

Consciousness and its function

David M. Rosenthal*

The City University of New York, Graduate Center, Philosophy and Cognitive Science, 365 Fifth Avenue, New York, NY 10016-4309, United States

Received 2 March 2007; received in revised form 18 October 2007; accepted 15 November 2007

Available online 22 November 2007

Abstract

It is plain that an individual's being conscious and an individual's being conscious of various things are both crucial for successful functioning. But it is far less clear how, if at all, it is also useful for a person's psychological states to occur consciously, as against those states occurring but without being conscious. Restricting attention to cognitive and desiderative states, a number of suggestions are current about how the consciousness of those states may be useful. It has been held that such consciousness enhances processes of rational thought and planning, intentional action, executive function, and the correction of complex reasoning. I examine these and related proposals in the light of various empirical findings and theoretical considerations and conclude that the consciousness of cognitive and desiderative states is unlikely to be useful in these or related ways. This undermines a reliance on evolutionary selection pressures in explaining why such states so often occur consciously in humans. I propose an alternative explanation, on which cognitive and desiderative states come to be conscious as a result of other highly useful psychological developments, some involving language. But on this explanation the consciousness of these states itself adds no significant function to that of those other developments.

© 2007 Elsevier Ltd. All rights reserved.

Keywords: Theory of consciousness; Rationality; Intentional action; Executive function; Higher order thoughts; Function of consciousness

1. The Issue about Function

It is widely held that we understand something only when we can explain it, and explaining a natural phenomenon typically if not always means locating it in its distinctive causal nexus. When those phenomena are biological, moreover, locating a phenomenon often means specifying what function it has for the relevant organism. And that is usually a matter of isolating something about the phenomenon that tends to benefit organisms of that type or confer on them some adaptive or reproductive advantage.

This is especially so with mental phenomena and their associated brain processes. We expect that a satisfactory explanation of perceiving, thinking, and other mental occurrences will involve coming to know how those processes and the brain events that subserve them function to benefit the organism.

These functions are often obvious. Perceiving enables organisms to interact successfully with its environment, and thinking enables it to figure out what to do and what to avoid in

novel or complex circumstances. And the various emotions help organisms react appropriately to characteristic beneficial and dangerous situations. More specific functions for perceiving, thinking, and the emotions are easy to come by, often simply by reflecting on our folk-psychological understanding of the relevant psychological states.

Psychological states such as thinking, perceiving, and feeling often occur consciously, but by no means always. Masked priming (Marcel, 1983a, 1983b) reveals the occurrence in normals of subliminal perception, as does blindsight in individuals with lesions in primary visual cortex (Covey & Stoerig, 1995; Weiskrantz, 1986, 1997). And it is widely acknowledged that thoughts, desires, and emotions all sometimes occur without being conscious.

Because all these states occur in both conscious and non-conscious forms, an additional question about function arises: What, if any, function do conscious versions of these states have that nonconscious versions lack? It is plain that we will not fully understand perceiving, desiring, and thinking without knowing what functions they serve. So it may also seem obvious that we also cannot fully understand the consciousness of these states without understanding the function of such consciousness. Understanding consciousness, it seems, requires knowing

* Tel.: +1 212 222 6186; fax: +1 212 817 1530.
E-mail address: davidrosenthal@nyu.edu.

how the consciousness of psychological states contributes to the well-being or procreation of the organism. As Humphrey (2002, p. 70) puts it, “when a modern Darwinian biologist asks [what difference consciousness makes, the] Darwinian’s answer has to be that it has evolved because and only because it is serving some kind of useful biological function.” Humphrey sees the appeal to function as so compelling that he sees no alternative but a flat denial that consciousness can be explained and understood at all.

It is important to distinguish the function that thinking, perceiving, and desiring have independently of whether they are conscious from the function, if any, that is added when these states occur consciously. In what follows, I often refer to this added function as the function of consciousness *tout court*. The present discussion focuses on that specific function.

The function of consciousness just described concerns the consciousness of psychological states. But there is something else that is often referred to as the function of consciousness. Not only do we distinguish between psychological states’ being conscious or not; we also distinguish the conscious from the unconscious condition of individual organisms, themselves. A creature is conscious if it is awake and responsive to sensory input, as against being asleep, anesthetized, or otherwise unconscious. There is no doubt that an organism’s being conscious has many important functions; its being conscious enables it to interact with its environment in ways that greatly enhance its well-being and survival.

But even when an organism is fully conscious, many of its psychological states may fail to be conscious states. Fully conscious humans often have many thoughts and desires that are not conscious states, and sometimes have subliminal perceptions, which are also not conscious. Doubtless the same holds for other organisms as well. So we cannot infer from the function of an organism’s being conscious to a function of the consciousness of its psychological states.

The difference between these two functions is sometimes overlooked (e.g., Merker, 2005; Morsella, 2005; Rossano, 2003), perhaps because it is assumed that the psychological states of an awake organism are invariably conscious, or at least psychological states of a particular type. But since that is not the case, it is crucial to distinguish these two questions about function.

There is a second issue that is sometimes confused with that about the function that accrues specifically to psychological states’ being conscious. Thinking, perceiving, and the emotions all have significant functions independently of whether they are conscious. But some have assumed that psychological states never occur without being conscious (e.g., Nagel, 1974), or at least that qualitative psychological states never do (Block, 1995, 2001). If one assumes that, one will not distinguish between the function these states have independently of being conscious and the function that is added by their being conscious. One would see the function of consciousness as simply the function that conscious states have, ignoring the question of what function might be added specifically by those states’ being conscious.

But not all psychological states are conscious, and there is little reason to think that only conscious psychological states tend

to benefit the organism in some significant way. So we must distinguish the function that is due specifically to a psychological state’s being conscious from the function that that state has independently of its being conscious.

Even for states that are conscious, we must distinguish the function that is specifically due to its being conscious from the function that results from others of its psychological properties. Indeed, this distinction would be crucial even if all psychological states were conscious. Psychological states have many properties that contribute to their causal powers, and so may result in the states’ having some distinctive function. Being conscious may be one such property, but it is by no means the only one. Psychological states also differ in respect of representational content and other psychological properties. So even for conscious states, we must distinguish that aspect of their function due specifically to their being conscious from those aspects which are due to other psychological features of the states.

This distinction also bears on the so-called neural correlate of consciousness. Theorists sometimes identify the neural correlates of particular psychological states with the neural correlates of those states themselves, on the assumption that the states cannot occur without being conscious; an example is Block (2005, 2007), in discussing qualitative psychological states, which he calls phenomenal consciousness (cf. Rosenthal, 2002a). But it may be that the neural events that subserve the qualitative character of such states are distinct from the neural events that subserve those states’ being conscious. Indeed, there is evidence in some cases at least that those neural states are distinct; so they be in all case (for a useful discussion, see Lau, *in press*).

In what follows I restrict attention to the function, if any, that is due specifically to psychological states’ being conscious. I also restrict attention to the consciousness of thinking, planning, intending, and related psychological states, often called *intentional states*. So I do not consider here the function that may result from the consciousness of perceiving or sensing.

It is worth noting that the benefits that being conscious is usually thought to add to qualitative states are different from the benefits that being conscious is typically thought to add to intentional states, such as thinking and desiring. Thus, Gray (2004, chs. 7–8) has argued that the consciousness of perceiving enables late error detection; if so, that function would be distinctive to perceiving, as against cognitive or desiderative intentional states. But late error detection may be possible even when perceiving is not conscious. There is evidence that change detection occurs even when the relevant perceptual states are not conscious (Fernandez-Duque & Thornton, 2000, 2003; Laloyaux, Destrebecqz, & Cleeremans, 2003; Silverman & Mack, 2006); so late error detection may as well.

Weiskrantz (1997, ch. 7) has argued that it is beneficial for perceiving to occur consciously because thinking flexibly about perceptual content occurs only when the perceiving is conscious. Weiskrantz appeals mainly to the failure of flexible thinking to recruit the nonconscious perceptual contents that occur in such clinical disorders as prosopagnosia, blindsight, and amnesia. But because this important proposal concerns the possible function of perceiving’s being conscious, I cannot do justice to it here.

In what follows I argue that the consciousness of intentional states has no significant function. I consider a number of the significant functions that researchers have suggested result specifically from the consciousness of thinking, desiring, and intending, and argue that none of these suggestions is defensible. But it is possible, I urge, to explain the widespread occurrence of psychological states independent of any appeal to some significant function of psychological states' being conscious. In a final section, I offer such an explanation.

Lack of function does not imply that the consciousness of these states has no causal impact on other psychological processes, but that causal impact is too small, varied, or neutral in respect of benefit to the organism to sustain any significant function. So my conclusion about function for does not imply epiphenomenalism. Consciousness would be epiphenomenal only if it had no causal impact at all on psychological functioning (Huxley, 1896; cf. James, 1890, p. 129, who adapted the medical term 'epiphenomenalism' for this use), not if that impact is simply too varied and insignificant to yield stable beneficial effects. (I am grateful to Mortimer Mishkin, personal correspondence, for pressing this issue.)

2. Consciousness and rationality

It is frequently suggested that when psychological states such as thinking and desiring are conscious, they have some special tie to rationality, or to intentional action or executive function. On this suggestion, the conscious occurrence of the states enables rationality in inference and making plans or control of action that is not possible when those states fail to be conscious, or at least that those states' being conscious enhances such rationality or control. And this, it is often thought, might help explain what difference an intentional state's being conscious makes in the psychological functioning of humans and other animals, and so help isolate a function that is due specifically to such states' being conscious.

I consider in this section whether consciousness does facilitate rationality, raising doubts about such facilitation. In the subsequent two sections I do the same for the possible facilitation by consciousness of intentional action and executive control, respectively. I conclude that there is little significant function that results specifically from the consciousness of intentional states.

Because the consciousness of thinking and perceiving is subjectively central in our lives, it is intuitively inviting to assume that their being conscious has some significant function. But that commonsense assumption is usually not specifically about the function that results specifically from those states' occurring consciously, but rather the overall function of the states together with their being conscious. So it does not isolate some function due specifically to the consciousness of those states, as against their representational and other psychological properties.

Still, some discussions that respect that distinction do seek to isolate a specific function for consciousness. Thus, Armstrong (1968) has argued that problem solving requires consciously considering several responses to the problem and consciously choosing which of them best suits one's goals. He concludes

that "any animal that solves problems mentally must" be aware of the relevant mental states (1968, p. 163).

Shoemaker (1996) advances a more elaborate argument to similar effect. Actions are caused by beliefs and desires, and beliefs and desires have representational content. An action is rational, Shoemaker urges, if the beliefs and desires that cause the action jointly represent that action as rational. By parity of reasoning, those beliefs and desires are themselves rational only if there are second-order beliefs and desires about the first-order beliefs and desires. And Shoemaker argues that these second-order beliefs not only must figure in causing the first-order beliefs and desires, but also must represent those first-order states as being rational. One will have second-order desires that one's planning and inferences be rational, consistent, and otherwise optimal, and those second-order desires will interact with second-order beliefs in determining which first-order beliefs and desires to have.

One feature of rational thinking is the adjusting of plans and inferences to make them more rational. Shoemaker sees an especially pressing need for second-order beliefs and desires in that context. The "one thing [that first-order beliefs and desires] do not rationalize is changes in themselves" (1996, p. 33); so second-order beliefs and desires are needed to do that. Shoemaker does not himself claim that these second-order beliefs result in the consciousness of the first-order states they are about. But he does hold that the second-order beliefs give one knowledge of one's first-order states, and the beliefs and desires one knows one has are normally conscious. So the process of rationalizing one's thoughts and desires results in those thoughts and desires' being conscious.

A tie between consciousness and rationality also underlies Block's (1995) influential notion of access consciousness, which is a type of consciousness that specifically figures in reasoning and rationality. A state is access conscious, Block maintains, if its content is "poised for use as a premise in reasoning, . . . [and] for [the] rational control of action and . . . speech" (1995, p. 231; cf. Block, 2001).

The idea that consciousness has some essential tie to rationality also inspires the well-known global-workspace theories of consciousness (Baars, 1988, 1997; Dehaene & Naccache, 2001; Tononi, 2004; Van Gulick, 2004). On that view, a state is conscious just in case it has global ties to many other cognitive states, ties that enhance the rationality of many psychological processes.

But there is reason to doubt that any such essential connection actually holds between the consciousness and rationality of thinking. Rational thinking is not always conscious, and behavior is often rational even when the mental processes that lead to it are not conscious. We sometimes rationally solve problems and work out plans even when we are not thinking consciously about those problems or plans. Intuitively it seems that rational solutions "just come to us"; that is our introspective impression. The best explanation is that these solutions actually come to us as a result of thinking that is not conscious.

One might think, as Shoemaker urges, that when we correct or adjust reasoning, at least the thoughts and desires that figure in that reasoning would have to be conscious. But such

adjusting often takes place without any conscious monitoring, and indeed even when the thoughts that lead us the adjustment are not themselves conscious. Indeed, it is relatively rare that we adjust our reasoning by consciously rehearsing the steps, and when we do, that process is typically awkward, slow, and sometime inefficient.

There is some experimental confirmation of these commonsense observations in recent findings of Dijksterhuis, Bos, Nordgren, & van Baaren (2006; cf. Bargh, 2002; Dijksterhuis, 2004) that deliberating is often more successful when it is not conscious. Dijksterhuis et al. tested rationality in consumer choices, both in the laboratory setting and outside it. Their striking finding was that, whereas conscious deliberation yields more satisfactory choices with simple choices, complex decisions are more rational when the thinking that led to them was not conscious. These findings fit well with the observation that problem solving and decision making are often most rational when they are due to thinking that is not conscious, and that conscious monitoring of those thought processes frequently results in their being relatively awkward and inefficient.

The classic results of Nisbett and Wilson (1977; see also Wilson, Dunn, Kraft, & Lisle, 1989; Wilson, Hodges, & LaFleur, 1995; Wilson & Schooler, 1991) and related findings provide additional, albeit indirect reason for a dissociation between the rationality of thinking and its being conscious. Nisbett and Wilson's study focused not only on cases in which subjects confabulate the causes of their being in particular cognitive states, but also on cases in which they confabulate stories about what states they are actually in. Our consciousness of the thinking that guides our preferences is often confabulatory, resulting from a desire to have the things we do make sense from our own perspective or from that of others. Since the thoughts we are conscious of ourselves as having in these cases differ from those we actually have, consciousness is independent of the rational thinking that guides choices.

The foregoing rehearse various commonsense and experimental reasons to doubt that the consciousness of thoughts and desires enhances their rationality. But there are also theoretical reasons to expect that consciousness is independent of rationality. The rationality of thoughts and desires is a matter of their intentional content. Rational thoughts and desires have intentional content that reflects rational connections among those states, and with behavior and the environment as the organism perceives it. These connections are both necessary and sufficient for the relevant thoughts and desires states to be rational. They are what rationality consists in.

Thoughts and desires occur without being conscious. And since intentional content is a necessary feature of thoughts and desires, these states exhibit intentional content independently of being conscious. Thoughts and desires also interact causally, both with one another and with stimuli and behavior, in ways that reflect their intentional content. A thought about a favorite food may cause a desire to eat some, and thoughts about where to get it. Intentional content inevitably tracks the causal connections that thoughts and desires have with one another and with behavior and perception. Such tracking is required for thoughts

and desires to function as they do, independently of whether they are conscious.

Indeed, on many theories of intentional content, a state's content actually consists at least partly in the propensity those states have to cause and to be caused by other psychological states, behavior, and perceptual input (e.g., Block, 1987; Field, 1977; Harman, 1987, 1982; Peacocke, 1992). On that view, a state has its particular content only if it has suitable causal connections, actual and potential, with other relevant thoughts, perceptions, and desires. Thus, a thought might have the content that it is raining, for example, in virtue of its causal connections with possible perceptions of rain and possible desires not to get wet.

Even apart from such theories, however, thoughts and desires can play their role in the psychological economy of creatures only if the causal interactions among them reflect connections among their intentional contents. And since rationality is a matter of connections among intentional contents, causal connections among thoughts and desires must to some extent reflect rational connections among their intentional contents. The tie between causal connections and intentional content by itself ensures that thoughts and desires will be largely rational. And since thoughts and desires have intentional content independently of being conscious, rational connections among them will tend to occur independently of whether they are conscious.

Because rationality is a function of intentional content, not of consciousness, even when thinking is rational, we have no reason to think that its being conscious contributes much, if anything, to its being rational. So even when thoughts and desires are conscious, the rational ties among them are due mainly or even entirely to the propensities those states have to cause and be caused by other thought and desires, not to their being conscious.

This conclusion may seem to offend against pretheoretic intuition. It seems subjectively that we sometimes control our thinking and planning by taking note of what thoughts and desires we have and figuring out on that basis what to think next. As Armstrong (1968, p. 163) puts it, "if our mind is to work purposively . . . we must have awareness of our minds." Only by being aware of one's "current mental state . . . can we adjust mental behaviour to mental circumstance"; "[o]nly if we do become so aware we will we know what to do [i.e., what to think] next" (p. 327).

The question of how consciousness does figures in the adjusting and fine tuning of sequences of thoughts will come up again in Section 4. For now it is enough to note that Armstrong's claims and the commonsense ideas that underlie them appear to rely on an analogy between the way we think about the objects we perceive and the way we operate psychologically on our thoughts and desires. This echoes Armstrong's (1968, pp. 95–99, 323–348) view that the process whereby we are conscious of our psychological states is itself quasi-perceptual.

But the analogy is at best misleading. We seldom if ever operate psychologically on our own thoughts and desires as we think about the objects of our perceptions. The sequences of thoughts and desires that occur in planning and problem solving rely mainly, if not entirely, on the tendency of each thought and desire to cause others in ways that reflect rational connections. When thinking and planning are conscious, we note consciously

the rational ties that constituent thoughts and desires have with one another. But we have no independent reason to think that the consciousness of those thoughts and desires figures in the direction our thinking and planning takes.

3. Consciousness and intentional action

Just as it is intuitively appealing to see a connection between consciousness and rationality, so it is inviting to hold that consciousness is essential to intentional action. This view gets support from introspection. We are introspectively aware of our thoughts as being rational only when those thoughts are conscious, and introspectively aware of our actions as being intentional only when the relevant volitions and desires are conscious.

But introspection itself cannot tell us whether these connections actually hold. Introspection has no access to thinking that fails to be conscious; so it cannot reveal whether rational thinking or volitions occur without being conscious. And the considerations raised in the foregoing section also apply to the question of whether actions are intentional only if the relevant volitions are conscious.

An action is intentional if it is initiated by a volition to perform that action. And volitions tend to cause the particular actions they do in virtue of the intentional content of those volitions. Since volitions, like cognitive states, can occur without being conscious, a volition having intentional content is independent of its being conscious. So an action will be intentional even when the volition that causes it fails to be conscious. Though we are aware of our own actions as being intentional only when the relevant volitions are conscious, volitions need not be conscious for the actions to be intentional.

These considerations apply equally to an action's being deliberate or voluntary. An action is deliberate if its occurrence results from the agent's prior deliberation, and voluntary if it results from or at least conforms to the agent's prior desires. And as Dijksterhuis (2004; cf. Bargh, 2002; Dijksterhuis et al., 2006) shows, the relevant deliberations need not be conscious; Nisbett and Wilson (1977; cf. Wilson et al., 1989, 1995) show the same for antecedent desires.

Jacoby's (1991; Debner & Jacoby 1994; Jacoby, Toth, & Yonelinas 1993; Jacoby, Toth, Yonelinas, & Debner 1994) use of an exclusion task is sometimes held to support a connection between consciousness and intentional action. Subjects are presented with a word, say, 'reason', then presented with a word stem, say 'rea-' and asked to complete that stem with any word other than that which was presented. When the word is presented in a way that allows subjects to see it consciously, subjects mainly succeed in following the instruction. But when the word is presented subliminally, using masking, divided attention, or brief presentation, subjects report seeing no word, but tend then to complete the stem with the word that was presented.

Because subjects intentionally exclude a word only when they consciously see it, it may be tempting to infer a tie between intentional action and consciousness. But these results do not show that intentional action must result from conscious volitions, or even that it is facilitated by the volitions' being conscious. Sub-

jects are instructed not to complete the stem with a word they just saw. When that word is subliminally presented, subjects do not see the word consciously. Since they are not conscious of having seen the word, they consciously think that they did not see it. So completing the stem with that word, which the subliminal presentation primes them to do, does not consciously seem to them to be in violation of the instructions. What matters is how subjects consciously process the instructions, not whether the relevant volitions are conscious.

Acting intentionally does require perceptual awareness of environmental objects. That perception can be subliminal, and so can occur without conscious awareness, as in blindsight and masked priming. Exclusion-task subjects do not, however, make use of their subliminal perception to complete the stem with some word other than the subliminally seen target. But that does not show that intentional action requires that perception of relevant environmental conditions be conscious. Since subjects' volitions derive solely from consciously processed instructions to exclude a seen word, we should expect that only consciously seen words will figure in the execution of that intention.

Intentional actions can result from volitions that are not conscious. But even when the volitions are conscious, their being conscious itself arguably plays no role in initiating those actions. There is evidence for this in the findings of Libet (1985; Libet, Gleason, Wright, & Pearl 1983), replicated and refined by Haggard (1999; Haggard & Eimer, 1999; Haggard, Newman, & Magno, 1999). These results show that, when subjects consciously decide to perform a simple action, the neural event (readiness potential) that initiates the action occurs significantly prior to the conscious volition. (Haggard & Eimer, 1999 identify the readiness potential in a slightly different way.) The conscious volition that seems relevant to an action does not figure in initiating the action.

The best interpretation of these results requires distinguishing between a volition's occurring and its being conscious (Rosenthal, 2002b). Subjects report when their volition occurs, and they can that only when the volition is conscious. We can conclude that the volition becomes conscious only after it initially occurs.

Once we see that a volition can occur without being conscious and only subsequently come to be conscious, we can identify the readiness potential isolated by Libet or Haggard with the volition, independent of its being conscious. The volition does initiate the action, but prior to its becoming conscious. The best interpretation of the Libet–Haggard results involve a lag between the initial occurrence of the volition and its becoming conscious.

It is natural to speculate that such a lag holds also for the cognitive states that constitute thought processes generally, that each thought and other cognitive state initially occurs without being conscious, has its causal impact, and only after that comes to be conscious. Each individual thought in a chain of rational inference would cause a subsequent thought in that chain prior to, and hence independently of, that thought's becoming conscious. This may be more difficult to test, however, since the neural location of the events makes them less susceptible to precise timing.

Libet (1985; Libet et al., 1983) finds that, although neural initiation of actions we consciously decide on occurs prior to our volitions' being conscious, subjects still retain some conscious ability to call the action off, or "veto" it. Conscious decisions to call off the action can occur after the nonconscious neural initiation and thereby prevent that nonconscious initiation from leading to action. Libet urges that this ability to veto mitigates the intuitive conflict of nonconscious neural initiation with our commonsense idea that voluntary actions result from conscious volitions.

That conflict results from the initiating neural event's occurring prior to any volition's being conscious. So the ability to veto will soften the conflict only if the conscious decision to veto is not itself preceded by some nonconscious neural event that results in calling off the action. Libet offers no evidence that there is no such earlier neural event. The Libet–Haggard findings suggest identifying neural events that initiate actions with nonconscious volitions that subsequently become conscious. By the same token, a neural event (cf. Brass & Haggard, 2007) may operate to veto before becoming conscious, and only become conscious thereafter. The psychological states that initiate actions or allow us to call them off would both occur and have their causal impact prior to becoming conscious. Absent some reason to think that such an antecedent nonconscious neural vetoing event does not occur, the ability to veto cannot help dispel the intuitive conflict with common sense.

The foregoing arguments do not preclude there being some distinctive types of behavior that occur in humans or other creatures only when the relevant volitions or desires are conscious. Almost certainly there are. But that by itself does not show that the consciousness of those motivational states plays any role in making those behaviors possible. The behaviors and the consciousness of the relevant motivational states may well be jointly caused; the very factors that result in those behaviors may also cause the volitions and desires to be conscious.

Indeed, the Libet–Haggard findings suggest that this may well be so. The actions relevant there, for example, moving a finger, often occur without any conscious volition. And since the volitions in the Libet–Haggard experiments come to be conscious only after they cause the relevant behavior, that may well also happen with behaviors that tend not to occur in the absence of conscious volitions.

The Libet–Haggard results shed light on other proposals about the possible function of volitions' being conscious. Wegner (2003; cf. Wegner, 2002) has urged that the occurrence of a conscious volition provides information, albeit fallible, that particular pieces of behavior are due to one's own agency. He concludes that conscious volition has the function of providing that information. Wegner assumes that the conscious volition does not itself cause the corresponding action, but that a third event causes both the conscious volition and the associated action. So he sees the consciousness of volitions as the mind's trick in informing us about our authorship of behavior.

But Wegner does not consider the possibility that each volition first occurs without being conscious, and that in that initial, nonconscious condition it causes both the corresponding action and the subsequent consciousness of the volition itself. The

Libet–Haggard findings suggest that alternative picture, since they show that actions are caused by brain events that we can identify with nonconscious volitions, and that conscious volitions occur only subsequently.

There is in any case reason to doubt Wegner's view that the function of volitions' being conscious is to provide information about the authorship of particular pieces of behavior, since nonconscious volitions would also provide that information. The information would not then be conscious, but it would still play a role in our psychological lives, even if not consciously. Wegner's hypothesis about the function of volitions' being conscious just pushes the question one step back, to the question why it is useful for information about the authorship of one's actions itself to be conscious.

Humphrey (2002) has argued that the consciousness of psychological states enables us to rationalize our behavior, and by extension to rationalize the behavior of others. We understand behavior as rational, he argues, by appeal to rationalizing beliefs and desires, and one can invoke those rationalizing beliefs and desires in one's own case only if they are conscious states (cf. Carruthers, 2000, ch. 8, esp. p. 225; Humphrey, 1980).

But it is arguable that Humphrey has the order of things reversed. The process of rationalizing one's own actions doubtless leads to one's making reference in thought to the beliefs and desires that cause those actions. So rationalizing one's actions leads to one's having thoughts about the relevant beliefs and desires in virtue of which they come to be conscious. It is the rationalizing that leads to one's being conscious of oneself as having those beliefs and desires, not the consciousness of those states that makes that rationalizing possible.

4. Consciousness, higher order theories, and executive function

The Libet–Haggard findings provide evidence that psychological states initially occur without being conscious and become conscious, if at all, only later. This in turn lends support to an increasingly prevalent type of theory about what it is for a psychological state to be conscious.

On these theories, known as higher order theories, a psychological state's being conscious consists in one's being conscious of that state in some suitable way. I will call this basic idea, which is shared by all higher order theories of consciousness, the *Transitivity Principle (TP)* (Rosenthal, 1997), since it seeks to explain what it is for a state to be conscious by appeal to one's being conscious of that state. TP is compelling, since a state of which one is in no way aware does not intuitively count as conscious.

Though higher order theories of consciousness all endorse TP, TP does not itself specify the way one must be aware of a state for that state to be conscious. The various higher order theories differ mainly in advancing different views about how TP is implemented.

I have argued (most recently in Rosenthal, 2005) that TP is implemented by having thoughts about one's psychological states that are roughly concurrent with, but distinct from, those states. A psychological state is conscious, on this theory, if one

has a thought, distinct from the state itself, to the effect that one is in that state. Because these thoughts are about other psychological states, it is convenient to call them *higher order thoughts* (HOTs). These thoughts are higher order only in respect of having intentional content that is about other psychological states. They are not higher order, as Weiskrantz (1997, p. 72) notes, in being somehow more abstract in content than lower order thoughts. Nor, as Weiskrantz (1997, pp. 72–73) notes, is there reason to expect that disorders that affect consciousness will all be due to a single type of breakdown in the hierarchy of thoughts.

We are seldom conscious of having any such HOTs. But that is to be expected. A HOT will itself be conscious only if one also has a third-order thought that one has the second-order thought in question. We can assume that such third-order thoughts are rare, and so HOTs are seldom conscious. The reason to invoke HOTs is not that introspection reveals them, which it rarely does; HOTs are states posited to help explain the data of consciousness. There is, however, some evidence that states are conscious when, but only when, a distinct neural state occurs in mid-dorsolateral prefrontal cortex (area 46) (Lau & Passingham, 2006), and it is reasonable to explore identifying these neural occurrences with the posited HOTs.

It may sometime happen that one consciously infers that one is in a particular psychological state. One might infer from consciously noting something about one's own behavior, from what others tell one, or from good theoretical evidence that one is in the state. States that one infers consciously or from conscious observation of oneself are not, however, conscious states. A state is conscious only if one would be aware of it independently of conscious inference or self-observation. So the HOTs in virtue of which one psychological states are conscious, cannot rely on conscious inference or observation. This captures the common-sense idea that we are aware of conscious states in a way that is intuitively unmediated.

The HOT theory fits well with the Libet–Haggard results. Since the HOT in virtue of which a psychological state is conscious is distinct from that state, we can expect that each volition would initially occur without being conscious and there would be a slight lag before a HOT occurs and it becomes conscious. The theory also fits with the foregoing explanation of the Jacoby exclusion results. Subjects exclude only words they see consciously because their conscious intention to follow instructions interacts mainly with HOTs that they have seen those words.

TP specifies that a psychological state's being conscious consists in one's being in some suitable way conscious of that state. It therefore predicts the widely acknowledged tie between a state's being conscious and its being reportable, on which much experimental work relies. Reporting a state is simply asserting in some way that one is in that state; so one can report something if, but only if, one is aware of it. Given that a state's being conscious consists in one's being in some way conscious of it, a state's being conscious coincides with its being reportable.

This conclusion fits with Weiskrantz's (1997, p. 76; cf. p. 167) observation that "it is the very . . . ability to make a commentary of any particular event that gives rise to awareness" (cf. Weiskrantz, 1995, 1998, 2001). Making a commentary on an event is reporting its presence; Weiskrantz's commen-

tary methodology reflects the principle that consciousness and reportability coincide, at least absent independently established blocks to reporting.

Since TP underwrites the reportability test for a state's being conscious as well as Weiskrantz's related commentary methodology, their success provides support for TP. But they also suggest that TP is implemented by HOTs. Reports are assertions, and like other assertions, each report expresses a thought whose intentional content matches the meaning of the report; a report that it is raining, for example, expresses one's thought that it is raining. So a report or commentary that one is in a particular psychological state expresses one's thought that one is in that state.

Reports of one own psychological states, moreover, typically occur independently of conscious inference or self-observation. So the thoughts those reports express are the very HOTs that the theory posits. Both the reportability test and Weiskrantz's commentary methodology support the theory that a state is conscious in virtue of one's having a HOT about that state. (cf. Weiskrantz, 1997, pp. 71–75, 167 and Weiskrantz, 2001, p. 174.)

The HOT theory also fits well with Weiskrantz's (1997, ch. 7) observation that flexible thinking does not recruit the non-conscious perceptual contents that occur in disorders such as blindsight, prosopagnosia, and amnesia. When flexible thinking does enlist perceptual contents, it very likely brings along HOTs about those contents, resulting in the perceptions' being conscious. So if their being conscious is blocked, flexible thinking about them is likely to be as well. Benefit in the normal case, however, may then be due just to the flexible thinking, and not to the perceptions' being conscious.

The theory that each state's being conscious is due to the occurrence of a higher order state, goes, however, suggest a possible function for psychological states' being conscious. Higher order states and processing are sometimes invoked in connection with executive function (e.g., Shallice, 1988), which involves the adjusting and fine tuning of behavior and hence the adjusting of the relevant intentions and desires. So perhaps a benefit of such states' being conscious is that their being conscious enables executive function.

Since executive function operates on first-order intentions and desires, it is natural to speculate that it proceeds by way of states whose higher order content is directed on those first-order volitions. Still, it is doubtful that any connection holds between executive function and consciousness. For one thing, the adjusting that occurs in executive function need not involve states with higher order intentional content. Such adjusting is often just a matter of resolving or eliminating conflicts among competing or dissonant first-order desires and beliefs. That process need not involve any higher order monitoring of those first-order states; it can simply be a matter of causal interactions among the first-order states themselves. As noted above, beliefs and desires interact causally in ways that reflect their intentional content, resulting in rational connections among those states. Those same causal interactions, which reflect the intentional content of the states, will also serve to iron out dissonance and competition among beliefs and desires. Causal connections that reflect the content of first-order beliefs and desires very likely

suffice for rational adjustments and fine tuning among our beliefs and desires. Such adjustment doubtless often simply involves a competition among mutually dissonant first-order beliefs and desires, in which the winners are those that already have stronger causal ties to other first-order states. No higher order states are needed.

There is a second reason to question whether the beliefs and desires on which executive function operates have to be conscious. Even if executive processing does, sometimes involve states with higher order content, those states may not be the kind needed for the relevant first-order states to be conscious. For a state to be conscious, the relevant HOT must have the content that one is in that state. But the higher order states that might occur in executive processing need not have that content. Those higher order states might simply register occurrences of conflict or dissonance among particular first-order states, and indicate possible compensatory adjustments. And the content of the metacognitive states might do that without explicitly representing that one is in any first-order state. The metacognitive state might simply have the content that a first-order state with some particular content conflicts with other beliefs, and that a state with a particular alternative content would not so conflict. The metacognitive state would play its adjusting role by referring to the state without explicitly representing that one is in that state.

Similar considerations apply to a proposal by [Rolls \(2004, 2005\)](#) of a type of HOT theory on which the HOTs enable one to correct errors in multistep chains of reasoning. A mistaken step in such a multistep chain, [Rolls](#) argues, can be located only by means of HOTs about each of the steps in that chain. This locating, he further argues, requires that those HOTs represent the syntactic ties among steps in such chains. [Rolls](#) concludes that HOTs, and hence the consciousness of psychological states, have a function that links them to rationality. Such correcting and adjusting of multistep chains of inference is, moreover, akin to the adjusting that occurs in executive function. But [Rolls's](#) argument involves specific appeal to the locating of erroneous steps in chains of reasoning and the syntactic character of the HOTs he argues would be required to do that.

But if a step in a multistep chain of reasoning is erroneous, that step will by itself likely result in some first-order dissonance with other antecedent beliefs. That dissonance will serve to locate the error, and so make possible the adjusting of the multistep chain at that point. Interactions among first-order states that reflect the intentional content of those states can iron out errors independently of any higher order monitoring, and hence independently of those first-order states' being conscious.

[Rolls](#) argues that the need for HOTs to locate erroneous steps pertains only to multistep chains of reasoning, and not to simple inferences. But [Dijksterhuis \(2004\)](#) and [Dijksterhuis et al. \(2006\)](#) seem to show the opposite. Their finding was that multistep deliberating is more effective when it occurs without being conscious, but that conscious deliberating is often more effective in making simple choices.

[Dijksterhuis et al.](#) find that when multistep reasoning involves the explicit application of rules, for example, the working out of arithmetic sums, the reasoning is more efficient when it is con-

scious. So perhaps the consciousness of psychological states has a function at least in these cases. But the success of such reasoning is likely not due to its being conscious; rather the efficiency and consciousness likely both result from another factor.

Some rules, such as those which govern arithmetic sums, are typically learned by applying them verbally. Since one learns the rules by verbally applying them, their application will be more efficient when it is verbalized. The thoughts that occur in applying the rules most efficiently are expressed verbally. Moreover, verbally expressed thoughts are, at least in the human case, conscious. whenever we say anything, the thought we thereby express is conscious.

Rule-governed reasoning is more efficient when it is verbalized, and in the human case verbally expressed thoughts are conscious. But the increased efficiency need not on that account be due to the thoughts' being conscious; rather the efficiency and consciousness of the reasoning are likely both due simply to the verbalizing.

Saying something has the same pragmatic force as saying that one thinks that thing; saying that it is raining, for example, is pragmatically interchangeable with saying that one thinks that it is raining. So whenever one says that it is raining, therefore, one could as easily have said that one thinks that it is raining. But the second assertion expressed a HOT that one thinks that it is raining. Since one would as easily have said that, the HOT must itself have been available to be expressed. Whenever one says anything at all, one has a HOT about the thought expressed by what one says. So whenever one says anything, the thought one thereby expresses is conscious ([Rosenthal, 2005](#), ch. 10). Verbalizing rules enhances the efficiency of the application, and verbalizing thoughts results in their being conscious. It is the verbalizing that is responsible for both the consciousness and efficiency.

Two other hypotheses about the function of consciousness can be considered briefly. [Nisbett and Wilson \(1977; cf. Wilson et al., 1989, 1995\)](#) show that subjects sometimes confabulate desires and beliefs that would make their behavior appear rational to themselves and others. The HOT hypothesis fits well with those results. On that theory, subjects in these cases have HOTs that represent them as having beliefs or desires that they do not actually have. Since subjects believe their confabulations, these confabulatory HOTs might themselves influence their subsequent thinking; these HOTs might therefore play a significant role in one's thinking. (I owe this suggestion to [Zoltán Dienes](#), in correspondence.) But the apparent goal of such confabulation is to make one's first-order states seem more well-founded, either by one's own lights or those of others, and this is typically an ephemeral matter, with little impact on one's thoughts and desires.

A different type of function of consciousness is sometimes suggested in connection with psychoanalytic theory. Psychoanalysis is said to ameliorate neurosis by a process in which troubling unconscious desires become conscious. So their being conscious might function to relieve neurosis. But neurotic symptoms are said on that model to result not from desires' failing to be conscious, but from their having been repressed ([Freud](#),

1957). Amelioration results from the undoing of repression, not simply from the desires' coming to be conscious.

A higher order theory of consciousness predicts that the consciousness of psychological states would be minimal, to little to sustain selective pressures needed for an evolutionary explanation of consciousness. The major function is carried by the first-order states, and the benefit of higher order states, if any, would be minimal relative to that of the first-order states. But the foregoing considerations that point to a minimal function for psychological states' being conscious are independent of any higher order theory. Since those considerations independently sustain a prediction that higher order theories make, they lend support to such theories.

An alternative to higher order theories advanced by Dretske (1995, ch. 4; cf. 1993) holds that a state is conscious simply if one is conscious of something in virtue of being in that state; no awareness of the state itself is needed. So on this theory, the function of a state's being conscious is simply the ordinary function of an organism's being conscious of things, which is very substantial (Dretske, 1997)(1995, ch. 4; 1997).

Since all perceiving results in one's being conscious of the things perceived, Dretske's theory seems to dictate that perceiving is always conscious. This conflicts with robust findings about blindsight (Covey & Stoerig, 1995; Weiskrantz, 1986, 1995, 1997, 1998) and masked priming (Marcel, 1983a, 1983b), among many others. Dretske (2006, 167–176) seeks to meet this difficulty by urging that perceiving is conscious only if one can cite one's perceiving as a justifying reason for doing something. Subjects cannot do that when the perceiving is subliminal.

But this modification of the theory tacitly reintroduces the machinery of higher order theories. To cite something is to express a thought one has about that thing. The ability to cite one's perceiving requires that one be conscious of the perceiving. So if that ability to cite is necessary for the perceiving to be conscious, as Dretske suggests, the perceiving is conscious only if one is conscious of it, as a higher order theory requires.

5. Why are any psychological states conscious?

On the HOT theory, psychological states are conscious in virtue of being accompanied by distinct, occurrent HOTs that represent the individual as being in those first-order states. Carruthers (2000, 221–222, 225) has argued that it is unlikely that organisms would have evolved to have occurrent HOTs. Distinct, occurrent HOTs would so be costly in cognitive and neural overhead, he urges, that some powerful adaptive value would have been needed for them to have evolved. He concludes that the consciousness of psychological states must be due not to occurrent HOTs, but merely to a disposition for such HOTs to occur.

Carruthers assumes that the occurrence of actual HOTs would roughly double the demand on cortical capacity. But it is unclear that that is so. And whatever the cognitive and neural overhead of distinct HOTs, biological processes with little adaptive value do sometimes arise and become widespread without being adaptive (Gould & Vrba, 1982), occasionally even at some cost to the organism.

In any case, the dispositional theory that Carruthers offers cannot explain consciousness (Rosenthal, 2004, §4). For one thing, it cannot implement TP; since being disposed to have a thought about something does not make one conscious of that thing, merely being disposed to have a HOT about a state would not make one conscious of that state. Nor is it clear that the cost in cortical overhead of being disposed to have HOTs would be significantly less than the cost of having actual HOTs.

It is important to understand why some psychological states are conscious (Weiskrantz, 1997, ch. 7). But an evolutionary explanation is not the only possibility. Though many psychological states and processes confer significant benefit on the organism and may therefore have evolved in response to suitable selection pressures, it is by no means obvious that that is so for all significant psychological states and processes. Some states and processes may arise not because they afford reproductive or other advantage, but because they are by-products of psychological processes already in place. Given the difficulty of finding any credible function for the consciousness of psychological states, it is likely that such consciousness arises in that second way.

The HOT theory suggests just such an explanation. That explanation proceeds differently for the case of qualitative psychological states, such as perceptions and sensations (Rosenthal, 2005, pp. 218–219), and nonqualitative states, such as thoughts, desires, and intentions (Rosenthal, 2005, pp. 303–305; slightly revised in what follows). In keeping with the foregoing discussion, I restrict attention to nonqualitative, intentional states.

An organism cannot have HOTs about psychological states unless it has a concept of those states that characterizes them in the way relevant to HOTs. To have HOTs about thoughts and desires, the organism must have a concept of a state that has intentional content.

Thoughts, desires, and other intentional states are states that are expressed by the things we say as well as by various forms of nonverbal behavior; in each case the state expressed causes the relevant verbal and nonverbal behavior (Rosenthal, 2005, ch. 3). Any creatures capable of speech must have thoughts and desires that are expressed by the things they say, though many creatures doubtless have thoughts and desires but lack any concept of those states.

One way that creatures might come to have concept of thoughts and desires is by positing such states as the causes of their verbal and nonverbal behavior. They would conceive of desires as inner causes of much nonverbal behavior, and thoughts as inner causes of ordinary assertions. These inner states would be folk-theoretic posits, common among curious, prescientific peoples.

We individuate the things we say by their meanings. So the internal folk-theoretic posits that cause verbal behavior must themselves differ in ways that parallel differences of meaning. Since the intentional contents of thoughts and desires corresponds to the meanings of the things we say, inner states with properties that parallel differences of meaning are inner states with intentional content. To posit such states is for therefore to develop the concept of inner states with intentional content (Sellars, 1963, pp. 46–59).

Occasionally we consciously infer our own thoughts and desires from self-observation or theory. But because we are seldom aware of those thoughts and desires in a way that is subjectively unmediated, the inferred thoughts and desires are rarely conscious. Creatures that have posited such inner states to explain their verbal and nonverbal behavior would ascribe those states not only to others, but also to themselves, by inferring from their own behavior.

Those inferences would not be conscious, since these creatures have as yet no HOTs about intentional states. Still, thoughts and desires these creatures infer themselves to have will not be conscious. Those their inferences are not conscious, they rely on no on conscious observations of their own behavior, since perceiving comes to be conscious independently of nonqualitative states, such as thoughts and desires (Rosenthal, 2005, pp. 218–219). And inference from conscious self-observation prevents the resulting awareness from being subjectively unmediated. These creatures awareness of their own thoughts and desires will be on a par with their awareness that others are in those states.

As these curious, proto-theoretical creatures become practiced in inferring from their verbal and nonverbal behavior to thoughts and desires that their behavior expresses, their inferring would become relatively automatic. So it might even sometimes rely on self-observation that is not conscious. The resulting awareness of their own thoughts and desires would be subjectively unmediated, and the self-ascribed thoughts and desires would thus be conscious.

Sometimes, moreover, these creatures will be merely disposed to say something without actually saying it. And once the inference to their own thoughts and desires becomes sufficiently automatic, simply being disposed to say something could by itself, and independently of any inference whatever, prompt a HOT about the thought that saying that thing would have expressed. Since these HOTs would again be subjectively unmediated, the self-ascribed intentional states would be conscious. These creatures would have come to have the very HOTs that result in the consciousness of thoughts and desires.

The foregoing sketch of how creatures with particular linguistic abilities would be likely to come to have HOTs about their thoughts and desires makes no appeal to selective advantage or other benefit that such HOTs might confer. Subjectively unmediated HOTs are prompted by nonconscious observations and by dispositions to express first-order thoughts and desires. The dispositions and nonconscious observations themselves benefit the organism, but there is no added benefit conferred by the resulting HOTs. Rather, creatures that develop folk-theories about the internal causes of their verbal and nonverbal behavior will likely end up not only ascribing thoughts and desires to themselves, eventually in a way that results in those states' being conscious. It is reasonable to speculate that this process lies behind the human ability to have HOTs about our own thoughts and desires.

Thoughts about one's own thoughts and desires initially occur by inferring in a folk-theoretical way from conscious observations of one's own behavior. These inferences do serve a useful purpose, since they give rise to a general theory of mind,

which in turn enables and enhances elaborate social interaction (Carruthers, 2000, 225; Whiten & Byrne, 1997).

But the thoughts that occur in applying a theory of mind to oneself need not be the type that results in the self-ascribed states' being conscious. One can apply a theory of mind to oneself by having thoughts about one's own psychological states that rely on consciously observed behavior and body language. Since such thoughts do not result in subjectively unmediated awareness of one's own states, those states will not be conscious.

But subjectively immediacy confers no added benefit. The thoughts that self-ascribe beliefs and desires on the basis of conscious self-observation would facilitate all the relevant social purposes. Subjectively unmediated awareness might occur more rapidly and more often, but subjectively immediacy itself adds no additional function. Self-ascription of thoughts and desires, moreover, would have already had to be very good for creatures to come to have subjectively unmediated HOTs, further undermining the possibility that such HOTs could add any significant benefit to that already in place. The consciousness of thoughts and desires builds on a process that facilitates social interactions without those states' being conscious.

Nonlinguistic animals have many thoughts and desires, which they express in their nonverbal behavior. So one might expect that nonlinguistic animals could infer from such nonverbal behavior to the thoughts and desires that behavior expresses. (I am grateful to Chris Frith, personal communication, for pressing this possibility.) Still, it is unlikely that nonlinguistic animals will be sufficiently sophisticated conceptually to infer from their behavior to internal states that cause that behavior. The folk-theoretic inference from behavior to inner states will occur only in creatures with fairly elaborate conceptual abilities, abilities of the sort that very likely only develop with language use.

Pretheoretic intuition and experimental findings both suggest that the perceptions of nonlinguistic animals are often conscious (e.g., Cowey & Stoerig, 1995). And nonlinguistic doubtless also have many thoughts and desires. But there is little reason to think that the thoughts and desires of nonlinguistic animals are also conscious. Pretheoretic intuition is arguably silent about that question, and there are no experimental findings that bear on it.

In any case, since the perceptions and sensations of nonlinguistic creatures are often conscious, an explanation is needed of how those states come to be conscious, an explanation that is independent of language. The HOT theory suggests an explanation that appeals to perceptual error and to the conceptual connection between the mental qualities that occur in perceiving and the physical properties that perceiving makes available (Rosenthal, 2005, pp. 218–219). Like the explanation just sketched for the consciousness of thoughts and desires, the explanation for conscious perceiving proceeds independently of any benefit that the relevant HOTs might confer. But that is beyond the scope of this discussion, which has been limited to purely intentional states, such as thoughts and desires.

Still, it should not be surprising that the factors that lead to qualitative states' being conscious would differ from those responsible for the widespread consciousness in humans of

thoughts and desires. Indeed, the need for independent explanations may help explain our divergent pretheoretic intuitions about the two kinds of consciousness. We intuitively regard the consciousness of qualitative states as somehow more basic than that of purely intentional states, even to holding that the consciousness of qualitative states is intrinsic to those states. This is unsurprising given that language is needed for the consciousness of intentional states, but not the consciousness qualitative states.

But even if the need for two kinds of explanation explains the appeal of those intuitions, they do not support them. Since both explanations appeal to HOTs, there is no difference in what it is for intentional and qualitative states to be conscious. The difference is simply in the process that leads to the consciousness of the two kinds of state.

6. Summary

The consciousness of thoughts, desires, and volitions adds little if any benefit for rational thinking, intentional action, executive function, or complex reasoning. Nonetheless, an explanation is available of why those states are often conscious that makes no appeal to beneficial effects or evolutionary adaptive value.

References

- Armstrong, D. M. (1968). *A materialist theory of the mind*. New York: Humanities Press.
- Baars, B. J. (1988). *A cognitive theory of consciousness*. Cambridge: Cambridge University Press.
- Baars, B. J. (1997). *In the theater of consciousness: The workspace of the mind*. New York: Oxford University Press.
- Bargh, J. A. (2002). Losing consciousness: Automatic influences on consumer judgment, behavior, and motivation. *Journal of Consumer Research*, 29(September (2)), 281–285.
- Block, N. (1987). Functional role and truth conditions. *Proceedings of the Aristotelian Society*, LXI, 157–181.
- Block, N. (1995). On a confusion about a function of consciousness. *The Behavioral and Brain Sciences*, 18(June (2)), 227–247.
- Block, N. (2001). Paradox and cross purposes in recent work on consciousness. *Cognition*, 79(April (1–2)), 197–219.
- Block, N. (2005). Two neural correlates of consciousness. *Trends in Cognitive Sciences*, 9(February (2)), 46–52.
- Block, N. (2007). Consciousness, accessibility, and the mesh between psychology and neuroscience. *The Behavioral and Brain Sciences*, 30(August (4)).
- Brass, M., & Haggard, P. (2007). To do or not to do: The neural signature of self-control. *The Journal of Neuroscience*, 27(August (34)), 9141–9145.
- Carruthers, P. (2000). *Phenomenal consciousness: A naturalistic theory*. Cambridge: Cambridge University Press.
- Cowey, A., & Stoerig, P. (1995). Blindsight in monkeys. *Nature*, 373(January (6511)), 247–249.
- Debnar, J. A., & Jacoby, L. L. (1994). Unconscious perception: Attention, awareness and control. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 20(March (2)), 304–317.
- Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: Basic evidence and a workshop framework. *Cognition*, 79(April (1–2)), 1–37.
- Dijksterhuis, A. (2004). Think different: The merits of unconscious thought in preference development and decision making. *Journal of Personality & Social Psychology*, 87(November (5)), 586–598.
- Dijksterhuis, A., Bos, M. W., Nordgren, L. F., & van Baaren, R. B. (2006). On making the right choice: The deliberation-without-attention effect. *Science*, 311(February (5763)), 1005–1007.
- Dretske, F. (1993). Conscious experience. *Mind*, 102(April (406)), 263–283.
- Dretske, F. (1995). *Naturalizing the mind*. Cambridge, MA: MIT Press.
- Dretske, F. (1997). What good is consciousness? *Canadian Journal of Philosophy*, 27(March (1)), 1–15.
- Dretske, F. (2006). Perception without awareness. In Tamar Szabó Gendler & John Hawthorne (Eds.), *Perceptual experience* (pp. 147–180). Oxford: Clarendon Press.
- Fernandez-Duque, D., & Thornton, I. M. (2000). Change detection without awareness: Do explicit reports underestimate the representation of change in the visual system? *Visual Cognition*, 7(January–March (1–3)), 324–344.
- Fernandez-Duque, D., & Thornton, I. M. (2003). Explicit mechanisms do not account for implicit localization and identification of change: An empirical reply to Mitroff et al. (2002). *Journal of Experimental Psychology: Human Perception and Performance*, 29(October (5)), 846–858.
- Field, H. (1977). Logic, meaning, and conceptual role. *The Journal of Philosophy*, 74(July (7)), 379–409.
- Freud, S. (1957). The unconscious. In James Strachey (Ed. & tr.), *The complete psychological works of Sigmund Freud* (Vol. 14, 1957, pp. 166–215). London: The Hogarth Press. Original work published in 1915.
- Gould, S. J., & Vrba, E. S. (1982). Exaptation—A missing term in the science of form. *Paleontology*, 8(January (1)), 4–15.
- Gray, J. (2004). *Consciousness: Creeping up on the hard problem*. Oxford: Oxford University Press.
- Haggard, P. (1999). Perceived timing of self-initiated actions. In Gisa Aschersleben, Talis Bachmann, & Jochen Müsseler (Eds.), *Cognitive contributions to the perception of spatial and temporal events* (pp. 215–231). Amsterdam: Elsevier.
- Haggard, P., & Eimer, M. (1999). On the relation between brain potentials and awareness of voluntary movements. *Experimental Brain Research*, 126(1), 128–133.
- Haggard, P., Newman, C., & Magno, E. (1999). On the perceived time of voluntary actions. *British Journal of Psychology*, 90(May (Part 2)), 291–303.
- Harman, G. (1982). Conceptual role semantics. *Notre Dame Journal of Formal Logic*, XXIII(April (2)), 242–256.
- Harman, G. (1987). (Nonsolipsistic) conceptual role semantics. In Ernest LePore (Ed.), *New directions in semantics* (pp. 55–81). New York: Academic Press.
- Humphrey, N. (1980). Nature's psychologists. In B. D. Josephson & V. S. Ramachandran (Eds.), *Consciousness and the physical world* (pp. 57–80). Oxford: Pergamon Press.
- Humphrey, N. (2002). The uses of consciousness. In N. Humphrey (Ed.), *The mind made flesh: Essays from the frontiers of evolution and psychology* (pp. 65–85). Oxford: Oxford University Press.
- Huxley, T. H. (1896). On the hypothesis that animals are automata, and its history (1974). In T. H. Huxley (Ed.), *Methods and results: Essays* (pp. 199–250). New York: D. Appleton and Company.
- Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory & Language*, 30(October (5)), 513–541.
- Jacoby, L. L., Toth, J. P., & Yonelinas, A. P. (1993). Separating conscious and unconscious influences of memory measuring recollection. *Journal of Experimental Psychology: General*, 122(June (2)), 139–154.
- Jacoby, L. L., Toth, J. P., Yonelinas, A. P., & Debnar, J. A. (1994). The relationship between conscious and unconscious influences: Independence or redundancy? *Journal of Experimental Psychology: General*, 123(June (2)), 216–219.
- James, W. (1890). *The principles of psychology*. New York: Henry Holt and Company.
- Laloyaux, C., Destrebecqz, A., & Cleeremans, A. (2003). Implicit change identification: A replication of Fernandez-Duque and Thornton (2003). *Journal of Experimental Psychology: Human Perception and Performance*, 32(December (6)), 1366–1379.
- Lau, H.C. (in press). Are we studying consciousness yet? Draft of chapter to appear in Lawrence Weiskrantz & Martin Davies (Eds.), *Frontiers of*

- Consciousness: The Chichele Lectures 2006*. Oxford: Oxford University Press.
- Lau, H. C., & Passingham, R. E. (2006). Relative blindsight in normal observers and the neural correlate of visual consciousness. *PNAS*, *103*(December (49)), 18763–18768.
- Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *The Behavioral & Brain Sciences*, *8*(December (4)), 529–539.
- Libet, B., Gleason, C. A., Wright, E. W., Jr., & Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness potential). *Brain*, *106*(September (Part III)), 623–642.
- Marcel, A. J. (1983a). Conscious and unconscious perception: Experiments on visual masking and word recognition. *Cognitive Psychology*, *15*(April (2)), 197–237.
- Marcel, A. J. (1983b). Conscious and unconscious perception: An approach to the relations between phenomenal experience and perceptual processes. *Cognitive Psychology*, *15*(April (2)), 238–300.
- Merker, B. (2005). The liabilities of mobility: A selection pressure for the transition to consciousness in animal evolution. *Consciousness & Cognition*, *14*(March (1)), 89–114.
- Morsella, E. (2005). The function of phenomenal states: Supramodular interaction theory. *Psychological Review*, *112*(October (4)), 1000–1021.
- Nagel, T. (1974). What is it like to be a bat? *The Philosophical Review*, *LXXXIII*(October (4)), 435–450.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *LXXXIV*(May (3)), 231–259.
- Peacocke, C. (1992). *A study of concepts*. Cambridge, MA: MIT Press.
- Rolls, E. T. (2004). A higher order syntactic thought (HOST) theory of consciousness. In R. J. Gennaro (Ed.), *Higher-order theories of consciousness* (pp. 137–172). Amsterdam & Philadelphia: John Benjamins Publishers.
- Rolls, E. T. (2005). *Emotion explained*. Oxford: Oxford University Press.
- Rosenthal, D. M. (1997). A theory of consciousness. In Ned Block, Owen Flanagan, & Güven Güzeldere (Eds.), *The nature of consciousness: philosophical debates* (pp. 729–753). Cambridge, MA: MIT Press.
- Rosenthal, D. M. (2002a). How many kinds of consciousness? *Consciousness & Cognition*, *11*(December (4)), 653–665.
- Rosenthal, D. M. (2002b). The timing of conscious states. *Consciousness & Cognition*, *11*(June (2)), 215–220.
- Rosenthal, D. M. (2004). Varieties of higher-order theory. In R. J. Gennaro (Ed.), *Higher-order theories of consciousness* (pp. 17–44). Amsterdam & Philadelphia: John Benjamins Publishers.
- Rosenthal, D. M. (2005). *Consciousness and mind*. Oxford: Clarendon Press.
- Rossano, M. J. (2003). Expertise and the evolution of consciousness. *Cognition*, *89*(October (3)), 207–236.
- Sellars, W. (1963). Empiricism and the philosophy of mind. In W. Sellars (Ed.), *Science, perception and reality* (pp. 127–196). London: Routledge & Kegan Paul.
- Shoemaker, S. (1996). On knowing one's own mind. In S. Shoemaker (Ed.), *The first-person perspective and other essays* (pp. 25–49). Cambridge: Cambridge University Press.
- Silverman, M. E., & Mack, A. (2006). Change blindness and priming: When it does and does not occur. *Consciousness & Cognition*, *15*(June (2)), 409–422.
- Tononi, G. (2004). An information integration theory of consciousness. *BMC Neuroscience*, *5*(November (42)).
- Van Gulick, R. (2004). Higher-order global states (HOGS): An alternative higher-order model of consciousness. In R. J. Gennaro (Ed.), *Higher-order theories of consciousness* (pp. 67–92). Amsterdam & Philadelphia: John Benjamins Publishers.
- Wegner, D. M. (2002). *The illusion of conscious will*. Cambridge, MA: MIT Press.
- Wegner, D. M. (2003). The mind's best trick: How we experience conscious will. *Trends in Cognitive Sciences*, *7*(February (2)), 65–69.
- Weiskrantz, L. (1986). *Blindsight: A case study and implications*. Oxford: Clarendon Press.
- Weiskrantz, L. (1995). The problem of animal consciousness in relation to neuropsychology. *Behavioural Brain Research*, *71*(November (1–2)), 171–175.
- Weiskrantz, L. (1997). *Consciousness lost and found: A neuropsychological exploration*. Oxford: Oxford University Press.
- Weiskrantz, L. (1998). Consciousness and commentaries. *International Journal of Psychology*, *33*(June (3)), 227–233.
- Weiskrantz, L. (2001). Commentary responses and conscious Awareness in humans: The implications for awareness in non-human animals. *Animal Welfare, Special Issue: Consciousness, Cognition & Animal Welfare*, *10*(April (Supplement)), S41–S46.
- Whiten, A., & Byrne, R. W. (Eds.). (1997). *Machiavellian intelligence, II: Extensions and evaluations*. Cambridge: Cambridge University Press.
- Wilson, T. D., Dunn, D. S., Kraft, D., & Lisle, D. J. (1989). Introspection, attitude change, and attitude-behavior consistency: The disruptive effects of explaining why we feel the way we do. In Berkowitz & Leonard (Eds.), *Advances in experimental social psychology* (pp. 287–343). San Diego, CA: Academic Press.
- Wilson, T. D., Hodges, S. D., & LaFleur, S. J. (1995). Effects of introspecting about reasons: Inferring attitudes from accessible thoughts. *Journal of Personality & Social Psychology*, *69*(July (1)), 16–28.
- Wilson, T. D., & Schooler, J. W. (1991). Thinking too much: Introspection can reduce the quality of preferences and decisions. *Journal of Personality & Social Psychology*, *60*(February (2)), 181–192.