

The Phenomenology of Agency and Freedom: Lessons from Introspection and Lessons from Its Limits^{*}

Terry Horgan^{**}
thorgan@email.arizona.edu

ABSTRACT

The paper will focus on three interrelated matters. First is the phenomenology of agency, the “what it is like” of experiencing oneself as an agent – and more specifically, the experiential aspect of freedom that is an integral part of the phenomenology of agency. Second is the extent to which introspection is, or is not, a reliable way to answer questions about the phenomenology of agency and freedom. Third is the import of these first two matters for philosophical debates about agency and free will.

Briefly, my overall position goes as follows. The phenomenology of free agency has features that are well and aptly described by language of the kind that is traditionally employed by advocates of metaphysical libertarianism concerning the free will issue – language like “self as ultimate source,” and “agent as cause”. This is something that is reliably detectable by introspection. However, introspection by itself cannot reliably ascertain whether or not the satisfaction conditions for free-agency phenomenology require, for example, the falsity of state-causal determinism or the presence of the metaphysically heavyweight attribute that metaphysical libertarians call “agent-causal freedom”.

Moreover, the best overall theoretical position about the nature of free agency – the one that emerges by abductive “inference to the best explanation” of all pertinent, evidentially relevant, factors – is compatibilist. Among the considerations that underwrite this abductive conclusion is the fact that a suitable version of compatibilism can provide a full accommodation of the phenomenology of free agency; i.e., the right kind of compatibilism entails

* My thanks to Michael McKenna and Mark Timmons for helpful comments and discussion.

** University of Arizona

that normal humans do indeed exercise free agency, and also entails that their agentic experience does not misrepresent the nature of free agency itself.

1. SOME RELIABLY INTROSPECTIBLE ASPECTS OF AGENTIVE PHENOMENOLOGY²

I begin by describing some features of agentic phenomenology which, I submit, are readily ascertainable just on the basis of introspective attention to such phenomenology.

What is behaving like phenomenologically, in cases where you experience your own behavior as action? Suppose that you deliberately do something – say, holding up your right hand and closing your fingers into a fist. What can you ascertain about the phenomenology of this item of behavior, on the basis of introspective attention to this phenomenology? To begin with, there are of course the purely bodily-motion aspects of the phenomenology – the what-it’s-like of being visually and kinesthetically presented with your 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 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*

² This section is adapted, with some modifications, deletions, and additions, from similar sections in Horgan 2007b and in Horgan *et al.* 2003.

*passive phenomenology of psychological state-causation of bodily motion.*³ People often do passively experience causal processes *as* causal processes, of course: the experience of seeing the collision of a moving billiard ball with a motionless billiard ball is an experience as-of the collision causing the latter ball's subsequent motion; the experience of observing the impact of the leading edge of an avalanche with a tree in its path is an experience as-of the impact causing the tree to become uprooted; and so on. Sometimes people even experience their own bodily motions as state-caused by their own mental states – e.g., when one feels oneself shuddering and experiences this shuddering as caused by of a state of fear. 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.⁴

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 psychological 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 generated by *yourself*.

The language of causation seems apt here too, but differently deployed: 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* apt 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

³ 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.’

⁴ For discussion of a range of psychopathological disorders involving similar sorts of dissociative experience, see Stephens and Graham (2000).

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.

The phenomenology of doing typically includes another aspect which will be especially important in the context of the present paper: what I will call *core optionality*. (More presently on the reason for the modifier ‘core’). 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 optionality 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 effortfully exerting one’s will is really only one, quite special, case of the much more pervasive phenomenology of optionality⁵).

The core-optionality aspect of agentive phenomenology is intimately bound up with the aspect of self-as-source, in such a way that the former is an essential component of normal agentive self-source experience.⁶ In experiencing one’s

⁵ 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 core optionality 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).

⁶ I say that the aspect of core optionality is an essential component of *normal* self-source experience because I mean to leave open the possibility of unusual self-source experiences that lack

behavior as emanating from oneself as its source, one experiences oneself as being able to refrain from so behaving – or at any rate, as being able to refrain from willfully producing such behavior. This is so even when acts under extreme coercion or duress – e.g., handing over one’s wallet or purse to a thief who is pointing a gun in one’s face. It also is so even when one acts with an extreme phenomenological “imperativeness” – e.g., a mother’s unhesitatingly leaping into the river to save her drowning child, Luther’s acting out a sense of moral requirement (as expressed by his declaring “Here I stand, I can do no other”), the compulsive hand-washer’s act of washing hands for the third time in ten minutes. The *core* phenomenology of optionality that is essential to ordinary agentic experience remains present in all such cases, even though there are further, superimposed, phenomenological aspects (duress, moral-obligation experience, intensely strong irrational desires, or the like) whose presence can render appropriate, in context, a judgment that the agent “could not have done otherwise,” or “had no other option,” or “did not act freely”. Because the phenomenology of core optionality remains present even in such cases, it also can be contextually appropriate to use ‘could’ and ‘option’ and ‘free’ in a way that reflects this fact (rather than in a way that reflects the presence of one or another kind of superimposed non-optionality phenomenology). For instance, one might say this: “I could have refrained from giving the thief my wallet, and thus I gave it to him freely and with the option of refraining – even though refraining would have been quite stupidly irrational”. Hereafter I will use the expression ‘free-agency phenomenology’, in order to refer to the experience of self-as-source in a way a way that underscores the aspect of core optionality that is an essential component of normal self-as-source experience.

A few words are in order at this point about thought-experimental “Frankfurt scenarios” inspired by Frankfurt (1969). One such scenario is this: one’s body would have moved the same way even if one had not willed it to move that way, because a device implanted in one’s motor cortex would have triggered that same motion had one not willfully produced it; but in that case the motion would not have been experienced as willfully generated, and indeed

this aspect – for instance, self-source experiences in which one firmly believes that one is in a “Frankfurt scenario” in which one’s circumstances are such that were one about to will to refrain from performing the act one is about to perform, an evil scientist would cause the pertinent bodily motions to occur anyway and would also cause these motions to be accompanied by (epiphenomenal) experience-as-of willfully performing that action. More momentarily on Frankfurt scenarios.

would not have been experienced as one's own action. A different Frankfurt scenario is this: one's body would have moved the same way even if one had not willed it to move that way, because a device implanted in one's motor cortex would have triggered that same motion had one not willfully produced it; in addition, that device would have triggered the phenomenology of willing to move one's body in just that way – with the dual triggering operating in a manner that renders the phenomenology itself completely epiphenomenal vis-à-vis the bodily motion. As far as free-agency *phenomenology* is concerned (and that is the present topic), the main thing to stress is the following: in both of these scenarios (and in most Frankfurt-style scenarios), one's free-agency phenomenology is at least partially *non-veridical*, because the phenomenology includes not only the self-as-source aspect but also the could-do-otherwise aspect that is an essential component of normal self-as-source experience. The agent's phenomenology is as-of being a *full-fledged* self-source of the behavior, where full-fledgedness includes being such that one could have acted otherwise instead; but in Frankfurt scenarios, the agent is not a full-fledged self-source of the kind that the agent experiences himself/herself to be.⁷ These remarks about agentive phenomenology leave various moral and metaphysical questions still open – e.g., (i) whether the agent in a Frankfurt scenario is morally responsible for the action, (ii) whether the agent is a genuine self-source of the behavior even though the agent could not have done otherwise, and (iii) whether the agent acts freely even though the agent could not have done otherwise. Whatever one might say about those questions, the key point is that the self-as-source aspect of normal agentive *experience* includes the core optionality (core “could-do-otherwise”) aspect as an essential element.

Agentive phenomenology is more closely akin to perceptual/kinesthetic experience than it is to discursive thought. (Many higher non-human animals, I take it, have some 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

⁷ What about the Frankfurt scenario envisioned in note 6, in which one firmly believes that one is in a scenario in which core optionality is absent? Perhaps here one's agentive phenomenology would be as-of *non*-full-fledged self-as-source-hood in which the core-optionality aspect is lacking. But that would be extremely unlike ordinary agentive phenomenology. (Alternatively – as I myself suspect would be the case – perhaps even here the core-optionality aspect still would be present in one's agentive *phenomenology* despite one's *belief* that core optionality itself is absent. Compare experiences of the Muller-Lyer illusion, in which one horizontal line still looks longer than the other even when one firmly believes the two lines are the same length).

animals do); but thoughts employing these concepts are not to be conflated with agentive phenomenology itself.

2. SOME LIMITATIONS OF INTROSPECTION VIS-À-VIS AGENTIVE PHENOMENOLOGY⁸

The phenomenal character of one’s current experience is self-presenting to the experiencing subject. Self-presentingness is an especially intimate form of direct acquaintance between the experiencing subject on one hand, and the phenomenal character of some aspect of the subject’s current state of phenomenal consciousness; the state’s appearing a certain way, acquaintance-wise, is constitutive of the state’s actually being that way.

Let a *purely phenomenological question* be a question that (i) is about some aspect of the intrinsic phenomenal character of one’s present experience, and (ii) is such that the answer is entirely determined just by the intrinsic phenomenal character of one’s present experience. (The point of clause (ii) is to exclude questions that bring in some extrinsic aspect while still being in some sense “about” intrinsic phenomenal character – e.g., “Am I now undergoing an experience with the phenomenal character that I was writing about last Tuesday?”).

In light of the fact that phenomenal character is self-presenting, one might be tempted to think that any purely phenomenal question can be reliably answered directly on the basis of introspection. More specifically, one might be tempted to think that introspection alone can reliably determine whether or not free-agency phenomenology has metaphysical-libertarian satisfaction conditions. (Having metaphysical-libertarian satisfaction conditions means this: the intentional content of one’s free-agency experience is veridical only if one is an “agent-cause” in the metaphysically heavyweight sense of this notion that is invoked by metaphysical libertarians – which entails, *inter alia*, that state-causal determinism is false).

I maintain, however, that this claim about the powers of introspection vis-à-vis free-agency phenomenology is false. (Hence the more general thesis – that

⁸ This section is adapted, with some modifications and deletions, from section 3 of Horgan (in press-b). Other pertinent discussions of mine, sometimes collaborative, are Horgan (2007a, in press-a) and Horgan and Timmons (in press).

any purely phenomenal question can be reliably answered directly on the basis of introspection – is also false). In this section I will briefly say why.⁹

Let me begin by introducing some terminology. First, I distinguish two kinds of introspection concerning one's current experience. On one hand is *attentive* introspection: paying attention to certain aspects of one's current experience. On the other hand is *judgmental* introspection: the process of forming a judgment about the nature of one's current experience, and doing so spontaneously just on the basis of attending to the aspect(s) of one's current experience about which one is judging – without any reliance on collateral information or evidence. (Judgmental introspection thus deploys attentive introspection, while also generating a judgment about what is being attended to).

Second, I call a purely phenomenal question *conceptual-competence amenable* (for short, CC amenable) just in case it can be correctly answered by simply introspectively attending to one's current experience and then spontaneously exercising one's conceptual competence with the pertinent concepts. By contrast, a purely phenomenal question is *conceptual-competence transcendent* (for short, CC transcendent) just in case it *cannot* be correctly answered this way.

With these distinctions at hand, consider now the following three pairwise-incompatible claims about the satisfaction conditions of free-agency phenomenology.

- (1) Free-agency phenomenology has satisfaction conditions that (i) are fully fixed by intrinsic phenomenal character alone, and (ii) are metaphysical libertarian.
- (2) Free-agency phenomenology has satisfaction conditions that (i) are fully fixed by intrinsic phenomenal character alone, and (ii) are compatible with state-causal determinism (and hence are not metaphysical-libertarian).
- (3) Free-agency phenomenology has satisfaction conditions that (i) are not fully fixed by phenomenal character alone, (ii) instead are fixed by phenomenology in combination with extra-phenomenological facts about the experiencing agent's cognitive architecture, and (iii) are

⁹ For more extended elaboration and defense of the view, see Horgan (in press-b).

such that their being metaphysical-libertarian or not, and their being compatible with state-causal determinism or not, depends upon those cognitive-architecture facts.

Claims (1) and (2) both construe free-agency phenomenology as having “purely narrow” referential purport that lacks any constitutive externalistic elements, whereas claim (3) construes it as having “wide” referential purport that incorporates certain constitutive externalistic elements. For the phenomenology to have wide referential purport is for its reference-relation to its referent-property (if it has a referent-property) to depend constitutively not merely on the intrinsic character of the phenomenology itself, but also upon certain phenomenology-external facts about the nature of the experiencing agent – according to claim (3), facts about the agent’s cognitive architecture. On one potential view that comports with claim (3), the pertinent facts would concern the nature of the cognitive-architectural choice-generating and behavior-generating mechanisms that are normally operative in situations where the experiencing agent undergoes free-agency phenomenology, and meeting the satisfaction conditions would be a matter of exercising those cognitive mechanisms in the normal way.

Claims (1) and (2), on the other hand, construe free-agency phenomenology as referring, in the experience of all actual and possible creatures who are phenomenal duplicates of one another, to one and the same property – regardless of any differences in the cognitive architectures of different phenomenal duplicates.¹⁰ The essence of the property that constitutes free agency is entirely fixed by the intrinsic phenomenal character of free-agency experience alone. Claim (1) says that this phenomenologically fixed property has metaphysical-libertarian satisfaction conditions, whereas claim (2) says that it has satisfaction conditions that are compatible with state-causal determinism (and hence are not metaphysical-libertarian).

Consider now the following question, which pertains entirely to the intrinsic phenomenal character of agentive experience and whose answer depends only on that phenomenal character – and which is therefore a purely phenomenological question:

¹⁰This property need not actually be instantiated by the creature in order to be the referent-property of the creature’s free-agency experience. Indeed, it need not even be a property whose instantiation is metaphysically possible. (Maybe it is a metaphysical-libertarian property, and maybe – as some hard incompatibilists maintain – the instantiation of such a property is outright impossible regardless of whether or not state-causal determinism is true).

(Q) Which (if any) of the pairwise incompatible claims (1)-(3) is correct?

At the moment, the issue I am focusing upon is not what the answer is to question (Q), but rather this: whether or not one can reliably ascertain, just via judgmental introspection, what the answer is. I claim that one *cannot* do so, and that the reason why not is that (Q) is a CC transcendent question. Elsewhere (Horgan, in press-b) I defend these claims, and I also offer a proposed multi-component debunking explanation of the common judgmental-introspective beliefs that (a) one *can* reliably answer question (Q) just on the basis of introspection, and (b) that the answer is that claim (1) is true.

An explanatory task arises at this point that needs addressing – viz., the task of explaining credibly *why it should be* that (Q) is a CC-transcendent question. Since claims (1)-(3) all concern only the phenomenal character of free-agency experience, and since phenomenal character is self-presenting to the experiencing agent, something needs saying about why human agents are nonetheless unable to “read off” the answer to question (Q) just by directing their attentive introspection upon their own free-agency experience and then exercising their conceptual competence with concepts like the concept of state-causal determinism and the concept of free-agency phenomenal character.

I have addressed this explanatory task most extensively in in Horgan (in press-b); there is also pertinent discussion in Horgan (2007a, 2007b) and in Horgan and Timmons (in press). Although I lack the space here to rehearse my proposed account, let me just mention 3 key elements of the account. First, normal conceptual competence is mainly a matter of being able to correctly apply a given concept *to a concrete case* – or more precisely, to do so modulo one’s available evidence; consequently, conceptual competence alone is apt to be fairly limited as a basis for answering abstract general questions about the nature of satisfaction conditions. Second, these same facts about conceptual competence are in play when one introspectively attends to one’s agentic phenomenology with the goal in mind of forming an introspective judgment about question (Q): it is unreasonable and unwarranted to expect one’s capacity for concept-wielding to be that splendid when it is directed at general hypotheses concerning the intentional content of agentic phenomenology, just as it is unreasonable to expect it to be that splendid when it is directed at general hypotheses concerning the satisfaction conditions for concepts themselves. Hence third, general hypotheses about satisfaction conditions are

a matter for abductive inference – even when these hypotheses concern facts about the intentional content of self-presenting phenomenal character, facts that are fully fixed by that phenomenal character itself.

3. LESSONS

Let me draw out some lessons of the above discussion, with respect to philosophical debates about free agency. To begin with, participants in these debates need to explicitly acknowledge the existence of free-agency phenomenology – including its self-as-source dimension, and including the core optionality (core can/could do otherwise) aspect that is itself an essential component of normal self-as-source experience.¹¹

Second, it needs to be appreciated that there are intimate interconnections among these three matters: (1) the satisfaction conditions of free-agency phenomenology, (2) the satisfaction conditions of everyday statements and judgments that ascribe free agency or classify specific acts and decisions as the products of free agency, and (3) the metaphysics of free agency. Item (1) is apt to constrain item (2), in the following way: if free-agency phenomenology has metaphysical-libertarian satisfaction conditions, then thereby so do everyday ascriptions of free agency, whereas if free-agency phenomenology has compatibilist satisfaction conditions, then thereby so do everyday ascriptions of free agency. In addition, item (1) is apt to constrain item (3), as follows: if genuine free agency exists at all, then it fully conforms to the satisfaction conditions imposed on it by agentic phenomenology. (I will express these modes of constraint by saying that free-agency phenomenology *strongly constrains*, respectively, the concept of free agency and the metaphysics of free agency. And I will say that an overall position that treats the concept of free

¹¹ Some philosophers, notably Eddy Nahmias and his collaborators, do pay attention to free-agency phenomenology and yet deny that it really has an aspect of self-as-source. (See, e.g., Nahmias *et al.* 2004). But they appear to assume that if there were such an aspect, then (a) this aspect would have metaphysical-libertarian satisfaction conditions, and (b) its having metaphysical-libertarian satisfaction conditions would be reliably ascertainable introspectively. They thereby conflate two claims: (1) the claim that agentic phenomenology has a self-as-source aspect, and (2) the claim that agentic phenomenology has a self-as-source aspect with features (a) and (b). In my view they would be right to deny claim (2), but they are wrong to deny claim (1) – and they unfortunately muddy up the dialectical waters by conflating the two claims.

agency and the metaphysics of free agency as strongly constrained by free-agency phenomenology is a *strongly internally coherent* position).

Third, it is important to articulate various package-deal positions that simultaneously address items (1), (2), and (3), and it is important to subject such positions to comparative cost-benefit assessments *as* package deals. Concerning item (1), a package-deal position will embrace just one of these two (incompatible) claims: (1a) phenomenological libertarianism, asserting that free-agency phenomenology has metaphysical-libertarian satisfaction conditions, or (1b) phenomenological compatibilism, asserting that such phenomenology has compatibilist satisfaction conditions. Likewise, concerning item (2) there are two options: (2a) conceptual libertarianism, asserting that everyday free-agency ascriptions have metaphysical-libertarian satisfaction conditions, or (2b) conceptual compatibilism, asserting that such ascriptions have compatibilist satisfaction conditions. Concerning item (3) there are three options: (3a) metaphysical libertarianism, (3b) metaphysical compatibilism, or (3c) hard incompatibilism.

Fourth, barring powerful countervailing theoretical considerations, theoretical package-deal positions that are strongly internally coherent will be much more likely to be correct than those that are not. (The default theoretical presumptions are that free agency has the features it is experienced as having, and that the concept of free agency has satisfaction conditions that conform well to the satisfaction conditions of free-agency experience. People implicitly adopt these presumptions routinely, and people routinely implicitly take the presumptions to be epistemically well warranted. In principle, one could challenge these default presuppositions, but doing so in a credible way would require some heavy-duty, hard-to-envison, form of argumentation). A strongly internally coherent package-deal position will have these two features: first, it embraces (1a) if and only if it embraces (2a), and it embraces (1b) if and only if it embraces (2b); second, it asserts that if there is such a genuine phenomenon as free agency at all, then that phenomenon conforms to the satisfaction conditions laid down by free-agency phenomenology.

The fifth moral is conditional: if one can reliably ascertain, just on the basis of introspection, that free-agency phenomenology has metaphysical-libertarian satisfaction conditions, then there are only two package-deal positions that are strongly internally coherent, viz., (1a) + (2a) + (3a), and (1a) + (2a) + (3c). The first of these embraces phenomenological libertarianism, plus conceptual libertarianism, plus metaphysical libertarianism. This package deal is

libertarian through and through. The second view embraces phenomenological libertarianism, plus conceptual libertarianism, plus hard incompatibilism. This package deal asserts that there is no such phenomenon as free agency, on the grounds that (i) genuine free agency would have to conform to metaphysical-libertarian satisfaction conditions, and (ii) no real phenomenon conforms to such conditions.

The sixth moral is also conditional, and is a corollary of the fifth one: if one can reliably ascertain, just on the basis of introspection, that free-agency phenomenology has metaphysical-libertarian satisfaction conditions, then there is no viable compatibilist package-deal position that is strongly internally coherent. Thus the best one could do, by way of formulating a package-deal position that honors the introspectively manifest fact that free-agency phenomenology has metaphysical-libertarian satisfaction conditions, would be to adopt a partial-error theory asserting that although there really is a phenomenon of free agency, the nature of this phenomenon is very significantly misrepresented by free-agency experience. That kind of view is a very unattractive theoretical option for those who are inclined to reject metaphysical libertarianism. One reason to think so, *inter alia*, is that whatever phenomenon the account ends up treating as the one picked out by free-agency experience will be so different in reality from how it is experienced to be that there will be very little credible basis for claiming that it is an eligible referent of free-agency phenomenology (or of the concept of free agency).¹²

The six morals lately mentioned all draw upon the discussion in section 1 above, concerning reliably introspectible aspects of free-agency phenomenology. Let us now factor in the discussion in section 2, concerning the limitations of introspection with respect to free-agency phenomenology. That discussion yields this seventh moral: it is not the case that one can reliably ascertain, just on the basis of introspection, what the answer is to question (Q). This in turn brings an eighth moral in its wake, as a corollary: *viz.*, it is not the

¹² For taxonomic completeness, the following additional moral is worth mentioning, also conditional in form: if one can reliably ascertain, just on the basis of introspection, that free-agency phenomenology has *compatibilist* satisfaction conditions, then the only strongly internally coherent package-deal position that conforms with the introspectively ascertainable nature of free-agency phenomenology is package-deal compatibilism, *i.e.*, (1b) + (2b) + (3b). But it is extremely implausible to claim that it is introspectively *obvious* that self-as-source phenomenology has compatibilist satisfaction conditions, and I know of no compatibilist who does claim this. Rather, compatibilists tend either to ignore free-agency phenomenology altogether (the more typical tendency), or else to deny that agentive phenomenology has a self-as-source aspect at all (as do Nahmias and his collaborators).

case that one can reliably ascertain, just on the basis of introspection, that free-agency phenomenology has metaphysical-libertarian satisfaction conditions.

This leads to a ninth moral: there is another package-deal position that is consistent with what is reliably introspectively ascertainable about agentic phenomenology – viz., the position (1b) + 2(b) + 3b). This view is thoroughly compatibilist – phenomenologically, conceptually, and metaphysically – and is therefore strongly internally coherent. It begins with the contention that free-agency phenomenology has compatibilist satisfaction conditions. It then claims that free-agency phenomenology constrains both the concept of free agency and the metaphysics of free agency – in such a way that the concept has compatibilist satisfaction conditions too, and in such a way that genuine free agency is a phenomenon that is compatible with state-causal determinism (and hence is not correctly characterized by metaphysical libertarianism).

A tenth moral, also grounded in my discussion in section 2 of the limitations of introspection, is that there is an important role for abduction when one inquires about the satisfaction conditions of free-agency phenomenology – a role that is complementary to the roles of attentive and judgmental introspection, and that potentially can take up the slack left by introspection. That is good news for compatibilists, myself included.

4. SKETCH OF A VERSION OF PACKAGE-DEAL COMPATIBILISM

Let me now briefly sketch the version of package-deal compatibilism that I favor.¹³ I have defended various aspects of this overall approach in a number of prior writings, some collaborative (Horgan 1979, 2007a, 2007b, in press-a, in press-b, Graham and Horgan 1994, Henderson and Horgan 2000, Horgan and Timmons in press). The argumentation in those writings is largely abductive, and incorporates the contention that one cannot reliably ascertain the satisfaction conditions of free-agency phenomenology just on the basis of careful introspection.¹⁴

¹³ This section is adapted, with some modifications and deletions, from section 4 of Horgan (2007b).

¹⁴ I believe that there is significant work yet to be done by way of further elaborating my recommended approach – in particular, there is a need to say more about the satisfaction conditions of free-agency phenomenology, and about why and how these conditions can be met even if state-causal determinism is true. I am unhappy with possible-worlds satisfaction conditions according to which the possible worlds that are “accessible” to a freely choosing/acting agent include worlds in which a

As a prelude, let me distinguish two kinds of mental intentionality, which I call *presentational* content and *judgmental* content, respectively. Presentational intentional content is the kind that accrues to phenomenology directly – apart from whether or not one has the capacity to articulate this content linguistically and understand what one is thus articulating, and apart from whether or not one has the kind of sophisticated conceptual repertoire that would be required to understand such an articulation. Judgmental intentional content, by contrast, is the kind of content possessed by such linguistic articulations, and by the judgments they articulate. (Here I use ‘judgment’ broadly enough to encompass various non-endorsing propositional attitudes, such as *wondering whether*, *entertaining that*, and the like). Dogs, cheetahs, and numerous other non-human animals presumably have agentic phenomenology with presentational intentional content, although it is plausible that they have little or no sophisticated conceptual capacities of the kind required to undergo states with full-fledged judgmental content involving concepts like freedom or agency.

I do not mean to suggest that this distinction is a sharp one. It wouldn’t surprise me if the two kinds of content blur into one another, via a spectrum of intervening types of psychological state and/or a spectrum of increasing forms of conceptual sophistication in different kinds of creatures. Also, it may well be that the two kinds of content can interpenetrate to a substantial extent, at least in creatures as sophisticated as humans. It is plausible, for instance, that humans can have presentational contents the possession of which require (at least causally) a fairly rich repertoire of background concepts that can figure in judgmental states. One can have presentational experiences, for instance, as-of computers, automobiles, airplanes, train stations – all of which presumably require a level of conceptual sophistication that far outstrips what dogs possess.

“divergence miracle” occurs shortly before the agent chooses/acts otherwise than how the agent chooses/acts in the actual world. I am even more unhappy with satisfaction conditions according to which some “accessible” possible worlds are allowed to differ somewhat from the actual world at all moments in time prior to the agent’s non-actual choice/act. An idea that currently appeals to me is this: do the semantics of modals in terms of “scenario-specifications” that (a) are *epistemically* possible (relative to some contextually pertinent body of background information), and (b) need not be metaphysically possible. As regards modals about human agency, some such scenario-specifications will hold fixed the portion of the actual world that precedes a given agent’s choice/act, will specify some way the agent chooses/acts that differs from the agent’s actual-world choice/act, and will also specify that there are no violations of any actual-world laws of nature.

Briefly, the version of package-deal compatibilism that I favor comprises the following eleven theses. First, the presentational content of agentive phenomenology includes the aspect of self-as-source, which itself normally includes the aspect of core optionality (core “can/could do otherwise”) as an essential component.¹⁵ Second, the presentational intentional content of agentive phenomenology has satisfaction conditions that are compatible with state-causal determinism. Third, this compatibility is a non-manifest feature of agentive phenomenology; i.e., one cannot reliably tell, just on the basis of careful introspective attention to one’s own agentive experience and the exercise of one’s conceptual competence in judgment-formation, whether or not the compatibility hypothesis is true. Fourth, despite the compatibility of agentive phenomenology with state-causal determinism, a bodily event that is experienced as one’s action cannot also be *experienced* as state-caused, either by non-mental states or by mental states. Fifth, the presentational aspect of core optionality remains present as an essential component of normal agentive phenomenology even when one experiences oneself as acting under coercion or duress. Sixth, an essential aspect of experiences of state-causation, including experiences of one’s own bodily motions as state-caused, is the presentational aspect of *inevitability* – i.e., the aspect of inevitability *given the circumstances and the causing events*. Seventh, the two theses lately mentioned jointly explain the phenomenological mutual exclusion described in the fourth thesis: this exclusion results from the core optionality aspect of agentive phenomenology on one hand, and from the inevitability aspect of the

¹⁵ Many recent versions of metaphysical compatibilism about free agency not only ignore free-agency phenomenology altogether (including the phenomenological aspect of core optionality), but also presuppose both (a) that the capacity to choose otherwise and do otherwise is incompatible with state-causal determinism, and (b) that the “can/could do otherwise” feature is simply never required for genuine free agency. Compatibilists who affirm claim (b) typically do so because of the conceivability of Frankfurt-style scenarios – and they then go on to affirm (a) by conceding to the incompatibilists the latter’s own favored construal of ‘can/could do otherwise’. All this seems to me to be seriously mistaken. Even if there are possible scenarios in which one exercises free agency even though it is not the case (because of a preempted potential cause waiting in the wings) that one can/could do otherwise, it doesn’t *begin* to follow that the capacity to do otherwise is *never* required for genuine free agency. On the contrary, that capacity remains a *defeasibly* necessary condition for free agency, Frankfurt-style cases notwithstanding. My three biggest complaints about dominant versions of metaphysical compatibilism in the recent philosophical literature are (1) that they ignore free-agency phenomenology, (2) that they grossly overestimate the (quite limited) significance of Frankfurt-style scenarios, and (3) that they concede to incompatibilists the contention that if determinism is true then people can never choose or act differently than they actually do choose and act.

phenomenology of state-causation on the other hand. One cannot experience an item of one's own behavior both as inevitable and as something that one could have refrained from doing.

Eighth, at the level of *judgmental* intentional content, the concept of free agency involves a feature that is probably not exhibited by the free-agency aspect of *presentational* intentional content – viz., implicit contextual parameters that determine, in context-specific ways, contextually operative standards of satisfaction. For instance, in many contexts the standards operate in such a way that an action performed under extreme coercion – e.g., with a gun in one's face – do not count as free. I.e., under the contextually operative standards, the *judgment* that such an action is not free is correct.¹⁶ (In other contexts, however, the concept of freedom is correctly used in such a way that its satisfaction conditions coincide with those for the core optionality aspect of sensory-experiential intentional content – for instance, when one says “I could have refused to give the gunman my wallet, although that would have been a foolhardy thing to do; thus, I exercised freedom of choice in giving it to him”).

Ninth, the implicit contextual parameters governing the judgmental concept of free agency can take on a limit-case setting in certain contexts of judgment or conversation – i.e., a parameter-setting under which an item of behavior counts as a free action only if (i) it is not state-causally determined, and (ii) it comes about as a result of metaphysical-libertarian “agent causation” involving the self as a godlike unmoved mover.

Tenth, the satisfaction conditions for *presentational* free-agency intentional content – i.e., for free-agency *phenomenology* – coincide with certain non-limit-case, compatibilist, satisfaction conditions for *judgmental* free-agency intentional content. The satisfaction conditions for agentive phenomenology do *not* coincide with the incompatibilist satisfaction conditions that accrue to judgmental free-agency intentional content when the implicit parameters at work in the judgmental concept of free agency have extremal, limit-case, settings.

Eleventh, the metaphysics of free agency is constrained by the intentional content of free-agency phenomenology, and thus is also constrained by the (matching) intentional content of everyday, non-limit-case, ascriptions of free agency. So, since the phenomenological content and the conceptual content

¹⁶ Such judgments will normally be keyed to certain aspects of phenomenology too, aspects that are superimposed upon the underlying phenomenology of core optionality – e.g., the phenomenology of duress under threat, the phenomenology of moral imperativeness, and the like.

are compatibilist, free agency itself is a phenomenon that is compatible with state-causal determinism.

Elsewhere, sometimes collaboratively, I have set forth arguments in support of the various theses constituting this version of package-deal compatibilism. Contextualist compatibilism about the judgmental concept of freedom, in a form that acknowledges limit-case parameter-settings that are incompatibilist, is defended in Horgan (1979), Graham and Horgan (1994), Henderson and Horgan (2000), and Horgan (forthcoming). Other aspects of the full package-deal are defended in Horgan (2007a, 2007b, in press-a, in press-b), and in Horgan and Timmons (in press). I will not argue for the position here, because of space limitations.

I do recognize that when one attends introspectively to one's free-agency phenomenology, with its presentational aspect of self-as-source which itself includes the aspect of freedom as an essential component, and when one simultaneously asks reflectively whether the veridicality of this phenomenology requires one to be an "agent cause" in the sense espoused by metaphysical libertarianism, one feels *some* tendency to judge that the answer to this question is Yes. If the position I have sketched is correct, then this tendency embodies a mistake: the satisfaction conditions of free-agency agentive phenomenology do not require heavyweight, metaphysical libertarian, "agent-causal freedom," and do not require the falsity of state-causal determinism. I certainly acknowledge that a theoretically adequate version of package-deal compatibilism should provide a plausible explanation of this mistaken judgment-tendency – an explanation of why the tendency arises so strongly and so naturally, once the compatibility issue is explicitly raised. I have addressed this challenge elsewhere, e.g., Horgan (2007a, 2007b, in press-a, in press-b). Although I lack the space here to summarize the "respectful debunking" explanation I have offered for incompatibilist judgment tendencies, let me just say that my proposed explanation draws on two principal resources: first, that fact, already stressed, that agentive phenomenology and the phenomenology of state-causation are mutually exclusionary, and second, the contextualist element that I claim is operative in judgmental attributions of free agency.

So the version of package deal compatibilism I favor, which is contextualist about the concept of free agency, allows for a fairly plausible explanation of the incompatibilist-leaning judgment-tendencies that naturally tend to arise when one asks whether free-agency phenomenology is compatible with state-causal

determinism. When one factors this into the mix, alongside the various convergent forms of largely abductive evidence (not set forth here) that favor both phenomenological compatibilism and conceptual compatibilism, I think a strong case can be made in support of an overall position that is phenomenologically compatibilist, conceptually compatibilist about everyday free-agency ascriptions, and metaphysically compatibilist.

5. CONCLUSION

Although the rich and distinctive phenomenology of agency went largely ignored in mainstream philosophy of mind in the twentieth century, it is now receiving renewed attention in that branch of philosophy. Agentive phenomenology also received far too little attention in twentieth-century philosophical discussions of freedom and determinism – with advocates of compatibilism probably being the worst offenders. It is time to bring the phenomenology of free agency explicitly into the freedom/determinism debate, and to accord it significant weight. A complete treatment of the freedom/determinism issue should address three topics together: the phenomenology of free agency, the concept of free agency, and the metaphysics of free agency. All else equal, a package-deal treatment of these topics should be strongly internally coherent – i.e., it should treat the phenomenology of free agency as strongly constraining both the concept of free agency and the metaphysics of freedom. This theoretical desideratum would spell big trouble for compatibilism if one could reliably ascertain, directly on the basis of introspection, that free-agency phenomenology has metaphysical-libertarian satisfaction conditions. But there are strong reasons to think that introspection is simply not that powerful – a fact that opens up room for abductive considerations to enter the dialectical mix. Once such considerations are properly brought to bear and given their due epistemic weight, I maintain, the overall package-deal position that will look best in terms of theoretical cost-benefit evaluation will be phenomenologically compatibilist, conceptually compatibilist (yet also conceptually contextualist), and metaphysically compatibilist.

REFERENCES

- Chisholm, R. (1964). *Human freedom and the self. The Langley lecture* (University of Kansas). Reprinted in J. Feinberg & R. Shafer-Landau (Eds.) (2002), *Reason and responsibility: Readings in some basic problems of philosophy*, 11th Edition, (pp. 492-499). New York: Wadsworth.
- 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*, (pp. 95-100). New York: Oxford University Press.
- Frankfurt, H. (1969). Alternate possibilities and moral responsibility. *Journal of Philosophy*, 66(3), 829-839.
- Graham, G., & Horgan, T. (1994). Southern fundamentalism and the end of philosophy. *Philosophical Issues*, 5, 219-247.
- Henderson, D., & Horgan, T. (2000). What is a priori and what is it good for? *Southern Journal of Philosophy, Spindel Conference Supplement*, 38, 51-86.
- Horgan, T. (2007a). Agentive phenomenology and the limits of introspection. *Psyche*, 13(2).
- Horgan, T. (2007b). Mental causation and the agent-exclusion problem. *Erkenntnis*, 67(2), 183-200.
- Horgan, T. (in press-a). Causal compatibilism about agentive phenomenology. In T. Horgan, M. Sabates & D. Sosa (Eds.), *Supervenience and Mind: Essays in Honor of Jaegwon Kim*. Cambridge, MA: MIT Press.
- Horgan, T. (in press-b). Introspection about phenomenal consciousness: Running the gamut from infallibility to impotence. In D. Smithies & D. Stoljar (Eds.), *Introspection and consciousness*. Oxford: Oxford University Press.
- Horgan, T., Tienson, J., & 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*, (pp. 323-340). Exeter: Imprint Academic.

- Horgan, T., & Timmons, M. (in press). Introspection and the phenomenology of free will: Problems and prospects. *Journal of Consciousness Studies*.
- Nahmias, E., Morris, S., Nadelhoffer, T., & Turner, J. (2004). The phenomenology of free will. *Journal of Consciousness Studies*, 11(7-8), 162-179.
- Stephens, G. L., & Graham, G. (2000). *When Self-Consciousness Breaks: Alien Voices and Inserted Thoughts*. Cambridge, MA: MIT Press.

