

The Representational Theory of Consciousness

David Bourget

A thesis submitted for the degree of
Doctor of Philosophy
of the Australian National University

January 2010

This PDF contains the introduction of *The Representational Theory of Consciousness*. If you'd like to read more, please contact me. My details are on my web site at www.dbourget.com.

Introduction

A satisfactory solution to the problem of consciousness would take the form of a simple yet fully general model which specifies the precise conditions under which any given state of consciousness occurs. Science has uncovered numerous correlations between consciousness and neural activity, but it has not yet come anywhere close to this. We are still looking for the Newtonian laws of consciousness.

One of the main difficulties with consciousness is that we lack a language in which to formulate illuminating generalizations about it. Philosophers and scientists talk about "what it's like", sensations, feelings, and perceptual states such as seeing and hearing. This language does not allow a precise articulation of the internal structures of conscious states and their inter-relations. It is inadequate to capture relations of the kind we are looking for between conscious states and physical states.

In this thesis I refine and defend a theory of consciousness which promises to solve this regimentation problem: the representational theory of consciousness. I argue that the representational theory can solve the regimentation problem and smooth out other important obstacles to a fruitful study of consciousness. I also make a case for the theory independently of its payoffs, and I discuss the leading opposing theories at some length.

In the rest of this introduction, I will clarify what I mean by "consciousness", provide an initial characterization of the representational theory, and outline my project in more detail.

Consciousness

At this stage in the thesis, I cannot provide a precise and unambiguous definition of consciousness which will please everyone, but I will do my best to fix ideas. What matters is that you and I attend to the same phenomenon. Articulating the nature of this phenomenon is the project the rest of this thesis tackles.

The kind of consciousness this thesis is about is the kind which is widely thought to pose a unique challenge to scientific explanation. It is the kind which has come to be called "phenomenal consciousness". To be phenomenally conscious is to instantiate a *phenomenal state*. The paradigmatic phenomenal states are states we instantiate in the course of our sensory, emotional, and cognitive experiences. Here I am using "experience" in a sense which should be familiar from everyday talk. I will give some examples of each kind of experience just mentioned.

At the moment I have a mild back pain. My experience of pain is a paradigmatic sensory experience. There are many other kinds of sensory experience. When I look around myself, for example, I undergo a large number of visual experiences. Visual experiences, as I think of them, have a felt component a little bit like pain experiences. As Nagel (1974) puts it, there is something it's like for a subject to have a visual experience. There are also auditory, tactile, olfactory, gustatory, kinaesthetic, and proprioceptive experiences, and probably many other kinds of sensory experience which do not fall in any of these categories. There are probably infinitely many kinds of possible sensory experience possible creatures could have. All have felt components which are part of their essences as sensory experiences.

Emotional experiences also come in a wide variety. For example, one undergoes emotional experiences when one feels anxious, sad, relieved, or elated. Emotional experiences should not be conflated with emotions. On one common understanding of emotions, at least, they are states which can persist independently of how one feels. For example, one can be angry at a time without actually feeling angry at that time. We can leave the question of how emotions and emotional experiences

relate to each other largely open for now, but we need to recognize that emotional experiences, with their felt components, do not always accompany the emotions with which they tend to be associated. It is emotional experiences which are of interest here, not emotions.

I count as cognitive all conscious experiences which are not normally associated with emotions or sensory processes. Cognitive experiences are more elusive than sensory and emotional experiences. Goldman (1993) draws attention to cognitive experiences by comparing what it is like to hear and understand a statement with what it is like to hear the same statement without understanding it. There is a certain feeling of understanding missing in the second case—a feeling Strawson (1994) describes as an *understanding experience*. Goldman also mentions the tip-of-the-tongue feeling. There are many other kinds of "cognitive feeling", for example, the feeling that something is right (or not right), the feeling of being confused, and the feeling of *deja vu*. We constantly rely on cognitive feelings like these for guidance in everyday life.

Also central to one's mental life is sensory imagination (which I count as cognitive, but that is merely a convenient terminological choice). All of us have some capacity for visual sensory imagination. For instance, one can faintly visualize (experience) an object one is looking at being moved from its actual location to another location. Aural imagery is also very common. When you talk to yourself "in your head", you are experiencing aural imagery. I will leave other, more controversial forms of sensory imagery aside for now.

There might be cognitive experiences which are less sensation-like than emotional feelings and sensory imagery. We could describe these as *pure phenomenal thoughts*. I will remain neutral as to whether there are pure phenomenal thoughts.

Each kind of experience I have mentioned has a felt component: for each kind of experience I have mentioned, there is something it is like to have an experience of this kind. Phenomenal states are states of the kind one is in when undergoing

sensory, emotional, and cognitive experiences, in virtue of which there is something it is like to have these experiences.

Phenomenal state A state of the kind best exemplified by the states a) instantiated by individuals in sensory, emotional, and cognitive experiences b) individuated by the felt components they confer to such experiences.

It is noteworthy that experiences, in the everyday sense of “experience”, are not the same as phenomenal states or instantiations of phenomenal states. Consider the definition of the verb "to experience" given by the Oxford English Dictionary. According to the OED, to experience something is to "encounter" or "undergo" an "event or occurrence".¹ It is clear that this is how "to experience" is used when one says (for example) that Microsoft has experienced a slowdown. It is used in exactly the same way when talking about consciousness-involving sensory episodes in a lay context.² Take for example these three everyday statements:

- (1) I experienced pain
- (2) I experienced flu symptoms
- (3) I experienced powerlessness

The OED interpretation of "to experience" makes sense of all three: I encountered / underwent pain, I encountered / underwent flu symptoms, and I encountered / underwent powerlessness.

On this everyday interpretation of "experience", a sensory experience of red is a kind of encounter with redness. A sensory experience of an object (say, my kitchen table) is a kind of encounter with that object. Qua encounter, an experience requires the presence of what is encountered: one could not possibly encounter my kitchen table without being related to it in some way. The same goes for an experience of

¹"experience noun", The Oxford Dictionary of English (revised edition). Oxford University Press, 2005. Oxford Reference Online. Retrieved on 2 May 2009 from: <http://www.oxfordreference.com/views/ENTRY.html?subview=Main&entry=t140.e26250>

²Thanks to Alex Byrne for helpful comments here.

Sirius (the star): one could not possibly have a visual experience of Sirius in the OED's sense without being affected by Sirius in some way.

Many theorists would deny that there is such a thing as a Sirius-feeling. There is a phenomenal state one typically instantiates when one visually experiences Sirius in the OED's sense (when one visually encounters Sirius), but this state is not characteristic of Sirius experiences (as opposed to experiences of other stars). According to the theorists in question, one could instantiate exactly the same phenomenal state as part of a visual experience of (a visual encounter with) another star. For this reason, we cannot say simply that the phenomenal states are the states instantiated in sensory, emotional, and cognitive experiences. This would invite an understanding of phenomenal states as encounter states, and such states clearly involve more than what the philosophers in question take phenomenal states to be by definition, namely, states individuated by their felt components (their "what it's like" aspect). This is why I stipulate that phenomenal states are individuated by their felt components: because the felt aspect of an experience in the OED sense need not involve the particular object of the experience.

We will see later in the thesis that there are some who maintain (contrary to the aforementioned philosophers) that many if not all phenomenal states essentially have external particulars as components. This claim would be nearly trivial if phenomenal states were the states which constitute experiences in the everyday sense. On the other hand, it is far from trivial on my definition of phenomenal states. But my definition is not meant to exclude this view. The aim is to isolate the substantive issue which is at stake in the debate between proponents of this view and their opponents.

So far I have been using the term "experience" in its everyday sense. It is not used in this way by everyone, and the everyday sense is rather nebulous (though clear enough for the use to which I put it, I hope). For our purposes, it is helpful to restrict and regiment our use of "experience". From now on, I will use this term as

follows:

Experience An event which consists in instantiating a phenomenal state.

This definition of experiences makes them events of the kind described by Kim's (1976; 1991) theory of events. I think this is a common understanding of the noun "experience" among philosophers, but there are probably other uses. We must keep this in mind as we progress. We must also precisify the meaning of the verb "to experience". As I use it, it means instantiating a phenomenal state.

Theorists often talk about "phenomenal characters" and "phenomenal properties". Here I am going to use these terms interchangeably with "phenomenal state" (I take states to be properties). I take it that my use of these terms is common, but there are no doubt exceptions.

The representational theory

The representational theory of consciousness also goes by the names of "representationalism" and "intentionalism". The best known versions of this theory are illustrated by the following statements:

- (1) All mental facts are representational facts, and (2) all representational facts are facts about informational functions. (Dretske 1995: xiii)
- Phenomenal character is one and the same as representational content that meets certain further conditions. (Tye 2000: 45)
- The propositional content of perceptual experiences in a particular modality (for example, vision) determines their phenomenal character. (Byrne 2001)
- Phenomenal properties are identical to certain representational properties. (Chalmers 2004)

- Qualia are actually intentional contents, represented properties of represented objects. (Lycan 2005)

To a first approximation, the points which unify the representational theories of the preceding authors are a) that experiences essentially have representational or intentional contents and b) that their phenomenal characters are largely determined by their contents.

The recent history of the representational theory can be traced back at least to Anscombe (1965), Armstrong (1968), Hintikka (1969), and Pitcher (1971). It has attracted a significant following during the past decades, but it remains controversial.³ A large tangle of objections, counterarguments, counter-counterarguments, misunderstandings, and alternate formulations has developed. Unless some systematic way of simplifying and clarifying the debate is discovered, there is little hope for further convergence. This is part of what I aim to provide in this thesis.

It strikes me that the bulk of the objections to representationalism are only effective against specific versions of the theory. Some oppose the externalist commitments of Dretske's, Lycan's, and Tye's views, but there are internalist versions of representationalism. Others object that experiences do not have satisfaction conditions, but we don't have to think of the representational contents of experiences as satisfaction conditions. Some argue that experiences cannot be individuated by how the world "looks" or "seems" to one in them, but that experiences make the world look or seem a certain way in a non-trivial sense is not a general commitment of

³According to the PhilPapers Survey 2009 (<http://philpapers.org/surveys/>), the distribution of relevant views among the faculties of leading philosophy departments is as follows:

- representationalism: 31.4%
- qualia theory: 12.2%
- disjunctivism: 10.9%
- sense-datum theory: 3.1%
- other theories (e.g. the theory of appearing): 3.8%.

While representationalism is the most widely accepted view, the debate is far from settled. It is also noteworthy that the theory defended here is stronger than the average representational theory in many respects (though weaker in other respects).

representationalism. And on it goes.⁴

Since there seems to be no common basis for these objections, one naturally wonders whether we would be left with anything of interest after subtracting everything that is controversial from representationalism. Maybe not, but I believe that it is possible to articulate a theory which captures the essence of the representationalist program yet is free from the most controversial implications of current versions of the theory. Relatedly, I believe that it is both possible and necessary to clarify this program. One of the main reasons critics pick on the details of particular representationalist views is that there is no widely agreed upon statement of the core theory representationalists share. In particular, there is a bewildering variety of interpretations of the notion of representational content apparently central to the theory. As it is, one could reasonably doubt that there is such a thing as *the* representational theory of consciousness. I believe that there is a core representational theory, and I believe that this theory has few of the features of specific representational theories which have been found controversial. My goal with this thesis is to articulate this theory, exhibit its potential as a framework for the scientific study of consciousness, build a positive case for it, and show how well it fares on key issues which have surfaced in the debates surrounding representationalism.

I call the purified representationalism I defend *virtualism*. I use this new term for two reasons. Firstly, the names "representationalism" and "intentionalism" have by now received so many divergent interpretations that their linguistic meanings could well be beyond repair. Secondly, whether virtualism really is a kind of representationalism or not does not matter; I don't want to invite a debate on this by using the term. What matters (to me, anyway) is that virtualism captures what representational theories share with one another, including in particular their applications.

It would help to have in sight some of these applications before starting. I already mentioned representationalism's promise as a solution to the problem of regi-

⁴See section 1.2 for more on these objections.

menting consciousness. In my opinion, this is the most important application of the theory. Let suppose for the purposes of illustration that a phenomenal state is a state in which one phenomenally represents a certain proposition. Let us not worry about what it is to phenomenally represent a proposition; this will become clear later. If a representational thesis of this type were correct, all differences between phenomenal states would be accounted for by differences in the propositions one represents in them (their "contents"). Now, propositions can be described using logic, and general relations between propositions and others entities, be they other propositions or worldly entities, can likewise be drawn formally. For this reason, it is reasonable to expect this kind of representationalism to enable a new level of regimentation in describing phenomenal states, their inter-relations, and their place in nature. I will elaborate more on this in chapter 2.

Representationalism also promises to help solve (or dissolve) the problem of perception. The problem of perception is best seen as a tension between two observations. On the one hand, perceptual experience appears to afford us a direct, unmediated awareness of our surroundings. On the other hand, no one can deny the possibility of illusions and hallucinations. These two observations come in tension through arguments sense-datum theorists have offered which purport to show that the possibility of illusions and hallucinations implies that perceptual awareness is primarily awareness of mental "sense data" (c.f. Ayer 1940; Moore 1905; Jackson 1976,1977). If sense-datum theorists were correct, the two aforementioned observations would be inconsistent. For many philosophers of perception today, the problem of perception is to show that sense-datum theorists are wrong (Crane 2005). To show this, we need to make room for both the immediacy of perceptual experience and the possibility of illusion and hallucination. The sense-datum theory is sometimes referred to as "the representative theory of perception", but it should not be conflated with representationalism, because the latter is generally intended to provide an alternative to the sense-datum theory which achieves just this.

How does representationalism undermine the inference from the possibility of illusions and hallucinations to the conclusion that what we are immediately aware of in perceptual experience are mental sense data? Representationalists hold that experience is a species of representation. This does not mean that our experience of the world is "mediated" in the sense that an experience of the world can only be had by experiencing some internal entity which stands for or represents external facts. To a first approximation, all this means is that one can experience an F without there being an F. This is on the face of it perfectly compatible with our having experiences of everyday objects and qualities without those experiences being mediated by awareness of anything like sense data. Compare with the case of belief: one can believe that there is an F without there being an F, but this does not imply that belief is mediated by awareness of anything like sense data. As we will see in chapter 2, the argument from illusion relies on an assumption to the effect that one cannot experience an F unless one is suitably related to an F. Since representationalists' claim that experience is a species of representation is incompatible with this assumption, the representationalist view blocks the argument from illusion. Moreover, it does this without rejecting the immediacy of experience. It therefore seems to dissolve the apparent dilemma posed by the case for sense data.

In addition to the sense-datum theory, there are two main incompatible alternatives to representationalism today: the intrinsic qualia theory and disjunctivism.⁵ According to the former, consciousness is a matter of a state or person instantiating intrinsic qualitative properties of a special kind. The principal motivations for the intrinsic qualia theory will be discussed in chapter 6. To a first approximation, disjunctivists claim that there are two kinds of phenomenal state: those which can occur in veridical conditions only, and those which can occur in other conditions. I will discuss disjunctivism at some length in chapters 7 and 8.

⁵Of course, these are by no means the only alternatives to representationalism. For example, there is also the "the theory of appearing" (Langsam 1997; Alston 1999). See footnote 3 for the distribution of views in philosophy of perception.

Outline

Part I. The theory and its applications

Chapter 1 articulates the kind of representationalism I aim to defend (virtualism).

I begin by highlighting the difficulties which led me to concentrate on the virtualist view. The alternative representational theories all face one of two problems: either they are too vague to yield a good framework for the scientific study of consciousness or they are loaded with questionable commitments. Virtualism minimizes both difficulties, or so I try to show.

Chapter 2 argues that virtualism can play the key explanatory roles widely attributed to representationalism. This chapter also aims to provide an initial motivation for the theory by making these roles explicit.

Part II. A case for virtualism

Chapter 3 puts forward a framework I refer to as *the perceptual conception of sensory experience*. The perceptual conception of sensory experience enables us to describe sensory phenomenal states precisely in everyday language by using perceptual verbs intensionally. This will be important in the chapters which follow.

Chapter 4 builds on the perceptual conception of sensory experience to make a case for a restricted version of virtualism I call *sensory virtualism*. I offer an account of the semantics of perceptual verbs in the process.

Chapter 5 argues that we should not stop at sensory virtualism: if sensory virtualism seems plausible, so should the full virtualist theory. This chapter includes a discussion of objections to representationalism which turn on the role of sensory modalities.

Part III. Alternatives

Chapter 6 considers objections to representationalism which tend to support the case for a qualia theory. I discuss the objections from perspective, blurry vision, double vision, imagery, inverted spectra, and other cases or phenomena which have been held to pose a challenge for representationalism and require that we postulate intrinsic qualia. These matters have been extensively discussed in the literature already; I limit myself to novel responses the framework developed here enables.

Chapter 7 assesses the phenomenological evidence for and against disjunctivism. Disjunctivists have often claimed that disjunctivism is supported by the "naive" conception of experience we acquire through introspection. In this chapter, I argue that virtualism, not disjunctivism, is supported by introspection.

Chapter 8 complements the preceding chapter with a discussion of two central motivations for disjunctivism which are largely independent of the revelations of introspection. I first ask whether disjunctivism is justified by considerations pertaining to skepticism and the role of experience in grounding thoughts about the external world. I then argue that disjunctivism should be rejected on the ground that perceptual experience is massively illusory.