pain and can intelligibly communicate their pain thoughts. The concept these people have is not a *sensory* concept in our sense.

This is precisely what we find when we look at experiences from the point of view of *sensory concepts*. Experiences are venues for information entry to the central conceptual system. There is as much information generated at the sensory level for pick-up by the conceptual system as there are different venues (sensory modalities and submodalities), distinct dimensions within these venues (pitch, frequency, amplitude; color, geometry, light intensity, etc.), and different (usually continuous) values each of these dimensions can take (red, orange, yellow, etc.; loud, very loud, even louder, etc.). Not only can we discriminate reds from oranges, oranges from yellows, but we can also discriminate a color from the spatial expanse of which it is the color, as well as discriminate visual experiences from tactile, auditory, gustatory, and olfactory ones.

It is this multiplicity of information entry that allows us to treat sensory experiences and its parameters as information-generating sources. Of course, if there is information, there is no logical guarantee that there will be something carrying this information; but our conceptual system has evolved precisely to pick up this information and use it in the service of guiding behavior. When I token RED in response to a ripe tomato, my concept does carry information about which neurophysiological property[41] is instantiated in the relevant part of my visual cortex. Although the information carried by RED does not represent this property as having a complex structure, it does eliminate other possibilities relevant for the color dimension of my visual experience: it could have represented green or yellow, etc.

When we make similarity judgments like "*x*'s color is more like *y*'s than *z*'s" we are making similarity judgments about the colors of objects on the basis of a similarity ordering of our color experiences. Indeed, if we look at what color science seems to tell us, because of metameric phenomena most of the time we find no similarity in the particular spectral reflectances paralleling the similarities our experiences represent. It is the sensorially represented similarities that the conceptual system picks up in making these similarity

[41] As we have said, this may be a psychofunctional property, realized by a small disjunctive set of physical properties of the relevant sort.

judgments.[42] Here it is useful to appeal to a quality space generated by how the experiences represent their objects. The inferential regimen governing our sensory conceptual repertoire reflects or parallels the (inter-)relational structure of this quality space. But this space must be such that our conceptual system is able to pick it up from the similarity relations among color sensations. And the only way of doing that we can imagine from an engineering perspective is a structuring and ordering of the physico-functional properties of the relevant brain states in a certain way, i.e., so that it will act as an information-generator of the right sort — the sort that enables the conceptual system to end up with the particular set of sensory concepts and particular inferential structure which it actually has.

It isn't just the different values of a dimension of an experience that generate this kind of information exploited by the conceptual system. As we have mentioned, the conceptual system is also sensitive to variations along intramodal dimensions, as well as activations of the different modalities themselves. There should be no controversy about this: the information is there to be picked up, and our conceptual system does, as a matter of fact, register it.[43]

So what is the informational value of the *de re* judgment, "[the E-damage to which PAIN is informationally connected] is pain"? It consists of whatever other possibilities are eliminated by the instantiation of the relevant neurophysiological property constituting E-damage. It is not only *this* kind of pain I am having (different from *this* and *that* …), but also it is *pain* I am having, not an itch or tickle, or a sensation of warmth, or a sound for that matter. It may be that not all this information is being carried by PAIN: when I discriminate stinging pains from pricking ones, I seem to be deploying more specific sensory concepts (with a shorter abstraction distance) than just classifying my sensation as pain, eliminating only, say, the possibility of its being a tickle or a sensation of mild warmth, and so forth. The general point, however, should

[42] Cf. Shoemaker (1981/97, 1994).

[43] Hilbert (ms.) contains an insightful discussion of why having this information is an extremely useful thing for cognitive organisms like us with sophisticated and peculiar epistemological needs— not just for doing philosophy, but for adaptive behavior that has survival value, since we need information about the sources of our perceptual beliefs to assess their reliability.

be clear: the content of such judgments is determined strictly according to information-theoretic principles: elimination of the relevant alternative possibilities.

What needs to be emphasized here — and what may be obscuring the clarity and plausibility of this naturalistic picture — is that the way PAIN indicates / represents this neurophysiological property reflects the way E-damage (pain) indicates / represents the tissue damage. E-damage carries analog information about the tissue damage in a non-extractable format so that this information is not available to the conceptual system for further digitalization. That is to say, whatever most specific information E-damage is sensorially carrying about the tissue damage, none (or, very little) of the information nested in it is available to the conceptual system for extraction. But this amounts to the fact that E-damage does not indicate / represent the tissue damage as having a complex structure. The conceptualization of this information by the central system reflects this condition: PAIN[44] does not represent the neurophysiological property it detects as having a complex internal structure.

But a vertical tokening of PAIN does carry information about what other possibilities are eliminated. When we apply PAIN vertically we don't represent the property to which it is actually applied as physical (how could something be physical if it doesn't seem to have *any* internal complexity to it?). But, we don't represent it as non-physical or immaterial either (how could we locate something non-physical in the body?). PAIN is basically topic-neutral on this issue, as J.J.C. Smart (1962) insightfully pointed out a long time ago. There is nothing peculiar or mysterious about any of this, if we keep information-theoretic principles firmly in mind. Simple signals can carry information about quite complex properties without making this information, on their own, available for further extraction. This is precisely what happens with our sensory experiences and sensory concepts.

## 8. PHENOMENAL CONCEPTS

As we have seen, "pain perception" is, technically speaking, a form of introspection. It does, however, share its information-flow structure with other forms of sensory processing and concepts. How is it then that one counts as introspection while the perceptual categorizations accomplished with exteroceptive sensory concepts do not? The latter are, in the first instance, perceptions (i.e., vertical categorizations of external stimuli under concepts) of an external reality; they are not perceptions of brain states realizing the sensory representations mediating information flow — although they carry information about these states too. We expressed this asymmetry earlier by saying that although both kinds of sensory concepts have dual informational content, their semantic contents are differently anchored or focused. We have explained why this should be expected given the immediate informational and practical needs of organisms like us shaping our selectional history, and the way the sensory information is integrated at different levels of the abstraction process.

How do we introspect our exteroceptive experiences generally? How do we come to know what it is like to see red, to hear C sharp played by a clarinet, to smell sulfur dioxide, to taste dark corn syrup, to feel a warm and soft object touching our cheek? As should be obvious by now, we propose that introspection of such sensory states involves a different utilization of the very same sensory concepts deployed in the perception of the external properties that these sensory experiences represent. Introspection of exteroceptive experiences is that mechanism or capacity that capitalizes on the second element of the *information* content of a sensory concept by selecting it as the *semantic* content of the concept. Clearly, as we have seen in the case of concepts of bodily sensation, our cognitive system is capable of doing this — they are the existence proof for such a capacity. But we do seem to have the introspective capacity at large to pick up information generated intra-personally by the multitude of information entry channels and make it the semantic focus of our sensory concepts.

---

[44] Or, whatever most specific concept we can deploy in *re*cognition tasks.

Vertical tokenings of RED carry information about experiences of red. But sensory concepts carry information about brain states without carrying information about their constituent structure in an extractable format, just as sensory experiences of secondary qualities carry information about, say, colors, even though the information about colors' constituents is not extractable from them by the conceptual system. In this, vertical tokenings of sensory concepts like RED discriminate and classify the relevant range of brain states as (more or less) simple primitives with respect to their intrinsic nature but as having external relations to other such states paralleling the sensory quality space represented by them. More precisely, sensory representations of secondary qualities carry the most specific information about them by carrying the analog information nested in this information in a non-extractable format. Similarly for vertical tokenings of sensory concepts: they carry the (more or less) most specific information about the sensory representations of secondary qualities without carrying the analog information nested in it in an extractable format. Introspection is precisely that mechanism which takes the second element in the *information* content of sensory concepts and makes it their *semantic* content.

Here we need to introduce a further distinction for types of concepts, following the dictum "different extensions / denotations yield different concepts". Instead of talking of the *sensory concept* of RED (call it s-RED), whose semantic content is the property *redness*, being utilized in a different way, we can talk about the *phenomenal concept* of RED (call it p-RED), whose semantic content is the *experience of redness*.[45] As we have seen, concepts of bodily sensations like PAIN are already phenomenal concepts in this sense: they apply to token experiences.

But how does the semantic switch or shift occur? The source of the phenomenal concept, p-RED, that introspection utilizes is the very same structure underlying the sensory

[45] But we will be relaxed about the terminology as long as it is clear what we mean by "using the same concepts in a different way".

concept, s-RED, which the perceptual categorization of distal stimuli deploys. The significance of this can be captured by (ES):

(ES) When p-RED is applied to experiences of *red*, it is impossible not to categorize the experiences, by this very application, as the *epistemic source* of the perceptual judgment/categorization of a distal stimulus. But this is just to categorize these experiences as representations of *redness* of a certain kind, i.e., of the kind *this* [sort of brain state] subserves — if and when we have the necessary intentional concepts (see below).

The truth of (ES) is the source of the familiar claim that introspection is "transparent" (at least in exteroceptive modalities): i.e., the reason why the properties we encounter when we introspect our experiences seem all to be the properties that our experiences *detect* rather than *exhibit*, is that introspection uses the same sensory concepts in a different way — by choosing the second element in their *information* content as the *semantic* content of the very *same* concept used to classify *what* in fact is represented by the experience.

The extent to which this capacity or mechanism to shift the semantic focus of sensory concepts like RED is innately given, and the extent to which it depends on ontogenetic maturation processes or cognitive development, is an open question. We think that the evidence from developmental psychology indicates that this capacity is acquired only after the acquisition of intentional concepts and a modicum of folk psychology. Children seem to acquire these concepts and this mastery fairly early — approximately at the age of four. We believe it is no accident that the acquisition of the capacity to introspect one's experiences emerges only after this development.[46]

[46] While having intentional concepts is necessary for the semantic shift in introspection, the acquisition and development of this capacity apparently takes more time and requires some intellectual sophistication. John Flavell's recent work [[refs.]] on children's introspective capacities suggests that this semantic shift, coupled with children's ability to start making sense

Intentional concepts, such as the concepts of information, representation, belief, and so forth, are acquired through third-person channels, not from one's own case.[47] Indeed, information theory does not allow for their first-person acquisition. The same is true for the concept of experience qua representation. However, even though intentional concepts cannot be acquired from a first person perspective, once acquired, they can be vertically deployed, which is to say that we can apply them to our own experiences *because* they are experiences.[48]

The exact way in which the acquisition of intentional concepts facilitates the acquisition of introspective capacities (and vice versa?) needs to be worked out in further detail,[49] but we think that this is more or less an empirical job to be left to psychologists and neuroscientists. Our main point is that at some stage in cognitive development we acquire the capacity to selectively focus the semantic content of our exteroceptive sensory concepts (we already pointed out the existence proof of its feasibility), and the acquisition of this capacity draws upon intentional concepts.

That we need to possess intentional concepts in order to introspect our exteroceptive experiences and what they are like is the reason why we don't *normally* think of pain perception as a form of introspection (even though, as we have seen, technically it is), since when we "perceive" our pains, what we "perceive" is a token experience that is *not*, indeed, *need not* be

---

of the notion of having a mind that "contains" ideas which can be voluntarily examined, occurs as late as age eight, which is later than most abilities studied in human development.

[47] This acquisition process may not exactly be *learning*; it can come about by the *triggering* effect of external stimuli.

[48] There is actually a nice account of this available in what is involved in '*k*' (the variable standing for what is independently known about the source) in Dretske's original definition of information in his (1981: 65): "A signal *r* carries the information that *s* is *F* = The conditional probability of *s*'s being *F*, given *r* (and *k*), is 1 (but, given *k* alone, less than 1)". The acquisition of the intentional idiom and folk psychology is the acquisition of independent knowledge about what is happening at the source, i.e. at the experiential level in one's own case.

[49] Note the revealing analogy that exists between acquiring this introspective capacity and the way in which acquiring new concepts from a third-person perspective makes one aware of new experiential qualities in wine tasting or listening to classical music (etc.): similarly, acquiring intentional / representational concepts from a third-person perspective makes one vertically aware of one's own experiences, aware of what they are like and what it is like to be in them, by somehow making it possible to shift the semantic content of sensory concepts we had already

conceived of in intentional / representational terms. The concept of pain here already has the token experience as its semantic content as a simple / primitive representation of a certain kind of brain state located within a quality space. No wonder pains have always been thought as paradigm cases of mental objects that don't themselves seem representational at all: we didn't have and didn't need intentional / representational concepts to "perceive" or come to know about them. Indeed, young children, as every parent knows, can think about and communicate their pains even before they have acquired the intentional apparatus of folk psychology.

So our proposal is that when we vertically apply p-RED to our experience of *redness*: the semantic content of the introspective *de re* judgment involved is something like:

- This is how *redness* is [registered]

  (or, experienced, sensorially represented, etc.),

where 'this' picks out a certain brain state primitively (only eliminating the relevant alternative possibilities and thus locating it within a relationally defined quality space — so it is predicative, not just purely indexical), i.e., without revealing its constituent structure. More accurately, we could have expressed it as "p-RED is how *redness* is [registered]", except that 'p-RED' is not English — see below.

One important consequence of this is that we now have a purely naturalistic (partial) explanation of the much-debated "reflexive" and "self-intimating" character of sensory states. In the very perceptual recognition of *redness* we also cognize the sensory experience mediating the recognition, and vice versa. The sensory concept RED is necessary for generating a cognitive structure, p-RED with the semantic content displayed above, since p-RED is the very same structure as s-RED only used differently *because* it carries information about both the sensory experience of red and *redness*. The "reflexive" and "self-intimating" character of

---

acquired in the process of perceiving the world around us. The information about experiences, similarly, is already there in the tokening of relevant sensory concepts.

sensory experiences stems in effect not from the experiences themselves but from the dual informational / semantic nature of the sensory concepts directly acquired from them. Put differently, and to relate the point to (ES) that we highlighted above:

(ES') It is the very "same concept" that is used both in picking out the relevant brain state (thus, eliminating the relevant alternative possibilities and hence locating it within a relationally defined quality space) and "commenting" on it as a [sensory registration] of *redness*.

Of course, "sensory registration"[50] is the intentional concept involved in the semantic shift. We conceive of the nature of this job in such a way that the intentional concept at issue can be quite rudimentary and basic — to the extent that the acquisition of folk psychology permits it in its earliest phases.

Obviously, if (ES') is true, there is a curious sense in which it is as if the same cognitive structure were used twice over simultaneously (as p-RED applying to E-red, and as s-RED applying to *redness*) in the introspective judgement about a red experience. Perhaps it would be useful to see that the semantic content of the relevant introspective *de re* judgment would be logically equivalent to the content of an internal vehicle (when used from one's own perspective, so to speak) whose "deep" *logical structure* can be displayed thus:

(LF) (i) [E-red] is p-RED, and

(ii) p-RED is how s-RED is [sensorially registered].

(If it helps, you may think of this conjunction as expressing the logical form of a mental sentence in the "introspection-box" of an introspector having an experience of red — we do not of course

[50] Or its referential equivalent, as indicated by square brackets in the previous sentence.

think that the internal vehicle involved in an introspective judgment made by a vertical application of a phenomenal concept is actually complex in this way.)

In (i), there is a primitive vertical application of a phenomenal concept (which picks out the information generated by the elimination of relevant alternative possibilities, say, E-green, E-yellow and so-forth), i.e., a phenomenal categorization of a token brain state which is the experience of red occurring at that moment. Simultaneously in conjunction with this, there is, in (ii), a "commenting" on this token categorization by using the same structure as saying what it is that the applicandum of p-RED (i.e., E-red) represents in the world, i.e., *redness* as expressed by s-RED. (LF) should be taken lightly insofar as it is suggestive and helps one understand the logic and the truth conditions of the *de re* introspective judgment involved. For as far as we can tell, we are trying to convey something that has always puzzled philosophers, something about the self-intimating character of experiences, and we propose that this stems from the peculiar nature of sensory and phenomenal concepts directly and immediately acquired from those experiences.

We don't know any other naturalistic account that integrates so tightly the vehicle of introspection with the vehicle expressing what the introspected state represents without giving up representationalism. We have already explained how closely these conceptual vehicles are informationally related to the target of the introspection (i.e., E-red) and to what it represents (i.e., *redness*). In fact, the so-called transparency of introspection that externalists emphasize so much is simply the other side of the "same coin": it naturally falls out of our account because of this tight integration. At the same time, as we hope to have shown, this account does justice to internalist intuitions, which we find important.

Another important aspect of this sort of introspective vertical processing is its sensitivity to the temporal window, or duration, of the activation of perceptual channels and its particular values. This is probably one of the major intuitions behind the tradition (found in Locke, Kant, and more recently Armstrong and Lycan) that regards introspection as a sort of internal sensing or monitoring (introspection as *inner sense*). But again it is worth emphasizing that this

monitoring eventuates in discrimination and conceptual categorization in the way we have explained. This feature of our account makes it a synthesis of otherwise quite opposite accounts of introspection: introspection as internal monitoring (Armstrong 1968, Lycan 1996), and introspection as higher order *thought* à la Rosenthal (1995, 1997, 2001), Carruthers (2000), Dretske (1995), and Tye (1995).[51]

## 9. CONCEIVABILTY ARGUMENTS AGAINST PHYSICALISM

Following Dretske's seminal work (1981), we have provided an information-theoretic account of sensory concepts, of how they are directly and immediately acquired from sensory experiences and how they are vertically deployed. We have also provided an account of the introspection of experiences based on the special informational / semantic nature of these concepts. We are now in a position to address some long-standing vexing philosophical problems. We will show how to reconstruct so-called conceivability arguments against physicalism from within the account of sensory concepts we have provided. Once we do that, the proper physicalist response will be self-evident.

Start with what has been pointed out thus far. The acquisition of sensory concepts from their sensory bases is not mediated by any consciously available more specific information: the sustaining mechanisms for these concepts are non-cognitive. They can also be vertically applied as such, without cognitive mediation. Notice that none of this implies that sensory concepts don't have conceptual / functional roles. They do — as we have seen when discussing the qualitative space generated by the multitude of their sensory bases and their similarity comparisons. Sensory concepts acquired from a given sensory quality space reflect a rich set of conceptual interrelations. But we have also seen that these conceptual relations reflect the external relations of each quality to others whose conceptualizations they are. They don't say

[51] However, unlike Rosenthal, Dretske (at least in print) thinks of introspection as more like a theoretical inference (inferentially mediated displaced perception) rather than perceptual *de re* knowledge — see also Shoemaker (1994), whose views on introspection are similar to Dretske's in certain respects.

anything (in an extractable format) about the internal nature of the secondary qualities represented by these sensations. The main point, however, is that even though sensory concepts have conceptual / functional roles in this sense, they are not part of the sustaining mechanisms mediating their acquisition and vertical deployment. Their semantics is fixed independently of such roles, by a direct and immediate informational link to sensory experiences. To say that this link or sustaining mechanism is direct and immediate is to say that it is non-cognitive, which is to say that these sensory concepts digitalize the most specific information carried by sensory experiences about the relevant values of a secondary quality dimension like color. And this is to say that they have an abstraction / digitalization distance that is maximally short.

It follows from this that our sensory concepts can pick out the qualities they denote directly and immediately, and that they are independent of any other concepts in this sense. In particular, they are independent of any physical / functional concepts, and therefore not only cannot be defined in terms of them but also no such concepts are even involved in fixing their reference: i.e., none of them is involved in the sustaining mechanisms that determine their semantics.[52] We have also shown that no concepts except sensory ones work this way: all others involve *cognitive* sustaining mechanisms. This is true even for what we have called perceptual concepts. This means that sensory concepts cannot be derived *a priori* from any other concepts or theories couched in them. As long as the introspection of sensory states requires redeployment of sensory concepts as phenomenal concepts, the same will be true of phenomenal concepts. This in turns means that, insofar as conceivability is a matter of concept use, it will be possible to genuinely conceive a zombie replica of a person, a creature with exactly the same physical / functional organization as the person's which lacks sensory experiences with conscious phenomenal qualities.

[52] This is why, on our view, the primary and secondary intensions of sensory concepts collapse into one on a two-dimensional semantics found in Chalmers (1996).

We use "genuinely conceivable" in a technical sense that we should clarify. Consider the standard way in which the apparent conceivability of $H_2O$ without water (or vice versa) is dispelled or explained away. It consists in showing that the conception of such a situation is only apparent — this is why Kripke claims that there is an *air of apparent contingency* in scientific identity statements. This is done by showing that what the conceiver actually conceives is not the situation expressed by the statement

(a) it is not the case that water = $H_2O$,

but rather, one expressed by

(b) it is not the case that the watery stuff = $H_2O$,

where 'the watery stuff' is a definite description contingently picking out a substance on the basis of the superficial qualities we normally use to identify water (or fix the reference of 'water'). It is in this sense that we would like to claim that (a) is only apparently conceivable. This sense requires the availability of a (commonsense) description / conception associated or connoted by 'water' / WATER that contingently picks out the same substance denoted by the scientific term / concept. If no such description / conception is available, we will say that the statement in question expresses a situation that is genuinely conceivable.[53] The association or

[53] Such statements need not be identities. Whatever form they take, the point concerns the availability of an associated definite description which contingently picks out the same thing picked out by a term used in the expression of the statement whose conceivability is in question. Here the claim is not that the conceiver should actually associate such descriptions. The point is about whether any such description or conception exists and is potentially available to the conceiver. The information contained in sensory arrays and used in reference fixing should be available for further conceptualization, but this does not mean that the conceiver has actually conceptualized every bit of information used in fixing the reference of a non-sensory concept.

connotation relation we have in mind does not require semantic or conceptual connections,[54] but requires the use of those very concepts that are involved in fixing the reference, or as we prefer to put it, the use of information (whether or not actually conceptualized) supplied by conscious experiences and used in the sustaining mechanisms for concepts. Notice that the concepts of most superficial properties cited in the description 'watery stuff' are part of the sustaining mechanisms for WATER.

It is in this sense we would like to claim that my zombie replica is genuinely conceivable. In other words, it is genuinely conceivable

(Z) that a complex property (my physical replica) expressed by a purely physical/functional predicate/concept is instantiated without the instantiation (by it) of an (apparently simple phenomenal) property that I pick out with a phenomenal predicate/concept I possess.[55]

When we fix the fact that the physical/functional property is instantiated, what we have to find out, in order to show that (Z) is not genuinely conceivable, is a description (expressing a complex concept) that is associated or connoted by the phenomenal concept (in the way we specified — i.e., involved in its sustaining mechanism) that contingently picks out the same property. But that there is no such description should be clear from the way we analyzed the direct and immediate acquisition and vertical deployment of sensory concepts from sensory experiences, and their introspective uses.[56] So (Z) is genuinely conceivable in our sense.[57]

---

[54] So, no definition or semantic/conceptual analysis is needed, contrary to what Chalmers (1996), Jackson (1994), and Levine (1993) assume. In this sense our reconstruction of their conceivability argument makes it even stronger by weakening one of its controversial premises.

[55] (Z) describes a phenomenal zombie. There are various non-equivalent ways of describing zombies (e.g., on the basis of local vs. global supervenience; token vs. type identities; state-based vs. individual- vs. species- based scope; by including propositional attitudes vs. not, etc.). Some of them involve important nuances. But we will not bother to be more specific here.

[56] We don't envision the need for intentional concepts in the introspective use of sensory concepts as part of their reference-fixing or sustaining mechanisms. The case of PAIN and

The most interesting aspect of modern conceivability arguments consists in novel attempts to draw metaphysical conclusions from this fact, i.e., the fact that (Z) is genuinely conceivable. They all harbor a premise, which has been implicit until very recently, that we will call the "bridging premise":

(B) For any proposition *P*, if *P* is genuinely conceivable, then *P* is metaphysically possible.

This principle, if true, would bridge the gulf between an epistemological / psychological claim and a metaphysical one, allowing one to validly draw a metaphysical conclusion from epistemological premises. This would be detrimental for physicalists, since if (Z) is metaphysically possible, then physicalism is false.[58]

But is (B) true? Or, true in a sense that can be used against physicalism? We submit that it is not, and our story about the acquisition of sensory / phenomenal concepts makes clear why it is not.

What is the status of (B)? As far as we can see, (B) would be implausible if read as a logical entailment. Conceivability is purely a matter of epistemology or psychology, i.e., the capacity of cognitive organisms to represent reality one way or the other; as such, any reflection on what is conceivable and what is not has no *logical bearing* on the constitution of metaphysical reality (necessity / possibility). To think otherwise is to risk an unacceptable form of

---

concepts of other bodily sensations show this. Rather the intentional concepts somehow help the semantic content to refocus on the proximal information already there.

[57] There is certainly a *transcendental* sense in which (Z) is not genuinely conceivable. Suppose that a completed science of cognitive organisms vindicates our information-theoretic account of concept acquisition — and makes it nomologically impossible (as we suspect) to lack sensory concepts (not necessarily the ones we actually happen to have or now think we have) in the way and sense we have proposed. Then a complete physical / functional description of me will entail that I have sensory / phenomenal concepts that *I* cannot derive from this description. Nevertheless, I will know that my sensory / phenomenal concepts as *I* have acquired them pick out the same physical / functional properties that some of the terms in the physical / functional description pick out. See below.

[58] There is actually some room to contest this conditional even if (B) is true in the relevant sense (see below). But we won't dwell on this point here.

verificationism. So we reject (B) on this strong reading.[59] But we don't think this rejection should be controversial, given our naturalistic story about phenomenal concepts.

However, there is no doubt that conceivability is often our only, and often quite reliable, guide to possibility. As such it would be foolish to reject (B) outright. We take it that (B) should be read as stating a reliable but *defeasible* rule of inference. Alternatively, it could be read as a *ceteris paribus* generalization. In this weaker sense, (B) may be accepted by physicalists, in which case they need to find good reasons to block the licensing force of the rule embodied therein; that is, they need to find defeaters in its application to the case at hand, namely to (Z), and the like. They need to tell a principled story about why we must refrain from applying (B) in *this* particular and tendentious case where phenomenal consciousness is at issue.

We certainly incur this debt, but we think that we have already discharged it by providing solid and principled reasons to suspend the rule in cases where conceivability involves sensory / phenomenal concepts. In other words, we have already provided the defeater. The deafeter is the fact that the nature of our sensory and phenomenal concepts, in terms of which we conduct thought-experiments about what is conceivable and what is not, is such that they inevitably support the intuitions about the genuine conceivability of (Z) and the zombie worlds.[60] But if our purely naturalistic story is correct, this is to be expected: no

---

[59] We think that Chalmers's arguments for reading (B) as a logical entailment fail mostly because of reasons given by Yablo (1993), Levine (1993, 1998), and Byrne (ms.). But even if Chalmers is right about the damage that would be done to the epistemology of modality unless (B) is read as an entailment, still, we think that this damage is not worse than the damage done to the metaphysics of mind if (B) is read as an entailment. If necessary, we are prepared to make whatever adjustments are needed in the epistemology of modality to save a physicalist metaphysics of mind — well … almost (we don't want to be dogmatic about this). We are aware that some (e.g., W.D. Hart, George Bealer, and Chalmers, of course) order their priorities the other way around. But we don't think this will be necessary: (B) can do all the work required by a proper epistemology of modality even when read as a defeasible rule — see the next paragraph in the main text.

[60] As we said, we use the terminology of 'apparent / genuine conceivability' in a technical sense which we have just characterized. This choice of terminology is not entirely a happy one, especially in the light of the fact that we are committed to the claim that zombie-worlds are *metaphysically impossible*. This entails that the genuine conceivability of zombies, as this notion is developed and used here, must still be a cognitive illusion — albeit a very different one than the ones created by the denial of standard a posteriori necessary scientific identities. Levine's (2001: 87ff) terminology of 'thin / thick conceivability' might have been a better choice here, but

metaphysical conclusion follows. Thus, our story in effect says that when conception involves sensory and phenomenal concepts in a certain way, (B) must be suspended. Or put differently, (B) may not apply to cases where what is conceived is conceived through sensory and phenomenal concepts.

It is interesting to note that if our story is right, the defeater we present against (B) and ultimately against the a prioristic conceivability arguments is empirical in nature. For we take it that the story we have told is ultimately a form of philosophically informed theoretical psychology to be vindicated by findings from empirical science. Indeed, consider Jackson's thought experiment about the omniscient color scientist, Mary, who has spent all her life in a black and white room until her release one day, when she sees colors for the first time. Before her release, suppose that Mary knows all there is physical to know not only about color vision but also all about introspection and concept formation. Then, supposing that something like our account is true, the details of this account are what she would know exhaustively. But then she would automatically be in a position to know about the curious asymmetry involved in the epistemic access to phenomenal / physical facts. This body of knowledge she has before her release would not of course remove her curiosity (the surprise element) about coming to know in a first-person way facts she already knew under their scientific description. On the contrary, we would expect that she would be even more curious and intrigued to *instantiate* those phenomenal / physical states herself, which are necessary to acquire the peculiar perspectival concepts, and thus first-person knowledge. Knowing all the scientific facts would also make her know that she lacks certain kinds of concepts necessary to know facts in a perspectival way, different from the way she already knew them from a third-person stance. We would expect her not to be moved by the familiar conceivability arguments at all. Given her scientific omniscience, she would be in a position to know better.

---

his distinction is not exactly the same as ours — although very similar. But having said this, and explained what we mean by the distinction, we will stick with this terminology until we find a better one.

Of course, our information-theoretic account and the way we use it to block conceivability arguments is no knockdown argument *against* anti-physicalism. Strictly speaking, our account of concept formation and introspection may be compatible with anti-physicalism. But this is as it should be: what we have provided is a naturalistic account that will make a non-demonstrative but extremely strong case against views like epiphenomenalism only when combined with general considerations about causality and methodological considerations about explanation and theory building. But of course, before all that, it is certainly good to know — and, important to underline — that there is no knockdown argument against physicalism and naturalism in general.[61]

[61] The arguments from absent qualia and spectrum inversion are species of conceivability arguments. To this extent, our response to these should be predictable from what we have said so far. We intend to elaborate on these elsewhere. But very briefly: we obviously deny that absent qualia cases are possible on the ground that any creature that instantiates a certain information-processing architecture (of the sort we have partly specified) will have qualitative sensory states that are conscious. We accept the possibility of inverted spectrum cases, but think that our account actually predicts the possibility of such cases, and that therefore they don't threaten physicalism/functionalism of the sort we envision. There are complications, though, about interpersonal inversions that need to be addressed with care. Our position here is close to that of Shoemaker (1994, 2001) in certain ways.

Intellectually this paper owes a lot to the previous groundbreaking work of Armstrong (1968), Dretske (1981), Loar (1990/97), Levine (1993, 1998, 2001), Lycan (1987, 1996), Rosenthal (1991, 1997), and Shoemaker (1981/97, 1994). We thank them all for their insights and influence on us. Block's insistence (1980, 1995) over the years that materialism must come to grips with the phenomenal in a serious way was also influential in developing a materialist information-theoretic account which we claim can do justice to internalist intuitions. Also, we would like to express our deep gratitude to Fred Dretske for his help, patience, and generosity for the long hours of discussion of this material. It was no easy task to convince him that his own extreme externalism is not warranted by his own information-theoretic account, and as far as we can tell, we failed in that. Also many thanks to Fred Adams, José L. Bermúdez, Paul Castle, John Kulvicki, Kirk Ludwig, Philip Robbins, Wade Savage, Brian C. Smith, Leopold Stubenberg, and Jonathan Weinberg for their thoughtful comments and criticisms. Portions of this work were delivered at the Eastern APA meeting in New York (December 2000), and the 27th SPP meeting in Cincinnati (June 2001). We thank the audiences for their comments and questions.

## REFERENCES


Armstrong, David (1962). *Bodily Sensations*, London: Routledge and Kegan Paul.

Armstrong, David (1968). *A Materialist Theory of the Mind*, London: Routledge and Kegan Paul.

Aydede, Murat (forthcoming). "Naturalism, Introspection, and Direct Realism about Pain," *Consciousness and Emotion*, Vol. 2, No. 1, 2001. Available as a PDF file at: http://web.clas.ufl.edu/users/maydede/pain.pdf

Aydede, Murat, and Güven Güzeldere (in prep.). "Introspecting Pain" to appear in *The Puzzle of Pain: Scientific and Philosophical Essays*, M. Aydede, G. Güzeldere, and Y. Nakamura (eds.), forthcoming from the MIT Press.

Barsalou, Lawrence (1999). "Perceptual Symbol Systems" in *Behavioral and Brain Sciences*, Vol. 22, No. 4, 1999.

Biro, John (1991). "Consciousness and Subjectivity" in *Consciousness* (Philosophical Issues, V.1), Villanueva, Enrique (ed).

Block, Ned (1980). "Troubles with Functionalism" in Readings in Philosophy of Psychology, Ned Block (ed.), Cambridge, MA: Harvard Univeristy Press, 1980.

Block, Ned (1995). "On a Confusion about a Function of Consciousness," *Behavioral and Brain Sciences*, 18(2).

Block, Ned, and Robert Stalnaker (1999), "Conceptual Analysis, Dualism and the Explanatory Gap," *Philosophical Review*, 108(1): 1–46.

Byrne, Alex (ms.). "Chalmers' Two-Dimensional Argument Against Physicalism", MIT.

Carruthers, Peter (2000). *Phenomenal Consciousness*, ??

Chalmers, David (1996). *The Conscious Mind*, Oxford, UK: Oxford University Press.

Churchland, Paul M. (1985). "Reduction, Qualia, and the Direct Introspection of Brain States," *Journal of Philosophy*, 82(1).

Churchland, Paul M. (1989). "Knowing Qualia: A Reply to Jackson" in *A Neurocomputational Perspective: The Nature of Mind and the Structure of Science*, Cambridge, MA: MIT Press.

Dennett, D.C. (1994). "Dretske's Blind Spot", *Philosophical Topics* (The Philosophy of Daniel Dennett). Vol. 22, Nos. 1&2.

Dretske, Fred (1981). *Knowledge and the Flow of Information*, Cambridge, MA: MIT Press.

Dretske, Fred (1986). "Misrepresentation" in *Belief: Form, Content, and Function*, R. Bogdan (ed.), Oxford: Clarendon Press.

Dretske, Fred (1994). "Differences that Make No Difference", *Philosophical Topics* (The Philosophy of Daniel Dennett). Vol. 22, Nos. 1&2.

Dretske, Fred (1995). *Naturalizing the Mind*, Cambridge, MA: MIT Press.

Farrell, B.A. (1950). "Experience", *Mind* 59: 170–198. [Reprinted in V.C. Chappell (ed.), *The Philosophy of Mind*, Englewood Cliffs, New Jersey: Prentice-Hall, 1962.]

Feigl, H. (1967). *The "Mental" and the "Physical": TheEssay and a Postscript*, Minneapolis: University of Minnesota Press.

Flavell, John ??

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

Fodor, Jerry A. (1981). "Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology" in *RePresentations: Philosophical Essays on the Foundations of Cognitive Science*, Cambridge, Massachusetts: MIT Press. [Originally appeared in *Behavioral and Brain Sciences* 3:1, 1980.]

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

Fodor, Jerry A. (1990). "A Theory of Content" (I & II) in *A Theory of Content and Other Essays*, Cambridge, Massachusetts: MIT Press.

Fodor, Jerry A. (1994). *The Elm and the Expert*, Cambridge, Massachusetts: MIT Press.

Hart, W.D. (1988). *The Engines of the Soul*, Cambridge, UK: Cambridge University Press.

Hilbert, David (1987). *Color and Color Perception*, Stanford, CA: CSLI Publications.

Hilbert, David (ms.). "Why Have Experiences?" draft, UIC.

Hill, Christopher (1991). *Sensations: A Defense of Type Materialism*. London, UK: Cambridge University Press.

Hill, Christopher (1997). "Imaginability, Conceivability, Possibility and the Mind-Body Problem", *Philosophical Studies* 87: 61–85.

Hill, Christopher and Brian McLaughlin (1999). "There Are Fewer Things in Reality Than Are Dreamt of in Chalmers's Philosophy," *Philosophy and Phenomenological Research*, 59: 2.

Jackson, Frank (1982). "Epiphenomenal Qualia," *Philosophical Quarterly* 32: 127–136.

Jackson, Frank (1986). "What Mary Didn't Know," *Journal of Philosophy*, 83(5): 291–95.

Jackson, Frank (1994). "Finding the Mind in the Natural World" in *Philosophy and the Cognitive Sciences: Proceedings of the 16th International Wittgenstein Symposium*, R. Casati, B. Smith, and G. White (eds), Vienna: Verlag Holder-Pitcler-Tempsky, pp. 101–112.

Kripke, Saul (1970, lectures). *Naming and Necessity*, Harvard University Press, 1980.

Kulvicki, John (2001). *On Images: Pictures and Perceptual Representations*, Ph.D. Dissertation, University of Chicago, Philosophy Department.

Laurence, Stephen and Eric Margolis (forthcoming). "Concepts, Content, and the Innateness Controversy", *Cognition*.

Levine, Joseph (1993). "On Leaving Out What It's Like" in *Consciousness*, edited by Martin Davis and Glyn W. Humphreys, Oxford, UK: Basil Blackwell, 1993.

Levine, Joseph (1998), "Conceivability and the Metaphysics of Mind," *Noûs*, 32(4): 449–480.

Levine, Joseph (2001), *Purple Haze: The Puzzle of Consciousness*, Oxford, UK: Oxford University Press.

Loar, Brian (1990 / 1997). "Phenomenal States" in *The Nature of Consciousness: Philosophical Debates*, edited by Ned Block, Owen Flanagan, and Güven Güzeldere, Cambridge, Massachusetts: MIT Press, 1997.

Loar, Brian (1999). "David Chalmers's *The Conscious Mind*", *Philosophy and Phenomenological Research*, 59(2): 465–72.

Lycan, William G. (1987). *Consciousness*, Cambridge, MA: MIT Press.

Lycan, William G. (1996). *Consciousness and Experience*, Cambridge, MA: MIT Press.

Margolis, Eric (1998). "How to Acquire a Concept," *Mind and Language*, Vol. 13, No. 3, pp. 347–69.

McGinn, Colin ( 1991). *The Problem of Consciousness*, Oxford, UK: Basil Blackwell.

Nagel, Thomas (1974). "What Is It Like to Be a Bat?," *Philosophical Review*, 83: 435-450.

Papineau, David. (1993). "Physicalism, Consciousness, and the Antipathetic Fallacy," *Australasian Journal of Philosophy*, 71: 169–83.

Pereboom, Derk (1994). "Bats, Brain Scientists, and the Limitations of Introspection," *Philosophy and Phenomenological Research*, 54(2): 315–29.

Perry, John (2001). *Possibility, Consciousness and Conceivability*, Cambridge, MA: MIT Press.

Prinz, Jesse (forthcoming). *Furnishing the Mind:Concepts and Their Perceptual Basis*, Massachusetts: MIT Press.

Rey, Georges (1997). *Contemporary Philosophy of Mind*, Oxford, UK: Blackwell Publishers.

Robinson, H. (1982). *Matter and Sense*, NY: Cambridge University Press.

Rosenthal, David (1991). "The Independence of Consciousness and Sensory Quality" in *Consciousness* (Philosophical Issues, V.1), Villanueva, Enrique (ed).

Rosenthal, David (1997). "A Theory of Consciousness" in *The Nature of Consciousness*, edited by Ned Block, Owen Flanagan, Güven Güzeldere, Cambridge, Massachusetts: MIT Press, 1997.

Rosenthal, David (2001). "Introspection and Self-Interpretation" forthcoming in a special volume of *Philosophical Topics* on introspection.

Shoemaker, Sidney (1981/97). "The Inverted Spectrum," *Journal of Philosophy*, 74(7): 357–81. Reprinted in *The Nature of Consciousness: Philosophical Debates* (with a postscript), edited by Ned Block, Owen Flanagan, and Güven Güzeldere, Cambridge, Massachusetts: MIT Press, 1997.

Shoemaker, Sidney (1994). "Phenomenal Character," *Noûs*, 28: 21–38.

Shoemaker, Sidney (2001). "Introspection and Phenomenal Character" forthcoming in a special volume of *Philosophical Topics* on introspection.

Siewart, Charles (1998). *The Significance of Consciousness*, Princeton, New Jersey: Princeton University Press.

Smart, J.J.C. (1959). "Sensations and Brain Processes", *The Philosophical Review*, 68: 141–156.

Stubenberg, Leopold (1998). *Consciousness and Qualia*, Amsterdam & Philedelphia: John Benjamins Publishing Company.

Sturgeon, Scott (1994). "The Epistemic View of Subjectivity," *Journal of Philosophy*, 91(5): 221–35.

Sturgeon, Scott (2000). *Matters of Mind: Consciousness, Reason and Nature*, Routledge (International Library of Philosophy), London.

Tye, Michael (1995). *Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind*, The MIT Press.

Tye, Michael (1999). "Phenomenal Consciousness: The Explanatory Gap as a Cognitive Illusion," *Mind*, 108: 705–25.

Tye, Michael (2000). *Consciousness, Color, and Content*, Cambridge, Massachusetts: MIT Press.

Yablo, Stephen (1993). "Is Conceivability a Guide to Possibility?", *Philosophy and Phenomenological Research*, Vol. 80, No. 1.