



Agentive Phenomenal Intentionality and the Limits of Introspection

Terry Horgan
Department of Philosophy
University of Arizona
Tucson AZ 85721-0027
thorgan@email.arizona.edu
© T. Horgan 2007

PSYCHE 13/1, April 2007

KEYWORDS: Agentive Experience, Agent-Exclusion Problem, Introspection, Phenomenal Intentionality

Abstract: I explore the prospects for overcoming the prima facie tension in the following four claims, all of which I accept: (1) the phenomenal character of experience is narrow; (2) virtually all aspects of the phenomenal character of experience are intentional; (3) the most fundamental kind of mental intentionality is fully constituted by phenomenal character; and yet (4) introspection does not by itself reliably generate answers to certain philosophically important questions about the phenomenally constituted intentional content of experience. The apparent tension results from the following initially plausible thought: if indeed the answers to such questions are entirely fixed by phenomenal character by itself, then presumably these answers should be directly available introspectively. I focus the discussion specifically on certain questions about the phenomenal intentional content of agentive experience—e.g., the question whether the content of agentive experience is compatible with causal determinism. I consider three alternative possibilities for explaining why the answers to such questions are not directly available introspectively—two of which I argue against and the third of which I embrace.

1. Introduction

Suppose one embraces the following claims.

First, the phenomenal character of experience is *narrow*, in this sense: it is not constitutively dependent upon anything “outside the head” of the experiencing creature.

Thus, creatures who are physically and functionally just alike must also be phenomenally just alike.

Second, virtually all aspects of the phenomenal character of experience are intentional: phenomenal character *represents* the world as being various ways.

Thus, phenomenal character virtually always has satisfaction conditions, because having satisfaction conditions is a hallmark of intentionality.

Third, the most fundamental kind of mental intentionality is fully constituted by phenomenal character.

This makes such intentionality narrow, given the preceding claims. George Graham and John Tienson and I call this narrow content *phenomenal* intentionality.

Fourth, for certain philosophically important questions about the phenomenally constituted intentional content of experience, introspection by itself does not reliably generate an answer.

There is a *prima facie* tension in these views, which can be expressed as follows. Phenomenal character—what-it’s-like-ness—is self-presenting in one’s mental life: it is directly present before the mind. Thus, phenomenal character is an aspect of mentality that is eminently accessible to introspection. But if the most fundamental kind of mental intentionality is narrow and phenomenally constituted, then presumably it too should be likewise self-presenting, and likewise accessible to introspection. The narrow intentional content of a current experience, being wholly fixed by the experience’s phenomenal character, should itself be directly ascertainable introspectively. But if so, and if the phenomenally constituted narrow content that is thus introspectively luminescent entirely fixes the answers to philosophically important questions about that very content, then *prima facie* it seems that the answers to those questions should *themselves* be directly accessible introspectively. As long as one has sufficient conceptual sophistication to understand the philosophical questions, it is natural to think, one should be able to ascertain their answers just by attending introspectively to the phenomenal character of one’s experiences.¹ Yet in fact one sometimes cannot.

My general goal in this paper is to explore the prospects for overcoming the *prima facie* tension—i.e., the prospects for explaining this initially surprising kind of limitation in our capacity for introspectively ascertaining phenomenally constituted facts about our experience. More specifically, I will address the version of this generic issue that arises if one also accepts the following thesis, in addition to the four mentioned above:

Fifth, one rich and distinctive aspect of the overall phenomenal character of experience is *agentive*: the what-it’s-like of (experiencing oneself as) *doing*.

An example of a philosophically important question concerning the narrow content of agentive phenomenology is this: Is the content of one’s agentive experience compatible with the hypothesis of causal determinism? Given the first three hypotheses and the fifth one, the answer to this question is fixed entirely by agentive phenomenology itself. But the fourth hypothesis is applicable here, I would claim: introspection alone does not reliably deliver an answer to the compatibility question.

For present purposes I will take for granted all five theses.² I seek not to call any of these claims into question, but rather to articulate and compare some potential ways of smoothly reconciling them. I will consider three different proposals for doing so, and I will explain as I go why I recommend rejecting the first two and

embracing the third. (I believe that the discussion to follow is potentially generalizable to other aspects of phenomenal intentionality beyond the phenomenology of agency, but I will not pursue that matter here.)

Since the issue I am addressing only arises for someone who both (1) rejects strongly externalist views about mental intentionality and (2) holds that phenomenal character is epistemologically special by virtue of its self-presenting character, one might be tempted to think that I am presupposing that direct introspective knowledge about the content of one's mental states would be precluded by externalism about mental content. Before proceeding, let me say something about my present topic vis-à-vis contemporary philosophical debates about whether or not direct knowledge of one's own mental states is threatened by externalism. In a nutshell, I view those debates as largely orthogonal to my topic here. Suppose that one now wishes to drink some water. One certainly cannot tell, just by introspection, whether or not the stuff one now wishes to drink is H₂O; for, the referent of one's word 'water' (and of one's concept WATER) is constitutively determined in part by matters external to one's own mental life. But can one tell, just by introspection, that the stuff one now wishes to drink is *water*? Nothing I will say in the present paper commits me to a negative answer to this question, as far as I can see. I am not assuming that direct introspective knowledge about one's mental life is incompatible with externalism. I am not assuming this with respect to a *modest* form of externalism (which I espouse) claiming that for some terms and concepts (e.g., those expressible by names and by natural-kind terms), reference (as distinct from phenomenal intentionality) is constitutively determined in part by externalistic facts. (On the contrary, it seems obvious to me that one *can* know directly by introspection that one wants some water, even though one cannot know just by introspection what water is composed of.) Nor am I assuming that *strong* forms of externalism (all of which I reject)—viz., views claiming that all mental intentionality is constitutively determined, at least in part, by externalistic facts—are incompatible with direct introspective knowledge about one's present occurrent mental states. Although I do believe that strong externalism is profoundly mistaken, my reasons for this belief seem to me largely independent of the claim that strong externalism precludes privileged self-knowledge. (And I do not espouse that claim.³)

2. Background

Before turning to the business at hand, I will set out some background. I will say something by way of elaboration of, and motivation for, the claims that I mentioned at the outset. I begin by summarizing the general picture of mental intentionality I favor along with George Graham and John Tienson (Horgan and Tienson 2002, Horgan, Tienson and Graham 2004, Graham, Horgan & Tienson 2007; in press a). Next I summarize some views I hold, many of which I also share with Graham and/or Tienson, about agentive phenomenology (Horgan, Tienson and Graham 2003, Horgan and Tienson 2005). Then I describe a philosophical problem about agentive phenomenal content that I call the "agent exclusion problem," and I summarize my views about the limits of introspection in addressing this problem.

It is not my present purpose to address the agent-exclusion problem directly; I have done so elsewhere (Horgan forthcoming a, forthcoming b). Rather, here I use the

problem as a basis for highlighting what I take to be some important limitations in our capacities of introspection, vis-à-vis certain aspects of phenomenal consciousness. Those limitations are what need explaining, within the framework of phenomenal intentionality.

1.1. Phenomenal intentionality and externalistic intentionality⁴

Lately Graham and Tienson and I have been articulating a general approach to mental intentionality that is very much at odds with the various versions of strong externalism that have recently been fashionable in philosophy of mind. (Strong externalism is the view that all mental intentionality depends constitutively on certain connections—e.g., causal, and/or covariational, and/or historical, and/or social—between the cognitive agent and the agent’s external environment.) Although our approach repudiates strong externalism, we claim that it nonetheless does justice to the genuine insights about matters of intentionality and reference in the work of philosophers like Putnam, Kripke, and Burge. Central in our approach is the role of *phenomenology* or *phenomenal consciousness*, by which we mean those aspects of one’s mental life such that there is “something it is like” to undergo them. Briefly, the position goes as follows.

Phenomenology is *narrow*: it is not constitutively dependent upon anything “outside the head” (or outside the brain) of the experiencing subject. Indeed, it is not constitutively dependent upon anything outside of phenomenal consciousness itself; in this sense, it is *intrinsic*. Your phenomenology, being narrow and intrinsic, supervenes at least nomically upon physical events and processes within your brain.

Phenomenology is also richly and pervasively *intentional*: there is a kind of intentionality that is entirely constituted phenomenologically (*viz.*, *phenomenal intentionality*), and it pervades our mental lives. Among the different aspects of phenomenal intentionality are the following. First, there is the *phenomenology of perceptual experience*: the enormously rich and complex what-it’s-like of being perceptually presented with a world of apparent objects, apparently instantiating a rich range of properties and relations—including one’s own apparent body, apparently interacting with other apparent objects which apparently occupy various apparent spatial relations as apparently perceived from one’s own apparent-body centered perceptual point of view. Second, there is the *phenomenology of agency*: the what-it’s-like of apparently *voluntarily controlling* one’s apparent body as it apparently moves around in, and apparently interacts with, apparent objects in its apparent environment. (This is a central focus of the present paper, of course.) Third, there is *conative and cognitive phenomenology*: the what-it’s-like of consciously (as opposed to unconsciously) undergoing various occurrent propositional attitudes, including conative attitudes like occurrent wishes and cognitive attitudes like occurrent thoughts. There are phenomenologically discernible aspects of conative and cognitive phenomenology, notably (i) the phenomenology of *attitude type* and (ii) the phenomenology of *content*. The former is illustrated by the phenomenological difference between, for instance, *occurrently hoping* that Hillary Clinton will be elected U.S. President and *occurrently wondering* whether she will be elected—where the attitude-content remains the same while the attitude-type varies. The phenomenology of content is illustrated by the phenomenological difference between occurrently thinking that Hillary Clinton will be

elected U.S. President and occurrently thinking that she will *not* be elected—where the attitude-type remains the same while the content varies.)

Since phenomenal intentionality is entirely constituted phenomenologically, and since phenomenology is narrow, phenomenal intentionality is narrow too. Hence, there is an *exact match* of phenomenal intentionality between yourself and your brain-in vat (BIV) physical duplicate. This exactly matching, narrow, intentional content involves exactly matching, phenomenally constituted, *narrow truth conditions*. But whereas the narrow truth conditions of your own beliefs are largely satisfied, those of your BIV physical duplicate's matching beliefs largely fail to be satisfied; thus, the BIV's belief system is systematically nonveridical.

On the other hand, exact match in narrow content between your own intentional mental states and the corresponding states in your BIV physical duplicate does not require or involve exact match in *referents* (if any) of all the various matching, putatively referring, thought-constituents. For instance, certain of your own occurrent thoughts that you would express linguistically using certain proper names—say, the thought that Bush is not a genius—involve singular thought-constituents whose referents (if any) are determined partly in virtue of certain external relations that obtain between you and those referents. Thus, your occurrent thought that *Bush is not a genius* involves a singular thought-constituent that purports to refer to a particular specific person (viz., Bush); its *actually* referring, and its referring to the specific individual to whom it does refer, depends upon there being certain suitable external relations linking you to a unique eligible referent (viz., Bush). A Twin-Earthly physical duplicate of yourself, in a Twin-Earthly duplicate local environment, would refer to a *different* individual (viz., Twin-Bush) via the corresponding singular thought-constituent of the corresponding occurrent thought. And in the case of your BIV physical duplicate, the matching singular thought-constituent *fails to refer at all*, because the BIV does not bear suitable externalistic relations to any suitably reference-eligible individual in its own actual environment. (Parallel remarks apply to thought-constituents that purport to refer to natural kinds, such as the thought-constituent that you yourself would express linguistically with the word 'water'.)

For mental states involving thought-constituents for which reference depends upon externalistic factors, there are two kinds of intentionality, each involving its own truth conditions. First is the kind of intentionality already mentioned above: *phenomenal* intentionality, with truth conditions that are phenomenally constituted and narrow. Second is *externalistic* intentionality, with wide truth conditions that incorporate the actual referents (if any) of the relevant thought-constituents. Your own thought that Bush is not a genius, and the corresponding thoughts of your BIV physical duplicate and your Twin Earth physical duplicate, have matching phenomenal intentionality, with matching truth conditions. (These truth conditions are satisfied in your case and in the case of your Twin Earth duplicate, but not in the case of your BIV duplicate.) On the other hand, your own thought that Bush is not a genius and your Twin Earth duplicate's corresponding thought do not have matching *externalistic* intentionality, because the externalistic truth conditions of these respective thoughts do not match: the truth value of your own thought depends upon the intelligence level of *Bush*, whereas the truth value your Twin Earth duplicate's corresponding thought depends upon the intelligence level of an entirely

different individual, viz., Twin-Bush. (Each thought's wide truth conditions are indeed satisfied.) As for your BIV duplicate's thought, it lacks externalistic intentionality and wide truth conditions, because its singular thought-constituent purporting to refer to a person called 'Bush' does not actually refer at all.

This conception of mental intentionality rests heavily and essentially upon two key contentions. First, mental reference to many properties and relations—including various spatiotemporal-location properties, shape-properties, size-properties, artifact-properties, and personhood-involving properties—is wholly constituted by phenomenology alone. Even systematically *nonveridical* phenomenology, as in the case of the BIV, provides *reference-constituting experiential acquaintance* with such properties and relations. It makes no difference to such experiential acquaintance with such properties—and hence it makes no difference to mental *reference* to such properties—whether or not the properties with which one becomes experientially acquainted are ever actually instantiated in one's ambient environment.

Second, in the case of thought constituents whose reference (if any) depends constitutively upon certain externalistic elements, the mechanisms of reference-fixation crucially involve phenomenally constituted *grounding presuppositions* (as we call them). Thus, phenomenal intentionality is more basic than externalistic intentionality, since the latter depends in part upon the former (as well as depending in part upon externalistic factors). Suppose, for example, that you have an occurrent thought that you could express linguistically by saying "That picture is hanging crooked," where the singular thought-constituent expressible linguistically by 'that picture' purports to refer to a picture on the wall directly in front of you. This thought-content involves phenomenally constituted grounding presuppositions that must be satisfied in order for the singular thought-constituent to refer: roughly, there must be an object at a certain location relative to yourself (a location that you could designate linguistically by a specific use of the place-indexical 'there'), this object must be a picture, there must not be any other picture at that location that is an equally eligible potential referent of 'that picture', and this object must be causing your current picture-experience. If these grounding presuppositions are satisfied by some specific concrete particular in your ambient environment—some specific object that is a picture and is uniquely suitably located—then your singular thought-constituent thereby refers to that very object. *Which* object your thought-constituent refers to, if any, thus depends jointly upon two factors, one phenomenally constituted and one externalistic: on one hand, the phenomenally constituted grounding presuppositions, and on the other hand, the unique actual object in your ambient environment that *satisfies* those presuppositions.

Let me introduce another distinction, also to be invoked below, between two kinds of mental intentionality. (This distinction may well be vague, and/or a matter of degree; and it is orthogonal to the distinction between phenomenal intentionality and externalistic intentionality.) On one hand is *judgmental* intentionality—the kind possessed by full-fledged, occurrently conscious, thoughts, intentions, suspectings, wonderings, and the like. This kind of intentionality requires considerable conceptual sophistication, and it is plausible that the mental lives of non-human animals have much less of it than do the mental lives of humans. On the other hand is *presentational* intentionality—the kind possessed, for instance, by ordinary perceptual experience. This

kind of intentionality presumably occurs in some form even in non-human creatures—although it probably occurs in richer, conceptually colored, forms in humans. (Dogs, for instance, surely cannot visually experience an automobile *as an automobile*, or a computer *as a computer*.)⁵

1.2. The agentive phenomenology of first-person agency⁶

Graham and Tienson and I also have been urging specific attention to the phenomenology of first-person agency—the “something it is like” to experience oneself as behaving in a way that constitutes action.⁷ We maintain that agentive phenomenology is richly intentional, presenting in experience a self that is an apparently embodied, apparently voluntarily behaving, agent. Because agentive experience is intentional, it has *satisfaction conditions*—which raises two philosophically important questions. First, what are those satisfaction conditions? I.e., what is required of the world, including oneself and one’s own body, in order for one to be an agent of the kind one experiences oneself as being? Second, are those conditions actually satisfied? I.e., are humans *in fact* agents of the kind they experience themselves as being? Such questions have received very little attention in recent philosophy of mind, largely because the phenomenology of agency itself has received very little attention. But Graham and Tienson and I have been arguing that this needs to change.

In this section I will briefly summarize some of what we have had to say descriptively about the phenomenology of doing—about what this kind of “what it’s like” is like. Issues about satisfaction conditions will be central in the remainder of the paper.

We employ the term ‘behavior’ in a broad sense, one that is neutral about whether or not any particular instance of behavior counts a genuine *action*. Paradigmatic behaviors are certain kinds of bodily motions. (Although there can be other forms of behavior, such as remaining still or remaining silent, I largely set them aside for simplicity.) The point of using ‘behavior’ in this broad sense is to remain neutral about the question whether the bodily motions called behavior really meet the satisfaction conditions imposed upon them by the phenomenology of doing.

What is behaving like phenomenologically, in cases where you experience your own behavior as action? Suppose that you deliberately perform an action (or anyway you have an experience as of doing so)—say, holding up your right hand and closing your fingers into a fist. As you focus on the phenomenology of this item of behavior, what is your experience like? To begin with, there is of course the purely behavioral aspect of the phenomenology—the what-it’s-like of being visually and kinesthetically presented with one’s own right hand rising and its fingers moving into clenched position. But there is more to it than that, of course, because you are experiencing this bodily motion *as your own action*.

In order to help bring into focus this specifically actional phenomenological dimension of the experience, it will be helpful to approach it in a negative/contrastive way, via some observations about what the experience is *not* like. For example, it is certainly not like this: first experiencing an occurrent wish for your right hand to rise and your fingers to move into clenched position, and then passively experiencing your hand and fingers moving in just that way. Such phenomenal character might be called *the*

phenomenology of fortuitously appropriate bodily motion. It would be very strange indeed, and very alien.

Nor is the actional phenomenological character of the experience like this: first experiencing an occurrent wish for your right hand to rise and your fingers to move into clenched position, and then passively experiencing a causal process consisting of this wish's causing your hand to rise and your fingers to move into clenched position. Such phenomenal character might be called *the passive phenomenology of mental state-causation of bodily motion*.⁸ People often do passively experience causal processes as causal processes, of course: the collision of a moving billiard ball with a motionless billiard ball is experienced as causing the latter ball's subsequent motion; the impact of the leading edge of an avalanche with a tree in its path is experienced as causing the tree to become uprooted; and so on. But it seems patently clear that one does not normally experience one's own actions in that way—as passively noticed, or passively introspected, causal processes consisting in the causal generation of bodily motion by occurrent mental states. That too would be a strange and alienating sort of experience.⁹

How, then, should one characterize the actional phenomenal dimension of the act of raising one's hand and clenching one's fingers, given that it is not the phenomenology of fortuitously appropriate bodily motion and it also is not the passive phenomenology of mental state-causation of bodily motion? Well, it is the what-it's-like of *self as source* of the motion. You experience your arm, hand, and fingers as being moved *by you yourself*—rather than experiencing their motion either as fortuitously moving just as you want them to move, or passively experiencing them as being caused by your own mental states. You experience the bodily motion as caused by *yourself*.¹⁰

The phenomenal character of actions also typically includes aspects of *purposiveness*: both a generic what-it's-like of acting *on purpose*, and often also a more specific what-it's-like of acting *for a specific purpose*. The phenomenology of purposiveness can work in a variety of ways.¹¹ Sometimes, for instance (but not always), the action is preceded by conscious deliberation. In one variant of deliberative action, the process involves settling into reflective equilibrium prior to acting: the overall phenomenology includes, first, the what-it's-like of explicitly entertaining and weighing various considerations favoring various options for action, then the what-it's-like of settling upon a chosen action because of certain reasons favoring it, and then the what-it's-like of performing the action for those very reasons. (Examples range from the weighty, such as deciding which car to buy or which job offer to accept, to the mundane, such as deciding what to order for lunch in a restaurant.) In another variant, the action is preceded by the occurrence in experience of an explicit psychological syllogism: the overall phenomenology includes, first, the what-it's-like of mentally going through a particular piece of practical reasoning, and then the what-it's-like of performing an action because doing so is the upshot of that reasoning. (A familiar example of such an action is a deliberative version of the philosopher's workhorse of belief/desire explanation: at a party you consciously experience a desire for a beer and a perceptually generated occurrent thought about where the beer is located; you consciously form an intention to walk to that location and grab a beer; and then you act, with the explicit purpose in mind of getting yourself a beer.)

Actions are very often performed without prior deliberation, however. Here the tinge of purposiveness, within the phenomenology of doing, is typically more subtle. For example, as you approach your office you pull your keys out of your pocket or purse; then you grasp the office key; then you insert it into the lock; then you twist it in the lock; and then you push the door open. All of this is routine and automatic: no deliberation is involved. Nonetheless, the what-it's-like of doing these things still certainly includes an on-purpose aspect, and indeed an aspect of doing them for specific purposes both fine-grained and coarse-grained: getting hold of your keys, getting hold of your office key in particular, activating the door lock, getting into your office, etc. In some cases of non-deliberative action, it appears, certain specific purposes for which one acts are explicitly conscious but not salient. In other cases, it seems, certain specific purposes are not explicitly conscious at all, but nonetheless are accessible to consciousness. In still other cases—for instance, specific actions performed during fast-paced sports such as soccer and basketball—some specific purposes for which the agent acts in one specific way rather than another probably are neither explicitly conscious nor even consciously accessible after the fact, because of the way these specific purposes are linked to very short-lived, and very intricately holistic, aspects of the player's rapidly changing perceptual phenomenology. Nonetheless, even here the phenomenology still normally includes the what-it's-like of acting in a specific way *for a specific purpose*, whether or not one finds oneself in a position after the fact to tell what that purpose was. Purposiveness is phenomenologically present in all these types of nondeliberative action, with specific purposes coloring conscious experience even when they are not explicitly conscious themselves.¹²

The phenomenology of doing typically includes another aspect, distinguishable from the aspect of purpose: viz., *voluntariness*. Normally when you do something, you experience yourself as *freely* performing the action, in the sense that it is *up to you* whether or not to perform it. You experience yourself not only as generating the action, and not only as generating it purposively, but also as generating it in such a manner that you *could have done otherwise*. This palpable phenomenology of freedom has not gone unrecognized in the philosophical literature on freedom and determinism, although often in that literature it does not receive as much attention as it deserves. (Sometimes the most explicit attention is given to effort of will, although it takes only a moment's introspection to realize that the phenomenology of voluntarily exerting one's will is really only one, quite special, case of the much more pervasive phenomenology of voluntariness.¹³)

Associated with the voluntariness dimension of agentive phenomenology is the following fact about the experience of reasons as motives. Although often one does experience certain conscious reasons (e.g., occurrent beliefs, occurrent wishes, etc.) as playing a state-causal role in relation to one's action, this role is experienced as one's being *inclined* by those reasons to perform the given action; the role is *not* experienced as one's action being *necessitated* by those reasons. Such experiences of one's mental states playing a state-causal motivational role—viz., the role of state-causing an *inclination* toward an outcome—are importantly different from experiences of outright state causation of an outcome itself. For, when one undergoes experiences of state causation of an outcome, typically the cause *is* experienced as necessitating the effect (in the

experienced circumstances). The phenomenology of necessary connection between cause and outcome is introspectively palpable in such cases (although it seems to involve some form of necessity other than logical or conceptual). By contrast, it is palpably *absent* in the phenomenology of agency.

The intentional content of agentive phenomenology is *presentational* rather than judgmental. In this respect, agentive phenomenology is more closely akin to the phenomenology of perceptual experience than it is to the phenomenology of discursive thought. (Many higher non-human animals, I take it, have a fair amount of agentive phenomenology, even if they engage in little or no discursive thought.) Of course, we humans also wield *concepts* like agency, voluntariness, and the like (whereas it is questionable whether non-human animals do); but the judgmental phenomenology of thoughts employing these concepts is not to be conflated with agentive phenomenology itself, which is presentational.

2. The Agent-Exclusion Problem and the Limits of Introspection

Agentive phenomenology has presentational content that is intentional: such phenomenology has satisfaction conditions. Philosophical questions thus arise about the nature of these satisfaction conditions—including questions about whether the intentional content of agentive experience is compatible with various metaphysical hypotheses that are seriously entertained in philosophy. One such hypothesis is *state-causal determinism*, which asserts that the laws of nature are such there are no two nomically possible worlds that are exactly alike in some instantaneous total state, that differ in subsequent total states, and in which those laws are never violated. A second is the hypothesis of *physical state-causal closure*, which asserts that the state of the world at any moment in time, insofar as it is diachronically determined at all, is diachronically determined by prior physics-level phenomena, on the basis of the fundamental laws of physics. A third is the hypothesis of the *mental state-causation of behavior*, which asserts that the behaviors that one experiences as one's actions are always state-caused by certain mental states of oneself, such as occurrent wants in combination with occurrent beliefs.

The agent-exclusion problem, as I call it, is the (prima facie plausible) possibility that the presentational content of agentive experience is incompatible with one or another of these metaphysical hypotheses. If indeed there is such content-incompatibility, and if the given hypothesis is also true, then the apparent upshot is that people are not really agents of the sort they experience themselves to be: genuine agency is metaphysically excluded, and agentive experience is systematically nonveridical. (Agent-exclusion could arise in other ways too. Suppose, for instance, that (a) the content of agentive experience is incompatible with state-causal determinism, (b) determinism is false, but (c) to the extent that any phenomenon is not state-causally determined, it is just *random*. Arguably, if (a)-(c) obtain, so that human behavior is random to whatever extent it is not state-causally determined, then human behavior never qualifies as genuine action: it never really emanates from the *self* as its source, even though the agent *experiences* it as so generated.)

I myself favor blanket compatibilism about the agent-exclusion problem—the view that the content of agentive experience is compatible with all three above-mentioned metaphysical hypotheses. But as I said at the outset, my present aim is not to mount a

defense of compatibilism, but rather to use the agent-exclusion problem as a way of focusing on some important apparent limitations of introspection vis-à-vis presentational intentionality. For specificity, let us focus on the hypothesis of state-causal determinism. Consider the following question. Can one tell, just on the basis of introspective attention to one's agentic experience, whether or not the content of that experience is compatible with state-causal determinism?

The answer, I submit, is no. Although there is presumably a definite fact of the matter about the compatibility/incompatibility issue itself, this is not a fact that one can directly "read off" from one's agentic experience just by focusing one's attention introspectively on that experience itself, while bearing in mind the compatibility/incompatibility question. The fact one is seeking to ascertain is a fact *about presentational phenomenology*—viz., a fact about its intentional content—and yet it is a fact that is not directly accessible introspectively. This lack of introspective accessibility, concerning a feature of one's own phenomenal consciousness, is the central point about the agent-exclusion problem that I seek to emphasize here. The question is how to explain this inaccessibility phenomenon, given the claims summarized in sections 1.1 and 1.2 about phenomenal intentionality and about agentic phenomenology. If indeed the answer to the compatibility question depends solely on the phenomenal character of agentic experience, then *why isn't the answer directly accessible to introspection?*

Before proceeding to address this question, let me pause to comment on the fact that some will claim both (a) that they *can* tell by introspection what the answer is to the compatibility question, and (b) that introspection reveals that the content of agentic experience is *incompatible* with state-causal determinism (and incompatible with physical causal closure, and incompatible with the mental state-causation of behavior). These claims, I would argue, reflect a fairly subtle form of introspective confabulation. It is one thing for agentic experience to have this feature:

(A) *Not* presenting one's behavior *as* state-caused.

It is quite another thing for agentic experience to have this distinct feature:

(B) Presenting one's behavior *as not* state-caused.

Feature (A) is indeed introspectively accessible. (It is a feature I stressed in section 1.2 above, where I was describing introspectively accessible aspects of agentic phenomenology.) But I maintain that feature (B), in addition to being distinct from (A), also is *not* one that introspection reveals to be an aspect of agentic experience. Rather, I claim, the question of compatibility/incompatibility of agentic content with the state-causation of behavior is one whose answer—whatever it is—simply is not accessible to introspection. To think that introspection reveals that the answer to the compatibility question is negative is to be guilty of conflating the introspectable feature (A) with a distinct feature whose presence or absence is not introspectively accessible—viz., (B).¹⁴

Let me be clear about what I have *not* just said, concerning feature (B). (It is easy to get 'knotted up' by the current dialectic.) I have not denied that feature (B) is *in fact* instantiated by one's agentic experience; nor, of course, have I *affirmed* this. For all I have just claimed, agentic experience might indeed present one's behavior as not state-caused (even though, as a compatibilist, I do not think so.) What I have denied is that one

can tell by introspection whether or not feature (B) is instantiated in one's agentic experience.

3. First Proposal: Agentic Presentational Content Is Externalistically Intentional

What can be said, from within the framework of phenomenal intentionality, about why introspection cannot directly reveal whether or not agentic presentational content is compatible with hypotheses like state-causal determinism? One possible answer would go as follows. Certain aspects of presentational experience are *externalistically reference-purporting*, in this sense: they purport to refer to certain objects, properties, or kinds whose essence is something not revealed directly in experience itself. This is so, for instance, for visual experiences as of certain natural kinds, like the natural kind called 'water': it is an empirical question what the deep essence of such stuff is; the answer to this question is not something accessible to introspection, because the answer is not determined by phenomenal consciousness.

The same goes for agentic presentational content, according to the current proposal: this content too is externalistically reference-purporting. The experientially presented category *being a self-generated item of behavior* is a natural-kind category; it purports to pick out a kind whose essence is not itself determined by phenomenal consciousness. Rather, the essence of this natural kind is a matter of the actual normal etiology—whatever it is—of those behaviors that are experienced as actions. Possible exceptional cases aside, if a piece of behavior that one experiences as self-generated is produced in the way that such behaviors are normally produced, then that behavioral event thereby *really is* self-generated. Furthermore, it is an empirical question what the normal etiology of behavior is, in cases where the behavior is experienced as action. Much that is relevant to this question is not accessible to consciousness—for instance, facts about the cognitive architecture of the behavior-control system in humans, and about the neurobiological mechanisms that implement this cognitive architecture.

On this proposal, agentic phenomenology has two kinds of intentionality (as do the other aspects of mentality that are externalistically reference-purporting). The narrow, phenomenologically determined, intentionality has the truth condition that *there is some* natural-kind feature Φ such that (a) Φ 's instantiation normally produces those behaviors experienced as actions, and (b) one's present behavior results from Φ . The externalistic intentionality, on the other hand, has the truth condition that one's present behavior is produced by *the actual feature F* (whatever it is) whose instantiation normally produces the actual behaviors actually experienced as actions.

What the actual feature F is depends on non-introspectable facts about the world—in particular, facts about the neural mechanisms that normally produce the behaviors that people experience as their own actions, and facts about the functional cognitive architecture subserved by these neural mechanisms. Moreover, as regards the agent-exclusion problem, the correct answer to the relevant compatibility questions *also* depends on such non-introspectable facts. If the normal neural mechanisms producing the pertinent behavior operate deterministically, for instance—or virtually deterministically, without any quantum-level indeterminacies ever getting amplified in a way that could significantly affect the behavior of whole neurons—then the content of agentic

experience is compatibilist. I.e., if *F* is a feature whose instantiation is compatible with state-causal determinism, then the content of agentive experience is likewise compatible with state-causal determinism. But conversely, if *F* turns out to be a feature that is *incompatible* with state-causal determinism, then the content of agentive experience is also incompatible with state-causal determinism. It all depends on the nature of feature *F*—a matter that is not fixed phenomenologically, but depends instead upon certain extra-phenomenological facts.

On the plus side, this proposal has several attractive features. First, it provides a straightforward explanation of the non-introspectability of the correct answer to our compatibility question: the correct answer is not introspectable because it depends in part on the non-phenomenological facts. (It is an empirical matter, not fixed by phenomenology, what sorts of psychological and sub-psychological mechanisms generate behavior of the kind paradigmatically experienced as action.) Second, the proposal nicely vindicates the epistemological authority that people routinely accord to their agentive experience: we are well justified in trusting agentive experience, in taking it to be veridical rather than illusory, because there is surely *some* feature that constitutes the self-generation of behavior—viz., whatever feature is the one that *actually typically produces* those behaviors experienced as actions. Not only does one not need to know what this constitutive feature is, in order to be justified in believing that it is instantiated by one's own behavior, but one also does not need to know whether or not it is compatible with determinism.

But despite these theoretical advantages, the proposal strikes me as implausible, and as unattractive from the perspective of a fan of phenomenal intentionality. Consider, for instance, a Twin-Earthly phenomenal duplicate of oneself (a TEP duplicate) whose brain is silicon-based rather than carbon-based, and also is organized in such a way as to implement a behavior-control system that is rather different from the cognitive architecture of behavior control in humans. Intuitively, this TEP duplicate's presentational agentive phenomenology yields fully determinate agentive content all by itself—agentive content that exactly matches one's own, purely in virtue of intrinsic phenomenal character (just as much as the TEP duplicate's visual presentational experience, for instance, yields by itself a rich array of determinate content that exactly matches one's own: content involving apparent objects and certain of their apparent properties, e.g., shape properties, size properties, relative-position properties, and the like). Intuitively, that is, the presentational content of the TEP duplicate's agentive experience is just as impervious to matters externalistic as are those aspects of visual experience that pertain to properties like apparent shape and apparent relative position. Thus, one's strong intuition about this TEP duplicate is that the duplicate had presentational experience with truth conditions that fully coincide with the truth conditions of one's own presentational experience. In this important respect, the presentational category 'self-generated action' is like the presentational category 'spherical', and unlike the presentational category expressible linguistically by 'water'. It is *not* a natural-kind category that purports to refer to some feature whose essence is phenomenologically underdetermined. On the contrary, it is a category for a kind whose essence is given right there in experience, and is the same for both you and your TEP duplicate: viz., self-as-source-ness.

4. Second Proposal: Agentive Presentational Content Is Constituted by Broad-Based Phenomenology

A second approach to our problem backs away from the idea that presentational agentive content is externalistically reference-purporting. But this approach urges a broad way of understanding the idea of agentive content as being “constituted purely phenomenologically.” Perhaps the most natural construal of this idea is a non-broad one: viz., that the intentional content of one’s current experience is phenomenally constituted by one’s current, occurrent, phenomenology (where “current” should be understood as including the aspect of the specious present, and in any case certainly includes one’s current, occurrent, memory-experiences). But there are other aspects of phenomenology too. There is one’s history of occurrent phenomenology, for example. And there is also one’s *counterfactual phenomenal profile*: a body of facts about what courses of experience one *would* undergo if one *were* to get various kinds of phenomenal sensory-input experience while being in a given, initial, total phenomenal state.¹⁵ An effective brain-in-vat scenario, for instance, or an effective Cartesian deceiver scenario, underwrites a robust, systematically deceptive, counterfactual phenomenal profile: there is not just the experiencer’s *actual* course of experience (in the scenario as stipulated), but there is also a hugely rich range of counterfactual experiential processes, all of which would preserve perfectly the experiencer’s massive illusion about the external environment.¹⁶

Why and how might certain facts about one’s counterfactual phenomenal profile contribute constitutively to the content of one’s current, occurrent, agentive experience? It will be useful to approach this question by setting aside presentational content for a moment and focusing instead on certain kinds of judgmental content. (This does not mean setting aside phenomenology, of course; I am taking it for granted that occurrent *thought* has rich phenomenology—a key tenet of the phenomenal intentionality framework.) There are counterfactual psychological facts of various kinds that are relevant to the content of certain concepts. Take the concept of knowledge, for instance. Most people have never consciously thought about “Gettier cases,” but nonetheless do have robust counterfactual judgment-tendencies concerning them: most people are strongly disposed to judge, of such cases, that they are not instances of knowledge. These judgment tendencies belong to an agent’s counterfactual *phenomenal* profile, because occurrent thought is an aspect of one’s overall occurrent phenomenology. Furthermore, possession of such judgment tendencies is partially constitutive of conceptual competence vis-à-vis the concept of knowledge. (On a Twin Earth where conceptually competent people systematically apply the sign-design k-n-o-w-l-e-d-g-e to Gettier cases, that sign-design would not refer to *knowledge*.)

Return now to presentational agentive content. The current suggestion is that here too, in various and sundry ways, such content is fixed not by one’s current occurrent phenomenology alone, but rather in a way that depends constitutively upon various aspects of one’s counterfactual phenomenological profile.

What aspects of that profile might be constitutively involved, and why? Answering that question in a plausible way would be part of the burden of the present proposal—a task I will not try to pursue in any detail here, beyond the following brief remarks. Presumably the right treatment of presentational content would need to mesh in

a suitably smooth way with the right treatment of judgmental intentional mental content. So, since one's counterfactual phenomenal profile evidently plays a constitutive role in the fixation of occurrent *judgmental* content, perhaps it thereby also plays a constitutive role in occurrent *presentational* content. In particular, perhaps one's judgment tendencies vis-à-vis concepts like agency, freedom, choice, and the like play such a constitutive role. (Recall, once again, that I am assuming that thoughts have phenomenal character themselves, and thus that one's judgment tendencies are part of one's counterfactual *phenomenal* profile.)

If the present proposal could be made to work, then the upshot concerning the central problem I am addressing in this paper would be as follows. Even though current, occurrent, phenomenology is introspectively self-presenting (to a very great extent, at least), nonetheless one cannot introspectively hold before one's mind, all at once, all aspects of one's total counterfactual profile that figure constitutively in the fixation of presentational agentive content. In particular, one cannot introspectively hold before one's mind, all at once, all pertinent facts about one's relevant judgment tendencies concerning concepts like agency, freedom, choice, and the like. But this rich body of counterfactual phenomenological facts, *collectively*, plays a constitutive role vis-à-vis the very matters about agentive phenomenal intentionality that are of central concern in this paper—matters like whether or not the content of agentive experience is compatible with the hypothesis of state-causal determinism. Since one cannot simultaneously introspect all constitutively relevant counterfactual-phenomenological facts at once, one cannot tell by introspection alone whether or not compatibilism about agentive experience is correct.

So this approach nicely bears the burden of explaining the limits of introspection with regard to agentive phenomenology. But how plausible is the approach, from the perspective of a fan of phenomenal intentionality? Not very, it seems to me. There is a strong suspicion that it gets the cart before the horse. Intuitively speaking, the content of one's present agentive experience is thoroughly determinate in and of itself, apart from facts about one's counterfactual phenomenology. There may well be various aspects of one's counterfactual profile that would be *appropriate* to one's present agentive experience—e.g., tendencies one might possess to make various judgments like “I am now performing a genuine action” in various potential experiential circumstances. But intuitively, such matters of counterfactual judgmental phenomenology do not contribute constitutively to agentive presentational content. On the contrary, since the right treatment of presentational content would need to mesh in a suitably smooth way with the right treatment of judgmental intentional mental content, and since presentational agentive content seems to be constitutively dependent only on the phenomenal character of one's current, occurrent, experience, the proper use of judgmental concepts like agency and freedom is *itself* heavily constrained by presentational agentive content. Certain judgment tendencies involving these concepts are right *because* they respect the already-determinate presentational content of presentational agentive experience—which seems diametrically opposite to the idea that such judgment tendencies figure constitutively in *determining* presentational agentive content.

This point can be further elaborated by distinguishing between constitutive dependence and epistemic support. Consider the tendency to judge of oneself that one is acting, and acting freely, in various experiential circumstances in which the question

arises explicitly. To say that this judgment tendency is *appropriate* to presentational agentive content is to say that the presentational content provides *substantial epistemic support* for such a judgment. Moreover, this epistemic-support relation requires the presentational content of one's agentive experience to be linked *semantically* to the judgmental content of such a judgment, in a suitably tight way.¹⁷ But this kind of semantically grounded epistemic-support relation rests on the fact that agentive experience has determinate content solely by virtue of its *intrinsic* phenomenal character. The judgment-tendencies are appropriate because the judgmental content coincides with the presentational content that epistemically supports it—which seems diametrically opposite to the idea that these judgmental dispositions are partly *constitutive* of presentational agentive content.

5. Third Proposal: Agentive Presentational Content is Constituted by Occurrent Phenomenology

On a third view, the content of agentive experience is fully determined by current, occurrent, phenomenology. There is a fact of the matter, just on this basis, about whether or not agentive presentational content is compatible with determinism—or with physical causal closure, or with the mental state-causation of behavior. This approach seems intuitively the most attractive, within the framework that gives pride of place to phenomenal intentionality. But how can one explain, under this proposal, the limits of introspectability regarding these compatibility issues? I will first propose a fairly abstract answer to this question, and I will then propose two complementary ways of making the abstract answer more concrete and specific—of putting some meat on its bones.

5.1 Introspecting phenomenal content: The judgmental dimension and its limitations

To begin with, it is important to acknowledge that understanding and addressing the compatibility questions requires considerable *conceptual* sophistication. Consider non-human critters, such as dogs. Dogs, I take it, have agentive presentational phenomenology too. And under the present proposal, presumably, the occurrent phenomenal character of doggie agentive phenomenology fixes its intentional content so completely as to make it determinate whether or not that content is compatible with causal determinism. Nevertheless, dogs can't *begin* to contemplate the compatibility issue, or to introspect their own agentive phenomenology with that issue in mind. For, dogs don't begin to have the conceptual resources for such discursive thought; only humans do, surely.

Can dogs introspect, one might ask? Well, presumably the distinctive *self-presentational* nature of phenomenology is a feature of doggie phenomenology, just as it is for human phenomenology. This self-presentational nature includes the qualitative character of the phenomenology—the *what* of “what it's like.” It also includes the subjectivity of the phenomenology—what's it's like *for me*. What is *introspection* of one's phenomenology, though? In a thin sense, one might think of it as just *focally attending to qualitative character*—something that dogs might well be capable of (e.g., when they are in excruciating pain and can't do anything to alleviate it). But in a more robust sense, introspection is a cognitive capacity that involves not just attending focally

to some aspect of one's phenomenology, but also forming *judgments* about it on the basis of such focal attention. Robust introspection thus includes the deployment of concepts, and the formation of mental states that (i) have judgmental content, and (ii) are *about* the phenomenology to which one is attending. Dogs surely are not capable of that.

Return now from dogs to humans. Perhaps not all humans could even understand the compatibility questions under consideration in the present paper; but let us restrict our attention to those who can. Such people do possess a suitable conceptual repertoire to be able to address the compatibility questions, and to be able to try answering them by robust introspection. Nevertheless, I have claimed, they cannot arrive at a confident answer just by robust introspection. Why not? Briefly, the answer I propose is this: answering such questions solely via robust introspection would require a degree of cognitive skill in the deployment of the pertinent concepts, and in the formation of beliefs about one's phenomenology with the pertinent kind of judgmental content vis-à-vis that phenomenology, that far exceeds what is required for the conceptually competent use of these concepts. Indeed, it probably exceeds the cognitive capacities of humans altogether. Although the beliefs one would be seeking to generate, via robust introspection, would be about the nature of a certain kind of *presentational* content—viz., agentive presentational content—nevertheless these beliefs themselves would have very rich *judgmental* content—viz., judgmental content concerning the compatibility or incompatibility of that agentive presentational content with causal determinism, with physical causal closure, and with the mental state-causation of behavior. Even though agentive phenomenology itself is self-presenting to the experiencing agent, and even though the agent has conceptual mastery over all the pertinent concepts, nonetheless the agent lacks the capacity to determine, just via introspection, how those concepts would correctly apply to that self-presented phenomenology.

Here is an analogy from mathematics. Suppose you are visually presented with a vertically oriented isosceles triangle ABC, with point A at the vertical tip, point B at the lower left tip, and point C at the lower right tip. Each angle is overlaid with a translucent protractor, suitably positioned to display the size of the angle (in degrees): angle BAC is shown to be 20 degrees, and thus angles ABC and ACB are each shown to be 80 degrees. The presentation includes a line-segment BD, whose lower left tip is point B and whose upper right tip is a point D on line segment AC; angle CBD is 60 degrees (as indicated by the protractor at point B). And there is also a line segment CE, whose lower right tip is point C and whose upper left tip is a point E on line segment AB; angle BCE is 70 degrees (as indicated by the protractor at point C). Finally, line segment ED, sloping downward from point E at its upper left to point D at its lower right, connects points E and D. Angle CED is marked with an 'x'.

As should be introspectively obvious from the visual display, angle x in the diagram must have a unique, determinate, value—a value determined by the specified angles in the figure. But can one tell, just by visually inspecting the visual presentation and bearing in mind the question about the value of angle x, what that value is? Surely not—not if one is a normal human, anyway, as opposed to a mathematical savant.¹⁸ One might very well have a mastery over any pertinent geometrical concepts—concepts whose suitably clever deployment would yield a geometrical construction (or a series of linked geometrical constructions) that reveals the uniquely correct answer.¹⁹ But (in

normal humans, at least) such mastery, together with careful attention to the visual presentation itself while bearing in mind the question at hand, simply will not suffice to produce an introspectively obvious answer to the question about the value of angle x .

Return again to the matter at hand. As I have been stressing, in order to tell by robust inspection whether or not agentive presentational content is compatible with determinism (or with physical causal closure, or with the mental state-causation of behavior), not only would one need to attend carefully to one's agentive presentational content itself, but also one would need to form sophisticated *judgments* about the nature of that presentational content. These judgments would, of course, deploy one's concepts in a sophisticated way, and the judgments themselves thus would have sophisticated *judgmental* content (even though they would make claims about presentational content). This opens up the possibility of various kinds of fallibility in one's deployment of these concepts in forming such judgments, and/or the possibility of various kinds of limitation. In our mathematics example, normal human mastery of geometrical concepts won't begin to provide one with the capacity to tell, just by attentively gazing at the figure with the question in mind "What is angle x ?", what the answer is to that question.²⁰ Likewise, just because one understands the concepts needed to formulate abstruse compatibility-questions about agentive presentational phenomenology, and just because one has a humanly normal degree of mastery in deploying these concepts, this hardly guarantees that one is so splendidly skilled in their deployment that one can immediately tell, just by introspectively attending to one's agentive phenomenology with the pertinent compatibility questions in mind, what the answers are to those questions.

Can more be said about likely sources of fallibility or limitations in one's capacity to form accurate introspective judgments about one's phenomenology—sources that are especially apt to be operative in the case of the compatibility questions now under consideration? I think so. I will propose two plausible-looking sources.

5.2. Concrete classificatory judgments, general ideological hypotheses, and abductive inference

First, normal human conceptual competence is often mainly a matter of being able to correctly apply a given concept *to a concrete case*—or more precisely, do so correctly *modulo one's available evidence*. Typically one does this automatically and spontaneously. This is very different for simultaneously (i) holding consciously before one's mind a formulation of general satisfaction conditions for the given concept, and (ii) consciously comparing the concrete case against the general satisfaction conditions, "checklist" style, to ascertain whether or not the case falls under the concept. Even if the concept happens to have some tractably formulable, cognitively surveyable, set of satisfaction conditions that would constitute a "conceptual analysis" of the sort typically sought in the high-church analytic philosophy of the last century, no such analysis is consciously deployed in applying the concept. (In judging that some Gettier case is not an instance of knowledge, for instance, one does consciously deploy some general conceptual analysis; on the contrary, one directly and intuitively forms the judgment.)

Furthermore, a given concept may not have such a tractably formulable, cognitively surveyable, "conceptual analysis" anyway. (The sad history of attempts to provide such analyses in high-church analytic philosophy certainly provides ample basis

for pessimism here—with the poster child being the history of post-Gettier attempts to offer an adequate conceptual analysis of the concept of knowledge.) Perhaps, instead, the conditions under which an object A would satisfy a concept C could only be exhaustively specified by something like a cognitively non-surveyable specification of an infinite range of potential concrete scenarios—a specification that could not possibly be consciously entertained all at once by a human being.

The point is that, however exactly a competent human forms a classificatory judgment that brings a concrete case under a general concept, normally it is *not* done via conscious rehearsal of, and conscious deployment of, the concept's satisfaction conditions. Moreover, classificatory judgments probably *could not* be produced that way for many concepts, because many concepts probably have satisfaction conditions that cannot be captured in a cognitively surveyable formulation anyway. At the level of consciousness, the deployment of concepts to form concrete categorization judgments is normally a matter of applying the concepts directly and immediately to concrete cases.

Since competent concept-wielding is so much a matter of direct and appropriate application of concepts to *concrete* cases, conceptual competence alone is apt to be fairly limited as a basis for answering abstract general questions about the nature of satisfaction conditions. What typically happens, even when one is thinking of oneself as practicing high-church conceptual analysis, is that (i) some general claim about the semantics of a given concept or class of concepts is put forward as an *ideological hypothesis*, and (ii) various data are adduced as evidence for or against the hypothesis; in effect, the argument is that the data would be well explained (or would fail to be well explained) by the ideological hypothesis. (Thus, in effect the argumentation is a matter of inference to the best explanation.) One especially salient form of data, of course, involves spontaneous intuitive judgments about whether or not to apply a given concept to a given concrete scenario—e.g., whether or not to apply the concept *knowledge* to a given Gettier case, whether or not to apply the concept *water* to the substance XYZ on Twin Earth, and so forth. Thus, normally the intuitive judgments are about the concrete cases, whereas the reasoning about general ideological hypotheses is a matter not of direct intuitive judgment, but rather of abductive inference—inference to the best explanation.²¹

In light of these remarks, consider what is apt to happen when one introspectively attends to one's agentic presentational phenomenology with the goal in mind of forming an introspective judgment about whether or not this presentational content is compatible with determinism (or with physical causal closure, or with the mental state-causation of behavior). Those compatibility/incompatibility claims are *general ideological hypotheses*—albeit in this case, ideological hypotheses about presentational content rather than about judgmental content. So, given that competent judgmental deployment of concepts is so largely a matter of their correct use in forming *concrete* categorization judgments, should we expect a conceptually competent person to be able to tell directly, just by robust introspection, whether or not any of these general ideological hypotheses obtains vis-à-vis the person's own agentic presentational content?

No, we should not. It doesn't matter that agentic presentational content is self-presenting, in the manner of phenomenology quite generally. It doesn't matter because the task envisioned here, for robust introspection, is to form conceptually sophisticated *judgments* about this self-presented content—and, moreover, sophisticated judgments

concerning various general ideological hypotheses about such content. That is, the envisioned task is the *direct, intuitive*, formation of judgments about such general ideological hypotheses—just on the basis of introspection itself. It is unreasonable and unwarranted to expect one’s capacity for concept-wielding at the level of judgmental content to be that splendid when it is directed at general ideological hypotheses about agentive *presentational* content, just as it is unreasonable to expect it to be that splendid when it is directed at general ideological hypotheses concerning concepts that figure in *judgmental* content. When it comes to reliability and accuracy of direct, intuitive, judgments, the forte of conceptual competence is concrete judgments about specific cases. Ideological hypotheses are a matter for abductive inference—even if these ideological hypotheses pertain to presentational content rather than to judgmental content.

I have been describing an approach to agentive presentational content asserting that such content is constitutively determined by one’s current, occurrent, phenomenology alone; and I have been proposing an account of the limits of introspection about such content that is consistent with this approach. It bears mentioning, at this point, that one’s counterfactual phenomenological profile does still have an important bearing on our compatibility questions, even though I have denied that it plays a constitutive role vis-à-vis one’s occurrent phenomenal intentionality. Its bearing is not constitutive, but rather is epistemological. I have been maintaining that reasoning about general ideological questions—including general ideological questions pertaining to presentational content—is a matter of abductive inference. This being so, the evidentially relevant data includes, inter alia, various facts about what one *would* say or think about various concrete situations, if one *were* to find oneself in such situations (or one were to occurrently contemplate such situations). And all this data is part of one’s counterfactual *phenomenological* profile, since the counterfactual facts in question would remain unchanged even if one happened to be a lifelong brain-in-a-vat. The ideological hypotheses about our compatibility questions that have the highest overall abductive support—that best explain the full body of pertinent data—will be those hypotheses that best explain not just actual phenomenology but also counterfactual phenomenology. One’s counterfactual phenomenal profile thus emerges as very important evidentially vis-à-vis the compatibility questions—very important as a source of data to feed into abductive reasoning about the nature of agentive presentational content—even though it does not play a constitutive role in determining the presentational content of one’s actual agentive phenomenology.

Also worth mentioning is another important piece of data that feeds into the hopper of abductive reasoning about our compatibility questions. (This allows me a brief plug for compatibilism, even though defending compatibilism is not my concern in the present paper.) In ordinary life, we routinely and ubiquitously take for granted that we are agents of the sort that our agentive experience represents ourselves to be. (We also take for granted that other humans are such agents too.) In practice, we take for granted that we are *justified* in assuming that we are genuine agents (and that other humans are too). These facts are data, vis-à-vis the competing ideological hypotheses about the compatibility or incompatibility of agentive presentational content vis-à-vis determinism, physical causal closure, and the mental state-causation of behavior. Ceteris paribus, an ideological hypothesis is better than its competitors if (i) it renders actual epistemic standards epistemically appropriate, and (ii) its competitors render actual epistemic

standards woefully inadequate. (The hypothesis is better than its competitors because, *ceteris paribus*, it better explains actual epistemic practice—viz., in a way that avoids treating the use of those standards as a lapse in semantic and/or epistemic competence.) Well, compatibilism fares well with respect to this kind of data, whereas incompatibilism fares badly. For, if in fact the satisfaction conditions of agentive experience require the falsity of determinism, physical causal closure, or the mental state-causation of behavior, then actual epistemic standards are far too lax—since (let’s face it) we not only lack good evidence *against* all three hypotheses, but we possess rather good evidence in favor of at least two of them (viz., physical causal closure and the mental state-causation of behavior). So, the fact that compatibilism fits with actual epistemic standards we employ in our beliefs and assumptions about the reality of agency, whereas incompatibilism does not, is itself powerful abductive evidence in favor of compatibilism.

5.3. Interference effects from scorekeeping confusions

I submit that what was said in section 5.2 is sufficient by itself to discharge the burden of explaining the limitations of introspection vis-à-vis our compatibility questions, and to do so in way that comports with my claim that the presentational content of agentive experience is constitutively determined by occurrent phenomenology alone. But I also maintain that there are additional, complementary, factors that further contribute to the overall story about the limitations of introspection. Let me close with these. (My claims in section 5.2 could be accepted even by someone who has doubts about what I will now say, however.) The considerations I will now mention concern the sources of the temptation to think both (i) that one really *can* tell introspectively whether or not the content of one’s agentive presentational phenomenology is compatible with the three metaphysical hypotheses under consideration, and (ii) that introspection reveals that the answer to these compatibility questions is “No.”

Late in section 2 I pointed out that one might mistakenly think, when one actually introspectively ascertains feature (A) in one’s agentive phenomenology, that one introspectively ascertains (B):

- (A) *Not* presenting one’s behavior as state-caused
- (B) Presenting one’s behavior as *not* state-caused.

Such a mistake could arise, for instance, by fallaciously inferring the introspectibility of (B) from the introspectibility of (A)—a quantificational fallacy. Or it could arise from just conflating (A) and (B), without noticing the difference between them. But even though it might be fairly easy to commit such a fallacy or conflation, I suspect the temptation to think that one can introspectively ascertain feature (B) also has a deeper and more subtle root.

I maintain that many concepts that figure importantly in philosophical problems are governed by implicit, contextually variable, semantic parameters—and that some forms of philosophical puzzlement arise largely because (i) posing a philosophical problem can tend to shift the implicit parameters toward settings under which the claims made using a given concept are more “demanding” in their truth conditions than the claims that would normally be made using that concept, and (ii) one tends not to notice this shift of the “score in the language game” when one is contemplating the

philosophical problem. I favor a contextualist approach to the problem of knowledge and Cartesian skepticism, for instance.²² I agree with those who claim that the very posing of the possibility that one might be deluded by a Cartesian deceiver tends to alter the contextually operative settings on certain implicit semantic parameters that govern the concept *knowledge*—and tends to drive those parameter settings so high that, in the newly created context, virtually no belief counts as an instance of knowledge. I also agree with contextualists that in more *typical* contexts, the operative (and semantically appropriate) settings on the implicit parameters governing the concept *knowledge* are less demanding—and that under these parameter settings, numerous knowledge-claims are in fact true.

Likewise, I also favor a contextualist approach to the problem of freedom and determinism.²³ I maintain that the very posing of the question whether human freedom is compatible with state-causal determinism tends to alter the contextually operative settings on certain implicit semantic parameters that govern the concept *freedom*—and tends to drive those parameter settings so high that, in the newly created context, no item of behavior that is state-causally determined counts as *free*. I also maintain that in more typical contexts, the operative (and semantically appropriate) settings on the implicit parameters are less demanding—and that under these parameter settings, numerous freedom-attributions are in fact true.

Suppose that this contextualist orientation is right, with respect to the concept of freedom. (How to spell out the contextualist story about freedom in more detail would then be a significant task; different potential implementations might differ substantially from one another.) Then it probably extends to other, related, concepts too—in particular, the concept of *agency*. If so, then this concept too has these two features: (i) it has a limit-case setting of implicit contextual parameters under which an item of behavior counts as a genuine action only if it is *not state-caused* (and a fortiori is not *mentally* state-caused, and also is not subject to physical state-causal closure), and (ii) it has more typical, non-limit-case, parameter-settings under which an item of behavior can perfectly well be a genuine action even if it *is* state-caused.

What should be said about agentive *presentational* content, given this contextualist conception of judgmental content employing the concepts of agency and freedom? Three claims about agentive presentational content look very plausible to me. First, such content is not itself subject to implicit contextual parameters. One reason, *inter alia*, to think this is that non-humans probably have experience with some degree of agentive presentational content, whereas it is not credible that animal mentality is sophisticated enough to include the presence of implicit contextual semantic parameters.

Second, default settings on implicit contextual parameters for the concepts *freedom* and *agency* are probably such that the truth conditions for *judgments* attributing freedom and agency, under these default settings, are very tightly linked to the truth conditions for presentational agentive content. Although the implicit contextual parameters for judgmental content featuring these concepts can deviate in certain contexts from the default settings, nevertheless under the default settings the truth conditions for judgmental agentive content are closely aligned with the truth conditions for presentational agentive content.²⁴

Third, when one introspectively attends to one's agentic presentational phenomenology, while simultaneously asking oneself whether the intentional content of this phenomenology is compatible with state-causal determinism (or with physical causal closure, or with the mental state-causation of behavior), the judgmental dimension of such introspection is very apt to be infected by "scorekeeping confusions," in much the same way that such confusions arise when such compatibility questions are under consideration in a purely judgmental way and without attention to presentational phenomenology. That is, the very posing of the question one is introspectively trying to answer is apt to induce a shift in the implicit contextual parameters that govern the concepts *freedom* and *agency* away from their default values—and toward limit-case parameter settings, under which judgmental attributions of agency and freedom become incompatible with state-causal determinism (and with physical causal closure, and with the mental state-causation of behavior).

How do these three observations bear on the task of explaining, within the approach to agentic presentational content that treats it as constitutively determined only by current, occurrent, phenomenology, the limitations of introspection about compatibility questions? The first observation tells us that it would be too simplistic and flat-footed to try applying contextualism directly to presentational agentic content itself; dogs have agentic experience, but they don't traffic mentally in implicit contextual parameters.

The second observation eliminates any cart-before-the-horse problem like the one that confronted the approach to agentic presentational content that was canvassed in section 3 above: agentic presentational content constitutively constrains semantically appropriate *judgments* about agency and freedom (both actual judgments and counterfactual ones), rather than the other way around. (Default settings of implicit contextual parameters on judgmental agentic content are dependent on agentic presentational content, rather than agentic presentational content being constitutively dependent on counterfactual facts about how one would form judgments of agency and freedom in various counterfactual scenarios.)

The third observation is the one with the real explanatory punch. Introspective inquiry about one's agentic presentational phenomenology is subject to the following kind of subtle error. When one undertakes, by means of robust introspection about one's agentic presentational content, to answer compatibility questions about that content, one's judgmental deployment of the concepts of agency and freedom is apt to be unwittingly infected from the very start—by the unnoticed effects of implicit score-changes in contextual parameters governing these concepts. That is, one is apt to mistakenly think that the presentational content of one's free-agency phenomenology is self-evidently incompatible with state-causal determinism (and with the other two hypotheses)—when one's tendency to think so has actually arisen because the judgmental aspect of one's introspective processing has unwittingly been prodded into a scorekeeping confusion by the posing of the very question one is seeking to answer. Introspection about agentic presentational content thus gets perverted by a scorekeeping confusion—even though agentic presentational content *itself* is not even subject to contextually variable implicit parameters. Although one seems to oneself to be introspectively "reading off" a negative answer to the compatibility questions directly

from presentational agentive phenomenology itself, what one is really doing is deploying beliefs about that phenomenology using an introspective process whose judgmental dimension has been unwittingly tripped up by a subtle scorekeeping confusion. That unnoticed confusion could easily make one think that one is introspectively ascertaining feature (B) in one's agentive presentational phenomenology, when in fact one is doing no such thing.

6. Conclusion

Within the framework of phenomenal intentionality, the most natural and plausible thing to say about the constitutive determinants of presentational agentive content is that they are all a matter of one's current, occurrent, phenomenology. This does make it harder than it would otherwise be to explain the limits of introspection concerning questions about the compatibility or incompatibility of agentive presentational content with metaphysical hypotheses like state-causal determinism, physical state-causal closure, and the mental state-causation of behavior. But a plausible explanatory story is available even so—a story that stresses three facts. First, introspection about presentational intentional content involves the formation of beliefs that themselves have rich *judgmental* content, even though they are beliefs about the nature of presentational content. Second, normal competence in deploying the concepts that would figure in such beliefs simply does not provide one with the capacity to answer the compatibility questions just on the basis of robust introspection; this is because conceptual competence primarily involves the capacity to apply concepts to concrete cases, whereas conclusions about truth conditions, compatibility issues, and the like normally must be reached abductively rather than introspectively. Third, the attempt to answer the compatibility questions by robust introspection alone is especially liable to suffer from interference effects from scorekeeping confusions; such confusions are apt to arise, within the judgmental dimension of introspective inquiry, by virtue of the very posing of the compatibility questions one is seeking to address.²⁵

References

- Chisholm, R. (1964). Human freedom and the self. The Langley Lecture (University of Kansas). Reprinted in J. Feinberg and R. Shafer-Landau, eds., *Reason and Responsibility: Readings in Some Basic Problems of Philosophy*, 11th Edition (Wadsworth, 2002), pp. 492-99.
- Chisholm, R. (1995). Agents, causes, and events: The problem of free will. In T. O'Connor (ed.), *Agents, Causes, and Events: Essays on Indeterminism and Free Will*. Oxford University Press.
- Cohen, S. (1987). Knowledge, context, and social standards. *Synthese* 73, 3-26.
- DeRose, K. (1995). Solving the skeptical paradox. *Philosophical Review* 104, 1-52.
- Graham, G., Horgan, T., and Tienson, J. (2007). Consciousness and intentionality. In S. Schneider and M. Velmans (eds.), *The Blackwell Companion to Consciousness*. Blackwell.

- Graham, G., Horgan, T., and Tienson, J. (In press a). Phenomenology, intentionality, and the unity of mind. In A. Beckermann and B. McLaughlin (eds.), *The Oxford Handbook of Philosophy of Mind*. Oxford University Press.
- Henderson, D. and Horgan, T. (2000). What is a priori and what is it good for?. *Southern Journal of Philosophy* 38, Spindel Conference Supplement on the Role of the Empirical and the A Priori in Philosophy, 51-86.
- Henderson, D. and Horgan, T. (2002). ‘The a priori isn’t all that it’s cracked up to be, but it is something. *Philosophical Topics* 29, 219-50. Issue honoring Alvin Goldman.
- Horgan, T. (1979). “Could,” possible worlds, and moral responsibility. *Southern Journal of Philosophy* 17, 345-58.
- Horgan, T. (forthcoming a). Causal compatibilism about agentive phenomenology. To appear in T. Horgan, M. Sabates, and D. Sosa (eds.), *Supervenience in Mind* (MIT Press), a festschrift for Jaegwon Kim.
- Horgan, T. (forthcoming b). Mental causation and the agent-exclusion problem.
- Horgan, T. and Graham, G. (1994). Southern fundamentalism and the end of philosophy, *Philosophical Issues* 5, 219-47. Reprinted in M. DePaul and W. Ramsey (eds.), *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophy*, Rowman and Littlefield, 1998.
- Horgan, T. and Tienson, J. (2002). The intentionality of phenomenology and the phenomenology of intentionality. In D. Chalmers (ed.), *Philosophy of Mind: Classical and Contemporary Readings*. Oxford, 520-33.
- Horgan, T. and Tienson, J. (2005). The phenomenology of embodied agency. In M. Saagua and F. de Ferro (eds.), *A Explicacao da Interpretacao Humana: The Explanation of Human Interpretation*. Proceedings of the Conference Mind and Action III—May 2001. Lisbon: Edicoes Colibri, 415-23.
- Horgan, T., Tienson, J., and Graham, G. (2003). The phenomenology of first-person agency. In S. Walter and H. D. Heckmann (eds.), *Physicalism and Mental Causation: The Metaphysics of Mind and Action*. Imprint Academic, 323-40.
- Horgan, T., Tienson, J., and Graham, G. (2004). Phenomenal intentionality and the brain in a vat. In R. Schantz (ed.), *The Externalist Challenge*. Walter de Gruyter, 297-317
- Horgan, T., Tienson, J., and Graham, G. (2006). Internal-world skepticism and the self-presentational nature of phenomenal consciousness. In U. Kriegel and K. Williford (eds.), *Self-Representational Approaches to Consciousness*. MIT Press, 41-61.
- Lewis, D. (1996). Elusive knowledge, *Australasian Journal of Philosophy* 74, 549-67.
- Stephens, G. L. and Graham, G. (2000). *When Self-Consciousness Breaks: Alien Voices and Inserted Thoughts*. MIT Press.

Notes

1. Admittedly, it doesn't *follow* from the four theses that one should be able to ascertain answers to such questions. But there is *prima facie* tension among the theses nonetheless, because of the naturalness of the thought that the introspective accessibility of the factors that determine the answers to the philosophical questions should make introspectively accessible the answers themselves. If that thought is mistaken, then there is a felt need for an explanation of why and how it goes wrong. I maintain that the thought is indeed mistaken; what I am after is just such an explanation.
2. For some of my own writings in defense of these theses, often collaborative with George Graham and/or John Tienson, see Horgan (forthcoming a, forthcoming b), Horgan and Tienson (2002, 2005), Horgan, Tienson, and Graham (2003, 2004).
3. I do espouse a different way of harnessing knowledge-related considerations to argue against strong externalism and in favor of pervasive phenomenal intentionality. I maintain that advocates of strong externalism cannot adequately explain why "internal-world skepticism" (i.e., radical skepticism concerning one's beliefs about one's own current mental states) is psychologically non-gripping, and why such non-grippingness is rationally appropriate. An adequate explanation, I maintain, must eschew strong externalism and must appeal to phenomenal intentionality. Cf. Horgan, Tienson, and Graham (2006).
4. This section is adapted, with some modifications and additions, from section 1 of Horgan, Tienson, and Graham (2004).
5. I might have marked this distinction using the phrases 'conceptual content' and 'nonconceptual content'. But there seem to be almost as many ways of using that terminology as there are philosophers who use it. Also, depending in part on what one means by 'conceptual', presentational intentionality perhaps qualifies as conceptual anyway.
6. This section is adapted, with some modifications and additions, from section 2 of Horgan, Tienson, and Graham (2003).
7. Also important is the third-person phenomenology of agency, the "something it is like" to experience others as agents who are acting for reasons. For discussion, see Horgan and Tienson (2005).
8. Here and throughout I speak of 'state-causation' rather than 'event-causation'. More below on my reasons for this choice of terminology. States can be short-lived, and often when they do they also fall naturally under the rubric 'event.'
9. For discussion of a range of psychopathological disorders involving similar sorts of dissociative experience, see Stephens and Graham (2000).
10. The language of causation seems apt here: you experience your behavior as caused by you yourself, rather than experiencing it as caused by states of yourself. Metaphysical libertarians about human freedom sometimes speak of "agent causation" (or "immanent causation"), and such terminology seems phenomenologically appropriate regardless of what one thinks about the intelligibility and credibility of metaphysical libertarianism. Chisholm (1964) famously argued that immanent causation (as he called it) is a distinct species of causation from event causation (or "transeunt" causation, as he called it). But

he later changed his mind (Chisholm 1995), arguing instead that agent-causal “undertakings” (as he called them) are actually a species of event-causation themselves—albeit a very different species from ordinary, nomically governed, event causation. Phenomenologically speaking, there is indeed something episodic—something temporally located, and thus “event-ish”—about experiences of self-as-source. Thus, the expression ‘state causation’ works better than ‘event causation’ as a way of expressing the way behaviors are not presented to oneself in agentive experience. Although agentive experience is indeed “event-ish” in the sense that one experiences oneself as undertaking to perform actions at specific moments in time, one’s behavior is not experienced as caused by states of oneself.

11. The points made in this and the next paragraph, about different ways the phenomenology of purpose can work, are closely connected to the typology of different kinds of phenomenology of doing in Horgan and Tienson (2005).

12. With respect to successively more fine-grained details of action, specific purposes tend to be progressively less explicit phenomenologically, and progressively less accessible to consciousness—even for actions that result from conscious deliberation. For instance, when you consciously and deliberately decide to get yourself a beer by walking to the fridge in the kitchen and removing a beer from the fridge, the specific purpose in virtue of which your perambulatory trajectory toward the fridge angles through the kitchen doorway, as opposed to taking you directly toward the fridge and smack into the intervening wall, normally will color the phenomenology of your action without becoming explicitly conscious at all. And in some cases, sufficiently fine-grained aspects of one’s action might lack even this kind of subtle, non-explicit, phenomenological tinge of specific-purpose phenomenology. For instance, when you grab a can of peas from the grocery shelf, there might be nothing in the phenomenology that smacks even slightly of a specific purpose for grabbing the particular can you do rather than any of several other equally accessible ones. (Indeed, maybe there is no specific purpose for grabbing this can rather than any of the others, let alone a purpose that leaves a phenomenological trace.)

13. This is not to deny, of course, that there is indeed a distinctive phenomenology of effort of will that sometimes is present in the phenomenology of doing. The point is just that this aspect is not always present. A related phenomenological feature, often but not always present, is the phenomenology of trying—which itself is virtually always a dimension of the phenomenology of effort of will, and which often (but not always) includes a phenomenologically discernible element of uncertainty about success. (Sometimes the phenomenological aspect of voluntariness attaches mainly to the trying dimension of the phenomenology of doing. When you happen to succeed at what you were trying to do but were not at all confident you could accomplish—e.g., sinking the 10 ball into the corner pocket of the pool table—the success aspect is not experienced as something directly under voluntary control.)

14. On the other hand, it seems that the following higher-order feature of agentive presentational content is indeed introspectively accessible, when one introspects in a suitably careful way: being such that the presence or absence of feature (B) is not itself introspectively accessible.

15. This counterfactual phenomenological profile might be nondeterministic, even if one’s mental life supervenes on underlying neurobiological processes that are themselves

deterministic. For, how one's phenomenal experience would unfold might depend in part upon the specific way one's total mentality is physically realized on a given occasion. To take a simple example, how one's experience will evolve temporally, from a given initial phenomenal state and a given body of sensory input, depends on whether or not one is (phenomenally undetectably) on the verge of a sudden stroke.

16. Normally the inner workings of one's mind, including whatever implementational mechanisms are involved (a working brain supplied with suitable nutrients, for instance), are thought of as not being tampered with. In a brain-in-vat scenario, the tampering is all "outside the transducers." Likewise, a Cartesian deceiver confines itself to providing misleading sensory input; it does not otherwise interfere with the experiencer's normal mental processing.

17. I address the nature of this semantic connection in Horgan (forthcoming b). There I propose (a) that the judgmental concepts of agency and freedom are governed by contextually variable parameters, (b) that presentational agentive content is not subject to such parameters, (c) that under the default settings of the judgmental parameters, the satisfaction conditions for judgmental agentive content coincide with those for presentational agentive content, (d) that under these default settings, attributions of agency and freedom are compatible with determinism, and (e) that there are limit-case settings of the implicit contextual parameters under which judgmental attributions of agency and freedom are not compatible with determinism. This approach is compatibilist; and it is semantically contextualist about judgmental agentive content, but not about presentational agentive content. (See also section 5.3 below.)

18. You will also find, if you try it, that angle x cannot be calculated just by filling in values for the remaining angles of all the subtriangles in such a way that for each subtriangle, the sum of the values assigned to its three interior angles is 180 degrees. There is no unique way to do that. The problem is much harder than it initially looks.

19. Leave aside trigonometry—which is easy for me to do myself, since I only studied it for about two weeks in high school and have long since forgotten almost all of what little of it I once learned. (I do have in hand a trigonometric solution that was sent to me by a student and was generated by a commonly used piece of mathematics software; it employs what appear to be differential equations containing trigonometric expressions. But I don't remember enough trigonometry or calculus even to understand what the equations mean, let alone to understand why they are true.) Try solving the problem just via geometrical construction(s). I myself have been playing with this problem for over two years now, and at present my complicated sequence of interconnected geometrical constructions is still not yet complete. I curse the day I encountered this problem—in a South American airline magazine (of all places). My brother tells me that his son's honors-math teacher in high school speculated that the problem has no unique solution. A philosopher I know who specializes in philosophy of mathematics has confidently insisted to me that the problem has no unique solution. But I persist, because I find it (gulp!) introspectively obvious that there must be a unique solution. David Chalmers, formerly the multi-year high school math champion of Australia and later a graduate student in mathematics at Oxford University before switching to philosophy, thought about the problem briefly when I posed it to him, but then wisely set it aside and returned to thinking about philosophy.

20. One can imagine someone who has visual-recognitional skill concerning the size of angles that is comparable to auditory “perfect pitch” concerning musical notes. Such a person would be able to tell the size of an angle at a glance, confidently and accurately. But this would be an unusual and super-normal skill, and would certainly exceed what constitutes normal human mastery of geometrical concepts.
21. On the idea of “ideology” as a form of abductive inquiry in which intuitive judgments about concrete hypothetical scenarios figure as empirical data, see Horgan and Graham (1994). For further discussion that classifies this form of inquiry as “low grade a priori,” see Henderson and Horgan (2000, 2002).
22. See, for instance, Cohen (1987), DeRose, (1995), Lewis (1996). Although I am not completely happy with the details of any specific version I have seen of contextualism about the concept of knowledge, I am very sympathetic to the general approach.
23. See Horgan (1979, forthcoming a), Horgan and Graham (1994), Henderson and Horgan (2000).
24. Think, for instance, of turning over one’s wallet or purse to a thief who is holding a gun in one’s face. In terms of presentational agentive content, one’s experience still has the aspect of agency, and thus the ‘could do otherwise’ aspect—even though refusing to comply would be palpably irrational. Accordingly, under the default settings of implicit contextual parameters for the judgmental concept of freedom, it is correct to say or think, “I freely complied with the thief’s demand, even though it would have been irrational to refuse.” This is so even though the parameter-settings can deviate from their defaults. The judgment that one’s behavior was not free because it was coerced, for example, invokes a more demanding setting of the implicit contextual parameters for the concept freedom—a setting that becomes contextually appropriate, for instance, when issues of moral blameworthiness arise. (Thus, both judgments are correct in context, and they do not directly conflict with one another.)
25. Thanks to Michael Gill, George Graham, Uriah Kriegel, Keith Lehrer, Sean Nichols, John Pollock, Susanna Siegel, John Tienson, and Mark Timmons for ongoing discussion and feedback.