Understanding Phenomenal Consciousness

WILLIAM S. ROBINSON

Iowa State University
Contents

Acknowledgments xi

Part I

1. INTRODUCTION 3
   Perception and Our Basic Question 5
   Organization of This Book 9
   A Note on Terminology 11

2. QUALITATIVE EVENT REALISM 14
   Three Clarifications 15
   Afterimages and Events 21
   Structure and Location 25
   What Is Consciousness? 32
   The State of the Argument So Far 34

3. DUALISM 36
   Reduction and Scientific Progress 43
   Property Identity without Explanatory Reduction 46
   Our Limited Minds 50

4. REPRESENTATIONALISM 53
   The P/T Problem 56
   Differences of Content 59
   Differences of Representational Type 62
   Conclusion 72

5. TRANSPARENCY 73
   Reasons against Three-Dimensionality of Experiences 76
   A Variant of Transparency 83
   The True Relation between Experience and Representation 84
6. HIGHER ORDER THEORIES 87
   The Stone Objection and Some Replies 92
   A Dilemma for HOT Theory 96
   Unconscious Pains and T astes? 98
   Burdens 105

7. MONITORING 107
   Armstrong's Drivers 107
   Empirical Considerations 113
   Quadrangle Bells and Other Puzzling Cases 117
   Future Empirical Support for Monitors? 123
   Conclusions about Higher Order Theories 128

8. FUNCTIONALISM 130
   Behaviorism 133
   Homuncular Functionalism 134
   Microfunctionalism 138
   Identity and Supervenience 140

9. SKEPTICISM AND THE CAUSES OF QUALITATIVE EVENTS 143
   Skepticism versus QER: How May We Proceed? 145
   Excessive Determinateness? 148
   Internalism 153
   Conclusions 156

10. EPIPHENOMENALISM 158
    Antiepiphenomenalist Intuitions 160
    Self-Stultication 165
    Contingency and Knowledge 167
    Qualitative Events and Knowledge: A Dilemma 174
    Loss of Control 178

Part II

11. UNIFIED DUALISM 183
    Knowing Matter from Inside 187
    Quantum Mechanics 193
    Why There Will Always Be One Kind of Explanatory Gap 201
    Conclusion 205

12. PATTERNS AS CAUSES OF QUALITATIVE EVENTS 207
    A Thought Experiment 207
    Homogeneity 214
Contrast with Chalmers's Principle of Organizational Invariance 217
Homogeneity, Computation, and “Information” 224

13. A POSSIBLE FUTURE 227
Independence of Pattern from Medium 227
Panpsychism 233
Isn’t This Still Mere Correlation? 235
Conservation 241
The Nontransitivity Problem 243
Conclusion 249

References 251
Index 259
1

Introduction

The central problem of this book can be introduced by thinking about even the simplest cases of phenomenal consciousness. Let us, therefore, begin with afterimages, and let us fix our ideas by reference to a novelty item that some readers may have encountered. This item is similar to an American flag, only it is printed in green, black, and yellow. A caption instructs users to stare at the oddly printed flag for 10 seconds, in good light, and then to look at a white wall or a white piece of paper. The predicted, and actual, result is that one will then see an American flag in its correct colors.

The principle behind this novelty item is that afterimages brought about by brightly colored things will have the complementary colors of the objects that cause them. This principle can easily be verified by staring at some brightly colored thing in strong light, then looking at a white surface. Red or green will each produce the other, as will blue or yellow. This little experiment will also demonstrate that the distance of the afterimage will be the same as the distance of the surface upon which one's eyes are focused.

I want to know how red comes into the situation we are in when we have stopped looking at the printed “flag” and are having an afterimage in which we see red. The reason for puzzlement here is that there isn’t anything red in front of me, where a flag appears to be. Unless I have some strange disease, there isn’t anything red in my brain, either. (Well, there is oxygenated blood; but this is always present, and so its color has nothing special to do with the redness in my afterimage, i.e., nothing that it doesn’t equally have to do with green or blue afterimages.) But to say that there is nothing red at all anywhere in the situation would seem to
deny an obvious fact, and to make it difficult to give an account of the difference between the redness of the stripes in the afterimage and the blueness of its upper-left-hand corner.

There are, of course, differences among the neural events that cause our afterimages. Color stimuli increase the activity in some of our cells, and different stimuli increase the activity in different sets of cells. Increase of activity in some cells results in inhibition of other cells, and when a stimulus is removed, there can be a “rebound” effect in the inhibited cells—an effect that “overshoots” the neutral state for a short time and thus produces a reverse color impression. These neural activations, however, are not something of which we are ordinarily conscious. For example, Aristotle knew about afterimages almost as well as we do, but he had no knowledge whatever about cells in our visual systems. If we confine ourselves to the neuroscience of how afterimages are produced, we will never have any reason to use color words, and thus we will not have a full answer to our question of how colors come into the afterimages that neural activations bring about.

It would be natural to say, at this point, that having an afterimage of the stripes in the American flag is a situation in which it looks like there is something red before us. This is surely correct, but it leads to the further questions “What exactly is looking like?” and “What is the difference between its looking like there is something red before us, when there isn’t, and its looking like there is something blue before us, when there isn’t?” One can give a short answer, that the difference is the difference between red and blue; but as there need be no red or blue things present, this answer just returns us to the question of how colors come into situations in which we have afterimages.

In this book, I shall consider several answers to this surprisingly complex question. This question, however, is only one, readily understood, member of a family of related questions. For example, other senses could have been considered. Removing a tight hat can leave us with what we may call an “afterfeel” of pressure along a circle around the head. Some foods, or medicines taken by mouth, can leave an aftertaste. I want to know how these sensory qualities come into the situations described. I want to know what is happening that is different when (due to different hats) afterfeels are in different places on the head, and what constitutes the taste difference between two aftertastes.

1 Aristotle, *De Somniis* (On Dreams). I say “almost”, because Aristotle seems to have missed the fact that afterimages often have colors complementary to those of their external causes.
Bodily sensations, e.g., pain, itch, sexual pleasantness, or nausea, do not present “after sensations”, but they do lead to questions similar to those about afterimages. For example, pain is caused by tissue damage, and our neurons have to be working in order for us to feel pain. But having pains does not give us knowledge about cells or about the neural causes of our painful feelings. Thus, we can sensibly wonder just how we should describe the way in which painfulness comes into a situation in which we are suffering. Similar remarks hold for a wide variety of feelings that are associated with emotions, e.g., fear, the feeling of butterflies in the stomach, the feeling one has when angry, and so on.2

Visual images (not afterimages) can be formed by most people upon request. For example, one can be asked to imagine a pink elephant or the face of some famous actor. Auditory images are likewise familiar – imagine, for example, the sound of some sinister remark by Darth Vader. Imagery of this kind is not the most typical example of the occurrences to which the arguments of this book are intended to apply, and there are many aspects of imagery that will not be investigated here. Nonetheless, the views to be considered have some bearing on the question of what kind of difference occurs when, on one occasion, we have an auditory image of, say, “red” and on another occasion we have an auditory image of “blue”.

PERCEPTION AND OUR BASIC QUESTION

Some years ago, while lost in thought, I happened to stare out through a window with a venetian blind for a considerable time without moving. When I did move, my eyes chanced to fall on my trouser leg, where, to my surprise, I saw that the brown fabric had a pin stripe in it that I had never noticed before. Or did it? How could I have never previously noticed such a thing? Yet, there it was. It was at least 15 seconds before I was able to convince myself that the unnoticed pin stripe was in fact merely the afterimage of the bright spaces between the slats of the blinds.

This anecdote indicates a close relation between what happens in cases of afterimaging and cases of ordinary perception; namely, they are so alike

2 Besides feelings, emotions often include a cognitive component. For example, to be remorseful, one not only has to have a certain sort of bad feeling, one also has to believe that one has failed to live up to a standard one accepts. To be embarrassed, one must believe that one has done something socially unacceptable. This book will not attempt a theory of emotions; they are mentioned here only because the feelings involved in them are examples of phenomenal consciousness.
that one may be mistaken for the other. My puzzlement about afterimages is thus naturally entwined with the question of how colors come into our ordinary perceptions of colored things. To pursue this question, and to fully understand the motivation behind it, it will be essential to have an account of some basic facts about how we see. Fortunately, these facts are very familiar, and not controversial until the end of the account, so a brief and sketchy overview will be sufficient. As it will be helpful to consider a particular example, let us begin by telling the story of what happens when Eve sees a ripe, red apple.

Eve will not see anything if it is pitch dark, so let there be light, and let us assume it is sunlight. The light must strike the apple if Eve is to see it, and that light must be reflected into at least one of Eve’s eyes. At the point of reflection, that is, on the surface of the apple, something distinctive must happen—something that transpires differently, depending on whether the reflecting surface is red, green, or some other color. This difference arises because of the particular molecular structure of the surface of the skin of the apple, which results in some of the wavelengths present in sunlight being reflected more efficiently than other wavelengths. The light reaching Eve’s eye from the direction of the apple thus has a composition different from sunlight, that is, the ratios of the amounts of light at various wavelengths differ from the ratios found in sunlight.

The propensity to make this difference in the wavelength composition of arriving light and reflected light is a property of the apple’s surface that it acquires when it becomes ripe and that remains with it for a good while (that is, until it spoils). It is convenient to have a name for this property, and we shall use “reflectance profile”. The apple has its particular reflectance profile because of its molecular structure; but things with a different molecular structure might have the same reflectance profile, so we should count the molecular structure and the reflectance profile as two properties that the apple possesses.

Let us resume our story at the point where reflected light is traveling toward one of Eve’s eyes. This light enters her eye, becomes focused by her lens, and falls on the cells that compose her retina. Some of these cells (the cone cells) contain chemicals that change their state when light falls on them. There are several of these chemicals, and they differ in the wavelength of light that is most likely to cause them to change their state. Suppose that a collection of cone cells is illuminated by light reflected from the apple, and suppose that this same set of cone cells is illuminated on another occasion by light reflected from a daffodil bloom. The reflectance profile of the daffodil bloom is different from that of the apple, so the
wavelength composition of the light entering Eve’s eye from the daffodil bloom will be different from that of the light that enters her eye from the apple. The result will be that the chemical changes in the set of cone cells will be different in the two cases.

Cone cells, like neural cells, can be regarded as having an input end and an output end. We can take the chemical changes to be the inputs. The outputs are releases of molecules called “neurotransmitters”. Releases of neurotransmitters, in turn, raise the probability that other cells, namely neurons, will undergo a certain kind of change. This change commonly goes by two names, “action potential” and “firing”. What these terms refer to is a process that results in a cell releasing neurotransmitters from its output end.

Our story now becomes exceedingly complex and fascinating in detail, but, fortunately, it admits of a simple summary if we take a somewhat abstract point of view. Neurotransmitters from cone cells cross the tiny spaces that separate cone cells from neurons, called “synapses”. Neurons receive these neurotransmitters and, under certain conditions, undergo internal changes that lead to the release of neurotransmitters at their output ends. These neurotransmitters cross synapses and are received by other neurons, which again may undergo internal changes that lead to their release of neurotransmitters. This story is repeated again and again; in general, millions of neurons in the brain may have their activity altered as a result of the light reflected from the apple falling on Eve’s retina.

In due course, Eve may respond to the presence of the apple. Perhaps she will utter the words “What a nice red apple!” or perhaps she will reach for it. In such cases, more neurons will undergo changes, release neurotransmitters, and raise the probability of other neurons firing and releasing neurotransmitters. At the end of this process, some of the released substances will encounter muscle cells instead of neurons. In this case, the muscle cells will contract, and this contraction will result in the movement of Eve’s lips and tongue or her arm.

The foregoing account has been selected and emphasized for philosophical purposes, but it is the account that, so far as it goes, can be found in any introductory psychology textbook. At this point, however, further questions arise that do not have noncontroversial, textbook answers in either psychology or philosophy. We can see how they arise by quickly listing the elements we have mentioned: wavelength composition of sunlight, reflection, altered wavelength composition of the reflected light, lens, retina, chemical changes, neurotransmitters, synapses, neural activities, neurotransmitters, synapses, neural activities, neurotransmitters,
synapses, neural activities . . . muscle-stimulating substances, muscle con-
tractions, movements. Nowhere in this list do we find red or any other
color. But Eve sees red. The apple she sees is red, and if conditions are
normal, it looks red to her. How or where, we may ask, does Eve’s experi-
ence of red come into our account at all? How is what normally happens
when Eve sees a red apple like, or not like, what happens when she has a
red afterimage? It will be convenient to summarize these questions into
one general formulation, which will serve as a Basic Question in the early
stages of our discussion.

(BQ) How does color come into a full accounting of what normally happens
when a person sees a red apple?

The theories to be considered in the next few chapters do not all approach
this question from the same angle, nor do they accord it equal prominence.
They would differ in their ways of making the question more specific.
Despite these differences, they all imply answers to the Basic Question
as just formulated, and considering these answers will provide a way of
comparing the commitments and success of several important theories.

The Basic Question is most easily discussed by reference to particular
examples, but there is nothing special about the color red or the sen-
sory modality of vision. The label (BQ) or the phrase “Basic Question”
may thus stand indifferently for the preceding formulation or for a still
more general formulation. To give the more general formulation, we will
have to have a term that will cover qualities of all the kinds we have
mentioned, such as colors, pitches, tastes, smells, degrees of pressure
and warmth, shapes (e.g., of afterimages or apples), pain qualities, itches,
sexual pleasantness, nausea, and other qualities like these. This term is
“phenomenal qualities”. The resulting general formulation of (BQ) is
“How do phenomenal qualities come into a full accounting of what hap-
pens when a person is having a perceptual experience or sensation?” Since
some philosophers have special understandings of the terms used in this
general formulation, I stipulate that it is to be understood simply as a way
of encompassing the example given and others that are like it in the way
the items on this list are alike:

How does flavor come into a full accounting of tasting a spoonful of honey?
How does sound come into a full accounting of hearing a harp string?
How does warmth come into a full accounting of feeling the brow of a fevered
patient?
How does pain(fulness) come into a full accounting of what happens when someone stubs a toe?
How does red come into a full accounting of what happens after staring at a flag printed in the complementaries of its usual colors?

We can see red things that, like heated metals, fireflies, and the sun, emit light rather than reflect it. We can also see colors contributed by transmitting media, such as stained glass windows, or ice cubes made from pink lemonade, and colors produced by refraction, as in rainbows, diamonds, and oil slicks. Differences among these cases, however, will not make any substantive difference in the arguments to follow. I shall thus avoid the tedium of repeating these other possibilities and conduct the discussion of visual examples almost entirely in terms of objects that reflect light.

**ORGANIZATION OF THIS BOOK**

Let us imagine one answer that might be given to (BQ). This answer, which I will call “Minimalism”, begins by noting that we learn to apply the word “red” to a large set of objects that have a variety of reflectance profiles. Because of the structure of normal human perceptual and cognitive systems, normally sighted and normally trained English speakers almost always agree about when “red” should be applied to things they have never previously seen. According to Minimalism, the class of red things is the class of things that normally sighted people who have had the training typical of English speakers call “red”. (Of course, nonnative speakers can learn the English habit. Further, many languages have a word that is applied to all and only the same things to which “red” is applied by normal English speakers. So, despite some circularity, Minimalists may also say that the red things are those that normal English speakers call “red” and that speakers of other languages call by their term that translates as “red”.) The property red that a thing may have, according to this way of thinking, is the property of being such that normally sighted and trained English speakers would call it “red” in normal conditions of seeing if they are being sincere and are not incapacitated by drugs or disease. And, according to Minimalism, the way in which red enters into a full accounting of what normally happens when Eve sees a red apple is exactly this: the apple that Eve is seeing has the property red (i.e., the property red as understood by Minimalism).
We may put this claim as the view that Eve is seeing an apple that is M-red, i.e., that has the property red as that property is understood by Minimalism. But a Minimalist holds that there is no other property red, no other redness in any sense anywhere to be found in the full accounting of what normally happens when Eve sees a red apple. So, Minimalists have no reason to add a special flag to their uses of the word “red”; for them, M-red just is red.

Most philosophers are not Minimalists. But if something needs to be added to what Minimalism provides, there are still many questions about what kind of addition should be made. This book will argue for a particular kind of addition. To anticipate, it will be argued that an adequate answer to the Basic Question requires us to recognize that there are nonmaterial events that consist in occurrences of phenomenal qualities, and that such occurrences are nothing more or less than episodes of phenomenal consciousness. I call this view “Qualitative Event Realism” (QER, for short), and I will begin to explain what this label means in the next section. Chapter 2 will begin the substantive argument for the view.

QER shares its dualism (i.e., its commitment to nonmaterial events) and its focus upon phenomenal qualities with many other theories. Thus, part of what is to be done in this book is to justify the particular version of dualism presented here. Since there are several distinctive aspects of QER, this is a task that will continue throughout the book. Some of my initial efforts to motivate QER will support a variety of views, but many of the arguments to come are reasons for adding certain specific commitments to a view that begins by accepting the cogency of the initial motivations. The overall aim will be to support dualism by articulating the most defensible form of it and arguing for that.

In Chapter 3, I will explain why dualism cannot be demonstratively established in a simple and direct way. This fact entails that QER can be accepted only if it proves to be the best theory of phenomenal consciousness, all things considered. This conclusion, in turn, forces the argument for QER to have a certain structure: not only must its virtues be identified, the inadequacies of its rivals must also be shown. That is why several of the chapters that follow are devoted to the discussion of alternatives to QER.

The alternatives to QER that we must consider fall into two kinds, which can be roughly characterized as the partially sympathetic and the downright hostile. The latter are versions of materialism, and these are considered in several places, mostly in Part I. While the ultimate conclusion of these discussions is negative, it should be borne in mind that
these views have been very thoroughly worked out. Many ideas that we must understand in order to understand phenomenal consciousness are either parts of the materialist views we will examine or are most easily understood in the context of explaining their inadequacies.

Part II has a somewhat different character. If nonmaterial qualitative events are recognized, their relation to the neural events that cause them must be investigated. Chapter 11 contains discussions of two views that appear to have some promise but are not adopted here. Chapter 12 presents the distinctive answer of this book, namely, that patterns (probably of neural activation) are the causes of qualitative events. This proposal naturally leads to a number of objections, and the view is developed in reference to these in Chapter 13. This chapter is necessarily speculative, but goes as far as seems possible toward envisaging a future unified, intellectually satisfying perspective on the problem of understanding phenomenal consciousness.

A NOTE ON TERMINOLOGY

“Phenomenal quality” was introduced earlier as a general term for qualities that enter in some way into perceptual situations, bodily sensations, emotions, and imagery. This is the meaning that this phrase will have throughout this book. In the works of some of the authors to be discussed, we will also encounter another term that has much the same use, namely, “qualia” (singular: “quale”).3

One example of a phenomenal quality is pain. This is a general property, of which shooting pain, dull pain, pain in the left leg, and so on are species. Besides having species, however, the property pain has particular occurrences. “I have a shooting pain in my left leg” says that an event of a certain kind is occurring, and the kind in question is shooting pain in the left leg. I could have other occurrences of the same kind on other occasions, and in each case I would be correctly said to have a pain. “A pain” can, of course, also be used to indicate a species of pain; for example,

3 The term “quale” goes back to Peirce’s work in the 1890s (see Peirce, 1938), where it is introduced in connection with a number of difficult doctrines. The root meaning of contemporary usage is best understood by reference to C. I. Lewis. “There are recognizable characters of the given, which may be repeated in different experiences, and are thus a sort of universals: I call these ‘qualia.’ But although such qualia are universals, in the sense of being recognized from one to another experience, they must be distinguished from properties of objects” (Lewis, 1929, p. 121; emphasis in the original). Lewis’s examples in nearby text were red, blue, round, and loud.
arthritic pain is a pain and headache is another. But in one of its standard uses, “a pain” indicates a certain kind of occurrence, which has a time of onset and lasts for some particular duration.

There is no term in common use that unambiguously functions like “a pain” but can be used in a general way for an occurrence of any phenomenal quality. “A quale” might serve, by analogy with “a pain”, but the term seems so wedded to its use as a genus of properties that confusion would probably result. It seems best, therefore, to coin a term for the explicit purpose of indicating occurrences of phenomenal qualities, and this is the meaning of the term “qualitative event”.

Just as one can ask whether a pain is identical with an event in some part of one’s brain, so one can ask whether qualitative events are identical with events in brain parts. But we should note that an answer to the first question cannot be gotten out of the grammar of “a pain” alone. A pain has to be an occurrence of (some specific) pain property, but whatever else may truly be said of it has to be argued for. Analogously, a qualitative event has to be an occurrence of a phenomenal quality, but the grammar of the term alone neither includes nor excludes further characterization.

By contrast, the theory that this book will offer will make substantive claims about qualitative events that do not follow merely from the meaning of the term. According to this theory, qualitative events are caused by, but not identical with, some brain events, namely, occurrences of patterns of activity (most likely, of neurons). Many other claims will be made that need some preparation for their understanding. These substantive claims must be (and will be) argued for and defended.

The name of the theory to be offered is, as noted previously, “qualitative event realism” (or QER). Now, one could think that qualitative events are real, and even real and nonmaterial, and still not accept other parts of the theory of this book. Thus, strictly, this book offers a particular version of QER – QER plus a number of distinctive, specifying theses. However, at present, neither the term “qualitative event” nor the phrase “qualitative event realism” is in use. I will thus use “qualitative event realism” as the name of the particular theory, with all its specifying theses, that I present in this book.

The occurrences to which “qualitative event” is intended to refer go by many names, of which “experiences” and (episodes of) “phenomenal consciousness” are the most familiar. To maintain connection with the thought of others, I shall often use these terms, and it is to be understood that the theory of qualitative events presented here is a theory of what experiences, episodes of phenomenal consciousness, etc. actually are.
It will be convenient to have a term that indicates acceptance of qualitative events without taking a position on certain specific, additional theses of QER. I shall use "experiential realism" for this purpose. QER is to be understood as one among several species of experiential realism. Some of the arguments for QER will be arguments for experiential realism in many forms. Others will be concerned to support QER as the best form of experiential realism, to the exclusion of other species.

It remains to say something about "consciousness". It is likely that every reader will have some idea of what is meant by this term, and that this idea will not be completely off the mark. Moreover, it is likely that most will agree that rocks and trees are never conscious, that normal, awake people are conscious, and that it is a lot more plausible that dogs are conscious than that bacteria are conscious. “Consciousness” ought to be whatever it is that we are thinking of as making the difference between these classifications with respect to being conscious.

Beyond this understanding it is not possible to go in a preliminary, definitional, noncontentious way. The work to follow, however, is most assuredly relevant to further understanding of consciousness. As was briefly indicated earlier, QER holds that qualitative events are episodes of consciousness; in clarifying our understanding of qualitative events, therefore, we are furthering our understanding of consciousness. Moreover, the speculative unification of thought attempted in Chapters 12 and 13 provides us with a glimpse of how intellectual satisfaction about consciousness and its relation to events in the brain might one day be achieved. At present, however, these claims can only be regarded as oracular and contentious; clarification and argument are promised.4

4 The structure of the phrase “phenomenal consciousness” may suggest a contrast with some other kind of consciousness. From the point of view of QER, however, “phenomenal consciousness” is almost redundant; what is phenomenal is always conscious, and where there is consciousness, there are phenomenal qualities. There is, however, no point in giving a preliminary argument for this statement, as it will become evident, if the argument of the rest of the book is accepted. For an interesting and completely different route to a similar conclusion, see Lormand (1996).