

ψυχη PSYCHE

A Review of Rocco J. Gennaro (ed.)

Higher-Order Theories of Consciousness: An Anthology

Michael Bruno
Department of Philosophy
University of Arizona
Tucson, AZ 85719-0027
U.S.A.
© Michael Bruno
mbruno@u.arizona.edu

PSYCHE 11 (6), October 2005

REVIEW OF: Gennaro, Rocco J. 2004. (ed.) *Higher-Order Theories of Consciousness: An Anthology*, John Benjamins, 371 pp. ISBN: 1588114953.

Introduction

Higher-Order Theories of Consciousness: An Anthology, edited by Rocco J. Gennaro, brings together fourteen new essays exploring the relative merits of and problems with higher-order representation (HOR) theories of consciousness. The anthology is divided into two parts. Part I contains articles by proponents of HOR theories arguing for their favorite version of the theory (Rosenthal, Gennaro, Van Gulick, Carruthers, and Lycan), responding to well-known objections (Gennaro, Van Gulick, and Lycan), and exploring potentially vindicating empirical results (Carruthers, Rolls, and Dienes and Perner). Part II contains critical articles which attempt to press both traditional objections (Seager) as well as new ones (Byrne, Lurz, and Robinson) to HOR theory, to undermine considerations usually put forth in its favor (Byrne, Lurz, Dulany, and Hardcastle), and to offer alternative theories which might appeal to HOR theorists (Lurz, Robinson, and Hill).

This is a fine volume. It should prove to be a useful resource for bringing readers up to date on the latest developments in one of the most promising attempts to provide a naturalistic explanation of consciousness. While the essays vary in terms of their quality, originality, and thematic relevance, there are many excellent articles that stand out. Much of the credit for this is owed to the editor, Rocco Gennaro, who managed to secure papers by many eminent philosophers and scientists. It is obvious that many of the authors shared their papers with one another. This makes for a good deal of conversation between

essays, which helps to unify the anthology. Additionally, Gennaro's introduction does a nice job of introducing readers in a very short space to the basic distinctions and unifying themes found in HOR theories of consciousness. It takes the reader through various versions of the theory, lists some of the major objections to it, and indicates which papers in the anthology discuss which issues.

That said, there were a few questionable editorial decisions. First, it would have been useful for this volume to have a more detailed index. The index itself is rather slim, omits some instances of indexed terms, and so isn't very helpful for readers wishing to track down some particular issue or debate. Second, one would have hoped that the volume would contain a comprehensive and organized bibliography on the subject of HOR theories, but none was to be found. While these points are rather minor, the third point is more significant. The papers by proponents of HOR theories are often solely concerned with adjudicating among various versions of the theory and/or answering objections. Of course, it makes sense that there were many such pieces given that the anthology aimed to provide the latest word on issues surrounding HOR theories. Theory development often requires a lot of working out the kinks, after all. But the anthology should have contained at least one essay whose primary target was to motivate HOR theories generally. Many express puzzlement over how HOR theory is regarded as even a *prima facie* plausible explanation of consciousness (see, e.g., Hardcastle). I think that such theories are in fact well motivated, and I will gesture towards explaining how in the sketch of HOR theory below. Nevertheless, what can be developed in a short review is unlikely to be adequate to convince skeptics that the view merits the serious and careful attention that it does. For this reason the volume would have benefited significantly from a substantial paper dedicated entirely to putting forward the positive case for HOR theory.

In what follows, I will briefly characterize and motivate HOR theory as an explanation of consciousness. After doing so, I'll provide a critical discussion of some of the essays in the volume by focusing on two emerging themes. The first emerges from the debates over various versions of HOR theory. I'll argue that these debates represent a positive direction for HOR theory generally because they facilitate further refinement and specification of the view. Explaining consciousness is no easy task, but there is reason to be optimistic that the further articulation of HOR theory's commitments and the nature of its theoretical apparatus will aid in making progress. Nevertheless, I'll emphasize that caution needs to be taken in these debates in order to avoid terminological disputes. The second theme I'll focus on is whether we should expect disputes over HOR theory to be resolved via conceptual or empirical means. The situation here is delicate. On one hand, there is a need to further explicate the empirical commitments of the view since it seems that certain issues surrounding it ought to be settled empirically. On the other hand, more conceptual care needs to be taken in determining whether and how various empirical results bear on HOR theories.

Sketch of HOR Theory

I see the principal *prima facie* motivations for HOR theory as coming from three distinct sources: conceptual, phenomenological, and methodological. The conceptual motivation for HOR theory hinges on the distinctions between the various notions of 'consciousness'

(explained below) and their interconnections. The phenomenological motivation stems from the fact the HOR theory takes seriously the subjective nature of conscious experience. Given that the higher-order representational state on such views has a content that includes the subject, HOR theory seems well positioned to explain of the centrality of subjectivity to consciousness. The methodological motivation stems from the fact that HOR theories have the resources for providing a naturalistic explanation of consciousness. This last motivation merits a few more words.

HOR theories are naturalistic theories in the following way. They attempt to explain an elusive mental phenomenon, i.e. consciousness, in terms of a more tractable one, i.e. intentionality. It seems likely that intentionality can be explained in purely physicalistic terms, and thus if HOR theory is correct, so can consciousness. Since we have good inductive reasons to believe physicalism, it speaks in favor of HOR theory that it reduces consciousness to something likely to be explained naturalistically. There are some further attractive features of this explanatory strategy that are worth mentioning. First, the strategy connects phenomenology and intentionality. This is independently plausible, as much recent work in the philosophy of mind has pointed out. Second, the strategy of HOR theory allows that it might still have significant explanatory value even if physicalism happens to be false. If, for instance, intentionality cannot be naturalized, HOR theory may still provide the tools for spelling out a non-naturalistic explanation of consciousness. Third, HOR theories take the task of explaining consciousness seriously since they attempt to give a principled account of which mental states are conscious and why they are.

In order to understand what it is that HOR theories attempt to explain and how they attempt to do so, it is necessary to draw a few distinctions. The word ‘consciousness’ is notoriously ambiguous. There are many different things we mean when we say that something or other is conscious. Despite this, there is at least a fairly clear distinction between attributions of consciousness to subjects and attributions of consciousness to mental states. Accordingly, HOR theorists draw a distinction between creature consciousness and state consciousness. Distinguishing further, there seems to be both transitive and intransitive uses of the notion of creature consciousness. Examples of attributions of intransitive creature consciousness include the following: chimpanzees are conscious, fish are not conscious, Mike is not conscious when he is asleep or after he has been hit in the head with a baseball bat. Hence, intransitive creature consciousness is an all-or-nothing notion; a subject (at a time) either has it or not simpliciter. On the other hand, attributions of transitive creature consciousness imply that the subject is conscious *of* something or other. For example, we might say that Mike is conscious of the fact that Canberra has a much cooler climate than Sydney. It’s worth noting that the notion of creature consciousness does not seem inherently mysterious on either of its uses. In fact, we should expect empirical science to shed a great deal of light on when something is creature conscious.

Most discussions of consciousness in philosophy, however, are concerned with state consciousness. We know that of all of subject’s mental states, some are conscious while others are not. Paradigm examples of conscious mental states include pains, emotions, visual perceptions, etc. Paradigm examples of unconscious mental states include Freudian desires, tacit beliefs, early perceptual representations, etc. The question

arises: how do we demarcate in a principled way the conscious from the unconscious mental states?

It is this question specifically which HOR theories seek to answer. The strategy is to account for intransitive state conscious (there do not seem to be clear-cut transitive uses of the notion of state consciousness) in terms of transitive creature consciousness. Conscious states are those that subjects are conscious of themselves as being in. Since the notion of transitive creature consciousness seems to be equivalent to the notion of awareness, we can reformulate this principle thusly: conscious mental states are those states a subject is aware herself as being in.

By taking this as starting assumption, HOR theorists reveal that they are primarily interested in explaining the subjective nature of consciousness. Philosophical discussions of consciousness do not always share this concern. For instance, first-order representationalist theories (such as those propounded by Michael Tye and Fred Dretske) focus on the qualitative aspect of consciousness. That is, their primary concern is to explain the nature of mental qualities such as phenomenal redness or greenness. This is not what HOR theorists attempt to do. Rather, HOR theorists often offer a separate account of mental qualities (in terms of discriminatory abilities, for example), which leaves it open that mental qualities may be present in unconscious as well as conscious mental states. Nevertheless, HOR theorists emphasize that their approach provides the tools for understanding the way in which a subject is conscious of her first-order state. In other words, they attempt to explain why there is something that it is like for a subject to be in some mental state.

We can now see how this works in the following way. Recall that HOR theories begin with the assumption that conscious mental states are states subjects are aware of themselves as being in. This assumption suggests a natural way of explaining consciousness in terms of intentionality. Roughly, HOR theories of consciousness hold that a mental state M is conscious just in case M is the target of some suitable meta-intentional mental state M^* , which has the content that one is in M . In virtue of this relation, we can say that M is a conscious state, but M^* , itself unconscious, determines what it is like for the subject to be in M . Different versions of HOR theory arise because there are different conceptions of the nature of M^* and the relationship between M and M^* . Higher-order thought (HOT) theories maintain that M^* is an actual occurrent thought. Higher-order perception (HOP) theories maintain that M^* is perception-like representational state. Additionally, there are theories that hold that a subject need only be disposed to token M^* , but needn't actually do so. Finally, there are theories that reject the characterization of M and M^* as distinct states, maintaining instead that conscious mental states are complex mental states with both a world-directed and a meta-intentional component.

Varieties of HOR: Development and Terminology

Readers won't be able to help but notice the alphabet soup of positions on offer by different authors in this volume. For example, the theories labeled with the following acronyms are all defended: HOT, HOP, WIV, HOST, HOGS, and SOR. This proliferation of higher-order views by and large represents a positive development. Many of these views have been worked out in order to respond to potential objections and/or to

expand the theory in order provide more detailed explanations of various conscious phenomena. This is exactly the sort of work we'd expect on a theory that has its basic structure already in place. Some may demur that all of this is pointless unless HOR theory has first been demonstrably shown to be on positive ground. While this reaction may have something to it, it is largely unjustified. To a significant extent, the best test of the theory of consciousness (as with any sort of theory) is its explanatory purchase. The nature of conscious experience is nuanced and variegated, and we cannot expect to know ahead of time what all will admit of explanation. It is unreasonable to insist that we must settle all of the foundational issues before getting on with the business of seeing where the theory can go. If HOR theory can offer plausible explanations of genuine phenomena, where other theories cannot, then this speaks greatly in its favor.

The debates over which is the best higher-order theory focuses on two different issues. First, there are debates about the nature of the meta-intentional state M^* . Second, there are debates about the relationship between M and M^* . I'll discuss these in turn. In both cases, the disputes are likely to help make progress in understanding consciousness, yet caution needs to be taken to avoid becoming entrenched in purely terminological debates that lack genuine substance.

Let's start with the dispute between HOT and HOP theorists. In his contribution to the volume, David Rosenthal argues for the superiority of his own actualist HOT theory over pretty much all other versions of HOR theory. Actualist HOT theory holds that the relevant meta-intentional state M^* is an actually tokened thought distinct from M (the state that is rendered conscious by being the target of M^*). In contrast to this, William Lycan, after rebutting a number of traditional objections to HOR theory, defends the HOP position that M is conscious in virtue of being represented by an inner sense or perceptual monitor. Rosenthal has in many places argued against HOP theory, claiming that it would be committed to the relevant higher-order representational states having sensory qualities. Lycan grants that Rosenthal is correct in insisting that this would be an implausible consequence and so accepts that the relevant higher-order states are unlike normal perceptual states on at least this point.

Granting this much of a disanalogy with perception, however, Lycan goes on to offer ten reasons for thinking that the relevant higher-order representational state employed is more perception-like than thought-like. Chief among these is that we can voluntarily control which areas of our phenomenal fields we want to make conscious in a way that we cannot control our thoughts. In response to Rosenthal's suggestion that some thoughts can admit this degree of voluntariness, Lycan insists that the relevant higher-order states are etiologically more akin to perception because they are normally produced by attention. Rosenthal responds to this and most of Lycan's other points in his own paper. In doing so, he explores how higher-order thoughts may be able to employ comparative conceptual resources to account for fineness of grain of experience and he argues that there are no purely recognitional concepts.

At this point, one might wonder whether some of the dispute between HOP and HOT theorists is overblown. It is agreed that the higher-order representational states do not have sensory qualities, unlike paradigmatic perceptions, and that they exhibit a degree of voluntariness uncommon to paradigm thoughts. Reaching consensus on the nature of M^* is important, but it leaves us wondering whether it really matters that we label M^* a

thought or a perception. This risk of losing track of substantive issues by focusing on how we label M might seem all the more salient when we consider that Peter Carruthers classifies his view as both a version of HOR theory and a dispositional version of HOT theory. If he's right to do so, this surely undermines the importance of how we label M*. The underlying nature it must have in order to render another mental state conscious remains, of course, a serious issue.

More of the papers in the volume which develop different versions of HOR theory are concerned with the relationship between M and M* rather than the nature of M* itself. These views all reject the actualist HOT view that M* is an actually tokened thought distinct from M. One paper argues that M* needn't be actually tokened and three papers argue that there must be some constitutive relationship between M and M*.

Peter Carruthers argues for his own dispositional dual content HOT theory in his paper. This is the view that a subject need only be disposed to token M* in order for M to be conscious. He claims that first-order representation (FOR) theories and actualist HOT theories face two similar problems. First, both theories are unable to account for purely recognitional concepts, which Carruthers claims are necessary for defeating anti-materialist arguments. Second, neither theory has a good explanation for why some mental states of a relevant kind, e.g. sensorimotor percepts, are phenomenal while others are not. He argues that a dispositional HOT theory faces neither problem. Rosenthal responds to both of these points in his own paper. Carruthers' theory also relies heavily on a theory of consumer semantics, i.e. the idea that the content of state is in part determined by what is done with it by downstream consumer systems. Rosenthal argues that this commitment reveals that Carruthers' theory is actually one that collapses the distinction between M and M*, which puts it on par with the sorts of views I discuss below.

In the second half of his paper, Rocco Gennaro puts forth his "wide intrinsicity view" (WIV), i.e. the view that conscious states are complex states composed of both higher-order and first-order elements, as an answer to the misrepresentation problem for HOR theories. The misrepresentation problem is one that occurs when M* misrepresents M in some way. For example, when M* represents some object as green, even though M represents it as red. The most salient version of the problem, also called the 'targetless HOTs' problem, arises because of the possibility of a subject being in a meta-intentional state M* without being in any first-order state M. It must be possible for M* to misrepresent the state of a subject in this way. But we may wonder whether the subject is in a conscious mental state at all in such a case. HOR theorists usually claim that M is conscious in virtue of being the target of M* in normal cases, but obviously M cannot be conscious if it does not exist. Furthermore, M* must itself be unconscious. Since M* is present, it seems as though the subject should be in some conscious state, but there do not seem to be any available candidate states. Gennaro's response to this problem is straightforward. If M and M* are intrinsic to the global conscious state, as the WIV holds, then such cases cannot arise.

Robert Van Gulick develops a new view of consciousness in response to several motivations and problems with traditional HOR theories. His view, called higher-order global state (HOGS) theory, incorporates bits from Daniel Dennett's account of consciousness as cerebral celebrity, from Chris Hill's view of introspection as volume

control, and from the idea that the neural correlates of consciousness are distributed. The basic idea behind HOGS is that a state is conscious when it is recruited into a dynamically linked content-sensitive global space. In the second part of his paper, Van Gulick becomes concerned to show that the reflexively meta-intentional or higher-order element is an integral part of his theory and so HOGS gets to qualify as a bona fide HOR theory. The case he makes for this is largely based on phenomenological considerations.

Robert Lurz's paper provides an interesting structuring of the various representational views on consciousness. Lurz lays out a general explanatory scheme and uses it to classify HOR, first-order representational (FOR), and same-order representational (SOR) theories. SOR theory holds that M is conscious in virtue of a subject being aware of its intentional content. Lurz argues both that FOR theories face problems having to do with conscious conative states and nonexistent intentional objects and that HOR theories face a dilemma arising from whether we think that small children can report their own conscious mental states. SOR, he argues, avoids these problems.

Rosenthal addresses these sorts of views on the basis of two considerations. First, he argues that these theories are unable to explain introspection. Second, he argues that they are unable to explain how, given that we distinguish mental states involving different psychological attitudes and that higher-order representations are always assertoric, we can have non-assertoric conscious mental states, e.g. desires.

There are, no doubt, responses that can be made to Rosenthal here. And actualist HOT theorists have their own different solutions to the problems that motivate some to embrace the idea that there is a constitutive relation between M and M*. This process of working out various candidate solutions to putative problems for HOR theory in this way is a good thing. Of course, just as with the issue over how we construe M* itself, these debates need to proceed with an eye towards avoiding terminological disputes. After all, the way in which we individuate mental states isn't always a clear-cut matter. It might be that different individuation criteria deliver different results, so that on one criterion we have two mental states whereas on another we have only one. If we want to be sure that these are debates that matter, we must be clear about what hangs on the different ways of individuating mental states.

Empirical v. Conceptual Issues

Another major theme that emerges in the anthology has to do with whether disputes over HOR theory are likely to be settled on the basis of conceptual considerations or empirical ones. It seems clear that both conceptual and empirical factors are important for determining the overall plausibility of the theory, but there does not seem to be any quick and easy way of classifying how each is important in settling more specific issues. For these reasons, in what follows I won't try to draw any very general moral. Rather, I'll simply review papers grouped together by following topics: (1) the role of parsimony in arguing for a theory of consciousness, (2) purely conceptual attempts to undermine motivations put forth in favor of HOR, (3) attempts to show that empirical results clearly count for or against HOR theory, and (4) the debate over HOR theory's treatment of animal consciousness. Proceeding in this way will hopefully shed light on how conceptual and empirical considerations intermingle in debates over HOR theory.

(1) Appeals to Parsimony

When we have multiple theories that explain some phenomenon equally well, it is reasonable to infer that the simpler theory is likely to be correct. There are two papers in the anthology that make such a move in order to argue against at least some versions of HOR theory.

Christopher Hill's article offers an intriguing positive account of pain according to which pains are bodily disturbances and awareness of pain is akin to forms of perceptual awareness. Hill then compares his view to HOP accounts. Both views allow for an appearance/reality distinction when it comes to pain, but they differ because HOP theorists do not identify pains with bodily disturbances. Rather, they identify pains with certain mental states and make a distinction between conscious and unconscious pains. Hill also argues for the superiority of his view on the grounds that his account is simpler.

William Robinson advances a similar type of argument. Robinson claims that subvocal speech exists and that in such cases (i) the subject has a conscious thought, and (ii) the auditory imagery of which the subvocal speech is composed is also conscious. This raises the question of how these two instances of consciousness are related. Robinson considers four answers and claims that the treatment HOR views offer loses out in terms of parsimony.

It seems that Hill and Robinson both present new challenges to HOR theory. Parsimony is indeed a theoretical virtue and so we should prefer simpler theories, all else equal. The problem with this, however, is that it matters a great what else we are looking at when holding all else equal. It might be that if we confine our attention to the case of subvocal speech or pain, HOR theory is indeed less parsimonious. But if we broaden things and look at a wide range of conscious phenomena, it might end up being more parsimonious because it can wield the same apparatus to explain a wide range of phenomena.

(2) Conceptual attempts to undermine the motivations for HOR theory

There are three papers, which attempt to undermine considerations put forth in favor of HOR theory.

Donelson Dulany's essay begins by elaborating his own theory of consciousness and then argues that philosophical rationales behind HOR theories are problematic. Dulany's own theory of consciousness, which he uses as a basis for his arguments, uses terminology in a way that is deeply disconnected with ordinary philosophical parlance. Since the requisite conceptual connections were not explicitly drawn out, it is difficult to assess the merits of his argument.

In her article, Valerie Gray Hardcastle attempts to diagnose problematic motivations for HOR theories by focusing on a conceptual argument advanced by Rosenthal. She goes through the argument and at each step in the argument she alleges that that step amounts to intuition mongering and repeatedly insists that we shouldn't prejudge what are, at the end of the day, empirical issues. Some of this seems quite fair. For example, her insistence that we really aren't sure whether the same mental state can occur consciously on some occasions yet unconsciously on others. It is likely that

figuring out the best criteria for type individuating mental states will largely be determined by psychology and neuroscience. Oftentimes, however, her criticism borders on mere insistence that one cannot make the theoretical move Rosenthal wishes to make without explaining why it is problematic.

The majority of Alex Byrne's paper is occupied with attempting to show that there is a genuine dispute between first-order and higher-order representational theories of consciousness. His main way of doing this is to elaborate a distinction between experiences and conscious experiences, which seems to be found in the work of Carruthers, Lycan, and Rosenthal. For each of these theorists there ends up being a special way of understanding the locution 'what it's like for so-and-so to experience ϕ '. Byrne argues that considerations about the semantics of 'what it's like' talk cannot be used as an argument against first-order theories. Whether or not Byrne succeeds, this is a great example of making important and useful conceptual distinctions, which promise to illuminate and advance further discussion of these issues.

(3) The bearing of empirical results on HOR theory

Results of empirical investigations will clearly play an important role in terms of our long-term assessment of the vitality of HOR theories. Many papers in the anthology bring various empirical results into their arguments in one way or another. My focus here, however, will be rather narrow. I'll consider two articles, which attempt to show that HOR theory provides a plausible explanation for some empirical phenomena.

Edmund Rolls argues that all emotions are states that induce various rewards and punishments. He goes on to advance a version of HOT theory he calls higher-order syntactic thought (HOST) theory and argues that higher-order syntactic thoughts provide an explanation of an adaptive solution to the 'credit assignment' problem, i.e. the problem of how the brain monitors what is responsible for good or bad outcomes to various stimuli.

Zoltán Dienes and Josef Perner claim that HOR theories provide useful theoretical constructs for distinguishing cases of knowing from cases of consciously knowing. This is because HOR theories help to vindicate 'subjective measures' of consciousness such as verbal reports. The main subjective criterion they discuss is the 'zero correlation criterion', which measures the correlation between confidence and accuracy of subjects' judgments. The zero correlation criterion has previously been used to study implicit learning and subliminal perception. And Dienes and Perner think that it may be used to infer the existence of a conscious mental state.

The intended force of the arguments in both of these papers is somewhat unclear. It could be that they simply aim at showing that HOR theories are consistent with some methods employed in empirical research. Or perhaps the aim is to show that HOR theories help to vindicate those methods. What they don't seem to do is show that there is something that HOR theories can explain, which other theories of consciousness cannot.

(4) Animal consciousness and HOR theories

One specific issue surrounding HOR theories concerns the nature of animal consciousness. It has been alleged that HOT theories in particular face a dilemma over

what to say about nonhuman animals. Given that thoughts are composed of concepts, HOT theories are committed to saying either that nonhuman animals are bona-fide concept possessors or that they are not conscious. But both of these consequences are problematic. While some proponents of HOT theory such as Peter Carruthers are willing to accept that nonhuman animals are not conscious, most are not. Furthermore, as William Seager discusses in his paper, there is good evidence that nonhuman animals are in fact creature conscious. Seager discusses various tests for consciousness, all of which attest to animals being conscious. Of course, HOR theories are primarily interested in state consciousness, not creature consciousness. Nevertheless it seems that there cannot be conscious creatures that never have conscious mental states.

If we grant that nonhuman animals are conscious, then the issue becomes whether it is more plausible to think that animals have concepts in the relevant sense or that HOT theory is false. In the first half of his paper, Rocco Gennaro argues for the first claim by insisting that the capacity for higher-order thought need not be very sophisticated. Unfortunately, he mainly argues by posing rhetorical questions as to why animals might not be employing some less sophisticated concepts. On the other hand, William Seager argues that the empirical evidence suggests that animals are not capable of attributing mental states to other animals. Seager takes this, combined with the assumption that animals do not possess cognitive abilities they never deploy in behavior, to support the claim that animals lack the conceptual abilities requisite for higher-order thoughts.

I'm generally sympathetic with Gennaro's position on this issue, but Seager's methodology is clearly to be preferred. In order to advance this debate, we need to figure out a positive account of concepts and concept possession, which doesn't prejudge the issue. Once such an account is in place, it becomes an empirical issue whether nonhuman animals have what it takes to have higher-order thoughts. Seager has convincingly argued that one way of cashing out concept possession points in the negative direction. Perhaps a different account of concept possession, which does not require a capacity to attribute of mental states to others, is in the offing and would do the work necessary for the HOT theorist. For all that has been said in the volume, the issue remains unresolved.

In sum, both empirical and conceptual factors play a large role in settling disputes over HOR theories of consciousness. The general lesson to take away from the discussion of this theme is that more care needs to be taken in the places where these factors intersect.

Conclusion

Rocco Gennaro has put together a wonderful collection, which should be read and studied by anyone interested in seriously coming to grips with and providing an explanation of consciousness. HOR theories offer a promising account of how and why our conscious mental states are conscious. The tools the theory brings to the task are continually being sharpened, altered, and put to different uses by its proponents. At the same time, its critics persist in pressing new and challenging objections. All this points to the vibrancy of these debates. Arriving at a fully satisfactory theory of consciousness may still be

somewhat off, but progress is being made, and the essays in *Higher-Order Theories of Consciousness: An Anthology* are a testament to that.¹

¹ I'd like to thank Tim Bayne, Uriah Kriegel, David Pereplyotchik, and David Rosenthal for very helpful comments on an earlier draft of this review.