

# 7

## Perceptual Consciousness

### I. Introduction

We speak of perceptual experiences as having qualitative or phenomenal character, and we also speak of them as being conscious. In Chapter 5 I proposed an account of perceptual phenomenology. In this chapter I will continue to be concerned with perceptual phenomenology, but the principal focus will be on perceptual consciousness. What does it consist in?

I will be concerned in this introductory section with two types of consciousness that are widely acknowledged in the literature—*introspective consciousness* and *access consciousness*. They are of interest here because they can both be possessed by perceptual experiences. After describing them in the present section, I will proceed in later sections to consider their relationship to perceptual phenomenology, eventually maintaining that although phenomenology is closely connected with these two forms of consciousness, it is metaphysically independent of them. I will then turn to consider the view that perceptual experiences possess another form of consciousness—*phenomenal consciousness*—that is independent of introspection and access. This view is widely accepted, and I am among its adherents, though my take on phenomenal consciousness stands apart from the mainstream. My goal in considering this form of consciousness will be to identify its metaphysical nature—that is, to explain what it ultimately consists in.

Introspective consciousness is the property of being an object of introspective awareness. It is therefore like the property of being seen—a property that an item possesses just in case it stands in an awareness-relation to a subject. It is the property that Freud seems to have had in mind in distinguishing between desires (or wishes or beliefs) that are conscious and desires (or wishes or beliefs) that are subliminal or repressed. As he saw it, the former are mental states that the subject is aware of, while the latter are mental states that lie outside the scope of awareness (Freud 1965). Contemporary therapists have inherited this usage.

What is introspection? For present purposes, it suffices to say that it involves mental states that represent other mental states—that is, second-order representations—and also a range of other things, including working memory, attention, and contextually defined parameters that determine whether tokening a second-order representation is task relevant. I will say more about the nature of introspection in Section IX.

When the expression “access consciousness” was introduced some years ago by Ned Block, he defined it as follows: a representation is access-conscious if it is poised for use as a premise in reasoning, poised for rational control of action, and poised for rational control of speech (Block 1995, 231). In more recent writing Block has liberalized the definition considerably, saying that a representation is access-conscious if it is available for use by a broad range of higher cognitive agencies, including the ones that are responsible for reasoning, problem solving, planning, producing verbal reports, creating perceptual beliefs, and forming episodic memories (Block 2022). Roughly, a representation is access-conscious if it is “globally available” for use by a variety of high-level cognitive agencies. Similar ideas have been promoted by Baars (1997) and Dehaene (2014), and have found favor with many writers on consciousness, both in science and in philosophy. Evidently, there are widespread and vivid intuitions to the effect that there is a form consciousness that is constituted by relations to high-level cognitive states.

It is useful to distinguish between occurrent and dispositional forms of the two forms of consciousness we have been considering. Here are some definitions: (i) Occurrent introspective consciousness is the property that a mental state has when it is actually being registered by the introspective faculty of the relevant agent. (ii) Dispositional introspective consciousness is the property of being available to the relevant introspective faculty. (iii) Occurrent access consciousness is the property of actually causing one or more high-level cognitive states in the relevant agent. And (iv) dispositional access consciousness is the property of being poised for causing such states. I will assume that it is appropriate to think of all of these properties as forms of consciousness. Certainly they have all been cited as such in the literature (Block 2002, Gennaro 2012).

It is clear that all of these properties can be exemplified by perceptual experiences. In view of this, we can see that there are at least four different answers to the question with which we started, “What does perceptual consciousness consist in?” Moreover, each of these answers has merit. Can we conclude, then, that our inquiry has been brought to completion?

This is one of the questions to be addressed in the present chapter. I will consider several ideas about additional forms of perceptual consciousness and offer evaluations of them. Before we consider that topic, however, we should consider the question of how the four forms on our present list are related to phenomenology. A number of writers hold that one or more of these forms is partially constitutive of phenomenology: that every phenomenal state necessarily stands in actual or potential causal relations to behavior and/or downstream cognitive states, including introspective states. This view is widely held, both by scientists and by philosophers (e.g., Tye 2002, Dennett 2005, Cohen and Dennett 2011, Dehaene and Changeux 2011, Dehaene 2014, and Ward and Scholl 2015), but it seems quite wrong to me. I will maintain in the next few sections that

phenomenology is metaphysically autonomous—that it is in principle possible for it to exist without being accompanied by any of the relational forms of consciousness we have just been considering. Indeed, I will argue for a proposition that is much stronger than this one. As I see it, there are good grounds for thinking that experimental work has actually managed to separate phenomenology from actual and potential relations to downstream phenomena.

After arguing that phenomenology is independent of introspection and access, I will consider the possibility that there is an additional form of perceptual consciousness that warrants the name *phenomenal consciousness* because it is more intimately linked to phenomenology than either introspective consciousness or access consciousness. I will argue that we do need to recognize a form of consciousness that answers to this description. More specifically, on the account I will defend, phenomenal consciousness is the intrinsic, categorical base for introspection and access, which are both relational in nature. As we will see, this view comes to the same thing as saying that phenomenal consciousness is realized by very high levels of activity in certain areas of the brain, such as the fusiform face area, which is known to be intimately associated with experiential awareness of faces. It is part of this picture that we may have to recognize phenomenology that is not phenomenally conscious. Unconscious phenomenology would be realized by lower levels of activity in those same regions of the brain.

## II. The Autonomy Thesis

The autonomy thesis is the proposition that phenomenology does not entail actual or potential relationships to downstream cognitive states. As noted, this proposition is highly controversial, with quite a few opponents ranged against it. It seems best, therefore, to assemble several arguments on its behalf. None of the arguments I will give is decisive by itself, but collectively they make a strong case, showing that the autonomy thesis can be defended from a variety of vantage points. I will give two arguments grounded in folk psychology in the present section, one that is based on introspective data, and another that is based on the role that phenomenology plays in the explanation of behavior. These two arguments are important because they explain the intuitive plausibility of the autonomy thesis. In the next section I will give a metaphysical, quasi-Leibnizian argument for the thesis, and in the section following that I will add an argument that stems from an important discovery in vision science. Hopefully every opponent of the autonomy thesis will find at least one argument in this set of four to have some lasting appeal.

Introspection provides us with our principal access to phenomenology, so it is natural to suppose that we can grasp some of the main properties of phenomenology by consulting intuitions about it that are grounded in introspection.

One of these intuitions is that perceptual phenomenology doesn't entail relations to concurrent cognitive states. Consider a case in which you are viewing a red object, and are therefore experiencing a reddish phenomenology. Suppose that you are introspectively aware of this phenomenology. Now it may be that you are also aware of a perceptual judgment to the effect that you are confronted by a red object, or a judgment that something looks red to you. You may also be aware of a decision to choose the red object to serve some purpose. It is clear, however, that these cognitive states are not in any way part of the phenomenology of your experience of red, though they may have their own proprietary phenomenologies. From the perspective of introspection, the phenomenology associated with your experience of the red object is exhausted by the appearance of red that you are enjoying. Relations to cognitive states have nothing to do with it.

This first intuition counts in favor of the independence of phenomenology from the occurrent forms of introspective and access consciousness. Another abiding intuition counts against the dispositional forms. According to this second intuition, phenomenology is *given* in introspection. We are *directly* aware of it. This calls into question the idea that phenomenology is partially constituted by dispositional properties. We are never directly aware of dispositions, though we may be directly aware of the non-modal properties that constitute their categorical bases, and are often aware of the concrete events that manifest them. We know about dispositions by inference from non-dispositional facts of these two kinds. But it is pretty clear that we do not have to resort to inference to apprehend phenomenology. The intuition of givenness is quite robust.

This brings me to the second reason for accepting the autonomy thesis, which focuses on the role that phenomenology plays in explanations of behavior. The key idea is that we sometimes need to appeal to phenomenology in explaining behavior even though the phenomenal states to which we appeal are not access conscious. The argument I will give is based on an example that originated with David Armstrong (Armstrong 1980, 59).

Consider a truck driver who "wakes up" to the realization that he has absolutely no memory of anything that has happened for the last 40 minutes. He has covered 30 miles of complex terrain, but his mind was fully occupied by other things—let us suppose that he was planning a vacation. Clearly, he must have been perceptually aware of the colors of traffic lights, the shapes, sizes, and distances of other vehicles, the curves in the road, and his own velocity and acceleration, for otherwise he would have collided with another vehicle or run off the highway. But, apparently, from beginning to the end, his states of awareness were not access conscious. Certainly they left no traces of themselves in memory. Indeed, the states of awareness really *could not* have been access conscious, since the driver's working memory was entirely occupied by plans for his vacation. Now in all normal cases, perceptual awareness that guides such complex behavior involves phenomenology. Accordingly, if we are to give an explanation of the driver's successful

navigation of the sort that we would give in other contexts, it is necessary to appeal to phenomenal states. And it seems reasonable to give an explanation of the same sort, because the behavior to be explained is of the same type.

As the reader may be aware from personal experience, unconscious behavior of this complex sort may occur even in contexts that are completely unfamiliar. Not all cases can be explained by appeal to learned dispositions that operate independently of high-level perception.

In evaluating this argument, keep in mind the fact that when drivers remember sequences of guiding perceptual states, they remember them as having proprietary phenomenologies. It seems that all that is missing in the case of Armstrong's driver is the memories. There is no reason to think that he lacked the kind of states that the memories are *of*.

### III. A Metaphysical Argument for the Autonomy Thesis

Another argument for the autonomy thesis is suggested by the fact that there can be various levels of activity in the neural areas that are believed to support phenomenology. It is a natural thought that all of these levels can support some form of phenomenology. The phenomenology supported by higher levels of activity in these areas would no doubt be more robust than the phenomenology supported by lower levels, but diminishing degrees of robustness are compatible with a common kind affiliation.

There are a number of different ways of interpreting the notion of a level of neural activity. For example, levels might be rates of firing, but they might also correspond to numbers of participating neurons. Or they might involve a weighted average of these two quantities. Another possibility is that levels might correspond to the degrees to which patterns of firing are stable, where stability is understood to entail an ability to resist perturbation or disruption by other, competing patterns. A more complicated hypothesis is that there is a difference between levels of activity in regions of the brain that support feature detectors and levels of activity in nerve fibers that connect such regions, thereby supporting representations that attribute multiple features to objects. The need for such a distinction is suggested by studies of the phenomenon known as *crowding*, in which subjects are aware of features without being able to assign those features to determinate objects. On this hypothesis, as on related ones, the total phenomenology associated with a state would be a function of levels of activity of different kinds. There are still other ways of interpreting the notion of a level of neural activity. I will not favor any one of these interpretations here. The line of thought I wish to pursue prescinds away from these more concrete proposals.

That there can be different levels of activity in areas known to support phenomenology is illustrated by a famous experiment by Moutoussis and Zeki

(2002). Their study involved stimuli that had contrasting colors but were otherwise identical. When pairs of such stimuli are shown, respectively, to the right and left eyes, they are invisible, even though each would have been consciously perceived if it had been presented alone. Using fMRI, Moutoussis and Zeki were able to show that invisible stimuli are able to give rise to low levels of activity in visual areas that are known to support phenomenology when they are highly active (specifically, the fusiform face area and the parahippocampal place area). Moreover, this is just the tip of an iceberg. There are a number of other studies that concern unconscious activity in areas known to support phenomenology, including studies of binocular rivalry, backward masking, and flash suppression. (The reader is referred to Dehaene 2014 for a review.)

The fact that there can be different levels of activity in areas relevant to phenomenology suggests that phenomenology may extend to lower levels, albeit in a form that is less robust than the form that exists at higher levels. And, by the same token, it suggests that there is phenomenology that is not associated with either introspective consciousness or access consciousness. This is because subjects in studies like the one Moutoussis and Zeki conducted show no evidence of either of these forms of consciousness when low levels of activity in phenomenology-supporting areas of the brain occur. Such subjects cannot describe or recall the stimuli that cause the activity. In general, lower levels of activity in those areas seem to lack higher-level cognitive effects.

I said that the experiments *suggest* that different level of activity in the given brain areas may be accompanied by phenomenology that has different degrees of robustness. Is there any reason to think that this idea may be correct? Yes. The idea is supported by the following *continuity argument*, which may remind the reader of some things that Leibniz says about *petites perceptions* (Leibniz 1996):

*First premise:* The pains that fall within the scope of introspection are ordered by a relation of comparative intensity. Some pains are extremely intense, some are extremely faint, and there is a range of intensities that lie between these extremes.

*Second premise:* There is a lower bound on the intensities of pains that are introspectable.

*Third premise:* But this is only an epistemic lower bound. There is no reason to think that it is also a metaphysical lower bound.

*Fourth premise:* In general, when a natural magnitude comes in degrees, there are low degrees of the magnitude that fall beneath the scope of human awareness. Thus, for example, there are volumes of sound that are too low for the human ear to discern, and textures that are too fine grained for haptic exploration to detect—they seem perfectly smooth, but actually the surfaces have tiny grooves. Epistemic lower bounds on magnitudes do not normally correspond to metaphysical lower bounds. Normally, the metaphysical lower bounds extend below the epistemic bounds, and often far below.

*Conclusion:* All of this being true, there is reason to believe that the ordering of pains in terms of intensity extends beneath the reach of introspective discernment.

Similar arguments can be constructed for other types of phenomenology. Thus, perceived lights, perceived sounds, perceived pressures, perceived aromas, and perceived tastes are all ordered by intensity relations. In all cases there are epistemic lower bounds to these intensity orderings, but there is no reason to think that these perceptual phenomenologies are exceptions to the general rule that the epistemic lower bounds of magnitudes do not track metaphysical lower bounds.

The continuity argument falls considerably short of being a bullet-proof demonstration. It is based on an extrapolative inference from the many domains in which magnitudes tail off gradually, without regard for human epistemic capacities. Pain and other types of phenomenology could be exceptions. It seems fair to say, however, that it at least raises the probability of its conclusion. As a result of the argument, it is more likely that phenomenal states are ordered by a relation of comparative intensity or robustness that is in lock step with the ordering of levels of neural activity in brain areas that are known to support phenomenology. This in turn makes it more likely that phenomenology can exist independently of introspection and access.

At all events, while I will continue to assume in the background that phenomenology comes in degrees of robustness, and that the lower degrees fall beneath the thresholds of introspection and access, I will be focusing in most of the rest of this discussion on phenomenology that is highly robust. When I speak of phenomenology in the next section and beyond, unless otherwise indicated, I will always mean to refer to phenomenology that is at the top of the scale, even if I don't explicitly say as much.

#### IV. A Fourth Argument for Autonomy

I conclude this case for the autonomy thesis by describing and defending the "overflow argument" that has been developed by Victor Lamme and Ned Block. I will describe the argument briefly, and then dwell for a while on the question of what exactly it shows. The reader is referred to Landman et al. 2003, Block 2011, and Block 2022 for detailed expositions.

The overflow argument is based on a classic experiment by Sperling (1960) in which subjects were shown a matrix of twelve letters at an initial time T1. When interrogated at a later time T2, after the letters had disappeared, the subjects were able to name four of the letters but no more. They could, however, recall having seen a full matrix of letters. That is, they could recall having seen letters at T1 that they could not name at T2. By using a special technique that I'll describe in a

minute, Sperling was able to find experimental confirmation for the subject's claim to have seen all of the letters. Now the ability to articulate perceptual states by naming stimuli is a hallmark of access consciousness. Accordingly, Block concludes that the letters that were seen at T1 but not specifically recalled at T2 lacked access consciousness. If, as is natural, we also suppose that the initial (T1) awareness of all of the letters had a proprietary phenomenology, we can infer that there are phenomenal states that are not access conscious.

Sperling confirmed the testimony of his subjects by a cueing procedure. The letters in the matrix were arranged in three rows. Sperling sounded tones after the letters had disappeared from the screen. A low tone meant that subjects were supposed to name the letters in the bottom row; a medium tone meant that they were supposed to name the letters in the middle row; and a high tone meant that they were to name the letters in the top row. Although the subjects were only able to name four letters, it turned out that they were able to name the letters in *any* row that was cued by a tone, provided that the tone followed the initial display by a short enough interval (around 300 milliseconds). Evidently, the subjects must have *seen* all of the letters in order to be able to name the ones in an arbitrary row, and they must have *remembered* all of the letters at least to the point when the tone sounded. We must accept these hypotheses in order to best explain the subject's success in naming the letters in the row that had been highlighted by a tone.

I believe that Block's line of thought is ultimately successful, but it is easy to get confused in evaluating it. It presupposes a criterion for access consciousness that is arguably too demanding. According to the criterion, a phenomenal state fails to count as fully access conscious unless one can recall and name the content of every one of its parts. But one might think it no less reasonable to say that a state can count as access conscious if one can recall and name the *gist* of its content. This is relevant because Sperling's subjects could in fact recall the gist of their original experience at T1. They remembered that they had seen a matrix of letters. Hence, according to a remembered gist criterion of access consciousness, their initial experiences of the matrix were access conscious. Moreover, assuming that their total experiences could be factored into "smaller" experiences of individual letters, we can say that these smaller experiences were also access conscious. This is because the *gist* of each of the smaller experiences was remembered. The subjects remembered the unnamed letters *as* letters. (For discussions of gist, see Oliva 2005, Oliva and Torralba 2006, Haberman et al. 2015, Cohen et al. 2016, and Ward et al. 2016.)

So there are two possible criteria for access consciousness, one that is based on total recall of an earlier representation and a less demanding one that is based on recall of gist. Are these criteria rivals, and if so, is there a standoff between their respective proponents? Are the criteria equally well motivated? If so, Block hasn't succeeded in prying phenomenology loose from access consciousness.

The conception of access consciousness that is based on recall of gist can be tempting, but I think that in the end there is a strong reason to prefer the conception that is based on total recall. The reason for having a concept of access consciousness in the first place is that we want to be able to distinguish between those aspects of the representational content of a state that are available for use by higher-level faculties and those aspects that aren't. Let R be a perceptual representation, let C be the content of R, and let G be the gist of C. It follows that we should not say that R is access conscious just because G is available for use by higher cognitive faculties. That might suffice for the claim that R is *weakly* access conscious, but for R to be *strongly* or *fully* access conscious, all of C must be available for such use. To put the point another way, for R to be fully access conscious, all of the details of C must be available to working memory. It isn't enough for a generic or gisty summary of C to be available.

Sperling's subjects were presumably in high-level visual states at T1 that represented individual letters as having specific properties like *having an "A" shape* and *having a "T" shape*. That is, at all events, the best joint explanation of (i) the fact that the subjects could recall the letters in any row after the tone had sounded, and (ii) the fact that at time T2 the subjects recalled that they had *seen* the letters at T1. But only four of those specific properties could be represented in working memory. It wasn't possible for all of them to be represented there. It follows that the original perceptions of the full matrix by Sperling's subjects were not access conscious, according to the understanding of access consciousness that best fits the motivation for having a concept of access consciousness, even though the gists of their perception were fully available to working memory.

There is nothing wrong with *supplementing* Block's original, full-strength concept of access consciousness with a less demanding concept based on gist. Indeed, it might prove theoretically useful to have such a concept available. But the concept based on gist cannot *replace* the full-strength concept. The distinction between full cognitive accessibility and partial accessibility is important in cognitive science and also in the metaphysics of mind. Block's full-strength concept corresponds to significant theoretical needs that the other one cannot fulfill.

In his early formulations of the overflow argument, Block maintained that Sperling's experiment indicates that *phenomenal consciousness* is independent of accessibility. (This is not true of his recent formulation in chapter 1 of Block 2022.) I should emphasize that I have maintained only that the experiment indicates the independence of *perceptual phenomenology*.

Perhaps it would be useful to underscore this point by reviewing the present version of the argument. I have assumed that Sperling's subjects remembered their original perceptions of the full matrix in terms of how the matrix *looked to them*. (This assumption is forced on us by the subjects' claim to have *seen* the full matrix.) Given this background assumption, and given also that Sperling was able to find powerful evidence that the subjects' memories were veridical, it is likely

that the subjects did in fact earlier enjoy phenomenally individuated perceptual representations of the full matrix of letters. The rest of the argument just consists in invoking the definition of access consciousness, and in pointing out that most of the subjects' phenomenally individuated representations did not meet the terms of the definition.

Note that this summary of the argument is concerned exclusively with *phenomenology*. There is no mention of *phenomenal consciousness*. That is, it is neither asserted nor presupposed that phenomenology does or even can possess a third form of consciousness that is independent of introspective consciousness and access consciousness. The argument is simply designed to show that phenomenology is independent of access. The question of whether there is a third form of consciousness that stands in some intimate relation to phenomenology will be taken up later.

In addition to presenting the overflow argument, which is based on an experimental tradition that began in the 1960s, Block has maintained that several more recent experiments in vision science point to the separability of visual phenomenology and access consciousness. In different ways, these experiments show that perceptual activity in the visual areas in the back of the brain can in certain special circumstances occur without being accompanied by cognitive activity that is supported by structures in the front (Pitts et al. 2014, Brascamp et al. 2015). The reader is referred to Block 2022 for exposition of several of the experiments, and for arguments aimed at establishing their relevance to the separability issue.

## V. The Categorical Base Hypothesis

We have found that phenomenology does not entail any of the four forms of consciousness that we considered in Section I. At the same time, however, it is evident that (sufficiently robust) phenomenology stands in intimate causal relations to those forms. There is no doubt, for example, that perceptual states with a phenomenal character can provide rational control of action, and have the power to cause episodic perceptual memories and perceptual judgments. Nor is there any doubt that phenomenal states play a role in generating the introspective states that are directed on them. But if relations of this sort are metaphysically contingent, as I have been maintaining, how can they be as intimate as they are?

A natural answer is that robust (that is, highly intense) phenomenology provides a *categorical base* for the causal powers that constitute dispositional introspective consciousness and dispositional access consciousness. (As is usual, I am thinking of the categorical base for a disposition to do X as a set of non-modal properties that are capable, given the laws of nature, of causing X. (For a rich discussion of categorical bases, see Choi and Fara 2018.)) To see the merits of this

view, notice that phenomenology seems to confer a selective advantage on the creatures that possess it. In general, the creatures that possess it are able to provide for their needs more successfully, and in a much wider range of contexts, than creatures like *Aplysia*, which, given the results of comparative neuroanatomy, pretty clearly lack it. Now, if this is right, then robust phenomenology has biological value. But what does this value consist in? The natural answer is that it consists in the downstream effects of phenomenal states, for, in general, it is what a biological item can *do*, its contribution to the *modus operandi* of the organism, that the biological value of the item consists in. But it could not be true that the enduring value of phenomenology consists in its downstream effects unless the dispositions to produce those effects were grounded in phenomenology. Since we have seen (in Sections II and III) that the dispositions in question are metaphysically independent of phenomenology, this “grounding” relationship is best described by saying that phenomenology provides the categorical base for the dispositions. (More precisely, it provides part of a categorical base. The total categorical base includes working memory and the various downstream cognitive agencies that working memories of phenomenal states can activate.) (The categorical base hypothesis has also been endorsed by Jesse Prinz. For further discussion, see chapter 3 of Prinz 2012.)

It is part of this view that robust phenomenology also supports the occurrent causal relations that constitute the non-dispositional forms of introspective consciousness and access consciousness.

It probably goes without saying, but I will say anyway, that the categorical base hypothesis does not claim that robust phenomenology is causally sufficient for *all* types of high-level cognitive activity. For example, a perceptual state representing an array of marks on paper may be sufficiently robust to support such cognitive operations as classifying the array as containing letters of the English alphabet and remembering the array as having this property; but it might still be the case that the letters are too small for a subject to read them easily, or too closely packed together for a subject to count them. Evidently access consciousness comes in degrees, where degrees are determined by the number and kinds of cognitive activity that are found in a given case. The categorical base hypothesis claims that robustness of phenomenology is sufficient for a certain base level of access consciousness, but it allows that factors other than robustness may be involved in supporting higher degrees of access.

## VI. Phenomenal Consciousness

As I said at the outset, it is widely acknowledged that perceptual states can possess the two forms of introspective consciousness that we have considered and also

the two forms of access consciousness. I will turn now to consider whether there is another form of consciousness that should be acknowledged. This inquiry will occupy us for several sections.

In addition to recognizing introspective consciousness and access consciousness, many philosophers maintain that perceptual experiences also possess a fifth form of consciousness, a form that is independent of occurrent and dispositional relations to second-order representations and to high-level cognitive faculties. This fifth form is often referred to as *phenomenal consciousness*—a designation that is meant to highlight the relationship between it and phenomenal character, which is thought to be more intimate than the relationships between phenomenal character and the other four forms of consciousness. (See, e.g., Block 1995, Chalmers 1997, and Block 2002.)

I will eventually argue that this view is correct, but the path to a full defense of the view is necessarily long and circuitous. For example, some care is needed to distinguish interesting forms of the view from the uninteresting conception of phenomenal consciousness that is suggested by the following passages from a book by James Mill:

Having a sensation, and having a feeling, are not two things. The thing is one, the names only are two. I am pricked by a pin. The sensation is one; but I may call it sensation, or a feeling, or a pain, as I please. Now, when, having the sensation, I say I feel the sensation, I use only a tautological expression: the sensation is not one thing, the feeling another; the sensation is the feeling. When, instead of the word feeling, I use the word conscious, I do exactly the same thing, I only use a tautological expression. (Mill 1869, 224)

It was of great importance, for the purpose of naming, that we should not only have names to distinguish the different classes of our feelings, but also a name applicable equally to all of those classes. This purpose is answered by the concrete term Conscious; and the abstract of it, Consciousness. (Mill 1869, 225)

With just a bit of anachronism, it is possible to summarize the semantic claim that Mill is expressing here by saying that the adjective “conscious” is a general term that stands for all and only mental states that have phenomenal character. A companion view is that the noun “consciousness” stands for the property *having some phenomenal character or other*. Suppose these claims are true—that there really are uses of the terms in question of the indicated sort. On this assumption, the thesis that there is a form of consciousness other than introspective consciousness and access consciousness is uninteresting. It is tantamount to the thesis that phenomenology is different than introspective consciousness and access consciousness, and this second thesis is truisitic. It would be conceded even by those mistaken authors who deny the autonomy thesis.

I hasten to add that there is room for doubt that Mill's semantic claims are correct. Perhaps he would have withdrawn them if he had noticed that there might be minimally robust forms of phenomenology of the sort suggested by the foregoing Leibnizian continuity argument. I have cited Mill's claims because I feel that many people have been tempted by a Millian view. (Certainly I have been.) I will not, however, be placing any dialectical weight on the claims.

The interesting thesis is that there is a use of "phenomenal consciousness" on which it is not pleonastic. Or, to set aside semantics for metaphysics, the interesting claim is that there is a form of consciousness *X* such that (i) *X* is distinct from both introspective consciousness and access consciousness, (ii) *X* is distinct from the property *having some phenomenal character or other*, and (iii) phenomenology + *X* = phenomenal consciousness. I will use the technical term "*p*-consciousness" for anything that satisfies these conditions.

## VII. Proposals about *P*-Consciousness

There are three possibilities as to what *p*-consciousness might be. It might be a property bestowed on phenomenal states by virtue of their relationship to certain other things, or it might be an entity of some kind that serves as a ground or foundation for phenomenal properties. That is, it could be a substance that instantiates them or a property or relation of which they are modes. Further, if *p*-consciousness isn't a relational property of phenomenal states or an entity that serves as a foundation for phenomenal properties, it might be a monadic property of phenomenal states other than the property *possesses some phenomenal character or other*. I will call these possibilities the *relational* view, the *inherence* view, and the *monadic* view. I will discuss a specific proposal from each of these categories, beginning with the first. I hope the reader will feel that the sample proposals I have chosen represent the best hopes for success of their respective categories.

*Example of the relational view:* Some philosophers maintain that there is a form of consciousness that is similar to introspective consciousness in that it involves second-order representations of mental states, but is different in that it lacks the other features of introspective awareness (Lycan 2004, Rosenthal 2005, Gennaro 2012). According to these philosophers, this form of consciousness does not involve attention, it does not involve working memory, and it is not controlled by mechanisms that determine whether second-order representations are task relevant. In other words, instead of occurring only in special circumstances, as is the case with introspective consciousness, it comes into existence spontaneously with great frequency. More concretely, its constitutive second-order representations are generated automatically whenever the mind is in a sufficiently full-bodied phenomenal state. To give it a name, the view of these philosophers is the

*second-order theory of p-consciousness.* They urge that the second-order theory is much better equipped than theories of consciousness based on introspection to explain fact that so many of our mental states count as conscious. There is an unbroken stream of conscious states throughout the day, and the second-order theory can nicely explain this fact; but introspection occurs too rarely for theories based on introspection to accommodate it. According to the theory, these conscious states count as conscious because they share the property *being an object of a second-order representation.*

*Example of the inherence view:* According to the *adverbial theory of p-consciousness*, *p-consciousness* is a generic, undifferentiated relation of awareness, and phenomenology consists of a range of phenomenal modes or determinations of that relation. The view implies, for example, that to be phenomenally conscious of an object *O* as blue is to be “bluishly” conscious of *O*, and that to be conscious of *O* as spherical is to be “spherically” conscious of *O*. *P-consciousness* is said to be the relation that is common to bluish consciousness of objects, spherical consciousness of objects, and so on. This theory derives from the *adverbial theory of perception* that was originally proposed by Ducasse (1952), and was strongly advocated by Chisholm (1957), among others. (We considered Ducasse’s theory of perception in an earlier chapter, but the evaluation of it there was quite brief. The criticisms that I will offer below of the adverbial theory of *p-consciousness* also apply to Ducasse’s proposal about perception.)

*Example of the monadic view:* The *categorical base theory of p-consciousness* begins by reaffirming the categorical base hypothesis that we considered in Section V. That is, it claims that phenomenology that is especially robust serves as a categorical base for the causal powers that constitute introspective consciousness and access consciousness. Here robustness is understood to be like the property of being loud, which is possessed by some noises and not by others. After making this initial claim, the categorical base theory goes on to assert that this sort of highly robust phenomenology counts as conscious in virtue of its role in supporting the various dispositions and causal relations that figure in these other forms of consciousness. It is a form of consciousness because it supports and explains these other two forms.

In summary, we can distinguish three general ideas about the nature of *p-consciousness*, the relational view, the inherence view, and the monadic view. There are reasonably plausible concrete proposals that belong to these categories. All three of these proposals insist on the metaphysical autonomy of phenomenology, but they diverge in their positive claims about the relationship between phenomenology and consciousness. They maintain, respectively, that *p-consciousness* is a relational property that phenomenal states have in virtue of being the objects of second-order representations, that it is a generic, undifferentiated relation of awareness, foundationally involved in relational facts of the form *X is F-ishly conscious of O*, and that it is a type of phenomenology that serves as the categorical

base for some very interesting causal powers. As far as I know, this set of views includes all of the concrete proposals belonging to our three broad categories that enjoy any degree of *prima facie* plausibility.

I will argue that the first and second of these ideas are dead ends, but that the third is probably valid.

### VIII. The Second-Order Theory of *P*-Consciousness

As we observed, the second-order theory has the virtue of being able to explain why so many of our mental states are conscious. It is also true, however, that its explanation is ontologically inflationary: for every conscious state *M*, the theory posits a representation of *M*. This would not be a flaw if there were no other way of explaining the broad distribution of consciousness, or if there was independent evidence that the posited second-order representations exist. But it is a serious flaw if neither of these conditions is fulfilled. I will argue that they are not.

First, the broad distribution of consciousness can be explained in terms of access consciousness. As we saw, it is plausible that robust phenomenal states are the categorical bases of the tendencies that constitute dispositional access consciousness. They are also constitutively involved in occurrent access consciousness. It follows from this that robust phenomenal states generally enjoy access consciousness. Hence, the explanatory power provided by the second-order theory is unnecessary—assuming, of course, that second-order awareness is distinct from access consciousness, as its proponents maintain.

Second, as far as I can see, there is no independent evidence that the representations posited by the second-order theory exist.

There is no introspective evidence—in my case at least. When I introspectively examine my conscious perceptual experiences, I am aware of phenomenology, but I have no impression that the phenomenology is accompanied by a second-order representation. Certainly there is no second-order representation with a distinct, proprietary phenomenology. There is no flashing or halo, no pointer or spotlight, that accompanies perceptual phenomenology but is not part of it. Nor is perceptual phenomenology duplicated, in a way that might suggest some sort of mirroring relation. Moreover, most of the time, there are no conscious thoughts that explicitly refer to the phenomenology. When introspection goes looking for second-order representations of perceptual states, it comes back empty handed.

Nor does folk psychology provide any grounds for embracing the second-order theory. Thus, for example, folk psychology clearly allows us to entertain the hypothesis that certain animals have conscious perceptual states but lack the ability to form second-order representations of them. This sort of claim is taken seriously by primatologists (Cheney and Seyfarth 1990, 2008). It can't be ruled out *a priori* by pointing to some more fundamental folk law. Nor, for that matter, can

the hypothesis that the energy costs would be too great for adult human beings to constantly form second-order states. Second-order representations of phenomenology are as invisible to folk psychology as they are to introspection.

Is there experimental evidence for second-order representations? Lau and Rosenthal argue that there is (Lau and Rosenthal 2011), but, in building their case, they appeal again and again to verbal reports of perceptual experiences, maintaining that they are indicative of second-order representations. This reliance on verbal behavior poses a problem, given that they want to distinguish second-order awareness sharply from access consciousness; for, as we have seen, the ability to issue verbal reports is part of access consciousness. It seems that the methodology of these authors is unable to pry second-order awareness loose from access.

This reflection leads to a dilemma. Either there is behavioral evidence for second-order representations or there isn't. If there is, then the relevant form of behavior is evidence that the representations are access conscious, for accessibility is required for behavior that is complex enough to serve as strong evidence for second-order representations. If there isn't behavioral evidence, then we must look to neuroscience for evidence of second-order representations. But in the absence of introspective and behavioral evidence for them, it is hard to see what grounds we could have for interpreting a particular type of neural activity as second order. Even if we could find a type of activity in the prefrontal cortex that always accompanied a certain type of activity in a perceptual system, why should we construe it as a second-order representation? After all, we lack independent evidence that it even *might* have that status.

It might be suggested that we should recognize second-order representations because they are needed to perform a vital cognitive function of some sort, a function that perceptual states could not perform on their own. But what could that function be? It would have to be quite important to justify the toil and trouble that would have been required of Mother Nature should she have equipped us with it. So it should be very much in evidence if it exists. As far as I know, however, no second-order theorist has ever explained what that function might be. (They could not respond by saying, "Well, of course consciousness is important," because what is at issue here is whether any importance attaches to the specific sort of consciousness that would come from second-order representations of first-order states.)

## IX. The Adverbial Theory of *P*-Consciousness

The adverbial theory posits an abstract relation of conscious awareness and claims that phenomenology consists of modes or determinations of this relation. It identifies *p*-consciousness with the abstract relation.

The problem with the adverbial conception is that it just seems false that phenomenology consists of modes of awareness rather than properties that count as objects of awareness. To be sure, we often describe phenomenology by propositions of the form *Object O looks F to agent A*. The only objects of awareness mentioned by such propositions are values of the variable O. Instead of picking out an object of awareness, F has the role of a modifier of the relation *looks to*. The fact that we rely on this construction lends some support to the adverbial conception of phenomenal consciousness. The support is minimal, however, for we also say things that imply awareness of appearances. As we noticed in Chapter 1, we acknowledge, in both discourse and thought, that we are aware of *how mountains look when seen from a distance*, of *how railroad tracks look as they recede towards the horizon*, and of *how tan things look when they are cloaked in shadow*. Moreover, we must suppose that we are aware of perceptual phenomenology if we are to explain how we have knowledge of its various forms. Surely we know what it is for an object to look blue or look small or look trapezoidal? If we could not be aware of phenomenal colors and sizes and shapes, how could we have knowledge of such things?

It might be held that we don't know about phenomenology by participating in facts of the form *O looks F to agent A* but rather by some sort of second-order awareness of such facts. It is, however, very hard to see what that awareness would consist in. It could not be perceptual, because there cannot be perceptual awareness of perceptual awareness. Nor could it be purely cognitive. One can have thoughts about a type of phenomenology (say, the phenomenology of a creature from a different species) without ever having been experientially aware of phenomenology of that type. In general, mere thinking is incapable of making us presentationally aware of perceptual phenomenology. That leaves introspective awareness. But surely we can be in touch with phenomenology when our introspective faculty is not engaged. Perceptual experience always has a phenomenological dimension (when we perceive an object, we are always aware of the appearance it presents), but introspection is comparatively rare. Phenomenology is experienced whenever we open our eyes. Introspection requires an additional faculty and additional energy.

This is not to deny that we can be aware of phenomenology by introspection. It's just that we need a theory of introspective awareness of phenomenology that is compatible with saying that we grasp phenomenology perceptually. There is growing support for such theories in the literature. Suppose a subject S is engaged in becoming introspectively aware of how a certain object O looks. A number of authors (e.g., Harman (1990), Byrne (2018), Gupta (2019), Wu (2022)) maintain that S will proceed in a world-to-mind fashion. Adopting this idea, I will here view S's procedure as involving three steps. (i) S will begin by using attention to boost the processing that determines the appearance of O, thereby enhancing the crispness and detail of the appearance, and also prolonging the experience of the

appearance so that it is easier for S to conceptualize and describe it. (ii) S's second step will be to form such a conceptualization of the properties of O under the guidance of the appearance, using concepts like *blue*, *spherical*, and *smooth*. (iii) Once such a conceptual representation of O is available, S will then form a judgment like *O looks blue, spherical, and smooth to me*. In this process, there is a direct, experiential grasp of phenomenology only at the first stage, which is perceptual. Introspection merely intensifies perceptual awareness of phenomenology and adds a layer of conceptual awareness.

In my view, this picture of introspection is strongly preferable to the principal alternative, which represents introspection as consisting in Russellian acquaintance (Chalmers 1997, Gertler 2011). Russell thought that epistemology requires a primitive cognitive relation between subject and object that is free from inference, interpretation, and all forms of representation, so he took over the commonsense notion of acquaintance and added these further conditions (Russell 1912). But there are a number of mainstream epistemological programs that have no use for this souped-up conception. Moreover, as far as I can see, there isn't a shred of empirical evidence that points to it. Cognitive scientists are baffled when Russellian acquaintance shows up in conversations with philosophers. They think of awareness as being constituted by representations. (As the reader may recall, I had more bad things to say about acquaintance back in Chapter 5.)

In sum, awareness of perceptual phenomenology is a perceptual affair, not a result of introspective awareness of perceptual states. It occurs whenever we are perceptually aware of appearances, which is whenever perceptual experience occurs, not only on those comparatively rare occasions when there is introspective awareness of perception. Moreover, reflection on the nature of introspection shows that the awareness of phenomenology that it provides already existed as perceptual awareness, albeit in a less attentionally enhanced form.

## X. The Categorical Base Theory of *P*-Consciousness

The categorical base theory makes two claims. First, highly robust phenomenology serves as the categorical base for the dispositions that are constitutive of introspective consciousness and access consciousness, and also serves as the monadic support for the causal relations by which those dispositions are manifested. And, second, highly robust phenomenal states count as *p*-conscious in virtue of underpinning these dispositions and relations. We have already seen that the categorical base hypothesis is plausible: it is the best explanation of the intimate relation between phenomenology and the other two forms of consciousness. What about the second claim? Is there a reason for embracing it?

Yes. Consciousness plays a role in laws that explain why introspectibility and accessibility are limited in certain ways. Thus, it is a law that only conscious states

can be objects of introspective awareness. It is also a law that only conscious states can fix perceptual beliefs, lay down episodic memories, serve as components of plans, and so on. As we all know, these laws provide valuable explanations of psychological facts of various kinds. For example, they explain why the early phases of perceptual and linguistic processing are not accessible to introspection, and why these processes lack the causal powers that are constitutive of access consciousness. Moreover, they clearly cannot be replaced by similar laws involving introspective consciousness or access consciousness. The principle that only introspectively conscious states can be objects of introspective awareness has no explanatory value. Thus, the foregoing laws of consciousness have a utility and an integrity which indicate that there is a causally important form of consciousness that is independent of the forms we have already recognized.

The laws we have been considering indicate that this form of consciousness can figure in explanations of why various mental states and processes lack introspective consciousness and access consciousness. I now add that it appears that this form of consciousness also figures in laws that can explain the downstream effects of these other two forms of consciousness. Thus, for example, it appears to be a law that conscious perceptual states can, in normal circumstances, fix perceptual beliefs, help to shape plans, and give rise to episodic memories. This law can clearly support highly substantive explanations of these downstream phenomena, unlike the principle that states that access consciousness has these same effects. It is at best a weakly informative explanation of a state *S1* to say that it is due to a state *S2* that has a disposition to produce states like *S1*.

It follows from what we have been noticing that this form of consciousness is a locus of causal powers and is therefore eligible to count as a natural kind. But it also follows that it is associated with the same causal powers as a kind that we were considering a few paragraphs back—the kind consisting of all highly robust phenomenal states. This is the best possible evidence that these two kinds are identical.

We have found a strong candidate for the role of *p*-consciousness. I suggest that we should accept both of the following identity statements: “*p*-consciousness = the form of consciousness that figures in the foregoing laws” and “the form of consciousness that figures in the foregoing laws = highly robust phenomenology.” Together these identities answer the question we have been considering since Section VI. *P*-consciousness is what occurs when phenomenology becomes highly robust.

To rephrase the points made in this section, it appears that we have a concept of consciousness that we use to explain certain limitations in the scopes of introspective consciousness and access consciousness. We can see this concept at work in laws such as the proposition that only conscious states can be objects of introspective awareness. It is clear that this concept is distinct from, and indeed independent of, the concepts of introspective and access consciousness. If we were to

substitute the latter concepts in the laws, the resulting propositions would be trivialities, altogether lacking in explanatory power. Further, it appears that we use this additional concept in giving deep explanations of the states that figure in the dispositions that constitute access consciousness. Finally, it seems quite appropriate to say that this additional concept is a concept of *phenomenal* consciousness, since it refers to phenomenology that is highly robust.

## XI. Conclusion

We can conclude, then, that there are several types of consciousness that perceptual experiences can enjoy: occurrent introspective consciousness, dispositional introspective consciousness, occurrent access consciousness, dispositional access consciousness, and substantial phenomenal consciousness (the property *being phenomenology that is highly robust*). Perhaps there is also a sixth type—Millian phenomenal consciousness (the second-order property *having some phenomenal character or other*). Another conclusion is that phenomenology is metaphysically autonomous, in the sense that it does not entail either of the forms of introspective consciousness and does not entail either of the forms of access consciousness. A more accurate account of the relationship between phenomenology and introspective consciousness and access consciousness is that highly robust forms of phenomenology—that is, *conscious* phenomenology—provide an intrinsic, categorical base for the dispositions and occurrent causal relations that constitute them.