

Presentation and the Ontology of Consciousness

Paul M. Livingston

University of New Mexico

pmliving@unm.edu

Abstract

The idea that we can understand key aspects of the metaphysics of consciousness by understanding conscious states as having a *presentational* character plays an essential role in the phenomenological tradition beginning with Brentano and Husserl. In this paper, the author explores some potential consequences of this connection for contemporary discussions of the ontology of consciousness in the world. Drawing on Hintikka's analysis of epistemic modality, the author argues that the essential presentational character of consciousness can be seen as accounting for the familiar failure of substitutivity of identicals in intensional contexts of conscious presentation and further provides for an ontology of conscious presentation that avoids metaphysically problematic commitments to "special" intentional objects. The author next considers how the irreducibility of intensional contexts, if grounded in the presentational character of consciousness, also may provide for a kind of irreducibility of the "first-person" perspective to third-person or objective description across possible worlds. Finally, drawing on Chalmers' "two-dimensionalist" semantic framework, the author argues that attention to the presentational aspect of consciousness as thus understood, while it provides significant motivation for the existence of the "hard problem" of the irreducibility of consciousness to physical and structural description, is nevertheless consistent with a globally monist, rather than dualist, ontology.

Keywords

consciousness – presentation – epistemic possibility – irreducibility – phenomenology

In this paper, I consider some potential implications for the metaphysics and ontology of consciousness of an idea that is crucial to many central developments of the phenomenological tradition beginning with Brentano and Husserl. This is the idea of an essential link between *consciousness* and

presentation, such that we can understand key aspects of the metaphysics of consciousness by understanding conscious states as having, essentially, a *presentational* character. As I shall discuss it in this paper, maintaining this idea involves, minimally, maintaining that (at least some) conscious states present something *as* something, or “give” or make available some particular thing as being some way. In addition to its basic importance for phenomenology, this idea also played an important role in motivating earlier accounts of mind and experience in the analytic tradition, underlying, for instance, Russell’s conception of acquaintance and Schlick and Carnap’s understanding of “protocol sentences.” Additionally (as I shall argue), it can be seen as bearing a close connection to Frege’s own favored metaphor for the sense of a singular term, that of a “mode of presentation” or “way of being given” [*Art des Gegebensein*]. Although recent analytic discussions of the ontology and metaphysics of phenomenal consciousness have not always considered its presentational character as centrally, I shall argue that it bears deep and important implications for these discussions. In particular, if considered in the context of recent modal and two-dimensional arguments for the falsity or limitations of physicalism, consideration of the presentational character of consciousness motivates a novel kind of ontological option for its placement in the world. This option vindicates the irreducibility of consciousness (in one sense of “irreducible”) to description or explanation in terms only of physical facts, and clarifies this irreducibility as resulting ultimately from broadly modal/semantic features of the presentation and individuation of entities across possible worlds. At the same time, it does not thereby require or invite anything like a substance or property dualism, since the features of consciousness that make for its irreducibility can also (as I shall argue) be accounted for by means of a global *monism* of substances and properties.

As Husserl often explains, phenomenology’s central method is to explicate and describe the content and structure of what is directly *presented* or *given* in consciousness. In his 1913 statement of the phenomenological “principle of all principles” in *Ideas* 1, Husserl (1913[1983], 43–44) invokes the methodological priority of what is directly presented in this sense as the basis for all phenomenological research:

No conceivable theory can make us err with respect to the *principle of all principles*: that every originary presentive intuition is a legitimizing source of cognition, that everything originarily (so to speak, in its ‘personal’ actuality) offered to us in ‘intuition’ is to be accepted simply as what it is presented as being, but also only within the limits in which it is presented there.

What this means is that the basis of all phenomenological reflection and analysis – the basic “material” for all its results, including those that bear on ontological and metaphysical issues – is what is directly “there” or present in consciousness. This does not mean that one must begin with specialized “inner” objects or representations, such as (as it might be) sense-data, ‘inner’ impressions, or empiricist “ideas”. Rather, the beginning point is the presentation of entities and phenomena in general, just *as* these are presented: as having, that is, the determinate characteristics and attributes that they are presented *as* having.¹

The idea of a characteristic structure of presentation, whereby something is presented as something, also remains central for ontological inquiry within other historically prominent developments of phenomenology, even when they do not foreground reflective, conscious intentionality in Husserl’s sense. These include Heidegger’s idea of a basic “hermeneutic” as-structure, underlying both linguistic and non-linguistic intentionality, and Sartre’s conception of the “for-itself” as the distinctive region of being defined by its presentational self-relation.² For Husserl himself, the most important methodological tool for phenomenological reflection on what is presented is the method that he terms that of phenomenological *epoche* or “bracketing.” In the *epoche*, a directly presented content is “bracketed” or “put in parentheses” by disregarding its existential bearing on the real world, so that reflective regard may be directed instead to the content itself. For example, if I perform the *epoche* with regard to my current perception of a flowering tree, I “bracket” or “put out of commission” the assumption of the existence of the tree itself so that I may consider explicitly the *content* of the perception itself, that whereby the tree itself is presented to me. This content is termed the “noema” and itself centrally characterized by what Husserl calls a “noematic sense,” the sense or meaning through or by which the tree is presented perceptually in consciousness. Noematic senses can thus be seen analogous to Fregean senses and can also be understood as “modes of presentation” of their objects; but the idea of a “mode

1 Though I will not argue for this in detail here, this last point is helpful in showing that phenomenology need not involve a “Myth of the Given” in the sense of Sellars’ *Empiricism and the Philosophy of Mind*: since phenomenological givenness is typically of *something as something*, it is not a “bare sensing” in the manner of the sense-data and empiricist theories that Sellars criticizes.

2 For Heidegger’s “hermeneutic” as-structure, see (Heidegger 1927, 158–60); for Sartre’s conception of the for-itself as based in a “pre-reflective” presentative *cogito*, see (Sartre 1943), Part II, Chapter 1, “Immediate Structures of the For-Itself;” especially Section 1: “Presence to Self.”

of presentation" is here generalized beyond the linguistic cases to which Frege typically restricts himself, taking in a wide variety of perceptual and other cognitive modes of intentionality.³

In further considering the implications of this kind of view of presentation for contemporary discussions, it is important to bear in mind two qualifications that distinguish it from others in the vicinity. The first is that this view of presentation should not be taken as implying *representationalism*, or the view that all conscious states are or involve 'internal' representations.⁴ The reason this should not be presupposed is that the kind of presentation that Husserl invokes may be (and in the most interesting cases, is) *direct*: it may proceed, that is, without requiring the mediation or existence of any kind of symbolic, internal, or cognitive *re*-presentation. My current perception may simply involve that I am presented with a flowering tree before me; as far as the direct phenomenological reflection of this presentation is concerned, there is no need to assume that this requires any 'internal' representation which mediates the presentation to me. Second, it is not necessary, in order to hold the view to be explored and defended here, to maintain either that *all* or *only* conscious states have a presentational or intentional content. Nothing about the view to be defended should be taken as excluding the possibility of conscious states that do not present anything, or at least that do not present anything other than themselves. At the same time, there is nothing about the view to be defended here that requires that presentation, in the relevant sense, be exhibited *only* or even "primarily" by conscious states as opposed to other kinds of intentional entities (such as spoken or written words or sentences, signs, pictures, etc.) Indeed, it is an important feature of the way that the current view characterizes the basic structure of presentation that it can plausibly be generalized to these other cases, and thereby seen as essentially neutral with respect to the medium (whether conscious, linguistic, pictorial or whatever) of the presentation itself.

3 This is a formulation of the so-called "sense-content model" of the noema, first proposed by Dagfinn Føllesdal in (Føllesdal 1969). Although others have contested its correctness as an interpretation of Husserl, for present purposes it can simply be taken as a coherent option in conceptual space, independently of the interpretive issues.

4 As an anonymous reviewer has pointed out to me, "representationalism" is perhaps more often used in contemporary discussions to characterize the view that experiential states represent *that* things are thus-and-so, rather than just the claim that conscious states involve internal representations at all. However, in this paper I will use the term in the older (and broader) sense.

1

Much contemporary analytic discussion of the metaphysics of mind takes place against the broad backdrop of the influential framework of “possible worlds” semantics that gained currency in the 1960s, 70s, and 80s. In its most prominent developments, reasoning about possible worlds has been used to establish conclusions about the metaphysics of modality, reference, and propositions, rather than being seen as bearing directly on the nature or metaphysics of consciousness. But an important line of analysis, mostly pursued earlier on in the history of the development of the relevant issues but again essential for the possibility of (what is today called) “two-dimensional” semantics, points to the possibility and significance of an analysis of the content of consciousness in terms of possible worlds. As I shall argue, this line of analysis, though most typically couched as a matter of “epistemic” possibilities or modalities as distinct from “metaphysical” ones, is actually more deeply and fundamentally motivated by considering the *presentational* aspects of consciousness. Thereby, it bears important consequences, which have not generally been appreciated, for the metaphysical place of consciousness in the world.

To begin to see some of these consequences, it is helpful to review the instructive application of modal reasoning to the semantics of propositional attitudes broadly, and perceptual attitudes specifically, made by Jaako Hintikka in a pair of articles from 1967 and 1969. Hintikka’s approach to these issues develops in the context of discussions of the problem of “quantifying in” to modal and propositional contexts, especially those that can be considered “intensional” or “non-extensional” in that they do not generally permit the substitution of co-referring terms while preserving truth.⁵ In particular, Hintikka considers the kind of quantification evidently involved in locutions such as “knows who”, “sees what”, “has an opinion as to the identity of,” etc. Such locutions apparently involve “de re” attitudes toward a specific individual, but may fail of substitutionality of co-referring terms; this makes it initially unclear how to apply existential and universal quantification with respect to them, and the unrestricted application of the quantifiers easily results in contradiction. But as Hintikka argues, these cases can after all be handled quantificationally, given the expanded resources of a semantics of *possible worlds* or *possibilities* (Hintikka 1969, 96–98).⁶ The key is to relax the assumption that *a*’s attitude

5 For the issue, see (Carnap 1947), (Quine 1953), (Quine 1956), (Quine 1960) (especially Chapter 4), and (Kaplan 1968).

6 Here and in other places (e.g. (Hintikka 1962)), Hintikka develops the relevant alternative possibilities as *model sets*, which differ from possible worlds in that their specification may

(of knowing, seeing, or whatever) apparently directed to an individual is indeed directed (only) toward the specific *real-world* individual, and instead require only that it be directed to the *same* individual in each of the worlds *compatible with a's attitude*. For example, on this kind of view, my *seeing that the person in front of me is wearing a hat* is not to be understood simply in terms of my relationship to the actual-world person I am seeing. Rather, it involves my relation to each of the individuals, *across* possible situations, compatible with my current perceptual state itself (including the relevant background information that helps determine the specific content it presents). Thus, on this view, such an attitude is not a matter only of my relation to an actual individual; rather, its explication involves specifying a unique individual in each of the possible worlds that are epistemically possible alternatives for me.⁷ The content of the attitude can then be identified with the set of possible worlds (or possibilities) compatible with the success, or truth, of the attitude with respect to the individual (in that world) so specified.

As Hintikka points out, however, this points to a further conceptual problem which must be addressed in order for the approach to be workable: that of the conditions under which we may treat as identical distinct individuals of different possible worlds. Under what conditions, and with what right, do we speak of individuals in different possible worlds as being identical (Hintikka 1969, 99)?⁸ As Hintikka suggests, we may naturally address this problem by postulating a set, F , of “individuating functions” defined over the set of possible worlds or models we are considering. Each of these functions f will pick out, for each situation or world μ , at most one individual from its domain of individuals $I(\mu)$. We allow that the functions may be partial – there may be, at some worlds (including possibly the actual one), no individual designated. These individuating functions then take the role of the “individuals” previously imagined as the objects of the relevant propositional attitudes (Hintikka 1969, 100–101).

be ‘partial’ in the sense that it specifies only a limited range of states of affairs or “situations”, rather than a “total possibility” or world as a whole. In what follows, I will typically ignore this distinction (and speak of “possible worlds” where Hintikka has “possibilities” or “models”), however, since it does not appear to bear essentially on the issues discussed.

- 7 This set is specified in terms of a relation of epistemic “alternativeness” holding between possibilities.
- 8 This is of course the familiar problem of (what David Kaplan called) “trans-world heir lines”. It has been addressed in a number of ways, including the one Hintikka suggests, Lewis’s “counterpart” theory, and several others.

This readily accounts for the failure of substitutivity of identity in “intensional” contexts such as those of knowledge and belief. For example, Smith may know who Jones is (by means of identifying information, perceptual acquaintance, or other means), and also (trivially) that Jones is Jones, without knowing that Jones is the spy. The reason is that the set of worlds (or scenarios) compatible with what Smith knows includes some in which Jones is not the spy (even if he is, in the actual world). Further, as Hintikka notes, this has the important consequence that we can naturally characterize the failure of substitutivity of identicals, which is characteristic of propositional attitudes, even when only singular terms and variables (rather than descriptions) are involved.⁹ In particular, in case there are two distinct individuating functions f_1 and f_2 , relevant to α ’s attitudes involving singular terms, such that $f_1(\lambda) = f_2(\lambda)$ but not $f_1(\mu) = f_2(\mu)$, the general rule of substitutivity will fail with respect to α ’s attitudes Q .

$$(\forall x)(\forall y)(x = y \rightarrow (Q(x) \rightarrow Q(y)))$$

As Hintikka points out, the availability of different individuating functions in this sense bears important consequences for the actual semantics of intentionality across possible worlds. Hintikka allows that, for certain limited purposes, one might sometimes speak of the apparently different individuals correlated by one of the “individuating” functions just as “being” one and the same individual. This is, after all, just what we ordinarily mean when we speak of “the same” individual as it is picked out or presented by means of a particular modality of individuation (say, descriptively, perceptually, in memory, etc.) (Hintikka 1969, 101). However, as Hintikka also notes, this “reification” of the individuating functions into actual individuals can also be misleading, for at least two significant reasons. First, and highly relevantly to the question of the epistemology and phenomenology of presentation, it has the effect of obscuring what is, as Hintikka says, “an extremely important non-trivial part of our native conceptual skills, namely, our capacity to recognize one and the same individual under different circumstances and under different courses of events” (Hintikka 1969, 101). This highly complex set of skills is, on the one hand, plausibly irreducible to any simple knowledge of an individual’s own characteristic or essential properties, or to whatever is thought to constitute its “metaphysical” identity across possible worlds; but on the other hand, it is

9 Thus, for instance, the framework can characterize the possible failure of substitutivity of identicals in “Jones knows that Hesperus is Phosphorus”, where “Hesperus” and “Phosphorus” function as singular terms rather than descriptions.

just what the individuating function f , for the relevant case of individuation, is intended to capture. There is thus, as Hintikka says, no reason to assume, in general, that the individuating functions f will individuate in such a way as to correspond simply to anything describable by speaking simply of the partial identities between the domains of individuals of the various possible worlds (Hintikka 1969, 101–102).

A second, and equally important, reason why the reification of the functions into individuals is misleading is that it obscures the significant possibility of *cross-individuation* between members of the same possible worlds. If cross-individuation in this sense is possible, as Hintikka notes, then there is in fact no real basis for the stability of our everyday sense of an “ordinary” individual; this stable sense is rather to be seen just as the reflection of one particular method of identification, enjoying *sub specie logicae* no more than a relative priority over others (Hintikka 1969, 102). Moreover, as Hintikka argues in detail in the earlier article “On the Logic of Perception,” the possibility of distinct cross-identification, whereby two individuating functions, both relevant to one’s attitudes, may cross-identify members of the same set of possible worlds, is plausibly actually realized in some cases of perception. When, for example, Smith perceives a man before him, but does not see who the man is, the content of Smith’s perception picks out, across possible worlds, what are really many different individuals, in the sense of material or physico-psychological individuals (Hintikka 1967, 171). Similarly, given descriptions of distinct states of affairs compatible with what Smith sees, we may ask whether two individuals involved in these distinct states of affairs are the same or different, as far as Smith’s perception is concerned (Hintikka 1967, 171–72). The cogency of these methods of identification and their relevance to phenomenological description demonstrates once more the arbitrariness – at least from the perspective of an analysis of phenomenological or intentional content – of any single method of individuation thought of as ontologically “real” in a privileged way or metaphysically absolute.

As Hintikka argues, it is also possible to see in the distinctive features of perceptual cross-identification the essential, though misinterpreted, root of traditional accounts of “sense-data” and other sorts of distinctively “intentional” objects (Hintikka 1967, 166–68; 177–78). In particular, perceptual identification may clearly be (and in fact typically is) incomplete: for example, I may perceptually identify the piece of chalk, *c*, on the table as being white even though I do not recognize that it is also the smallest object on the table (when in fact it is). Then the intersubstitutivity of identicals fails for the obvious reason: where *s* stands for “the smallest object on the table,” I thus do not recognize that *s* is

white, even though (in fact) $c = s$. Given this, it is easy (though also misleading) to conclude that the object actually perceived, since it does not bear the properties of the physical chalk, is a distinct one: namely, a peculiar object of perception such as a sense-datum or other “intentional” object (Hintikka 1967, 164–66). As Hintikka notes, traditional formulations of the “Argument from Illusion” also turn on this kind of conclusion: in cases of erroneous perception (so the argument goes), the object of perception cannot be the real physical object (if any) before one, so the real content of the (illusory) perception must rather be a special kind of non-physical object which is then assumed to be present in the general case, i.e. whether or not the perception is illusory with respect to its ‘external’ object (Hintikka 1967, 162–64). However, in both cases, given the existence of the relevant individuating functions, the inference to a special kind of perceptual or intentional object is highly misleading.

Given that the perceptual individuating function that characterizes, e.g., my (partial) perception of the chalk as white picks out, across possible worlds, a wide variety of (physically individuated) objects with varying physical properties, there is no need to suppose or countenance a distinct *kind* of entity as the “perceptual” or “intentional” one. To the extent that there is any reason to speak of such entities, these are just the “reified” forms of the relevant individuating functions. But since the ranges of these functions are just ordinary entities in the various possible worlds, there can be no question, as Hintikka emphasizes, of countenancing an *ontological* distinction, on the level of the entities themselves, between “intentional” and non-intentional (for instance “physical” or “material”) objects (Hintikka 1967, 172). The entities over which the variables range are, in each case, of the same type. What is different is just the functions that correspond to different methods of individuation across possible worlds; this is not a difference that appears *within* any one of the worlds, but rather only when entities within them are related to one another. As Hintikka also suggests, we may then think of the individuating functions themselves as in certain ways analogous to, or as replacements for, what Frege understands as the senses (*Sinne*) of singular terms (Hintikka 1967, 180). Just as Frege had held about senses, the individuating functions contain something that (in a certain way) goes beyond the various references themselves. Here, however, this “something beyond” is not an additional abstract entity, but just the method of individuation that is captured in the relevant function itself. We can then understand the function as itself capturing (at least for the relevant kind of singular term or presentation) the distinctive “mode of givenness” by which “the object” (i.e., the reified form of the individuating function) is given or presented. But there is no need to postulate a separate mediating object

or entity, beyond the function itself. Indeed, the tendency to invoke abstract entities such as *Sinne* or propositions to account for individuation can here be seen, Hintikka suggests, as of a piece with the invocation of sense-data in the perceptual cases. In both kinds of case, the real basis of the invocation is just the individuating functions across possible worlds; and once we have seen this, there is no longer any reason to suppose any *ontologically distinctive* type of “thing”, over above the functions and the ordinary (actual or possible) entities they refer to themselves.

2

It is clear that Hintikka’s analysis, if correct, has important implications for (what we might call) the “semantics” of conscious intentionality: for, that is, the question of how to understand what it is that conscious states present or presentationally make available, across possible worlds.¹⁰ These implications go beyond simply the semantics of “propositional attitudes” in the usual sense, for (as Hintikka’s examples concerning perception in particular make clear) the general picture here suggested is not restricted to the presentational content of linguistically expressed or even expressible propositions. Rather, it explicitly extends to cases of perception in which there need not be any specific linguistically articulated judgment or content, but occurrent conscious states may nevertheless be seen as, *qua* conscious, bearing presentational contents in

10 In *Husserl and Intentionality*, Chapters 6–8, David Woodruff Smith and Ronald McIntyre give a comprehensive and illuminating analysis of such a phenomenological semantics against the backdrop of Husserl’s own “horizon-analysis” of sense. Smith and McIntyre argue that the possible-worlds analysis can indeed be seen as an explication of Husserl’s notion of horizon, that a framework of possible worlds can thereby be seen as explicating Husserl’s conception of noematic *Sinn*, and that in particular, a Hintikka-style analysis facilitates the phenomenological treatment of *de re* (including, importantly, perceptual) intention. They disagree with Hintikka (as well as with the present analysis), however, by requiring (366–69) a perceptually acquainting sense to involve an indexical component which is “rigid” – i.e. which is to be understood as picking out the *same* individual in all possible worlds (where *same* is understood metaphysically rather than epistemically), and further suggest essentialist criteria for trans-world (metaphysical) identity in terms of natural kinds (372–73). On the current view, despite the far-ranging implications of Smith and McIntyre’s analysis as a whole, this particular approach to trans-world individuation threatens to forfeit most of the phenomenological richness and promise of Hintikka’s account, substituting for its analysis of the content of presentation criteria (i.e., those in terms of natural properties and kinds) that are essentially irrelevant to it.

the sense of allowing and excluding certain possibilities. Moreover, as Hintikka also notes, there is no evident reason to limit the relevance of the analysis to perception; other modalities of consciousness which involve methods of individuation depending on the situation of the individual (including memory and knowledge) may also be analyzed by means of it (Hintikka 1967, 181). Quite generally, wherever there are presentational states that present (actual or non-actual) objects as having certain determinations or features, we may take a Hintikka-style analysis to be relevant to the analysis of this presentational content, and in particular to the way that they individuate “the” entities they present. As Hintikka himself suggests, if the criteria of individuation are perhaps generally clearer in the perceptual case, this does not exclude the possibility of extending the analysis of the particular individuating functions characteristic, for an individual, of the other modalities of that individual’s consciously presentational states as well. Here, the overall shape of the content of each modality, and the specific contents of which it is capable, will be determined, in each case, by the ways in which it individuates its objects, or the individuating functions it deploys across worlds.

But even beyond these implications for the *semantics* of conscious intentionality, the possibility of a Hintikka-style analysis also has, as I shall argue in the next two sections, very significant implications for the *ontology* of conscious intentionality as well. We can begin to see some of these implications by considering how the analysis allows us to apply possible-worlds semantics to types of content (here, perception is a leading example) that are, in a distinctive way, given *only* from a specific point of view, and are thus, in general, plausibly *inaccessible* from a completely objective standpoint that abstracts from all specific points of view. This will be the case, in particular, if it can be shown that, for a presentational state bearing a particular presentational content, that content is, for principled reasons, directly accessible *only* to or for someone who is actually *in* or undergoing that state. There is here an evident, although only partial, analogy to what have been understood as indexical “characters” or (alternatively) modes of presentation: what is expressed, on an occasion, with “*Here* it is warm” or “*Now* it is 3 o’clock” has its determinate content only from the perspective of the person expressing it, or someone else who shares that perspective. Just as in the indexical cases, the content which is accessible to me when I am in a particular (say) perceptual state is plausibly available to me only *by means of* that state itself and *because* I am in it: it is only because I am enjoying a certain (say) perceptual state that I am able to individuate the “objects” presented in that state, across possible worlds, in just the way that I thereby do. By marked contrast with the indexical cases, however, the content to which I thereby do have access is (as we have seen) highly

complex and determinate, involving the complicated and distinctive pattern of cross-identifications between individuals in distinct possible worlds that is captured by the relevant individuating functions.

When this content is available to me on the basis of my current perceptual state, not only is it not *necessary* that I should be able to state in objective, third-person and propositional terms a rule or finite expression determining that function, but it may actually be *impossible* to do so in general. To see why, it is helpful to consider the problem of trans-world individuation from another direction. In "Individuation by Acquaintance and by Stipulation," David Lewis discusses individuation by acquaintance, within the framework of his own counterpart theory of cross-world identification. In two possible worlds, W_1 and W_2 , two otherwise similar entities, Y (in W_1) and X (in W_2) may both be objects of acquaintance (e.g., of perception), respectively, for subjects Z (in W_1) and U (in W_2). Are X and Y then counterparts by acquaintance for the subject Z ? The answer, Lewis notes, depends on whether U is Z (or, on Lewis's framework, which does not allow for actual cross-world identities but only counterpart relations, whether Z is U 's closest counterpart in W_1) (Lewis 1983, 13–14). In order, then, to describe how something is cross-identified by acquaintance for a certain subject, we must apparently first cross-identify the subject; and how are we to do this, in general and objective terms?

As Lewis argues, there is in fact no way to do so. We cannot identify the subject by acquaintance, i.e. by appealing to her relationship of acquaintance with *herself*: for while every subject presumably enjoys this kind of self-relation with herself *within* a world, we cannot assume that this relationship exists between subjects in different worlds without begging just the question of the subject's trans-world identity that we are trying to answer (Lewis 1983, 14). On the other hand, we also cannot cross-identify by description; for a subject's self-description may be badly mistaken, and in that case there may be, in some or many of the worlds compatible with the subject's belief, nothing sufficiently like that subject to be it (or its counterpart). In this case, there will be no way to identify the subject across some of her genuine doxastic alternatives, since in at least some of those alternatives she does not exist (Lewis 1983, 14–16). More generally, any attempt to cross-identify the subject by appealing to (any kind of) substantial information about it – including about its self-conception – founders on the realization that one must first show that the information really is about *that* subject, and hence that its cross-world identity, which is to be established, must rather be presupposed (Lewis 1983, 17). Finally, as Lewis correctly points out, it is no help simply to *stipulate* the cross-world subjective identities, in the way that Kripke suggests, in *Naming and Necessity*, we may do generally with respect to "cross-world" identities: if there is a genuine

question about whether something is an object of acquaintance for a subject across possible worlds, a question to which one should be able to give a truthful answer, then stipulation is not to the point (Lewis 1983, 19).¹¹

The problem is general, and does not depend on the particular details of Lewis's counterpart theory.¹² It appears to affect *any* attempt to describe the structure of perceptual (or other conscious) individuation from a perspective of description that does not involve the privileged "for-me" perspective discussed above: for any such attempt, the problem of how to cross-identify the subject will arise; and as Lewis notes, from any such perspective, the problem appears insoluble. Lewis's own resolution, from his own favored perspective of purely third-person description, is to relativize the acquaintance relationship between subjects and objects again to worlds: we should not ask (he holds) whether X and Y are counterparts for the subject Z, but only whether the acquaintance relationship between U and X (in W₂) is a suitable counterpart of the acquaintance relationship between Z and Y (in W₁).

On the current view, however, this is no solution, since with respect to the actual phenomenon of individuation by acquaintance, it is not an explanation but rather a denial of the existence of the phenomenon itself. For its consequence is that it is not really possible, after all, to individuate by acquaintance in Hintikka's sense: whereas each of the subjects, U and Z, enjoys a relationship of acquaintance to some particular object (respectively, X and Y) in its own world, there is no evident sense in which either subject's acquaintance (thus understood) itself includes the relationships to objects in *other* worlds which generally comprise the individuating functions as such. On the other hand, if individuation by acquaintance really *is* possible, it is possible for me genuinely to *identify* a "perceptual object" (in the sense described above) by means of my current perception, and this includes identifying *which* object it is in each of my perceptually possible worlds. For me to do this, it is not necessary for me to first identify *myself* – by means of description, self-acquaintance, or in any

11 Lewis also notes correctly (20–21) that the suggestion of using "haecceities" (in the sense of (Kaplan 1975)) to carry out the (apparently) requisite individuation of the subject will fail, for the same reason that individuation by description does: any amount of substantive information, going beyond just qualitative similarities, about the identification of the subject across possible worlds is consistent with the failure of the subject herself to believe that information, and hence with the possibility that she cannot be located at one or more of her belief worlds.

12 Though I will not argue for this here, similar issues appear to affect the attempt to use "centers" to identify a subject across possible worlds: if a subject's epistemic possibilities are to be identified with the *centered* worlds centered on her, how are we to identify in general terms which these are without presupposing her trans-world identity?

other way – in each of these worlds; all that I need to do is identify the relevant object, from a perspective (my own) that I presuppose. It is only if I abstract from this perspective that Lewis's question of the individuation of the subject arises; but then, as he argues, that question is unanswerable. If this is right in general, it then remains possible to preserve the phenomenon of individuation by acquaintance – and with it, the distinctive presentational content of consciousness itself – only by preserving as irreducible the “for-me” or first-person perspective from which this phenomenon is itself given.

If, then, one considers the *ontological* question of consciousness as the question of “what there is” from such an objective perspective, it is clear that the availability of the consciously intentional modes of presentation introduces a decisive kind of internal complication into this question. While, on the one hand, there is no reason to invoke peculiar objects (of the sense-data or ‘proposition’ type), the very possibility of conscious presentation “from a point of view” itself means that a global ontology that does not allow for this possibility is in a certain way incomplete. This incompleteness plausibly corresponds to the exclusion of the phenomenon of conscious intentionality itself; and if this is correct, then only an analysis that accounts for the relevant modalities of presentation itself will allow for a realist ontology of consciousness. The possibility of such an ontology will turn on that of the availability of facts characterizing the “modes of presentation” by which entities are presented to us, in perception and other conscious modalities, not only as facts or states of affairs *within* particular worlds, but also (and crucially) as facts characterizing cross-world identities and comparisons of the sort to which Hintikka's analysis points.

3

As we have seen, if we wish to characterize the presentational semantics of consciousness within a possible-worlds framework, we must do so in a way that preserves the perspectival aspects of actually presented contents as they are actually consciously presented. For this, it is not sufficient simply to characterize contents or senses in terms of “metaphysical” possibilities, whereby cross-world identities are typically presupposed, stipulated, or grounded on the “essential” properties of the (actual-world) objects; we must also use the possible-worlds framework in such a way as to facilitate the assessment of the identities determined *by* what is presented in consciousness, even when these cross-cut (as they often do) “metaphysical” identities (such as those determined by essential or other “modal” properties of actual-world objects). Nevertheless, questions of global *ontology* clearly turn on “metaphysical”

identities in that sense; so if we want to explore the ontological implications of conscious presentation, we need a framework that can accommodate both kinds of possibilities: both the usual “metaphysical” ones, and also, crucially, the “epistemic” (or more broadly “presentational”) ones.

The most prominent recent suggestion for doing so is the apparatus of “two-dimensional” semantics which has recently been developed by David Chalmers and others.¹³ This semantics associates with each concept both a *primary* and a *secondary* intension, differing according to how they are evaluated with respect to each possible world. For the secondary intension (the familiar one of classical possible-world semantics), the *actual* reference of the concept is held fixed, and what is evaluated in each case is *this referent* in each possible world, considered as counterfactual. For the primary intension, however, what is held fixed is just the initial profile or presentation associated with a term, and the other worlds are evaluated *as if* actual, in order to assess *what* referent – if any – they would have *if* that world were (i.e. were to ‘turn out to be’) actual. This plausibly allows for a treatment of *epistemic* possibility in terms of primary intensions: for instance, whereas it is metaphysically necessary (secondary intensions) that water is H₂O, it is epistemically possible (primary intensions) that water is (turns out to be) something else, for instance XYZ. Matters of epistemic possibility and necessity in this sense are, plausibly, typically *a priori*, whereas matters of metaphysical possibility and necessity seldom or never are.¹⁴ Further, it is plausible that some concepts (for instance, “ghost”) have determinate and well-defined primary intensions, even though the extensions of their secondary intensions are empty.

The idea of primary intensions itself descends in part from earlier treatments of the functioning of indexical and demonstrative terms in possible worlds frameworks, for instance those of Kaplan (1979) and (1989). According to accounts of this sort, the “character” or unitary meaning of indexical terms which vary their reference across different contexts of use is to be understood as a rule or function *from* contexts *to* something like a (conventional) intension or content. Thus, for instance “here” specifies a rule connecting contexts of use to particular (conventional) intensions which then determine, as their referents, particular locations. Similarly, “I” specifies a rule which connects each speaker’s use of it to a determinate intension picking out that particular individual in every world where it exists. The important point, in both cases,

13 For Chalmers’ development of two-dimensionalism, see (Chalmers 1996; 2002b; 2002c; and 2006).

14 Mathematical judgments and statements might reasonably be thought to be a counterexample to this.

is just that the primary dimension of the determination of meaning operates similarly to the determination of primary intensions in the two-dimensional framework: taking the actual-world “presentational” meaning of the term as fixed, scenarios (or contexts or worlds) are evaluated *as if* actual, to determine what referent the term *would* have in that case or context of use.

As Chalmers points out in his discussion of primary intensions in *The Conscious Mind*, the demand for the *explanation* of a phenomenon typically concerns the *primary* rather than the *secondary* intension (Chalmers 1996, 57). For example, if one asks for an explanation of water, before it is known that water is H_2O , what is wanted is an explanation of (roughly) the clear, watery stuff in the environment, rather than an explanation *of* H_2O . More generally, the demand for scientific explanation is very often the demand to “explain the phenomenon.” that is, to account for what is initially or pre-theoretically presented (very often, though probably not invariably, on a perceptual basis). As specifying *how* the reference of a concept depends on the way that the world turns out, primary intensions are also intimately and essentially related to the initial presentation of these concepts, in a way that secondary intensions are not. In this respect, intuitively, primary intensions are plausibly much more like (what we might think of as) “modes of presentation” than secondary intensions are. This is because of the way their evaluation across possible worlds depends, in each case, on their initial presentational profile, which is what is then evaluated in each world considered as actual. By contrast, secondary intensions are, as we have seen, evaluated at each world (taken as counterfactual) only once the reference is already fixed.

For this reason, the evaluation of primary intensions depends much more heavily on features of our initial grasp and use of a concept: here, by contrast with the case of secondary intensions, there is no substitute for repeated appeals to our own intuitions about *what we would say* in a variety of hypothetical cases. This already suggests good reason to think that the pattern of judgments embodied by primary intensions, including the sets of entities they identify across possible worlds, will often resist easy reduction to any simple or discrete rule. For there is no evident reason to think that the practical competence or intuitive judgment involved, in each case of the evaluation of a primary intension, in determining “what we should say” in that case, or how to “extend” our concept to the new case, must be formulable in terms of a discrete rule which is able to determine that extension completely and uniquely.¹⁵ Rather, it seems plausible that here, it is irreducibly non-extensional features of the concept

15 The issues here are evidently close to those involved in Wittgenstein's “rule-following” considerations.

itself, in relation to the variety of contexts, that will determine the extension in each case.

But a stronger and more general argument for the conclusion of the irreducible non-extensionality of primary intensions can be made by considering how the *presentational aspects* of primary intensions themselves affect how they are used and evaluated. As Chalmers notes, it is plausible that primary intensions typically have an *indexical* component: for example, the primary intension of “water” can be roughly specified as “the dominant clear, drinkable liquid in *our environment*” or, as we might also put it, “*this* dominant, clear drinkable liquid” (Chalmers 1996, 61). This essentially indexical aspect reflects the way in which primary intensions are initially given to us: not as abstract conceptual rules or functions, but rather as ways of actually (initially) presenting phenomena that are actually around us. The indexical aspect is then also reflected in how the primary intensions are evaluated *across* contexts: in each case, we ask ourselves how *we* would evaluate the reference of “water,” using the concept as we in fact (already) do, *if* that case turned out to be actual. This appeal to our own way of using the concept is irreducible in evaluating primary intensions at possible worlds or contexts.¹⁶ Clearly, for instance, there are all kinds of facts about the use of “water” (or a similar-sounding word) by others in other possible worlds that do not bear on the evaluation of the reference of (our) “water”; what matters is just what the reference of *our* term and concept, on the basis of the initial presentation, would be. Given this, it seems apparent that no amount of objective information about the use of (a word with the same sound as) “water” by speakers across possible worlds will suffice to determine the primary intension of “water” (as *we* use it); what is needed, instead, is an essential appeal to *our own* presentation of the stuff around us.¹⁷ Similarly, it appears that any wholly non-indexical presentation of the primary intension as a (wholly objective) function from worlds to referents will leave open the

16 I leave aside here the device of using “centered” possible worlds to account for indexical reference and indexical features of primary intensions. For reasons essentially connected to those considered here, I do not think that this device can succeed in reducing or eliminating the intensionality involved in indexical presentation, but I will not argue for this specific conclusion here.

17 This need not imply that all of the members of a community must initially have the same sensory or perceptual presentation in order to use the concept in the same way; only that their own individual initial presentations agree, at least roughly, in what they pick out across epistemic space. It is also compatible with this that the (full) primary intension associated with a given perceptual or initial presentation is in general only determinable on ideal rational reflection on the content presented (and so may not be fully initially accessible to the agent, prior to reflection).

question whether it captures the way *we* would use the concept, given the way it is initially presented to *us*.

Further, it is clear that both the perceptual individuating functions discussed by Hintikka and other functions characteristic of presentational contents of consciousness generally, will plausibly exhibit this kind of irreducibility. As we have seen (and as Hintikka himself emphasizes), given that the individuating functions will in many cases cross-identify individuals, across possible worlds, with respect to their “metaphysical” identities, there is no reason to suppose that the “perceptual objects” they define will be cleanly reducible to properties and features (including modal ones) of actual-world objects. Indeed, this typical cross-identification is, as we have seen, apparently at the actual basis of the familiar failure of “extensionality” (in the sense of the intersubstitutivity of identicals) in conscious presentational contexts generally. In this way, as we saw above, the familiar failure of intersubstitutivity in these contexts may apparently be seen as having its actual “metaphysical” basis in the presentational character of consciousness itself: in, that is, the capacity of consciousness to present (in perception or another conscious modality) “individuals” which, across possible worlds, do not correspond to, and cannot be reduced to, the characterization and properties of extant actual-world individuals.

This irreducibility of conscious presentation to “extensional” description may be seen as having important consequences for arguments about explanation, supervenience, and reduction. Familiarly, for objects which have their modes of presentation contingently, the (apparent) conceivability of cases in which those modes of presentation come apart from one another does not imply that there is actually a failure of metaphysical identity or supervenience in this case, but just that the case has been incorrectly described. For instance, the initial apparent conceivability of the case in which water – in the primary-intensional sense of “liquid stuff in our environment” – fails to be H_2O does not establish that water is not identical with H_2O , but only that this case is misdescribed: it is really one in which something else fills the role of the “liquid stuff in our environment”. Similarly, the right explanation for the apparent conceivability of the case is not that water might not have been H_2O , but only that something else could have turned out, on inquiry, to fulfill this initially presenting role. However, if there are any phenomena which have their mode of presentation *necessarily* – that is, they are presented, whenever they are, by means of *just those* modes of presentation – then it is plausible that the conceivability of these sorts of judgments *do* support actual (metaphysical) non-identity or nonsupervenience. The reason for this is that here there is no room to “explain away” the judgment of non-identity by reference to contingently differing modes of presentation, or to something else being (contingently)

presented by means of the same mode (or as another possible occupant of the same initially presenting role). As we shall see in the next section, this is plausibly the case with respect to the phenomena of consciousness: that is, that they are exhausted by their own modes of presentation, which they possess necessarily in a strong sense. If this is correct, it may be used as the key premise in an argument from conceivability to (actual metaphysical) possibility which has as a consequence the non-supervenience of consciousness on the totality of physical facts.

However, if the reasoning in this section is correct, the actual ground for this is not to be found in the metaphysically special character of consciousness itself but rather in its irreducibly *presentational* character: in the (meta-) logical structure, in other words, that qualifies conscious states, just as such, to operate as “their own” modes of presentation and thus to be defined and exhausted by this presentational character. It follows, on the one hand, that the argument for the nonsupervenience of consciousness can be cast in a much broader form than is typical: not only in terms of the irreducibility of consciousness to the physical, but in terms of the irreducibility of consciousness to *any* set of facts about objects or phenomena in the world that can be written down in natural-language sentences and preserve their truth-value when so written. On the other, this irreducibility is then not to be accorded in the first instance to any supposed difference in metaphysical composition or substance between the physical and the phenomenal, or between matter and mind, but rather to the “transcendental” difference between the facts of the world, on the one hand, and the perspective *from which* it is possible to present these facts in general.

4

In a number of places, Chalmers has formulated an influential general argument against materialism, or the claim that the totality of physical facts necessitates the facts about phenomenal consciousness.¹⁸ The argument exploits the framework of two-dimensional modal semantics, arguing from the *conceivability* of the combination of P with $\sim Q$ (where P is the conjunction of microphysical truths about the universe and Q is an arbitrary truth about phenomenal consciousness) to its real (metaphysical) *possibility*. In particular, the most crucial (and controversial) step of the argument is the move from the conceivability – or possibility in the sense of *primary* intensions

18 See especially (Chalmers 2002a) and (Chalmers 2010).

(‘1-possibility’) – of $P \ \& \ \sim Q$ to its real *metaphysical* possibility (or possibility in the sense of *secondary* intensions). If this move is successful, then the conceivability or ‘epistemic possibility’ of $P \ \& \ \sim Q$ is sufficient to establish its real possibility, and thus to establish that the facts about consciousness are not necessitated by the physical facts.

The first premise of the argument asserts the *conceivability* of $P \ \& \ \sim Q$. Familiarly, there are a variety of ways to support or illustrate this premise. One is to imagine a *zombie world*: one in which the physical facts are just as they are in our world, but no one is phenomenally conscious. Another is Jackson’s famous thought experiment of Mary, the neuroscientist who is confined to a black and white room and given access to the totality of information relevant to color perception and cognition (as we may imagine, including the totality of objective and non-indexical statements of a completed neuroscience), but still apparently does not learn *what it is like* to see red until after she leaves her room.¹⁹ Moreover, it is apparent that no amount of “physical” information – information, that is, which she could read in encyclopedias, or learn by viewing a (black-and-white) television monitor – will help her to know the phenomenal character of red, before she leaves the room. If this is correct, then the thought-experiment establishes at least that there is an “epistemic” gap between the totality of physical facts and phenomenal characters.

Before proceeding further, it is helpful from the current perspective, though, to consider carefully how the information that Mary has available to her in her room is specified. In Jackson’s original discussion, the information available to Mary is described as the totality of “physical” information. But as several commentators (Lewis, Churchland, Perry), have noted, this description is actually not essential, since we can imagine Mary being supplied with *any* information that can be written down without any essential use of indexical terms or tenses, and understood by her (given that she has never actually experienced any colors except for black, white, and shades of grey).²⁰ It is plausible, in particular, that she could be given any information that can be written down in the form of (what Carnap calls in the *Aufbau*) *structural definite descriptions*: that is, statements involving terms implicitly defined by means of their structural,

19 For the original development, see (Jackson 1982).

20 This may be seen as raising problems for the claim of the thought experiment to show anything uniquely bearing against *physicalism* (as opposed to, say, dualism): for if, for example, interactionist dualism were true, and the facts about physical-mental and mental-physical causation could themselves be written down in indexical-free and tense-neutral terms, then she could be supplied with these facts as well, and yet they still would not apparently give her any help in knowing the phenomenal character of red.

functional, causal or other systematic relations to each other, and without any essential use of indexicals or tenses.²¹ Statements of this form will capture the totality of information about structure and function, and there is no good reason to think that she cannot know them fully (and hence have access to the totality of such information) prior to leaving the room. What they will not do, though, is capture any aspects of the phenomena they characterize that underlie or go beyond the description of their structure and function. These phenomena plausibly include not only the “intrinsic” nature of the elements that are themselves structurally and functionally related, but *also* (as I shall argue) aspects of how they are *essentially presented* (if there are any such).

This last consideration is relevant both to the success and to the further implications of the crucial second premise of Chalmers’ argument. This is the premise that moves from the conceivability (or 1-possibility) to the real metaphysical possibility (or 2-possibility) of $P \ \& \ \sim Q$. As Chalmers notes, this move might be resisted in many different ways, but a chief source of resistance will come from those who hold that phenomenal facts are identical with or necessarily connected to physical facts, but this metaphysical identity or necessity is *a posteriori* (Chalmers 2002a, 255–27). For proponents of this strategy (Chalmers calls them “type-B” materialists), the obtaining of the physical (including all structural and functional) facts without the facts about phenomenal consciousness is indeed apparently conceivable – analogously to the apparent initial conceivability of water being something other than H_2O – but this combination is in fact not *metaphysically* possible. Even though, then, the initial conceivability (or 1-possibility) of the combination of $P \ \& \ \sim Q$ is admitted, the proponent of this view denies its real metaphysical possibility (2-possibility).

To make the premise go through against the objections of the type-B materialist, it is thus necessary to argue that any world that *verifies* $P \ \& \ \sim Q$ (i.e. is a world in which $P \ \& \ \sim Q$ holds according to their primary intensions) also *satisfies* $P \ \& \ \sim Q$ (i.e. is a world in which it holds according to their secondary intensions). (Chalmers 2002a, 256–57; 265–67; Chalmers 2010, 148–50). As Chalmers argues, in the case of concepts of phenomenal experience, it is plausible that their peculiar presentational features help make this inference go through,

21 More rigorously, we can imagine the totality of information available to her being put in the form of a single Ramsey sentence produced by means of repeated structuralization, with any remaining O-terms being ones whose meaning she already has cognitive access to. We can then argue that even the provision of this Ramsey sentence will not put her in a position to know what it is like to see red. For some related considerations, see (Livingston 2004), Chapters 2 and 5.

where it would not necessarily with respect to other concepts. In particular, in the case of the concepts of conscious phenomena, primary and secondary intensions appear to coincide. This captures the sense in which, for conscious phenomena, *esse* is *percipi*: that is, it is essential to their identity as the phenomena that they are that they are presented, and indeed that they are presented *as* they (in fact) “are.” In other terms, this may be put as the recognition that conscious phenomena are inherently “self-presentational”, or that they have their modes of presentation essentially or necessarily (i.e., in such a way as to be necessary to their identity). It plausibly follows from this that, in the case of conscious phenomena (as opposed to those which have their modes of presentation contingently) that primary and secondary intensions coincide. As Chalmers puts it (glossing a point already made by Kripke): “in the case of consciousness, there is no distinction analogous to that between water itself and mere watery stuff” (Chalmers 2002a, 256). That is, with respect to phenomenal states, there is no distinction between the presentations in terms of which they are initially identified (primary intensions) and what they are (secondary intensions). If this is right, it follows that any world which verifies “there is consciousness” is also one which satisfies “there is consciousness,” since such a world is one in which there is something that (at least) feels conscious, and this is sufficient for its being conscious. Thus, it is specifically the inherent presentational features of phenomenal states that qualifies them to underwrite the move from conceivability to actual possibility, or from epistemic issues to metaphysical ones, with respect to them.

But what about the parallel move with respect to P (the totality of physical facts)? Here, Chalmers considers the possibility that the type-B materialist might try to block the move from conceivability to possibility by arguing that there could be a world which *verifies* P but which does not satisfy P (Chalmers 2010, 150). This would be, in other words, a world W where things *look as they do* with respect to the actual physical facts (on their primary intensions) but in which these actual physical facts nevertheless do not hold. As Chalmers notes, however, if this is the case (i.e. if a world *verifies* P), it must have at least the *structure* of the real physical world: all the relevant facts will bear the same structural relationships that they do in our world, and the primary intensions of the physical concepts will pick out whatever properties play these roles in a *given* world, while the secondary intensions will pick out, across all worlds, the actual intrinsic properties underlying these structural relations. Given this structural identity, however, it is still possible, as Chalmers argues, for a type-B materialist to hold that W does not contain consciousness: she can do so by holding that, whereas W is identical to the actual world *structurally*, it differs with respect to the *intrinsic* characters or natures of what is structurally

related. If this difference is sufficient for W to lack consciousness, it follows that consciousness (as it is in the actual world) is not necessitated by the structural facts, but rather by the underlying intrinsic properties. These might be thought of, Chalmers suggests, as the *intrinsic categorial bases* for the structural relations of physics, and they would be, on this view, themselves responsible for the qualitative characters of consciousness. This sort of position – the position that consciousness is founded upon the underlying intrinsic properties of matter rather than their structural relationships – is the position that Chalmers terms “type-F” (or “Russellian”) monism (Chalmers 2002a, 255–57; Chalmers 2010, 151–52).

Since, as Chalmers agrees, this possibility cannot be ruled out on the two-dimensional argument, it is necessary to modify the conclusion of the argument slightly: it does not fully establish the conclusion that materialism is false (and dualism or some other non-materialist position is true), but rather only a disjunctive claim: materialism is false, or type-F monism is true (Chalmers 2002a, 256; Chalmers 2010, 152). As Chalmers points out, it is not in fact obvious whether or not the type-F monist position should be considered to be a physicalist one. If one holds that physical terms, as we already use them, already refer to the underlying intrinsic bases, then the type-F monist position has a good claim to be considered a physicalist rather than a dualist one (Chalmers 2002a, 265). On the other hand, it essentially introduces at least a conceptual and explanatory dualism between structural-dispositional properties, on the one hand, and their hidden categorial bases, on the other.

More generally, the considerations that come to bear in this argument are just a local version of the more general ones captured in the broader argument for an explanatory gap between the physical – if this is understood in terms of the totality of structural/dispositional facts – and the phenomenal. These considerations turn centrally on the relationship of the totality of structural and functional relations and explanations to what is outside or beyond them, or to what might be considered to vary while all structural and functional relations are held fixed.²² As Chalmers’ argument effectively points out, *one* such aspect of potential variation is indeed in the *intrinsic* or categorial bases of structural facts and relations. Indeed, the core of the argument for the possibility of type-F monism is the consideration that the specification of the *totality* of structural facts does not suffice to specify or determine the “intrinsic” natures of what is thereby structurally related. However, without gainsaying this consideration, we can reasonably ask whether this is in fact (as Chalmers suggests it is) the *only* relevant possible dimension of variation beyond what is specified by a

22 See (Livingston 2004), especially Chapters 1 and 6.

specification of total structure.²³ And indeed, in light of considerations we have already explored, it appears likely that there is another dimension of potential variation that is even more directly relevant to the question of the limits of structural explanation. This is variation in the total *presentational* aspects of the structure, or the position *from which* the total structure can itself be presented.

What, then, if we saw the root of both the conceivability of a structurally described world without consciousness, and the actual existence of consciousness in *our* world, as turning not on the presence or absence of “intrinsic” properties of a mysterious sort, but rather on the question of the presence or absence of conscious presentation itself? On the one hand, it is evidently conceivable for there to be a world in which all physical facts – in the sense of structural and functional facts – are as they actually are, but there is no conscious presentation: no actual conscious availability or accessibility of anything as anything. There is nothing evidently contradictory about such a possibility, since the specification of any physical, structural or functional fact within a world is compatible with the assumption that it is not in any sense consciously presented by anything or anyone *in* that world.²⁴ On the other hand, on this sort of view, the dimension of variation embodied by the presence or absence of consciousness would not primarily or exclusively characterize the “intrinsic” properties of (actual-world) physical entities, but rather the total variation between worlds in which there is, and worlds in which there is not, conscious presentation. Since, as we have seen, the individuation of entities by means of conscious presentation both cannot be reduced to the (actual or modal) properties of individual real-world entities and, in the most characteristic cases, cross-cuts the individuation of entities in terms of their “metaphysical” profiles, there is no necessity here for the relevant presentational “properties” to be identified with the intrinsic or categorial properties of just those entities. Since it is determined not by the individual properties of (actual-world or metaphysically defined) entities but rather by the individuating functions from worlds to their inhabitants, the relevant dimension of variation operates,

23 There may also be other relevant reasons to doubt the plausibility of type-F monism as an answer to the hard problem: for instance, if phenomenal properties do have an (intelligible and general) basis in *specifiable* intrinsic natures, then the information *about* these intrinsic natures could be provided to Mary, even prior to her release from her room, but it is unclear how this would help her in knowing what it is like to see red.

24 This just corresponds to the zombie world; of course, we must here suppose that zombies lack not only phenomenal consciousness but also many aspects of presentational intentionality that plausibly go along with it, such as belief and knowledge.

so to speak, on the level of *worlds as wholes* rather than simply on the level of these individual entities themselves. That is, in order to determine the presence or absence of the presentational properties in a world, we cannot simply look at the “intrinsic” properties of that world’s entities and facts – indeed, it must be insufficient to do so – but instead we have to look at whether and how these entities and facts are presented to inhabitants of that world. To do this, as we have seen, we will typically have to consider how this presentation individuates entities in ways that cross-classify individuals, across possible worlds, with respect to their “metaphysical” identities. But it is just here that the close connection we have seen between primary intensions – in the sense of what is primarily presented in consciousness – and the metaphysics of phenomenal consciousness itself, as plausibly constituted by just that presentation, comes directly to bear. Whereas this bearing is obscure on Type-F monism itself – which requires that we posit otherwise unknown categorial bases for the familiar structural and relational properties of physical matter, and then stipulate their identity (in our world) with phenomenal properties – on the kind of position recommended here, it is immediately and readily accounted for on simultaneously logical, metaphysical, and phenomenological grounds.

We can illustrate this position more vividly by means of a line of argument that Chalmers develops in “The Two-Dimensional Argument Against Physicalism” (Chalmers 2010, 153–54). As he argues there, it is not in fact necessary, in order to establish the conclusion of the general 2-dimensional argument, to assume that primary and secondary intensions must coincide with respect to Q. It is sufficient to add a “that’s all” clause or fact to P: a fact saying that P includes *all* the facts true at the world under consideration. Then we can readily argue (from the conceivability of $P \ \& \ \sim Q$) that there is a *minimal* world – a world which includes all the positive facts in P, and nothing else – that (at least) verifies P but in which the primary intension of Q is false. Then, if P’s primary and secondary intensions coincide, there is a minimal P-world (a minimal world *satisfying* P) in which Q is false, and thus physicalism is false about *our* world (where Q is true). As Chalmers says, on this alternative, it is possible that P necessitates Q, so that physicalism is actually true of the *facts* in Q, but given the existence of a minimal P world which fails to verify Q, physicalism is then (at any rate) false about the *modes of presentation* of these facts. Alternatively, if the primary and secondary intensions of P fail to coincide (as on the Type-F monist position), then the existence of a minimal world which verifies P but not Q leaves open the possibility that P necessitates Q, but this necessitation depends on both the structural *and* non-structural (Chalmers says ‘intrinsic’) profiles of P. In this case, as Chalmers says, we have a variety of monism in which the non-structural aspects of physical facts “are crucial for constituting

the properties associated with the modes of presentation of consciousness.” Chalmers characterizes these non-structural aspects as the “intrinsic properties” of physics; but with these considerations in place, it is clear that the reference to “intrinsic properties” is largely vestigial. What is essential to the case at hand – in which we are essentially considering the epistemic and metaphysical implications of the claim that the physical facts are *all* the facts (this is the result of adding the “that’s all” claim to P) – is rather just that we cannot thereby make room for some modes of presentation, in particular those that actually present phenomenal facts (in our world). This conclusion is already accessible, as soon as we consider the implications of adding the claim of totality with respect to the physical facts, and thereby considering the question of the position from which this totality can be presented, and it suffices to establish the disjunctive conclusion that dualism or *some* variety of non-structural monism is true. If we do take the monist alternative, however, it now becomes particularly clear that what is left out of the physical facts as structurally described is essentially related to – or perhaps identical to – the *presentational* aspects of consciousness themselves.

Despite the essential appeal to presentational aspects which are not inherently aspects of any “physical” or “material” object, this kind of position remains a monism rather than a dualism. Indeed, adopting it provides an important additional kind of motivation for monism, and thus for interpreting the general two-dimensional argument as supporting a monist position rather than any form of dualism. For as we have seen in the course of the discussion of Hintikka’s “individuating” functions, the temptation to assume that these functions introduce ontologically peculiar kinds of entities – for example sense-data, or indeed any kind of non-physical or non-material object – is readily countered by observing that the actual referents of the functions, across possible worlds, are just familiar entities of an ontologically single type. Thus the temptation to “reify” senses or other special intentional objects is shown to be simply an artifact of the way in which presentational individuation cross-cuts otherwise identified objects across possible worlds. If correctly analyzed, the phenomena of presentation thus suggest no reason, even fully granting the soundness of Chalmers’ two-dimensional argument, to adopt a dualism of substances or entities. It is true that the argument, as it stands, leaves open either dualism or monism; but given the availability of the monist alternative here suggested and the plausibility of the claim that it suffices to account for the presentational properties of consciousness, dualism now has no evident motivation. Something similar apparently holds, as well, with respect to property dualism. Just as there is no need to introduce an ontologically distinct kind of object if the presentational “properties” and “entities” work as suggested, there is also

no need to introduce any ontologically exotic types of properties of ordinary entities. The entities referred to in each of the worlds across which the individuating functions are defined, after all, just are the familiar ones, with their ordinary types of properties and relations. This importantly makes it evident that upholding the monist disjunct of Chalmers' disjunctive conclusion need not in any sense involve a property dualism, or anything resembling such a position on the level of global ontology.

Independently of this, there are other reasons to prefer the kind of presentational monism I am presenting here over type-F or "intrinsic" monism as an answer to the problem posed by absence of consciousness from a total structural and functional description of the world. One is the question that inevitably arises, if type-F monism is adopted, about the "intrinsic" natures themselves. If these "intrinsic" natures do indeed have determinate characteristics, such that their presence or absence could determine the presence or absence of consciousness from the world, then why should these characteristics themselves not be describable within a general objective description of the world? But if they were so describable, then Mary could apparently be given full knowledge of them, even before her release from the room, and we would again face the problem that this provision would apparently not give her knowledge of phenomenal properties. Conversely, moreover, type-F monism arguably does not help to account for the knowledge she *does* gain when she leaves the room: why should her visual perception present to her the (hitherto unknown) intrinsic *categorical bases* of the relevant structural properties, when presenting those structural properties in other ways does not? In both cases, the relevant phenomena are better accounted for by reference to the presence, or absence, of actual presentational properties, or of the actual occurrence or non-occurrence of an actual presentation of the relevant phenomenon. Finally, and for related reasons, type-F monism faces a difficult "combination" problem: how do the "proto-phenomenal" intrinsic properties that (on the view) actually underlie the physical structures "add up" to experienced phenomena? By contrast, the account in terms of presentational aspects does not involve any problem of composition or combination, since the possibility of a (phenomenal) presentation of a phenomenon is already seen as involved in the essential structure of a presentational perspective as such.

For all of these reasons, it appears that attention to the presentational aspects of consciousness motivates the novel kind of monist ontology I have argued for. On this ontology, though there is (as Hintikka emphasizes) only one ontological type of entities, there would nevertheless be a crucial *irreducibility* of consciousness to the purely extensional description of entities; and this explanatory irreducibility would be seen as producing an actual ontological

irreducibility of consciousness, in the sense of presentation, to “physical” (or, indeed, other purely extensionally described) facts. This irreducibility of consciousness would then be seen as an aspect, or ontological reflection, of its inherently first-personal or perspectival character, such that it picks out its referents, as classical phenomenology emphasizes, always *from* a specific position or point of view. Despite this essential invocation of a perspective or point of view, however, the recommended position is not an idealism; for the claim is not that the mind or subjectivity constitutes or produces the (actual or possible-world) phenomena it refers to, or even their sense. With respect to these entities and, indeed, “subjectivity” itself, it involves only the same ontological commitments to which standard possible-world semantics already makes recourse.

In closing, however, it must be admitted that given only these ontological commitments, it is not immediately obvious *why* consciousness should be irreducible in this sense. I have argued that there is reason to think this irreducibility can be connected to the irreducibility of the perspective from which primary intensions, or presentational individuating functions, are necessarily deployed across possible worlds. But we have not really seen why these functions, understood as such, must be essentially and irreducibly “non-extensional”: what, that is, that essentially *prevents* them from being cashed out as “functions in extension,” or in other words as (purely extensional) sets of ordered pairs of worlds and entities? After all, they are just functions: why could not any one, or all, of these functions just be given by means of finitely stated rules that are themselves accessible in principle from any point of view?

Though I will not develop these arguments here, however, I do think there are two broad ways in which one can argue for this irreducibility on principled grounds connected to what is plausibly the structure of these functions themselves. The first way would be to argue that because the presentational phenomena are, just as such, “semantic” in the sense first used by Tarski to characterize truth, the functions that characterize them exhibit an essential “meta-logical” irreducibility to (first-order) “syntactic” structures or systems.²⁵ On this sort of position, just as Tarski demonstrated that truth must be irreducible to the syntax of an extensional language, so, and for essentially similar reasons, the presentational phenomena might actually be seen as irreducible to the extensional description of facts. Monism on the level of these facts themselves could, however, naturally be preserved; and the metalogical implications of “diagonalization” (in the sense in which Tarski’s theorem applies it)

25 For a compelling recent argument for treating phenomenological intentionality as “semantic” in just this sense, see (Smith 2015).

would themselves suffice to guarantee the real irreducibility of consciousness as presentation. The analogy considered here – between the irreducibility of the mental to the physical, on the one hand, and the irreducibility of semantics to syntax, on the other – is actually offered by Davidson in his original defense of anomalous monism, in “Mental Events.”²⁶ But rather than applying it, as Davidson does, to considerations about law and causation, the present considerations appear to suggest its use to establish the actually ontological conclusion of the irreducibility of presentation to the totality of what is presented, while monism is nevertheless preserved.

The second way might be to appeal to broadly “Kripkensteinian” considerations about the *application* of the “content” of a presentation across cases, including (as we have seen) the variety of possible worlds, considered as actual. If, as Kripke interprets Wittgenstein as arguing, any attempt to capture this application by means of a finitely stated rule leaves open the skeptical possibility of a (purportedly) “non-standard” application in a new case, then the actual pattern of application that is embodied in this content cannot in general be reduced to such a finite statement (Kripke 1982).²⁷ This is perhaps why Wittgenstein says that, although any provision of a rule appears to demand another rule for interpreting that one, there is nevertheless a way of “grasping a rule” which is “not an interpretation” but rather turns on “what we” call following or going against the rule as we proceed from case to case (Wittgenstein 1953, Section 201). As I have suggested in connection with primary intensions and individuating functions, the collective first-personal “we” here may indeed be essential: it is not possible in general to account comprehensively for what is involved in a conscious presentation – that is, to account exhaustively for what it in fact determines, across possible worlds considered as actual – purely in third-person or indeed in simply extensional terms, and the irreducibility of presentational content as such to these terms would then once more be vindicated. It would be a further and welcome exegetical consequence of this that, far from repudiating or rejecting the idea of the essentiality of “inner” or consciously presented contents of thought, Wittgenstein’s considerations would rather be seen as pointing out, in a profound way, their real ontological character.

26 See (Davidson 1970), p. 119. I am indebted to John Bova for reminding me of this passage, and pointing out to me its suggestive implications for the ontology of consciousness, in the course of a joint re-reading of “Mental Events” during the summer of 2013.

27 (Kripke 1982).

References

- Carnap, Rudolf 1947. *Meaning and Necessity*. Chicago: University of Chicago Press.
- Chalmers, David 1996. *The Conscious Mind*. New York: Oxford University Press.
- Chalmers, David 2002a. "Consciousness and its place in nature." In: *Philosophy of Mind: Classical and Contemporary Readings*, D. Chalmers (ed.), Oxford: Oxford University Press, 247–72.
- Chalmers, David 2002b. "On sense and intension." *Philosophical Perspectives* 16, 135–82.
- Chalmers, David 2002c. "The components of content." In: *Philosophy of Mind: Classical and Contemporary Readings*, D. Chalmers (ed.), Oxford: Oxford University Press, 608–33.
- Chalmers, David 2006. "The foundations of two-dimensional semantics." In: *Two-Dimensional Semantics: Foundations and Applications*, M. Garcia-Carpintero and J. Macia (eds.), Oxford: Oxford University Press, 55–140.
- Chalmers, David 2010. The two-dimensional argument against materialism. In: *The Character of Consciousness*. Oxford: Oxford U. Press, 141–91.
- Davidson, David 1970. "Mental Events." Reprinted in *Philosophy of Mind: Classical and Contemporary Readings*, D. Chalmers (ed.), Oxford: Oxford University Press, 2002, 207–228.
- Føllesdal, Dagfinn 1969. "Husserl's notion of noema." *Journal of Philosophy* 66:20, 680–687.
- Heidegger, Martin 1927. *Being and Time*. Transl. by J. Macquarrie and E. Robinson. Oxford: Blackwell, 1962.
- Hintikka, Jaakko 1962. *Knowledge and Belief: An Introduction to the Logic of the Two Notions*. Ithaca, NY: Cornell U. Press.
- Hintikka, Jaakko 1967. "On the logic of perception." In: *Models for Modalities: Selected Essays*. Dordrecht, Holland: D. Reidel, 1969, 151–83.
- Hintikka, Jaakko 1969. "Semantics for propositional attitudes." In: *Models for Modalities: Selected Essays*. Dordrecht, Holland: D. Reidel, 87–111.
- Husserl, Edmund 1913. *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy: First Book*. F. Kersten (trans.) Dordrecht: Kluwer, 1983. Originally published in German as *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie, 1. Buch: Allgemeine Einführung in die reine Phänomenologie*. (Halle: Max Niemeyer, 1913).
- Jackson, Frank 1982. "Epiphenomenal Qualia." *Philosophical Quarterly* 32: 127–36.
- Kaplan, David 1968. "Quantifying In." *Synthese* 19:1/2, 178–214.
- Kaplan, David 1975. "How to Russell a Frege-Church." *Journal of Philosophy* 72:19, pp. 716–29.

- Kaplan, David 1979. "On the logic of demonstratives." In Peter A. French *et al.* (eds.), *Contemporary Perspectives in the Philosophy of Language*. Minneapolis: U. of Minnesota Press, pp. 401–12.
- Kaplan, David 1989. "Demonstratives." In: *Themes from Kaplan*, Oxford: Oxford University Press, pp. 481–563.
- Kripke, Saul 1982. *Wittgenstein on Rules and Private Language*. Cambridge, MA: Harvard University Press.
- Lewis, David 1983. "Individuation by acquaintance and by stipulation." *The Philosophical Review* 92:1, 3–32.
- Livingston, Paul M. 2004. *Philosophical History and the Problem of Consciousness*. New York: Cambridge U. Press.
- Quine, W.V. 1953. "Three grades of modal involvement." Reprinted in *The Ways of Paradox and Other Essays*. Cambridge, MA: Harvard U. Press, 1976, 158–76.
- Quine, W.V. 1956. "Quantifiers and Propositional Attitudes." *Journal of Philosophy* 53: 177–87.
- Quine, W.V. 1960. *Word and Object*. Cambridge, MA: MIT Press.
- Sartre, Jean-Paul 1943. *Being and Nothingness*. Transl. by H.E. Barnes. New York: Washington Square Press, 1956.
- Smith, David W. and Ronald McIntyre 1982. *Husserl and Intentionality: A Study of Mind, Meaning, and Language*. (Synthese Library, vol. 154). Dordrecht: D. Reidel Publishing Company.
- Smith, David W. 2015. "Truth and Epoché: The semantic conception of truth in phenomenology." In J. Bell, A. Cutrofello, and P.M. Livingston, ed., *Beyond the Analytic-Continental Divide: Pluralist Philosophy in the Twenty-First Century*. New York: Routledge, 111–127.
- Wittgenstein, Ludwig 1953. *Philosophical Investigations*. Rev. 4th ed., transl. by G.E.M. Anscombe, P.M.S. Hacker, and J. Schulte, trans. Malden, MA: Wiley-Blackwell, 2009.