

Robert Van Gulick

Consciousness and self-awareness seem intuitively linked, but how they intertwine is less than clear. Must one be self-aware in order to be conscious? Indeed, is consciousness just a special type of self-awareness? Or perhaps it is the other way around: Is being self-aware a special way of being conscious?

Discerning their connections is complicated by the fact that both the main relata themselves admit of many diverse forms and levels. One might be conscious or self-aware in many different ways, and to varying degrees. Thus the real questions of linkage must be posed more specifically. We need to ask not whether the two are bound in general, but whether and how being conscious in some specific sense and degree relates to some particular sort of self-awareness. Only those more specific questions are likely to have fully determinate answers.

However, I will argue that each form of consciousness seems to involve a corresponding type and degree of self-awareness. Self-awareness may not be the whole of consciousness, but it is one of its most important aspects. Though no single perspective by itself can give us a complete understanding of consciousness, viewing it through the lens of self-awareness can help us see a lot about its nature and place in the natural world that we might otherwise miss.

In that spirit, I will offer and defend a particular model of consciousness as self-awareness, the “HOGS” or higher-order global state model. I will aim to motivate the HOGS view both in its own right as a model of consciousness and by situating it within a more general and independently plausible telepragmatic view of mind.

I Higher-Order Theories of Consciousness

The link between consciousness and self-awareness has found prominent recent support among advocates of the so-called higher-order (HO) theory

of consciousness. The theory comes in many versions, but they all agree in treating the distinction between conscious and nonconscious mental states as a relational matter of whether or not the given state is itself the object of some higher-order (or meta-intentional) state directed at it. Conscious mental states are states we are aware of being in.

According to standard HO theory, what makes a mental state *M* conscious is not some difference in *M*'s intrinsic properties but the relational fact that *M* is accompanied by a higher-order state whose intentional content is that one is in *M*. For example, to have a conscious desire for a glass of champagne is to have both a desire for champagne and a higher-order awareness of having such a desire. "Higher-order" here means simply a meta-mental state, that is, an intentional state whose content itself concerns or is about a mental state: a thought about a thought, a belief about a desire, or an awareness of a mood or sensation. Unconscious desires, thoughts, or memories are simply those we are not aware of having.

Higher-order theories come in several forms, differing especially with regard to the psychological modality of the relevant conscious-making higher-order states. Higher-order thought (HOT) models treat those meta-states as thoughtlike, and higher-order perception (HOP) models regard them as at least quasi-perceptual and resulting from some form of inner monitoring or inner sense. The differences are potentially important, but for present purposes I will largely ignore them. My current concern is with the basic HO view, and with problems that affect it in both its main versions.¹

HO models in general, whether HOT or HOP, have at least three major strengths.

Common usage HO models accord well with at least one common use of "conscious." In everyday folk-psychological practice and in the neo-Freudian view that pervades contemporary culture, the conscious–unconscious distinction divides our mental states between those of which we are aware and those of which we are not. The fact that a desire, memory, or perception is unconscious in this sense is nothing more than our being unaware of it, which HO models analyze as our lacking the relevant HOT or HOP about it. Insofar as it is this notion of conscious mental state that the models aim to capture, they seem on target.

Scientific practice HO models also conform well with empirical scientific practice in determining when a subject does or does not have a particular conscious state, especially insofar as those practices rely on a first-person report criterion. For example, the method of judging subjects to have had a

conscious perception of a word just if they can report having had such a perception fits quite naturally with the HO view that what makes a perception conscious is our having a meta-awareness of it. When we are aware of it, we can report being in it, and when we are not we cannot. David Rosenthal has been explicit in claiming that the content expressed by such a first-person report is exactly the content of the HOT that would make the relevant state conscious (Rosenthal 1990, 1993b, 2002). Thus the HO theorist might plausibly claim that his analysis is not merely consistent with scientific practice but explains why those methods work.

Naturalism HO models, if successful, would effectively reduce the problem of consciousness to that of intentionality. Since many theorists of mind regard the latter problem as more tractable, indeed as having already been at least partially resolved in a naturalistically acceptable way, it would be a significant advance if the problem of consciousness could be subsumed within it (see, e.g., works by Dretske, Millikan, Van Gulick). This, of course, assumes that the meta-intentionality of the sort HO theory requires could be explained in much the same way as other sorts of intentionality—for example, in terms of covariance relations, tracking conditions, teleological relations, or whatever else that does not seem implausible. Thus HO models promise to dispel the air of apparent mystery about how consciousness might be explained as part of the natural world.

Despite its strengths, the higher-order approach has faced a great many objections, of which six deserve particular mention.

1 “Too Fancy” Objection

The meta-intentional conditions are criticized as being too sophisticated and demanding to be met by many beings we regard as conscious: dogs, bats, and small children. A newborn baby seems to have conscious perceptions, but is it plausible to regard it as having metathoughts with contents such as [I am now having a perception]? In response:

- i. Some HO theorists deny that the relevant meta-intentionality need be so sophisticated or conceptually demanding (Rosenthal 1993b).
- ii. Others “bite the bullet” and deny that small children and animals have conscious mental states (Carruthers 2000).

2 Extra Conditions Problem

Higher-order theories need to include further conditions to rule out obvious counterexamples to the sufficiency of the higher-order analysis (e.g., that the meta-intentional state be simultaneous with its lower-order object and that

it be arrived at noninferentially). But those conditions call into question the theory's basic idea of explaining consciousness solely in terms of meta-intentional content. The extra conditions may be required to sort the cases correctly, but it is not clear *why* they should matter, which suggests there may be some other factor (e.g., mode of representation) that makes the real difference and with which those conditions merely coincide.

3 Generality Problem

In general, having a thought or perception of some object of type *T* does not make it into a conscious *T*. My having a thought or perception of a stone or a pencil does not produce conscious stones or pencils. So why should having a thought or perception of a mental state transform it into a conscious state? (See Dretske 1995a; Byrne 1997.)

4 Phenomenal Adequacy Issue

Even if standard higher-order theory captures one sense of “conscious state” it does not so obviously capture another, namely, the phenomenal sense. Being a conscious state in this latter phenomenal sense involves the intuitive idea that there is “something that it’s like to be in that state” (Nagel 1974).

5 “Stranded Qualia” Problem

The transition from states that are not conscious in the “what-it’s-like” sense to states that are may seem to require the addition of qualia. But that seems to conflict with the basic higher-order claim that a state’s becoming conscious involves not changes in its intrinsic properties, but only the addition of an extrinsic factor, namely a meta-state directed at it. How can the addition of a merely relational element make a state into one that there is something that it’s like to be in?

Some higher-order theorists have a surprising answer. They distinguish having sensory qualities from having what-it’s-likeness. They restrict the latter to states that are conscious in the higher-order sense, but they allow sensory qualities to occur as properties of unconscious states that there is nothing it’s like to be in. What-it’s-likeness on this view involves not merely sensory qualities but an awareness of sensory qualities. Some such theorists (Lycan 1996) apply the word “qualia” to unconscious sensory qualities and others apply it only to conscious sensory qualities, but the basic explanatory move is the same: to account for what-it’s-likeness as a matter of having a higher-order awareness of sensory qualities that can and do occur as properties of unconscious states.

On this view, one can have qualitative but unconscious mental states such as unconscious color perceptions or pains, but there is nothing it is like to be in such a state. Only when one becomes aware of the state by having the requisite HOT or HOP is there something it is like to be in it. The transition from states that are unconscious in the “what-it’s-like” sense to those that are conscious does not involve adding qualia or sensory qualities to the state, but rather becoming aware via a meta-state of the qualia that the lower-order state already had.

The view requires unconscious qualia, which many find incoherent or contradictory, as they do the idea of unfelt pains. Moreover, it seems to strand the notion of phenomenal or experiential feel in a “no-man’s-land” unable to quite take hold at either the lower or the metalevel. According to the proposal:

- Qualia can be fully present in unconscious states that there is nothing it is like to be in.
- Conversely, and almost paradoxically, the meta-states whose addition supposedly accounts for first-person experiential feel are themselves devoid of any qualia.

Thus the proposal requires a commitment both to qualia of a sort that do not by themselves produce any “what-it’s-likeness” and to states that produce what-it’s-likeness but themselves lack any qualia. It is in that sense that the qualia seem “stranded.” Some HO theories (Lycan 1996) deny there is anything wrong, as opposed to merely unfamiliar, about such qualia, but the sense of a problem is not easily dismissed.

6 The Error/Illusion Problem

The possibilities of metalevel error or illusion create further problems:

- Error: What if the first-level mental state has quale Q_i (e.g., red) but the HO state represents it as having a different quale Q_k (e.g., blue)?
- Illusion: What if the relevant HO state occurs when there is no associated Q -state at all?

In such cases, what sort of what-it’s-likeness would the person experience? In the error case, would it seem red or blue? And how could it seem red (in the relevant *phenomenal* sense of “seem”) in the illusion case when there is no relevant lower-order state with qualia for the person to be aware of? None of the available answers seems satisfactory:

- If one opts for the what-it’s-likeness that is associated with the lower-order state, that would imply that the higher-order awareness is not really

responsible for what-it's-likeness. Contrary to the HO proposal, it seems to be the quale itself, not the awareness of the quale that matters.

- If one opts for the what-it's-likeness that is associated with the higher-order state, that would imply that what-it's-likeness does not derive from awareness of a quale as the proposal claims, since in the error and illusion cases there is no state with a quale (or none of the relevant type) of which to be aware.
- If the HO theorist tries to avoid the problem by claiming that error and illusion are impossible in such cases, he or she must provide some non-ad hoc support for the claim, and none is obvious as long as one follows standard HO models in treating the meta-state and the lower-order state as distinct.

Thus, despite their obvious strengths, existing higher-order models of consciousness face substantial objections. Their supporters may in the end be able to offer adequate replies, but it may be useful to explore alternative versions of the higher-order approach that aim to preserve its essential insights while avoiding the problems that affect the current models. I have elsewhere (Van Gulick 2001, 2004) tried to provide such an alternative in terms of higher-order global states or HOGS. My goal for the remainder of this chapter is to further develop that alternative model, with special emphasis on two linked features that are of special relevance to this volume:

A. The extent to which the HOGS model *weakens the distinctness assumption* found in the standard HO models.

Most existing HOT and HOP versions of higher-order theory treat the conscious-making meta-intentional states they posit as distinct and separate from their lower-order mental objects. Let us call this the *distinctness assumption*. On such models, a conscious visual experience of a clock face involves the occurrence of a visual perception VP of the clock face, which is made conscious by the simultaneous occurrence of a separate intentional state HO whose higher-order content concerns the occurrence of VP. However, as the HOGS model shows, there may be good reasons for not regarding the two states as fully separate and distinct, but rather as two interrelated intentional aspects of one and the same state.

B. The respect in which the HOGS model implies that the relevant meta-intentionality is implicit rather than explicit and thus *rejects the explicitness assumption*.

Standard HOT and HOP models treat the required meta-intentionality as the *explicitly represented* content of a distinct higher-order state, either thoughtlike or quasi-perceptual. Let us call this the “explicitness assumption”.

tion.” Along with weakening the distinctness assumption, the HOGS view differs from the standard models in analyzing the relevant meta-intentionality primarily in terms of information or understanding that is *implicitly embedded* in the structure, organization, and dynamics of the relevant global states. The HOGS view thus also rejects the explicitness assumption.

To bring these two features of the HOGS model into clear focus, we need to back up a bit and see the model within a larger context, specifically that provided by the teleopragmatic view of mind (Van Gulick 1980). Understanding consciousness requires us to situate it within a general account of mind and its place in nature. Thus in section II, I will sketch the basic teleopragmatic picture with special attention to those of its features most relevant to the HOGS model. Having filled in that background, I will address its application to reflexive intentionality in section III, and then return to the HOGS model itself in section IV.

II The Teleopragmatic View of Mind

The *teleopragmatic* perspective regards mind, at least in the first instance, as a biological aspect of reality. Though we may one day be able to create minded nonbiotic systems, mind as we know it is *biological in origin and nature*. Attending seriously to its biological dimension brings into relief important aspects of mind that might otherwise be missed. Indeed it highlights features that are directly relevant to understanding the HOGS model and the reflexive meta-intentional nature of consciousness.

The teleopragmatic view might be classed as a version of *functionalism*, but that latter term has been associated with so much extra and often dubious philosophical baggage (computationalist, reductionist, or analytic) that it may be best to avoid it altogether. Insofar as we do treat it as a version of functionalism, the view is unsurprisingly closest to so-called teleofunctionalism (Van Gulick 1980; Lycan 1981; Millikan 1984). A *teleological* aspect looms large in both, insofar as mental states are described in terms of the *roles they are meant or selected to play in contributing to the successful operation of a complex goal-directed system or organism*.

The “pragmatic” in the name reflects the fact that minds and the forms of understanding associated with minds are fundamentally anchored in *practice and practical use*, and “teleo” signifies that it is in the pursuit of *goals and ends* (Greek “telos”) that minds find their central application.² Minds, as well as their premental predecessors, evolved as means by which organisms might better adapt their action to the structure of their worlds. Organisms gather, store, integrate, transform, and act on information, but always in the

end the aim of any information process is to enhance the organism's chances for success in meeting whatever goals it has for growth, food, sex, survival, reproduction, and all the other sometimes far less basic needs and ends that organisms so purposefully pursue.³

The two interrelated notions of *information* and *understanding* are central to the teleopragmatic view, and both get anchored in a biological framework. An organism is informed about some feature of its world insofar as its own structure and organization are specifically shaped or adapted with respect to that feature in ways that enhance the realization of its goals. The bat is informed about the location and direction of the moth of which it aims to make a meal insofar as the echo signal that it processes guides its flight and attack in just the way that enables it to catch its moving target. But the bat's wing as well carries implicit information about the bat's world insofar as its structure has been specifically and adaptively shaped to reflect the dynamic properties of the air through which it flies. Indeed, from the teleopragmatic perspective, life itself is largely an informational process, one through which lifeforms—phylogenetically and ontogenetically—come to adaptively reflect their worlds in ways that meet their needs or ends.

Understanding is similarly grounded in successful pragmatic engagement. The bee understands the correlation between the flower's fragrance and its supply of nectar insofar as its system of behavioral control guides it along the gradient of scent that maximizes its foraging success.

The idea of understanding is particularly apt in several respects for expressing the teleopragmatic view of intentionality and cognition.

- In contrast with standard notions of knowledge, the concept of understanding emphasizes the *element of practical engagement* from the outset. The domain of knowledge is all too easily imperialized by its more intellectual form of knowing-that, demoting know-how to second-class status as a lesser and marginal mode of cognition. By contrast, the notion of understanding is linked at its core with the idea of enabling successful engagement. Thus one is far less likely to lose sight of the practical dimension of cognition when one talks of understanding than when one talks in terms of knowledge.
- There is also a tendency to think of knowledge as an all-or-nothing matter, whereas understanding is naturally regarded as *admitting of many forms and degrees*. As we will see below, such variations of degree are an important and biologically salient aspect of the teleopragmatic view. From an evolutionary perspective, they are just what one would expect.
- Some understanding is *essentially mutual and reciprocal*. In such cases, the practical engagement afforded by the relevant understanding depends on the

reciprocal causal interface between the understander and the aspect of the world that is being understood. A clear example is that of the mating dance of jumping spiders. Such spiders go through a complex multistage exchange of signal, response, countersignal and counterresponse with each action provoking just the response that in turn stimulates the appropriate next signal from the originator. This entire ritual *pas de deux* is choreographed on the basis of innately programmed mechanisms that are ready to run in spiders that have never before seen an opposite member of their species. What is crucial for present purposes is the respect in which the spider's biologically based understanding is embedded in those aspects of its sensorimotor interface that support its reciprocal causal engagement with its mate in what might be called a "dynamic lock-and-key" system.

- Some understanding is *constitutive*, or *self-constitutive*; the act or process of understanding in such cases at least partially underlies or helps create that which is understood. This aspect of understanding is connected with the reciprocal causal feature just discussed, but it goes beyond it insofar as it involves an ontological interdependence between understander and understood in such cases. Consider another example of social engagement, in this case human rather than arachnid. Long-married couples come to understand each other in rich and subtle ways, but more importantly each member of such a couple will often have developed into a person who plays a distinct and specific set of roles in the context of the marriage. Not only do such couples understand each other in the standard cognitive sense, but they understand each other in the more ontological sense that their respective natures and personalities are supported and maintained by their ongoing mutual engagement. In that latter respect, they "understand" each other in a way that reflects the literal etymology of the word; they "stand under" and support each others' distinctive personal nature. Some may think talk of "understanding" in this sense is just a verbal joke or a play on words, but I believe it is more than that, and as we will see below it may have relevance for explaining conscious minds as self-understanding systems.

For present purposes I need not give a comprehensive account of the teleo-pragmatic theory. I want to focus instead on one important aspect of the view, namely, the many variations in the type and sophistication of content that it permits, or indeed predicts. Organisms can be informed about or understand some feature of their world to many varying degrees. All such cases involve nonrandom correlation or tracking between the organism and the features about which it is informed, as well as ways in which that match enhances the organism's success in achieving its goals in ways that are

specifically responsive to those features. But within that general framework there is room for a wide diversity of cases.

Two parameters are of particular importance: the range of possible applications that the relevant information or understanding affords, and the degree to which that specific information is capable of being integrated with other information in appropriate content-sensitive ways. At one extreme would be cases in which some adaptive but narrow, fixed, and inflexible response is triggered and guided by the relevant information. The organism is able to apply the information, but only in the limited way provided by that single mechanism, and it is not able to bring it into play with other things it understands or about which it is informed. The relevant information or understanding involved in such a case will be crude in content insofar as it is locked into the specific isolated behavior-controlling process. At the other pole are cases in which the organism has a capacity to bring the information to bear in a more or less open-ended range of adaptive applications, including many that depend on combining it with other information in ways sensitive and appropriate to their respective contents. The wider the range of the possible applications and the richer the field of other contents within whose context the particular item's significance is actively linked, the more the organism can be said to understand the meaning of the relevant information.

Contrast, for example, the state you or I are in when we believe that a fly is buzzing round the table with the state that a frog is in when a similar fly is stimulating its retina in ways that link with the tongue-control system in its amphibian brain. The frog's visual system processes information about the location, direction, and speed of the fly as well as detecting its likelihood of being "foodlike," with the result that it triggers and guides the frog's tongue to intercept the fly in midflight and ingest it. The information the frog acquires and what it understands about the situation are undoubtedly adaptive and put to practical use in successful pursuit of its goals. Its behavior is informationally sensitive in appropriate ways. Yet the content of its intentional state is crude. It has only a limited understanding of what it sees and the content of the information it acquires is correspondingly crude in terms of both of the key parameters. The range of possible applications that the frog can make of that information is very narrow, and the frog has little or no ability to integrate it with other contents in inferentially appropriate ways or to situate it within a larger context of meaning. However, when you or I believe *that a fly is buzzing around the table*, there are all sorts of ways we can adapt our behavior in response to that information and a nearly open-ended range of other beliefs and

intentional contents with which we can connect it to better appreciate its significance.

It is important to note that the transition from states with crude content to those with more sophisticated intentionality is not a move from cases in which having information is linked with application to others that involve a purely intellectual way of being informed that is unconnected to application. Rather, the progression is from cases in which an organism has information that can be only narrowly applied in fixed ways, to cases in which the organism can apply the information in a far wider and diverse range of ways including many that require connecting it in content-sensitive ways to other information and other intentional contents. To put it in a slogan framed in terms of the hoary know-that/know-how dichotomy: knowing-that is still a form of knowing-how, just a much richer kind of know-how.

As noted above, what is true of knowledge and information is also true of understanding. With understanding, the practical dimension seems central from the start; what and how we understand some object, process, person, or other aspect of the world is inextricably bound up with the nature of our engagement with it. Moreover, understanding clearly comes in many forms, which each in its own way affords some partial mode of access for successful interaction.

Though much more would need to be said to spell out the complete teleo-pragmatic story, I hope what I have offered suffices to explain how such an account naturally generates a continuum of content and varying degrees of intentional sophistication based in part on variations along our two key parameters: range of application and richness of content-sensitive links to other contents.

III Implicit Self-Understanding and the Reflexive Loop

We are now in a position to ask how all this applies to the higher-order view of consciousness. The standard HOT and HOP models think of meta-intentional content as a highly sophisticated form of intentionality that arises only at the level of propositional states, specifically propositional meta-states. Whether thoughtlike or perception-like in mode, they are taken to have sophisticated that-clause type contents. The alternative I am proposing allows that much of the meta-intentionality that underlies and supports consciousness is much less sophisticated in content and often embedded in procedures that play important intramental roles in producing the sorts of self-modulation, self-regulation, and self-transformation that are essential to the structure and dynamics of consciousness. Insofar as those procedures are

specifically adapted to the intentional nature and content of the states and processes to which they apply, they can be taken to constitute the possession of information about or understanding of those mental items and their contents. But the content of the relevant meta-understanding may fall far short of the propositional level in its limited range of application and the poverty of the field of other contents to which it is sensitive.

Consider a parallel case involving learning rather than consciousness (Van Gulick 1988). An organism's ability to learn gives it the ability to adapt its responses not only to general regularities of its environment of the sort that might be gained through natural selection, but also to idiosyncratic patterns and regularities of its local situation. This obviously has lots of biological advantage. But the capacity to learn itself presupposes the existence of underlying processes and mechanisms that make learning possible, and they in turn must have been adaptively shaped to fit the causal structure and contingencies of the organism's world if they are to do their job successfully (Lorenz 1977). Most importantly for present purposes, those learning mechanisms will be effective only insofar as they have been adapted to fit the mental and intentional properties of the processes they modify. To be successful, a learning process must at least to some degree be informed about or implicitly understand the psychological significance and intentional contents of the processes it modifies.

Consider the famous example of Garcia's rats that learned not to eat foods whose ingestion had been after several hours followed by the injection of a nausea-inducing drug (Garcia and Koelling 1967). They did not acquire similar learned aversions to foods that were followed after a similar interval by equally aversive electric shocks. The relevant learning mechanism is tuned or adapted to the fact that unwholesome foods cause nausea but may not do so until several hours have passed. The learning mechanisms that underlie learned food aversions in rats are thus adapted to the causal fact of that delay. They are also adapted to the psychological significance of the feeding control systems that they modify. The learning mechanism does not make a random change in response to the induced nausea; rather it changes the feeding mechanism in just the way that alters the organism's future response to the very food that was eaten a few hours before the onset of illness. The learning mechanism is thus "tuned" in at least two respects: to the causal structure of the external world and to the mental or intentional structure of the organism's own internal organization. Its job indeed is to better harmonize the two in the context of the individual organism's local situation.

Insofar as at least some capacity to learn goes way down the phylogenetic scale, so too does the existence of at least implicit meta-intentional under-

standing. Indeed, the possibility of minds of more sophisticated kinds depends on the evolution and existence of such less sophisticated forms of meta-intentional understanding to provide the system of underlying processes that make more complex types of mind possible. You cannot have beliefs without a substantial capacity to learn; indeed beliefs are partly constituted by their links to learning. But the ability to learn, as we have just seen, presupposes and depends on the existence of subpropositional procedurally embedded forms of meta-intentional understanding. So beliefs must depend on the existence of such understanding as well.

From a teleopragmatic perspective, reflexive meta-intentionality is a pervasive and major feature of the mental domain. Once one recognizes the diversity of degrees and forms in which it can occur implicitly as well as explicitly, one finds it playing a key role in all sorts of contexts and organisms that one might not ordinarily associate with meta-intentionality. Rather than coming in only at the latest and most sophisticated levels of evolution and mentation, one finds some degree of reflexive meta-intentionality playing an important role at both the lower levels of the phylogenetic scale and the lower organizational levels of complex mental systems.

Indeed, the self-reflexive loop appears to be one of the basic processes driving the evolution of mind. Organisms come to better understand their worlds by coming to better understand themselves and the ways in which their own structures engage their worlds. The organism–world interface evolves in a reflexive feedback cycle that continuously tunes their engagement, enabling the organism to more fully and completely understand the reality of its world and itself. Thus looking at consciousness from the perspective of that larger reflexive process should help us better understand both its nature and its place in the natural realm.

IV The HOGS Model

Having gotten a basic grasp on the reflexive loop, we are now ready to consider the alternative version of the HO theory promised at the outset. The aim is to show that full-blown phenomenal consciousness depends on a high degree of implicit self-understanding and meta-intentionality that is built into the processes that underlie it and generate its organization and its exquisitely content-sensitive dynamics. The central claim is that we are able to be conscious in the phenomenal sense at the personal level (or whole organism level) only because we embody such a rich store of implicit and procedural self-understanding at the subpersonal level.

The basic idea of what I have elsewhere labeled the HOGS or higher-order global state model of consciousness is that transforming a nonconscious state into a conscious one is a process of *recruiting* it into a *globally integrated complex* whose organization and intentional content embodies a heightened degree of *reflexive self-awareness*. The meta-intentional content, on the HOGS model, is realized not by a distinct and separate external meta-state but rather by the organization of the complex global state itself, and the object state is a component of the global state into which it has been recruited.

The motivation for the HOGS model derives from many sources, but two ideas are especially important.

The first is Daniel Dennett's metaphor of consciousness as a form of "cerebral celebrity." The idea is that conscious mental states (and conscious intentional contents) are more widely and powerfully influential within the minds in which they occur. They are more broadly and richly connected, and their contents are more widely accessible by other subsystems throughout their containing mind-brain. Similar ideas are found in so-called global workspace models of consciousness such as that of Bernard Baars (1988, 1997). Conscious states are thus typically able to have more effect on the overall system's evolving mental state and on the organism's behavior.

Scientific evidence about the neural correlates of consciousness provides a second motive for the HOGS model. Based on imaging studies, most current empirical models do not treat consciousness as localized in any specific module or subsystem of the brain. Though the evidence is still far from certain, there does not appear to be any privileged region in which conscious events occur. Rather, the neural substrates seem to be distributed globally across the brain, and to involve large-scale patterns of integrated and reciprocal activation among disparate regions, whether they are entrained through synchronous oscillation or coherently bound through some other, and perhaps less direct, physical means.

The HOGS model departs from more standard HO models in just the ways proposed in section II. It weakens the distinctness assumption and explicates the meta-intentional aspect of consciousness in a more implicit way. The latter link is obvious and immediate. Insofar as the relevant meta-intentionality is located in subpersonal procedures that underlie phenomenal consciousness, it is realized implicitly by the way in which those procedures embody the relevant understanding in their organization, much as do the procedures that underlie our capacity to learn.

Divergence from the distinctness assumption similarly results from the systemic requirements for supporting consciousness and its dynamics. As

noted above, standard HOT and HOP models treat the transformation of a nonconscious mental state into a conscious one as a matter of adding a distinct and separate meta-state directed at it. The object state gains no new intrinsic properties. The transformation to conscious status involves only *extrinsic relational* changes; indeed it involves only added *intentional relations* directed at it by the conscious-making meta-state.

By contrast, the HOGS model of the basic transformation does not isolate the meta-intentional element in a distinct and separate state, but rather locates it in the organization of the global state into which the formerly nonconscious state is recruited and of which it itself becomes a component. Thus the distinctness principle is weakened, but not totally abandoned. Although the object state is retained as a component of the global state, it is typically altered somewhat in the process. In that respect the transition to conscious status involves some changes in the state's intrinsic properties, as well as its gaining a new systemic significance in virtue of the larger active context into which it is recruited. It's the same state, yet importantly different. Indeed, as we will see below the question of whether it is the same state or a different state gets somewhat blurry, and the answer largely turns on how we individuate states.

To get a better grasp of the HOGS model, consider a specific case such as the transition from nonconscious visual perception to conscious visual experience. The nonconscious state may register various features of the visual scene and make that information available for guiding at least some forms of adaptive behavior. It might for example register the presence of a bluish sphere just off to the left of my line of sight, and generate behavior that is situationally apt toward an object with those features. If that visual perception then gets recruited into the global state that is the transient substrate of my conscious experience, its contents [bluish, spherical, and off to the left] will all be preserved, but they will acquire a different sort of significance as features of an object in my phenomenal, experienced, lived-in world.

This is not accomplished by re-representing the relevant information again in another module of the mind or brain, and certainly not in the infamous "Cartesian theater." Rather, it involves incorporating the original representation into a new dynamic context that integrates it with a larger, richer network of other active states that bind it together in content-linking ways that reflect the unity and continuity of both the phenomenal world of objects and of the self to whom they are present in experience, about which more will be said below in section V, which also makes clear in just what respects this model is a higher-order model and not merely a global state theory.

This larger network of connections underlies both the richer and wider web of influences that constitute its enhanced cerebral celebrity and also its role as an intentional constituent of an integrated and unified phenomenal reality.⁴ The relevant network of connections is itself in turn realized or underlain by the global patterns of coherent reciprocal activation among the widely distributed regions that are constantly forming and reforming as the transient substrates of conscious experience. Or at least, that is what the HOGS model proposes.

V What's Higher Order about HOGS?

At this point, some readers may find the basic idea of the HOGS model attractive and even agree that a mental state's becoming conscious is largely a matter of its being recruited into a globally active and integrated state, yet doubt that the HOGS model is really a version of a higher-order theory. What, they might ask, is distinctly or essentially higher-order about it? Indeed why not just treat it as a global state (GS) model and drop the higher-order aspect altogether? Is the latter really needed, and does it do any real explanatory work? To speak in terms of acronyms, "Where is the HO in the HOGS? Why not just opt for a GS model of consciousness instead?"

The answer lies with the *particular nature* of the relevant global states and *how* they underlie conscious mentality. As noted above, the HOGS model treats the nonconscious–conscious transformation as a process of recruitment into a globally integrated complex *whose organization and intentional content embodies a heightened degree of reflexive self-awareness*. The requisite globally integrated states are able to serve as the substrate of consciousness only because their specific form of integration realizes and embodies the sort of reflexive intentionality and self-understanding that it does—or so the HOGS model claims in general. However, to make that claim plausible we must give some more specific and detailed account of just *how* that global organization embodies reflexive intentionality and *how* in doing so it contributes to the realization of distinctively conscious forms of mentality. Thus I will aim to do just that for the balance of this section.

The goal is to spell out some of the major explanatory links between the three main dimensions of the HOGS model: global integration, reflexive intentionality, and consciousness. In what ways must global integration involve reflexive intentionality in order to contribute to the realization of conscious mental states?

The nature of the link depends in large part on *which sort* of consciousness one is concerned to explain. If one's concern is with the commonsense notion

of a conscious mental state as simply a “state one is aware of being in” then the need for reflexive mentality is immediate. The substrate for being conscious in that sense, whether globally integrated or otherwise, would of necessity have to produce reflexive intentional content directed at the state that is being made conscious. If it did not, the basic conditions for being aware of the object state would not be met, and the state could not be a conscious one in the relevant sense.

However, if one’s concern is with states that are conscious in a phenomenal or “what-it’s-like” sense, the need for reflexive intentionality is less evident. It is not obvious why an explanation of how global integration contributes to phenomenal consciousness should need to pass through or rely on an account of how that integration contributes to reflexive intentionality. Though they are less obvious, those latter links between phenomenal consciousness and reflexive mentality are crucial to understanding the higher-order aspect of the HOGS model.⁵

Looking at four interrelated aspects of phenomenal consciousness may help reveal the connections. Each is a salient and essential feature of our phenomenal and experiential mental life, and each can be seen on examination to involve an important reflexive meta-intentional aspect, which might in turn be realized in part through global integration. The four key phenomenal features are:

- the presentational aspect in which objects seem immediately present to us in experience;
- the unity of experience;
- the autopoietic flow and dynamics of experience; and
- the affective and pragmatic aspect of experiential content.

There are many other important features of phenomenal consciousness, but these four are all quite central, and if we can see how they involve or require reflexive meta-intentionality that should go a long way toward making the case on behalf of the higher-order aspect of the HOGS model. I will consider each of the four in turn, though as will soon become apparent they are strongly interconnected.

Experiential presence Objects, scenes, and events are not merely *represented* in experience, they are phenomenally *present* to us. They *appear* to us, and we experience them as real and present to us here and now. We experience trees and chairs themselves as well as the flow of water from the tap as the objects and events they are, that is, we experience them as real parts of the world directly there before us. Whatever one may make of direct

realism as an epistemological or philosophical theory of perception, it is hard to deny that the phenomenal feel of normal experience is directly realist, and a large part of that realism consists in the way in which the objects of experience are *experienced as present* to us here and now.

However, the notions of presence and appearance require a second pole. They are coherent only insofar as objects are *present to* some conscious subject, or insofar as they *appear to* some self that experiences them in relation to itself. The notion of appearance thus implicitly supposes the existence of an experiential subject to which objects appear and to which they are present (Alston 1999; Kriegel 2003c).

It is partly in this sense that the intentional content of experience is reflexive. Insofar as that content involves objects being present or appearing to us, it essentially incorporates a reflexive aspect. The claim is not that reference to the experiencing self is an explicit element of the content of experience, and certainly not that it requires the sort of content that would be involved if one were to think to oneself, “I am having a visual experience of a cat sleeping on a carpet.” It is rather a claim about the intentional framework of experience. Insofar as that framework enables us to experience objects as present and appearing to us, its very structure must implicitly incorporate at least some understanding of the self or subject to whom objects can be present and appear. The required intentional structure could not support the experience of presence without at least implicitly incorporating some grasp of there being a self or subject to whom objects appear. Thus some degree of reflexive intentionality seems inescapably required by the experiential presence of phenomenal consciousness.

Having made that claim, let me make it clear that the relevant reflexive element need not involve explicit, sophisticated, highly conceptualized understanding. The teleoprismatic view allows for many types and degrees of understanding, and the minimum grasp of the reflexive intentionality required for the experience of presence needs to be set low enough to allow for the experiential lives of young children and nonhuman animals. Obviously, two-year-old toddlers and Fido the family dog do not have articulate propositional beliefs about themselves as conscious subjects of experience. Nonetheless, experience is not blind, and their capacity to experience a world of objects immediately present to them in itself manifests at least some implicit grasp of the reflexive structure of experience.

Unity of experience Two interdependent unities pervade the realm of phenomenal experience: the unity of the experienced world and the unity of the experiencing self.

We do not normally experience isolated sensory qualities or objects; we do not typically see mere colored regions, hear mere sounds, or feel mere shapes. Our phenomenal experience is of a world of independently existing, relatively stable objects located in a unified arena of space and time within which we ourselves are located as conscious perceivers. The objects we experience and the world within which they exist have a phenomenal and intentional “thickness,” not merely in the literal three-dimensional sense but also in the richness of their properties and relations and in the diversity and density of the connections that hold among them in constituting “a world” of objects.

On the correlative side of the experiential relation is the more or less unified self that coheres as a single subject of experience both at a time and diachronically, bound by links of content, inference, access, and memory. Our experience is not a mere sequence of distinct experiential episodes; rather it is the unfolding experiential flow of an ongoing self or subject, and its being experienced as such is a key part of its phenomenal feel as well as its structure. Nor is our experience at a moment the experience of isolated sensory items. When I stop to make a cup of tea, I hear the whistle of the kettle that I see before me on the stove as I also feel my hand reach out to turn off the burner on which it sits. All those experiences cohere in the moment as experiences of mine.

The two experiential unities—that of world and that of self—are mutually interdependent. The experienced world coheres as a unity largely from the point of view of a self located in it. Objects are experienced in terms of the relations they bear to each other from the perspective of the self, itself located in space and time. I experience the trees waving in the wind as out across the yard beyond the window through which I look, the pines to the left of the spruce and much taller than they were a few years back. My experience of them as *real objects of my world* is bound up with the spatiotemporal location and ongoing unity of the selflike point of view from which they are present to me. Conversely, the experiential unity of self coheres largely in relation to the world of objects present to it. The unity of the self is the coherent ongoing unity of a perspective on that unified world, indeed of a perspective located within it.

To put the matter in loosely Kantian terms, our experience of self is very much an experience of self-in-relation-to-world, as is our experience of world essentially an experience of world-in-relation-to-self. We experience each in relation to the other; so it should come as no surprise that their respective phenomenal unities are equally interdependent.

These two unities, and the thickness and richness of connection that underlie them, are just the sorts of intentional phenomenal features that

might be produced through global integration. It might occur within a sensory modality when seen colors, shapes, sheen, and use all get bound together as features of a single object. Or perhaps it may bind across sense modalities when sights and sounds are integrated as aspects of a single situation, as when the kettle on the fire sings while the toast pops up from the toaster off to the left. Or it may involve more distant forms of binding as between present experiences and past memories, or between current perceptual inputs and background information about the relevant actors or objects. Such integrative processes must be at work whenever as a matter of experience I see a small white sugar cube on a plate, reach for my ringing alarm clock, or recognize the man in the shop as the person who spoke up at last week's council meeting (Metzinger 1995; Cleermans 2003).

Given that interdependence, both unities obviously embody a substantial measure of reflexive intentionality. The unity of the self pervades both, and each requires a heightened degree of reflexive self-understanding. Given the intentional structure of phenomenal experience, one could not be able to have such experience without also having the necessary sorts of understanding that it presupposes. And much of that understanding with respect to the unity and reality of the experienced world turns out to be at least in part self-understanding and to be reflexively bound up with the phenomenal unity of the self. You can experience your eyeglasses sitting on a blank sheet of yellow paper only if you already grasp the reality of the world within which those glasses exist as objects. And understanding that reality, both of the world and of the things in it, is in part a matter of understanding them in relation to one's experience of them from one's perspective as an ongoing, continuing self within that world. Thus, given the interdependence of the two unities, there can be little doubt that both involve a substantial degree of reflexive higher-order intentionality and self-understanding. Insofar as the unity of the self requires higher-order intentionality, and the unity of the experienced world is in turn bound up with the unity of the self, it follows that the unity of the phenomenal world as well depends on reflexive higher-order intentionality.

Autopoietic dynamics and flow Experience is ever changing, and reflexive intentionality manifests itself in the dynamics of its flow as much as it does in its synchronic unity and presence. In part this reflects the two unities we have just been discussing. Our experience is that of an ongoing self within an ongoing world. So we would expect the unities of world and self to be expressed in the order of experience, and indeed they are. But consciousness is also active and self-organizing; its state and content at one moment typi-

cally shape what happens next. To use the apt Greek-derived term that has of late found its way from embryology into the science of mind, consciousness is an “autopoietic” process, that is, one that is self-making, self-creating, and self-organizing (Lorenz 1977; Varela, Thompson, and Rosch 1991; Van Gulick 2003).

The process is self-organizing and reflexive not only in the direct sense that early stages bring about and produce later ones in an orderly content-related way, but also in the sense that the intentional structure within which those patterns and relations are produced is essentially that of a self—indeed that of an active self—one that shapes and controls its own phenomenal flow, and which experiences that control as part of the feel of its own mode of conscious awareness. My visual experience can change because the cat wakes up and moves or because the doorbell rings, but it also changes when I shift my eyes or even the focus of my visual attention, as well as when I call up an image of a meal I ate last summer with friends in Florence, or when I think my way through the argument of this paragraph. In every case, my active experiential state at one moment is a major cause of what I experience next, and that fact of control is itself a key feature of the felt quality of my own phenomenology.

The autopoietic dynamics of experience provide obvious links to the two features of mind we have been aiming to connect: global integration and reflexive meta-intentionality.

The sort of control system required for experience to modulate and transform its own ongoing state as it does would clearly require a rich network of integrating connections and content-sensitive interactions to move it from one phenomenal intentional state to its appropriate target successor. On the HOGS model, what is needed is for each episodic stage in the transient flow of globally integrated states to be capable of organizing its successor by selectively recruiting and binding together its intentional constituents, continually adding and deleting elements from its global ensemble of reciprocally activated components. In many cases, of course, outside factors and stimuli will play a big part as well—doorbells do ring unexpectedly—but even in those cases the prior phenomenal state of the system and the globally integrated base on which it depends will still be major factors in shaping its subsequent evolution.

These experiential dynamics have an important reflexive aspect as well. The intentional structure within which one stage of experience produces another does not rely merely on links between the contents of those states but also essentially on their being states of one and same conscious subject or self. I aim to call to mind the trattoria in which last summer’s dinner took

place, and I thereby change the flow of my own visual phenomenology. The transition is apt not only because the image of the restaurant matches the content of what was aimed at, but also because it occurs as part of *my* experiential flow. Moreover, that reflexive aspect—*my* acting to shape *my* experience—is itself a central aspect of the intentionality and felt phenomenology of the dynamics of experiential flow. I experience it not as a passive observer but as an active agent shaping my own experiential path. Nor is it likely my experience would have the autopoietic control power that it does if it lacked that reflexive meta-intentional structure.

Philosophers concerned with phenomenal consciousness sometimes focus solely on its static or synchronic features, but its dynamic diachronic flow is also a central feature of how we experience ourselves and our world.

Affective pragmatic aspect of experiential content The role of the conscious mind as active autopoietic agent provides a transition to our fourth and final phenomenal feature, one that ties in nicely with the general teleopragmatic view of mind. From that perspective, mind is seen to have an essential pragmatic dimension, and notions such as information, understanding, and awareness are treated as having an important practical and contextual aspect. If minds evolved to enable organisms to successfully engage their worlds, then one would expect conscious minds to reflect that practical aspect in the intentional content and phenomenology of their experience. Indeed, one finds exactly that.

Although we may sometimes experience objects in ways that have little if any relation to our purposes or interests, our everyday experience is typically infused with practical significance. The intentional content of experience is shot through with modes of presenting and categorizing objects and situations in terms of their relevance to our needs or interests and the opportunities or affordances they present for interaction or engagement. This is especially true in social and culturally constructed contexts, but it applies across the board to pretty much all domains of experience. I see the jetway before me as a path to my plane, I hear the bark of the dog as a warning, and I feel the keyboard under my fingers as a means to produce the text of this sentence. Fresh coffee, angry scowls, can-openers, and even the layout of trees in a wood are typically present in experience in ways that at least partly incorporate some implicit recognition of their affective significance and their possibilities for practical engagement.

Once again, there are plausible links with the two key aspects of the HOGS model. Like our three earlier phenomenal features, the affective pragmatic

aspect of experience bears interesting explanatory connections both to global integration and to reflexive intentionality.

As we noted above in section II, the teleoprismatic view regards the move toward a more objective mode of intentionality and representation not as a move away from pragmatic engagement, but rather as a matter of setting representations within a wider and richer context that in fact allows for a larger and more open-ended range of possible relevant engagement. Our conscious experience of the world presents objects to us with a wealth of open possibilities. When I see the box on my desk, there is almost no end to the understanding ways I might respond to it. I can open it, throw it away, stack books on top of it, cover a stain with it, use it as a color sample to match some paint, drum a rhythm on its lid, or take it to class as a prop for a lecture on the *Investigations*. Indeed it is this very openness to such limitless possibilities of engagement that largely constitutes or contributes to the felt reality and objectivity of the experienced world (Merleau-Ponty 1945).

Producing an understanding system with such flexibility of response and application would obviously require a sophisticated underlying system of connections to integrate all the various modules and action-guiding subsystems that would have to interface and harmonize to realize such open-ended possibilities for purposive engagement. The most plausible way of freeing up practical understanding encapsulated within specific modules is by integrating and harmonizing those modules with others to increase the sphere within which that understanding might be applied. Insofar as phenomenal experience shows a high degree of such pragmatic openness, it likely requires just the sort of global integration the HOGS model proposes.

Reflexive intentionality comes in again as well in many ways, some obvious and some less so. Insofar as we experience objects in relation to our needs and wants and to our embodied means of interacting with them, the content of such experience must obviously incorporate some measure of reflexive intentionality. Understanding objects and situations in relation to our goals requires having at least some implicit understanding of those goals and of them as one's own.

However, the need to support pragmatic engagement may involve reflexive intentionality in other ways that are more subtle but potentially deeper and more important.

The processes through which organisms successfully engage and understand the world typically loop back and forth between the organism and its context, with each side generating signals and replies that pass back and forth in resonant harmonic cycles. Insofar as phenomenal consciousness in

part provides an interface for real-time causal interaction with the world, it likely generates not only efferent signals but anticipations and expectations of the reafferent response that it will receive back from the world. Successful engagement will typically require a good match between expectation and actual reafferent, or at least one good enough to establish an effective harmony between the two coupled sides of the interaction. Imagine how ineffective and clumsy one is when one's inner timing or sense of distance or direction fails to match the sensory feedback from one's actions. Try for example to cut a paper pattern with scissors using only a mirror-reflected image to guide your own action. One can learn to do it, but most of us at first can manage only some slow and awkward snipping.

Indeed, some have proposed (Gray 1995, 2004) that this system for matching predicted and actual reafferents is the core basis of conscious experience. However, even if one is not willing to go quite that far, it seems likely that such loops play at least some role in producing the felt experience of being and acting in the world. If so, then these loops present an important way in which consciousness, or at least one key aspect of consciousness, can be explained as an example of the teleopragmatic process through which organisms come to better understand objects in their worlds by coming to better understand the ways those objects affect them, as well as coming to better understand themselves by gaining a better understanding of their effects on those objects.

An even more radical option is suggested by another teleopragmatic feature of experience discussed back in section II. We noted there that understanding sometimes involves cases of reciprocal constitution, in which the understander and that which it understands mutually depend on each other and their resonant interaction for their nature and even for their existence as the sorts of things they are. Examples in the social domain seem most evident, but nonsocial cases are also surely possible. Perhaps phenomenal consciousness itself presents a crucial case of such "bootstrapping." If conscious minds are active autopoietic systems and they experience their objects in terms of their pragmatic potential for reciprocal engagement, then the existence of conscious minds as self-understanding systems may depend on a reflexive version of such mutual reciprocal constitution.

Conscious minds, on such a model, would have the properties they do in part because they understand themselves as having those very properties; they would in a more than punning sense be self-understanding systems in two connected senses: they would understand themselves in the *constitutive* sense by understanding themselves in the more ordinary *cognitive pragmatic* sense. The idea is suggestive if radical, and perhaps it could be successfully

developed, though perhaps not. For now, I leave it as just a suggestion whose exploration will have to wait another day.

Let us sum up what we have learned from our four phenomenal features about how the reflexive and higher-order aspects of the HOGS model might work together to help explain consciousness, and in particular about why one needs the HO in HOGS. We found that each of the four involves explanatory links in both directions. Global integration is required in every case to provide the rich network of connections that underlies the relevant phenomenal feature, whether it be the unity of the experienced world, the self-organizing dynamics of experience, or the open-ended pragmatics associated with the phenomenal experience of an objective reality. Similarly, we found that they all involved an essential, if sometimes implicit, aspect of reflexive higher-order intentionality, whether in the phenomenal feature of presence as presence *to self*, the unity of the experienced world *in relation to self*, or the autopoietic dynamics through which consciousness understands and organizes itself as the experience of a unified ongoing self. Moreover, in every case we saw that the two aspects of the HOGS model were not only both relevant but that they reinforced each other in explanatorily important ways.

VI HOGS and the Objections to HOT and HOP

Back in section I, we listed six common objections that have been raised against the standard HOT and HOP models of consciousness. I hope in the last few sections to have given a plausible account of the HOGS model and what it involves. Thus I will close by returning to those original objections, and asking how the HOGS model might deal with those challenges and whether it can give replies that are not otherwise available to more standard defenders of the HO point of view. Indeed, I believe it can.

Recall the six objections. HOT and HOP theories were jointly faulted on the basis of six difficulties:

- the “too fancy” objection
- the extra conditions problem
- the generality problem
- the phenomenal adequacy issue
- the “stranded qualia” problem
- the error/illusion problem

I believe the HOGS model has something distinctive and useful to say in response to each. In some cases, the relevant objection does not even really

apply to the HOGS model and thus can be simply dismissed. In others, the general objection is still relevant, but the HOGS model has additional resources to answer the challenge.

The generality problem and the extra conditions problem are two that fail to apply against the HOGS model. They get their bite against more standard HO models that explicate the nature of a conscious mental state as a relational matter of there being separate higher-order states directed at it. Whether the relevant HO states are thoughtlike or perception-like, it is still their distinct existence that supposedly explains the original state's transformation from nonconscious to conscious status. Thus the objections: Why doesn't the same apply to stones and pencils we perceive or think of? And why doesn't the transformation take place if the HO state is produced by inferential means?

On the HOGS model, the nonconscious-conscious transformation is a matter of its being recruited into a globally integrated state that crucially embodies a heightened degree of higher-order reflexive intentionality, and not merely a matter of having a separate HO state directed at it. Pencils and stones are never recruited into any such globally integrated states—indeed the idea of their being so does not even really make any sense. So the generality problem does not apply to the HOGS model. Similarly, there is no need to justify extra conditions such as the noninferential requirement to rule out counterexamples like those that would otherwise threaten more standard HO theories. Again the relevant cases do not involve the sort of recruitment the HOGS model requires, so they do not count as even apparent counterexamples to it, and thus there is no need to motivate extra conditions to exclude them.

The “too fancy” objection is also easily met by the HOGS model, especially if the model is placed within the larger context of the teleopragmatic view of mind, with its varying degrees of understanding and intentional content. On the HOGS model, much of the relevant reflexive higher-order intentionality is implicitly embedded in the structure and organization of the globally integrated states, and its content need not be as sophisticated as that associated with explicit higher-order propositional attitudes of the sort that seem too fancy for attributing to dogs or toddlers. No doubt the intentional contents associated with the experiences of infants and cats differ from those associated with adult human experience, and their conscious lives involve a less sophisticated understanding of themselves as mental subjects. Yet that provides no reason to doubt they could have the sort of implicit self-understanding that the HOGS model builds into the organization and structure of experience.

The phenomenal adequacy issue poses a genuine challenge to the HOGS model, as it would to any model of consciousness. No theory can be considered satisfactory as a general account of consciousness unless it successfully addresses the phenomenal and qualitative aspects of our conscious mental life, though what would count as success is itself an open question.

As noted in section I, standard HO models seem most plausible at explaining the commonsense notion of a conscious state as a “mental state of which we are aware.” Their supporters believe such models can be extended or supplemented with an independent account of qualia to provide an explanation as well of states that are conscious in the phenomenal or “what-it’s-like” sense. Perhaps they can indeed do so, but their attempts so far have failed to produce widespread optimism on that score. The HOGS model, though itself relatively new, may hold more promise for helping to explain states that are conscious in the problematic phenomenal sense. Indeed, as I tried to show in section V, the two key aspects of the HOGS model can be used together to build explanatory connections that help explain the nature and basis of key features of phenomenal experience such as its unity, presence, and autopoietic dynamics. I do not suppose the HOGS model will in itself explain everything we want to understand about phenomenal consciousness; given my telepragmatic commitment to explanatory pluralism I doubt any model could suffice for complete understanding. Nor is the HOGS model intended as a reductive theory; it aims only to provide some insights into what phenomenal consciousness is and how it could be realized that might otherwise be missed. And if it does so, that might be enough to count as at least a partial explanatory success.

Whatever general success or limits it may have in explaining consciousness, the HOGS model does not have to deal with the specific problem of “stranded qualia” that gets raised against more standard HO theories. The HOGS model does not separate qualia and “what-it’s-likeness” in the way that generates the objection. In particular, it does not treat qualia as *properties of mental states* at all, neither of lower-level states nor of higher-level states directed at them. Rather, the HOGS model treats qualia as *properties of experienced objects*, and thus it regards the problem of qualia as explanatorily posterior to the project of explaining how it is possible to experience a world of objects in the first place. Only when we have made progress on that prior issue is one likely to get a sense of how qualia might come in. If one gets off on the wrong foot by treating qualia as properties of our mental states rather than as properties of experienced objects, one may well fall into further confusions; indeed that may be how the “stranded qualia” problem

arises. In any case, the specific problem does not apply to the HOGS model.

Finally, there is the error/illusion problem. Standard HO models seemed to have no satisfactory choice about what would determine the specific “what-it’s-likeness,” if any, present in such cases. Nor did they have any non-ad hoc bases for ruling out the possibility of such higher-order errors or illusions. The HOGS model by contrast has a natural way to exclude them. On standard HO models the meta-state is a separate and distinct state directed at the object state that it makes conscious. Given their separate and independent existence, there must be at least the logical possibility of a mismatch between them. Hence the problem.

However, on the HOGS model the nonconscious-conscious transformation is a matter of recruiting the original state into a globally integrated complex, which retains it, or at least some near successor state, as a component of the overall integrated state. If one thinks in terms of recruitment, the possibility of a problematic mismatch does not arise. The original state is not the mere object of a separate higher-order thought or perception, but is itself an essential component of the integrated state, whose higher-order aspect is embodied in its organization through which it unites its component states into the experience of a real and ongoing world of objects. I do not claim that the HOGS model rules out all possibility of error about what qualities one encounters in experience. There are many interpretative metaprocesses that generate a diversity of explicit beliefs and attitudes about our mental lives. Surely they can sometimes go wrong, and they can probably err even about qualitative features of experience. But they would be mistakes of a different sort than those that provoked the error/illusion objection against standard HO models. The HOGS model does not attempt to explain the specific “what-it’s-likeness” of experience in terms of the contents of any such explicit meta-states, and so the possibility of error or illusion with respect to them causes it no special problems or dilemmas.

The HOGS model thus seems able to avoid or answer all six of the original objections raised against more standard HO theories. That should surely count in its favor, though of course it does not by itself establish the validity of the HOGS model nor even its overall superiority to HOT and HOP models. There could conversely be other objections that specifically tell against the HOGS model but not against the more standard HO views. Thus it is too soon to reach any conclusive judgment.

Nonetheless, I hope to have shown that the HOGS model offers new and interesting explanatory options, in particular for understanding how phenomenal consciousness might be linked to other functional and intentional

processes, such as those involving global integration and reflexive intentionality. The HOGS model provides an alternative slant on higher-order theory that clearly merits further consideration.

Notes

1. Those interested in their comparative merits can see Van Gulick 2001, which offers an evaluation that finds no clear winner; Lycan 2004, which argues for the relative superiority of HOP theory over HOT; and Carruthers 2000, which defends the opposite.
2. The appeal to teleology carries no nonnaturalistic implication; the only sort of teleology required is that which can be accommodated within standard theories of evolution and adaptation. Just which explication of teleology is best for which context remains open to debate, but there seems little doubt that at least some theoretically and explanatory useful notions of teleology can be explicated in naturalistically acceptable ways.
3. Indeed with humans the range of possible ends is more or less open-ended—whether racing bicycles up mountains or writing philosophical essays.
4. This may provide a reason for doubting that Ned Block's (1995) distinction between *access consciousness* and *phenomenal consciousness* marks a division between two different processes. It may instead merely distinguish two interrelated aspects of one and the same process.
5. Of course some higher-order theorists have denied the distinction between the two sorts of consciousness, or at least denied that consciousness in the what-it's-like sense can be separated from consciousness in the higher-order sense (Rosenthal 1991a; Carruthers 2000).