

## THEMES IN MY PHILOSOPHICAL WORK

Terry HORGAN  
University of Memphis

This paper is an overview of my philosophical work. It follows closely the structure of the handout I used as the basis for a talk on this topic at the 2000 meeting of the Austro-Slovene Philosophical Association. The section-headings mention major themes, and various key concepts are indicated by boldface terms in the text.

### 1. *Supervenience, in Relation to Reduction and Materialism*

I invoked the notion of supervenience in my doctoral dissertation, *Microreduction and the Mind-Body Problem*, completed at the University of Michigan in 1974 under the direction of Jaegwon Kim. I had been struck by the appeal to supervenience in Hare (1952), a classic work in twentieth century metaethics that I studied at Michigan in a course on metaethics taught by William Frankena; and I also had been struck by the brief appeal to supervenience in Davidson (1970). Kim was already, in effect, construing the relation between physical and mental properties as a supervenience relation – although he was not yet using the word ‘supervenience’. I assumed that a materialistic metaphysics was correct, and that integral to materialism is the idea that higher-level sciences (including psychology) are reducible to lower-level ones – ultimately to microphysics. One idea I pressed in the dissertation was that biconditional “bridge laws” would not suffice for genuine intertheoretic reduction if these inter-level laws were additional fundamental laws of nature alongside those of the reducing science; they would be what Herbert Feigl and J.J.C. Smart, in their writings on the psychophysical identity theory, called “nomological danglers.” I argued that the higher-level property in a bridge law should bear a relation of **strict supervenience**.

**nience** to its correlated lower-level property, rather than merely being nomically correlated with it. The basic idea was that there are no two physically possible worlds  $w_1$  and  $w_2$  – where a physically possible world is, roughly, a world in which the laws of microphysics obtain and in which there are no nonphysical substances like entelechies or Cartesian souls – such that the actual-world bridge laws obtain in world  $w_1$  but not in world  $w_2$ . (Thus, the bridge laws themselves are fixed relative to the fundamental physical facts and fundamental laws, rather than being fundamental laws themselves alongside those of microphysics.) Already when writing the dissertation, however, I was persistently bothered by the apparent conceivability of possible worlds that are physically just like ours but different with respect to the way phenomenal conscious properties are instantiated – e.g., worlds in which the color-experiences of humans are systematically inverted in comparison to the actual world. I realized that my version of materialism was committed to denying that such worlds are a genuine metaphysical possibility, despite their apparent conceivability. Ideas from my dissertation are presented in such papers as Horgan (1976, 1978b, 1981c, 1987d).

In subsequent work I moved away from the idea that materialism requires that the properties posited in psychology and the other special sciences, and in ordinary discourse, need to be correlated with (or identical to) physics-level natural-kind properties via universally quantified biconditional bridge laws (or via type-type identity statements). I called my position a version of **nonreductive materialism**, since such bridge laws and/or property-identities were widely considered a prerequisite for intertheoretic reduction; cf. Horgan (1993c, 1994b). I continued to stress that inter-level relations between properties should not be metaphysically *sui generis* within a materialistic metaphysical view, i.e., should not be fundamental laws alongside those of microphysics. One way I made this point, in Horgan (1984d), was in terms of the idea of **cosmic hermeneutics**: a Laplacean intelligence should be able to deduce, from a complete microphysical history of the world plus an understanding of meanings of non-physics-level terms and concepts, all the facts about the world. In order to accommodate the fact that higher-level facts and properties can sometimes supervene on lower-level facts and properties in a way that depends on non-local goings-on, in Horgan

(1982b) I introduced the idea of **regional supervenience** (although I only gave it this name in Horgan 1993b). Later Mark Timmons and I explicitly argued (Horgan and Timmons 1992a) that naturalism requires that supervenience relations be explainable in a naturalistically acceptable way; we articulated a format for such explanations, and we argued that the supervenience of mental properties on physical properties could be explained via this format to the extent that mental properties are functional properties. I stressed the explainability requirement in Horgan (1993b), where I introduced the term **superdupervenience** for materialistically explainable supervenience.

I have had ongoing, increasing, worries about the extent to which higher-order properties instantiated in our world are, or are not, superdupervenient on physics-level properties – and also about what would count as a materialistically acceptable explanation of supervenience (and why). One source of worry, which as I said was already present when I was writing my dissertation, is the conceivability – and hence the presumptive metaphysical possibility – of worlds that are physically just like ours but different with respect to the distribution of phenomenal mental properties. (This is closely related to what Joseph Levine (1983) calls the “explanatory gap” and what David Chalmers (1995, 1996) calls the “hard problem” of phenomenal consciousness.) Another source of worry, stressed for example in Horgan and Timmons (1993) and in Horgan (1994a), is that intentional mental properties seem to involve semantic normativity, a feature that threatens to prevent their supervenience on the physical from being explainable in a materialistically acceptable way. In Horgan (1994a) I argued that in light of this problem, a position worthy of serious consideration is what I called **preservative irrealism** about mental intentionality. Yet a third source of worry is my current belief, stressed in Horgan and Tienson (in press), that occurrent intentional mental states like thoughts and desires have a constitutive phenomenology – which means that the first worry applies to these mental states too.

## 2. *Conceptual Austerity and Contextually Variable Parameters of Concepts and Discourse*

The idea that philosophically interesting concepts are often relatively “austere” in the requirements for their correct application has been a persistent theme in my work – as have the contentions (i) that often these concepts are governed by certain implicit, contextually variable, parameters, and (ii) that in certain contexts – including contexts where philosophical puzzles are being raised – the contextual parameters tend to take on settings that considerably raise the standards for correct application of the concepts. These ideas are reflected in my compatibilist approach to the problem of freedom and determinism in Horgan (1979, 1985a); in the defense of folk psychology in papers such as Horgan and Woodward (1985), Horgan (1987a, 1993a), Graham and Horgan (1988, 1991), and Henderson and Horgan (in press b); and in my **causal compatibilism** about mental causation and causal explanation in Horgan (1989b, 1991a, 1993c, 1994b, 1998a, 2001a).

My collaboration with Jim Woodward that produced Horgan and Woodward (1985) began shortly after I joined the Memphis philosophy department; Woodward was visiting in the department, before he moved to the California Institute of Technology. Working with him on this paper convinced me that philosophical collaboration can be both fun and fruitful, and engendered in me the disposition to do philosophy collaboratively when possible. (The collaborative papers cited in the present paper are all thoroughly coauthored; order of authorship in the citations is alphabetical.)

## 3. *Minimalism in Ontology (Four Phases)*

My work has long been informed by a strong inclination toward a fairly minimalistic ontology. My pursuit of ontological minimalism has gone through four discernable phases, three of which will be described in later sections. The first phase involved the pursuit of “paraphrase projects”: ways of systematically paraphrasing or “regimenting” discourse about various putative entities into discourse that avoids Quineian “ontological commitment” to those entities –

often via appeal to nonstandard logico-grammatical constructions like adverbial predicate-modifiers and non-truth-functional sentential connectives. This approach was pursued to avoid events (including mental events and actions) in Horgan (1978a, 1981a, 1982a) and Horgan and Tye (1985, 1988); to avoid numbers in Horgan (1984c, 1987c); and to avoid objects of the propositional attitudes in Horgan (1989a). But in the course of trying to find ways of “paraphrasing away” discourse about such putative entities as symphonies, corporations, and sentence-types, I came to believe that paraphrase projects probably could not be successfully carried through for all the various kinds of putative entities that I was disinclined to countenance ontologically. Something different, and more radical, seemed needed.

#### *4. Contextual Semantics: Truth as Indirect Correspondence*

The idea that came to mind, as a way of eschewing ontologically problematic entities without either (i) paraphrasing away the discourse that posits them or (ii) claiming that such discourse is false, was to “go soft on truth” – i.e., give up on the idea that truth is always a direct correspondence-relation between language (or thought) and world. Instead, construe truth as **semantically correct affirmability**, under semantic standards that involve **implicit contextual parameters** and often require only **indirect correspondence** with denizens of the mind-independent, discourse-independent, world – rather than direct correspondence. (Direct correspondence is a limit case, in which the contextual parameters are maximally strict.) First I called this approach “language-game semantics,” and then “psychologistic semantics,” before Mark Timmons and I began to call it **contextual semantics**. Articulating and defending contextual semantics has been an ongoing project for me, and some of my work on it has been collaborative with Timmons. See Horgan (1986a, 1986b, 1991b, 1995b, 1998c, 2001b), Horgan and Timmons (1993, in press), and Timmons (1999). The turn to contextual semantics ushered in the second stage of my minimalism in ontology. The guiding idea here is that numerous statements have these features: (i) they are true under the semantic standards that normally govern

their correct use; (ii) they carry Quineian ontological commitment to entities that are not denizens of the mind-independent, discourse-independent, world; (iii) they cannot be paraphrased into, or “regimented” by, statements that eschew apparent ontological commitment to the offending entities; and (iv) they are not true under limit-case, direct-correspondence, semantic standards.

### *5. Metaethical Irrealism*

My interest in metaethics began in the course on this topic I took with William Frankena when I was a graduate student at the University of Michigan. Mark Timmons and I began discussing metaethics during our time together in Hilary Putnam’s 1986 summer seminar “Philosophical Problems about Truth and Reality,” sponsored by the National Endowment for the Humanities. Timmons was working on the metaethical dimension of Putnam’s thought, and one project I had underway was Horgan (1987b), a paper tentatively exploring the implications of contextual semantics (which I then called “psychological semantics”) for metaethics. That paper of mine staked out a version of moral realism, but ongoing discussion with Timmons led me to return to my earlier inclination toward moral irrealism. After Timmons joined the philosophy department at Memphis, he and I began collaborating on work in metaethics.

On the negative side, we have been attacking various contemporary versions of naturalistic moral realism, often deploying a thought experiment we call **Moral Twin Earth**. As a recipe for deconstructing a given version of moral realism, the basic idea is this: (i) assume, for argument’s sake, that humans bear relation R to natural property P, where P is whatever natural property the given form of naturalism identifies with moral goodness and R is whatever relation it claims is the reference relation linking the term ‘goodness’ to P; (ii) assume that some specific normative moral theory (say, some version of consequentialism) comes out true if ‘goodness’ refers to P; (iii) consider a Twin Earth scenario in which people bear relation R to a somewhat different natural property Q, where a different normative moral theory (say, some deontological theory) comes out true if ‘goodness’ refers to Q. When one confronts the given version

of moral realism with the appropriate kind of Moral Twin Earth scenario, we argue, then the view under scrutiny ends up exhibiting one of these two objectionable features: either it is committed to a chauvinistic form of relativism, or else it yields very little determinate moral truth. See Horgan and Timmons (1991, 1992a, 1992b, 1996a, 1996b, 2000a).

On the positive side, Timmons and I have been developing and articulating a version of metaethical irrealism that we now call **nondescriptivist cognitivism**. A version of the position is set forth in Timmons (1999), and a more recent version appears in Horgan and Timmons (2000b). The basic claim is that although moral judgments are genuine beliefs and moral statements are genuine assertions, their overall content is not descriptive content. There is a connection to contextual semantics: we hold that the distinction between descriptive and nondescriptive content turns on whether or not a given judgment or statement is governed by **tight semantic standards** – that is, semantic standards that conspire with the mind-independent, discourse-independent, world to fix determinate correct affirmability (i.e., truth); cf. Horgan (1996, 2001b), Timmons (1999). We claim that moral judgments and statements have nondescriptive content, in the sense that the semantic standards governing them are not tight. On our view, a moral belief is a certain kind of commitment state – an **ought commitment** – with respect to what we call a **core descriptive content**. For instance, the belief that the Tennessee legislature ought to pass a state income tax is an ought commitment with respect to the core descriptive content *the Tennessee legislature passes a state income tax*. The overall content of the belief – viz., that the Tennessee legislature ought to pass a state income tax – is distinct from the belief’s core descriptive content; and this overall content is not itself descriptive.

## 6. *Post-Analytic Metaphilosophy*

I began collaborating with George Graham during an invited visit to his department at the University of Alabama in Birmingham. In Graham and Horgan (1988, 1991) we articulated and defended an anti-eliminativist approach to folk psychology that we whimsically

named **Southern Fundamentalism**. One reason for the name was that we were both teaching in the American south. Our key contention was that the concepts of folk psychology are sufficiently non-committal about matters like the type-type reducibility of psychology to neuroscience, or the existence of language-like mental representations, that the kinds of scientific scenarios envisioned by eliminativists (e.g., the failure of type-type reducibility, or the repudiation of the language of thought hypothesis) would not impugn the integrity of folk psychology at all.

As we worked to articulate our reasoning as clearly and explicitly as possible, we came to appreciate that our arguments reflected a certain general conception of philosophical methodology – one that we believe is actually at work implicitly in much philosophical theorizing past and present, often without being recognized for what it is. In Graham and Horgan (1994), a metaphilosophical paper, we named this methodology **Post-Analytic Metaphilosophy**. One leading idea is that philosophical inquiry into the workings of philosophically interesting concepts and terms – **ideology**, as we called it – is a broadly empirical matter, even though often it can be comfortably pursued from the armchair because much of its data is very close at hand (*die vom Armchair aus zuhandenen Daten*, as we came to call such data after helpful linguistic consultation with my colleague Tom Nenon). Another key idea is that intuitive judgments about what is right to say about various concrete scenarios – e.g., Gettier cases of justified true belief that seem not to be knowledge, Twin Earth scenarios in which the word ‘water’ seems to have a different meaning because people use it to refer to some clear potable liquid other than H<sub>2</sub>O, etc. – really have the status of empirical data for ideological reflection, in much the same way that intuitive judgments about grammaticality and grammatical ambiguity count as empirical data for linguists who are constructing theories of natural-language syntax.

Graham and I also generalized our earlier use of the term **Southern Fundamentalism**, now employing it for the general ideological hypothesis that philosophically interesting terms and concepts typically are relatively **austere** in their ideological commitments, rather than being **opulent**. We also acknowledged, however, that such concepts often are philosophically puzzling because they exhibit **ideo-**



**logical polarity** that involves “opulence tendencies” – something that we suggested is often the result of (i) implicit, contextually variable, parameters that govern the concepts, and (ii) a tendency for these parameters to take on a maximally strict setting in contexts where philosophical problems are being posed. There are close connections here with contextual semantics, and specifically with the contextualist versions of compatibilism I favor concerning both the freedom/determinism issue and the mental causation issue – and also (although I myself have not written about this) with contextualist approaches to the problem of Cartesian skepticism in epistemology.

### 6. *The Sorites Paradox and its Implications for Semantics and Metaphysics*

I have long had a serious side-interest in paradoxes, and occasionally I find myself writing about them. I have addressed Newcomb’s problem in Horgan (1981b, 1985b), the Monty Hall problem in Horgan (1995a), and the two-envelope paradox in Horgan (2000b, 2001c) and in work now in progress. In general I am inclined to believe that the paradoxes discussed in philosophy are often more difficult, and more philosophically deep, than they initially seem to be. (I concur with the memorable closing statement in Quine (1966): “Of all the ways of the paradoxes, perhaps the quaintest is their capacity on occasion to turn out to be so very much less frivolous than they look.”) In my work on Newcomb’s problem, for example, I reluctantly came to believe that “one box” reasoning and “two box” reasoning are in stalemate with one another, and that this fact reflects a deep internal tension within our ordinary notion of rationality itself. And my work on the two-envelope paradox led me to the conclusions (i) that **epistemic probability is intensional** in a way that is widely unappreciated, and (ii) that because of this intensionality, there are forms of **nonstandard expected utility** some of which are rationally appropriate to maximize and some of which are not.

But perhaps the deepest and most philosophically potent paradox on which I have worked is the sorites paradox, which evidently

arises ubiquitously in connection with vagueness. The paradox and its morals are addressed in Horgan (1990, 1993d, 1994c, 1995b, 1997a, 1998b), in Horgan and Potrč (2000, in press), and in work of mine now in progress. According to my treatment of vagueness, which I call **transvaluationism**, vagueness by its very nature exhibits a certain kind of **benign logical incoherence**: vague terms and concepts are semantically governed by **mutually unsatisfiable semantic requirements** – a fact that surfaces explicitly in the sorites paradox. This kind of incoherence is benign because a form of **logical discipline** remains in force with respect to the semantically correct use of vague terms and concepts, and this discipline quarantines the incoherence so that it does not generate malignant results – such as rampant logical commitment to statements of the form  $(\Phi \ \& \ \sim\Phi)$ .

Transvaluationism about vagueness has important philosophical consequences, both for semantics and for metaphysics. Regarding semantics, transvaluationism leads to the conclusion that truth, for vague statements and vague judgments, must always be indirect correspondence to the mind-independent, discourse-independent, world – rather than direct correspondence. (Here my work on vagueness comes together with contextual semantics.) Regarding ontology, transvaluationism leads to the conclusion that **ontological vagueness is impossible** – i.e., the correct ontology, whatever it might be, cannot include vague objects, vague properties, or vague relations.

This repudiation of ontological vagueness is a third stage in the evolution of my ontological minimalism. I now find myself eschewing not only putative entities like symphonies, corporations, and numbers, but also putative entities like mountains, tables, and persons – because the latter, if they were real, would be essentially vague in certain respects (e.g., vague with respect to composition, and/or with respect to spatiotemporal boundaries). I do not regard my position as a radical repudiation of common sense, however, because in my view the semantic standards that usually govern discourse and thought are indirect-correspondence standards rather than direct-correspondence ones. Direct-correspondence standards, the limit case, operate only in fairly unusual contexts, e.g., contexts of serious ontological inquiry. (They do not operate even in scientific contexts, insofar as the posits of scientific theory are vague – as

they very often are). Thus, most of our discourse and thought does not carry **ultimate ontological commitment** to its vague posits, but only what I call **regional ontological commitment**. As such, many of our statements and judgments employing vague terms and concepts are indeed true – that is, they are semantically correct, under the contextually operative semantic standards that govern them.

What minimal ontology of concrete particulars is the correct one, given the impossibility of ontological vagueness? I believe that a strong case can be made for the view that there is really only one concrete particular – viz., the entire physical universe, which I call **the blobject**. I briefly argued for this position in Horgan (1991b), where I called the view “Parmenidean materialism.” But Matjaž Potrč persuaded me that this is not a good name, because evidently the object that Parmenides called “the one” was supposed to be entirely homogenous – whereas the physical universe surely exhibits enormous spatiotemporal complexity and nonhomogeneity. Potrč and I now call the view **blobjectivism**; we defend it in Horgan and Potrč (2000). In Horgan and Potrč (in press) we further elaborate the view by addressing certain questions for blobjectivism posed by Tienson (in press). Blobjectivism is the fourth, and most recent, stage in the evolution of my ontological minimalism.

Perhaps the only way to go any further in this minimalist direction would be to repudiate my metaphysical realism and join the ranks of the global irrealists who deny that there is a mind-independent, discourse-independent, world at all. But I doubt that I will ever take that step, partly because I find global metaphysical irrealism unintelligible; cf. Horgan (1991b, 2001b).

## *8. Connectionism and the Philosophy of Psychology*

When John Tienson joined the Memphis philosophy department, he and I together attended the weekly meetings of an interdisciplinary cognitive science seminar at the university. There we learned about the newly burgeoning approach in cognitive science known as connectionism. In 1987 Tienson and I directed our department’s annual Spindel Conference (named for its benefactor Murray Spindel, a great friend of the department until his death in 1999), on the topic

“Connectionism and the Philosophy of Mind.” As usual, the proceedings were published as a supplement to the *Southern Journal of Philosophy*: Horgan and Tienson (1988a). We also edited a collection, Horgan and Tienson (1991), that included numerous additional papers beyond those from the conference.

Tienson and I realized early on that we had similar ideas about what was wrong with the classical, computational, paradigm in cognitive science – and about what was potentially most philosophically interesting about connectionism. Beginning with Horgan and Tienson (1988b), the paper we wrote together for the 1987 Spindel Conference, we produced a series of collaborative papers and then a book that explained what we took to be in-principle problems for the classical computational view of cognition, and also described a connectionism-inspired nonclassical framework for cognitive science that potentially could overcome the problems faced by “classicism.” See Horgan and Tienson (1988b, 1989, 1990a, 1992a, 1992b, 1993, 1994a, 1994b, 1995, 1996, 1997, 1998, 1999a, 1999b); Horgan (1997b, 1997c); and Tienson (1997).

Our name for the recommended nonclassical framework for cognitive science was **noncomputable dynamical cognition**. Some key ideas are the following. First, human cognitive state-transitions are normally too subtle and complex to conform to any tractably computable transition-function over cognitive states – something that we claimed is a principal lesson of the family of related difficulties faced by classicism known collectively as the “frame problem.” Second, there is a form of mathematics that goes naturally with connectionist modeling, and is potentially more powerful as a basis for understanding cognition than the discrete mathematics employed in the theory and practice of computation – viz., dynamical systems theory. Third, in principle a suitably structured high-dimensional dynamical system, implementable by a connectionist network or a human brain, could subserve cognitive state-transitions that are too complex to conform to programmable representation-level rules. Fourth, in principle such a dynamical system could avoid the sources of the frame problem by accommodating much relevant information not in the form of explicit representations (as in classicism), but instead implicitly in the structure of the dynamical system itself. Such implicit information we called **morphological content**

– the idea being that this kind of content is present in the standing structure of the system, rather than being explicitly represented by occurrent cognitive states.

I presented the material later published as Horgan and Tienson (1989, 1990a) at a 1989 conference in Dubrovnik on philosophy of mind and epistemology. There I met Matjaž Potrč from Ljubljana. Potrč later spent the academic year 1996-97 in the Memphis philosophy department on a Fulbright Fellowship, where Tienson and I both had many useful discussions with him about our book, about related work of his, and about various other philosophical matters. He organized a conference on Horgan and Tienson (1996), which took place in Ljubljana in summer 1997; the proceedings were published as *Acta Analytica* 22, (1999). Since then I have had a close working relationship with him, including the co-authored papers mentioned above and our co-editorship of *Acta Analytica* 23 (1999), a collection of papers from the 1998 conference on vagueness in Bled.

### 9. *Epistemological Themes*

In the wake of my work with Tienson on connectionism and the philosophy of psychology, I began collaborating with my Memphis colleague David Henderson on a range of topics in epistemology. In our first paper, Henderson and Horgan (2000a), we focused on the important epistemological distinction between *having* justification for a given belief and *being* justified in holding that belief – where the latter is a matter of holding the belief *because* of the justification one has for it. We argued that much epistemological theorizing effectively presupposes an empirically questionable conception of this ‘because’-connection, and effectively ignores the role that morphological content very likely plays in the fixation of belief. We urged a new approach that we called **iceberg epistemology**, which emphasizes that cognitive states that are conscious or consciously accessible are typically only the “tip of the iceberg” insofar as the psychology of belief-fixation is concerned. We revisit this theme in Henderson and Horgan (forthcoming), where we spell out our recommended alternative conception of being-justified in more detail.

Iceberg epistemology has numerous potential implications, including the capacity to accommodate aspects of erstwhile competing epistemological theories like foundationalism, coherentism, and structural contextualism.

In Henderson and Horgan (2001), we argue that there is an important and insufficiently appreciated aspect of objective epistemic justification, over and above a belief's having been produced by a reliable belief-forming process. We call it **robustness of reliability** (or just robustness, for short). The basic idea is this: for a belief-forming process to be robustly reliable is for it to be reliable in a very wide range of "epistemically relevant" possible worlds – roughly, worlds in which the epistemic agent undergoes appearances much like the appearances experienced in the agent's actual world. The empirical beliefs of someone who is deceived by a Cartesian evil demon, for example, could be the products of belief-forming processes that are robust in this sense, and hence such beliefs could be objectively very well justified – even though those belief-forming processes happen to be thoroughly *unreliable* in the agent's actual world.

In Henderson and Horgan (2000b, in press a), we draw upon and extend the methodological ideas that George Graham and I had set out under the rubric 'post-analytic metaphilosophy'. We offer a nonclassical reconception of a priori reasoning and a priori knowledge – with a paradigm case being ideological inquiry within philosophy itself. The **low-grade a priori**, as we call it, is a form of inquiry and knowledge that is broadly empirical, and yet nonetheless is sufficiently distinctive – and sufficiently similar to the a priori as classically conceived – to still warrant the label 'a priori'. Such inquiry typically seeks to discover conceptually grounded necessary truths, and it typically relies heavily upon empirical data (including intuitions about how to apply key concepts to concrete thought-experimental scenarios) readily available by armchair reflection.

## 10. *The Limits of Systematizability*

Tienson and I, in our work on connectionism and the philosophy of psychology, argued that human cognitive state-transitions evidently are not fully systematizable by exceptionless, general, psychologi-

cal laws that take the form of programmable rules. Once this claim is taken to heart, one begins to appreciate that cognitive states need not be – and probably are not – fully systematizable by exceptionless psychological laws of *any* kind. This is consistent with their being *partially* systematizable by psychological generalizations that we called **soft laws** – that is, psychological laws that have *ineliminable* ‘*ceteris paribus*’ clauses adverting not only to lower-level exceptions like external interference or internal physical malfunction, but also to **same-level exceptions** characterizable at the psychological level of description; cf. Horgan and Tienson (1990b, 1996).

A theme that has been much on my mind lately, although not yet much articulated in my written work, is this: just as there are apparent limits to the systematizability of human cognitive processing by way of general laws, so too there could well be – and probably are – significant limits to the systematizability of various kinds of *normativity* to which humans are capable of conforming in their cognitive, linguistic, and behavioral practices – notably semantic normativity, epistemic normativity, and moral normativity. This idea is articulated and explored with respect to epistemic normativity in Potrč (2000). The same general theme also is explored in the work of two former doctoral students from Memphis: in the case of epistemic normativity, Stark (1993, 1994, 1999); and in the case of moral normativity, Haney (1999).

One way that this theme affects my current thinking is with respect to the further articulation and defense of contextual semantics. One might think that an adequately worked out version of contextual semantics should provide, at least in rough outline, a general and systematic formulation of the semantic normative standards that govern correct affirmability – in particular, correct affirmability of the indirect-correspondence variety. But increasingly I have come to think that such a formulation probably is not possible – not because contextual semantics is mistaken, but rather because semantic normativity is probably too subtle and too complex to be thus systematizable. To insist on such systematizability, on the grounds that otherwise human thought and human linguistic behavior could not conform to these semantic standards, would be to underestimate the capacities of human cognition.

## 11. *The Whole Hard Problem in Philosophy of Mind*

Ever since I wrote my doctoral dissertation, I have been worried about whether an adequate materialist account of phenomenal consciousness could be given, and about what it would look like. Things would go smoother if some version of functionalism were correct – perhaps a version that incorporates typical cause connections between inner states of the organism and features of the distal environment. As I remarked in section 1 above, it appears that materialistically acceptable explanations are available for the supervenience relations linking physical properties to functional properties. But for the phenomenal aspects of mentality, at least, functionalism has always seemed to me implausible. One familiar way to make the point is in terms of the apparent conceivability of “inverted qualia” and “absent qualia” scenarios, a topic I have written about myself; cf. Horgan (1984a, 1987d).

Also very potent dialectically is Frank Jackson’s famous thought experiment about Mary the colorblind neuroscientist. Although I wrote one of the first replies to Jackson on behalf of the materialist (Horgan 1984b), I still suspected that his thought experiment probably could be deployed somewhat differently than he had, to produce a serious challenge to materialist treatments of phenomenal consciousness. George Graham and I took up this gauntlet in Graham and Horgan (2000). We focused specifically on the account of phenomenal consciousness in Tye (1995), but we argued that our argument looks generalizable. The key line of reasoning was as follows. Let Mary Mary (the daughter of Jackson’s Mary) be a colorblind scientist/philosopher who not only knows all there is to know about the physics of color and the neurobiology of color perception, but also believes a specific materialist account of phenomenal concepts and the properties to which they refer – say, Tye’s account. She understands well what the materialist account says about the nature of color experiences, about the nature of phenomenal concepts, about the nature of the physical/functional properties these concepts supposedly refer to, and about the need to undergo color experiences in order to actually *possess* phenomenal concepts. (Basically, possession of such concepts is supposed to be a matter of being able to color-categorize things solely on the basis of seeing them, without



relying on collateral information.) Given her scientific knowledge and her philosophical beliefs, she has no rational reason to expect to be dramatically surprised should she begin to undergo color experiences in the future; for, she already knows that those experiences will allow her to begin color-classifying things just by looking at them. But of course, she *will* be dramatically surprised (and delighted) by color experiences – not because of the new classificatory capacities they provide her, but because of their *intrinsic phenomenology*. Such surprise, which evidently ought to be irrational given the philosophical theory she has accepted prior to having color experiences, is not irrational at all.

Some materialists, seeking to show full respect for the intrinsic phenomenology of phenomenal consciousness, advocate a type of position that John Tienson and I have dubbed “new wave materialism.” This view asserts that although “inverted qualia” and “absent qualia” scenarios are not metaphysically possible, they are indeed conceivable; it also maintains that although Jackson’s Mary deploys new concepts and acquires new knowledge upon first experiencing color vision, the properties she learns about are physical/functional properties that she knew about already. New wave materialists also claim that experience-based phenomenal concepts – for instance, the new color-concepts that Mary first acquires when she begins to have color vision – pick out their referents not via any *accidental* reference-fixing features of the referent, but instead via something *essential* about the referent-property. But in Horgan and Tienson (2001), it is argued that new wave materialism faces a theoretical problem that threatens to deconstruct the view. The problem is to explain intelligibly how the following three claims, each of which the view embraces, can be simultaneously true: (i) when phenomenal properties are conceived under phenomenal concepts, they are conceived essentially, as they are in themselves; (ii) when phenomenal properties are conceived under phenomenal concepts, they are conceived otherwise than as physical properties; (iii) phenomenal concepts refer to physical properties. Prima facie, it appears that the conjunction of claims (i) and (ii) is incompatible with claim (iii), given the plausible further claim (also embraced by new wavers) that when phenomenal properties are conceived under *physical* concepts, they are conceived essentially, as they are in themselves.

All of this underscores, and fleshes out in different ways, the so-called “hard problem” of phenomenal consciousness. I once thought, as many in philosophy of mind still do think, that the phenomenal and the intentional aspects of mentality are largely separable from one another. But lately, largely as a result of ongoing discussion with John Tienson and George Graham, I have come to believe that this “separatist” doctrine about phenomenology and intentionality is thoroughly mistaken. In Horgan and Tienson (in press) it is argued that phenomenology and intentionality are thoroughly and essentially intertwined – and in particular, that there is an essential “what it’s like” of having an occurrent thought, or of having an occurrent desire. This means that the **whole hard problem** – that is, the overall problem of explaining the phenomenal aspects of mentality – extends to paradigmatically intentional states too, thereby encompassing virtually the whole of human conscious (as opposed to unconscious) mental life.

I remain deeply attracted to materialism in philosophy of mind; I would like to believe that the mental is superdupervenient on the physical. But the whole hard problem looks very hard indeed, and I see no prospects currently in sight for dealing with it satisfactorily. Theories of mind that claim to do justice to phenomenology, and yet would be satisfiable by zombies who lack phenomenal consciousness altogether, do not seem credible. Much as I would like to be a materialist, at present I do not know what an adequate materialist theory of mind would look like.

### *Acknowledgements and Dedication*

My sincere thanks to Matjaž Potrč for initiating the Horgan Symposium that took place at the 2000 Austrian-Slovene Philosophical Euroconference, to Wolfgang Gombocz for organizing the conference, and to Johannes Brandl and Olga Markič for co-editing this collection. My deep thanks too to all the participating philosophers, for their very stimulating presentations. David Henderson and Mark Timmons each gave useful overview talks about their collaborative work with me. Others who gave presentations but did not submit papers to this volume were John Bickle, Bojan Borstner, Johannes

Brandl, Božidar Kante, Friderik Klampfer, Nenad Mišćević, and Diana Raffman. Thanks also to David Henderson, Mark Timmons, and John Tienson for discussion and feedback as I prepared the present paper and my reply to the papers in this volume.

My wife Dianne and I lost our son Alexander William Horgan on July 30, 2001. He died of suicide, at age 23. Alec was enormously talented in many ways, including as a musician, a song writer, and a fiction writer. He had a gigantic intellect, and was as brilliant a person as I have ever known. He suffered from severe depression from his teen years onward, and in the end it overcame him. I dedicate this volume to his memory.

## REFERENCES

- Chalmers, D. 1995. The Puzzle of Conscious Experience, *Scientific American* 273, 80-86.
- Chalmers, D. 1996. *The Conscious Mind*. Oxford University Press.
- Davidson, D. 1970. Mental Events. In L. Forster and J. W. Swanson, eds., *Experience and Theory* (University of Massachusetts Press), 79-101. Reprinted in D. Davidson, *Essays on Actions and Events* (Oxford, 1980).
- Graham, G. and Horgan, T. 1988. How to be Realistic About Folk Psychology, *Philosophical Psychology* 1, 69-81.
- Graham, G. and Horgan, T. 1991. In Defense of Southern Fundamentalism, *Philosophical Studies* 62, 107-34.
- Graham, G. and Horgan, T. 1994. Southern Fundamentalism and the End of Philosophy, *Philosophical Issues* 5, 219-47.
- Graham, G. and Horgan, T. 2000. Mary Mary, Quite Contrary, *Philosophical Studies* 99, 59-87.
- Haney, M. 1999. Dynamical Cognition, Soft Laws, and Moral Theorizing, *Acta Analytica* 22, 227-40.
- Hare, R. M. 1952. *The Language of Morals*. Oxford University Press.
- Henderson, D. and Horgan, T. 2000a. Iceberg Epistemology, *Philosophy and Phenomenological Research* 61, 497-535.

- Henderson, D. and Horgan, T. 2000b. What is A Priori and What is It Good For?, *Southern Journal of Philosophy* 38, Spindel Conference Supplement, 51-86.
- Henderson, D. and Horgan, T. 2001. Practicing Safe Epistemology, *Philosophical Studies* 102, 227-58.
- Henderson, D. and Horgan, T. In press a. The A Priori Isn't All That It's Cracked Up to Be, But It Is Something, *Philosophical Topics*, issue honoring Alvin Goldman.
- Henderson, D. and Horgan, T. In press b. What Does It Take to Be a True Believer? Against the Opulent Ideology of Eliminative Materialism. In C. Erneling and D. Johnson, eds., *Mind as a Scientific Object: Between Brain and Culture*. Oxford University Press.
- Henderson, D. and Horgan, T. Forthcoming. Morphological Content and Justified Belief.
- Horgan, T. 1976. Reduction and the Mind-Body Problem. In M. Marx and F. Goodson, eds., *Theories in Contemporary Psychology*, 2nd edition (Macmillan), 223-31.
- Horgan, T. 1978a. The Case Against Events, *Philosophical Review* 87, 28-47.
- Horgan, T. 1978b. Supervenient Bridge Laws, *Philosophy of Science* 45, 227-49.
- Horgan, T. 1979. 'Could', Possible Worlds, and Moral Responsibility, *Southern Journal of Philosophy* 17, 345-58.
- Horgan, T. 1981a. Action Theory Without Actions, *Mind* 60, 406-14.
- Horgan, T. 1981b. Counterfactuals and Newcomb's Problem, *Journal of Philosophy*, 331-56.
- Horgan, T. 1981c. Token Physicalism, Supervenience, and the Generality of Physics, *Synthese* 49, 395-413.
- Horgan, T. 1982a. Substitutivity and the Causal Connective, *Philosophical Studies* 42, 47-52.
- Horgan, T. 1982b. Supervenience and Microphysics, *Pacific Philosophical Quarterly* 63, 29-43.
- Horgan, T. 1984a. Functionalism, Qualia, and the Inverted Spectrum, *Philosophy and Phenomenological Research* 44, 453-69.
- Horgan, T. 1984b. Jackson on Physical Information and Qualia, *Philo-*

- sophical Quarterly* 34, 147-52.
- Horgan, T. 1984c. Science Nominalized, *Philosophy of Science* 51, 529-49.
- Horgan, T. 1984d. Supervenience and Cosmic Hermeneutics, *Southern Journal of Philosophy* 22, Supplement Issue on Supervenience, 19-38.
- Horgan, T. 1985a. Compatibilism and the Consequence Argument, *Philosophical Studies* 47, 339-56.
- Horgan, T. 1985b. Newcomb's Problem: A Stalemate. In R. Campbell and L. Sowden, eds., *Paradoxes of Rationality and Cooperation* (U. of British Columbia Press), 223-34.
- Horgan, T. 1986a. Psychologism, Semantics, and Ontology, *Nous* 20, 21-31.
- Horgan, T. 1986b. Truth and Ontology, *Philosophical Papers* 15, 1-21.
- Horgan, T. 1987a. Cognition is Real, *Behaviorism* 15, 13-25.
- Horgan, T. 1987b. Psychologistic Semantics and Moral Truth, *Philosophical Studies* 52, 357-70.
- Horgan, T. 1987c. Science Nominalized Properly, *Philosophy of Science* 54, 281-82.
- Horgan, T. 1987d. Supervenient Qualia, *Philosophical Review* 96, 491-520.
- Horgan, T. 1989a. Attitudinatives, *Linguistics and Philosophy* 12, 133-65.
- Horgan, T. 1989b. Mental Quausation, *Philosophical Perspectives* 3, 47-76.
- Horgan, T. 1990. Psychologistic Semantics, Robust Vagueness, and the Philosophy of Language. In S. L. Tsohatzidis, ed., *Meanings and Prototypes: Studies in Linguistic Categorization* (Routledge), 535-57.
- Horgan, T. 1991a. Actions, Reasons, and the Explanatory Role of Content. In B. McLaughlin, ed., *Dretske and His Critics* (Basil Blackwell), 73-101.
- Horgan, T. 1991b. Metaphysical Realism and Psychologistic Semantics, *Erkenntnis* 34, 297-322.
- Horgan, T. 1993a. The Austere Ideology of Folk Psychology. *Mind and Language* 8, 282-97.

- Horgan, T. 1993b. From Supervenience to Superdupervenience: Meeting the Demands of a Material World, *Mind* 102, 555-86.
- Horgan, T. 1993c. Nonreductive Materialism and the Explanatory Autonomy of Psychology. In Wagner & Warner, eds., *Naturalism: A Critical Appraisal* (Notre Dame), 295-320.
- Horgan, T. 1993d. On What There Isn't, *Philosophy and Phenomenological Research* 50, 693-700.
- Horgan, T. 1994a. Naturalism and Intentionality, *Philosophical Studies* 76, 301-26.
- Horgan, T. 1994b. Nonreductive Materialism. In R. Warner and T. Szubka, eds., *The Mind-Body Problem* (Blackwell), 236-41.
- Horgan, T. 1994c. Robust Vagueness and the Forced-March Sorites Paradox, *Philosophical Perspectives* 8, *Logic and Language*, 159-88.
- Horgan, T. 1995a. Let's Make a Deal, *Philosophical Papers* 24, 209-22.
- Horgan, T. 1995b. Transvaluationism: A Dionysian Approach to Vagueness, *Southern Journal Philosophy* 33, Spindel Conference Supplement, 97-125.
- Horgan, T. 1996. The Perils of Epistemic Reductionism, *Philosophy and Phenomenological Research* 66, 891-97.
- Horgan, T. 1997a. Brute Supervenience, Deep Ignorance, and the Problem of the Many, *Philosophical Issues* 8, 229-36.
- Horgan, T. 1997b. Connectionism and the Philosophical Foundations of Cognitive Science, *Metaphilosophy* 28, 1-30.
- Horgan, T. 1997c. Modelling the Noncomputational Mind: Reply to Litch, *Philosophical Psychology* 10, 365-71.
- Horgan, T. 1998a. Kim on Mental Causation and Causal Exclusion, *Philosophical Perspectives* 11, 165-84.
- Horgan, T. 1998b. The Transvaluationist Conception of Vagueness, *The Monist* 81, 316-33.
- Horgan, T. 1998c. Actualism, Quantification, and Contextual Semantics, *Philosophical Perspectives* 12, 503-09.
- Horgan, T. 2000a. Facing Up to the Sorites Paradox. In A. Anamori, ed., *Proceedings of the 20<sup>th</sup> World Congress of Philosophy. Volume 6: Analytic Philosophy and Logic* (Philosophy Documentation Center), 99-111.

- Horgan, T. 2000b. The Two-Envelope Paradox, Nonstandard Expected Utility, and the Intensionality of Probability, *Nous* 34, 578-602.
- Horgan, T. 2001a. Causal Compatibilism and the Exclusion Problem, *Theoria* 16, 95-116.
- Horgan, T. 2001b. Contextual Semantics and Metaphysical Realism: Truth as Indirect Correspondence In M. Lynch, ed., *The Nature of Truth: Classic and Contemporary Perspectives* (MIT Press), 67-95.
- Horgan, T. 2001c. The Two-Envelope Paradox and the Foundations of Rational Decision Theory. In B. Brogaard and B. Smith, eds., *Rationality and Irrationality: Proceedings of the 23<sup>rd</sup> International Wittgenstein Symposium* (öbv & hpt), 172-91.
- Horgan, T. and Potrč, M. 2000. Blobjectivism and Indirect Correspondence, *Facta Philosophica* 2, 249-70.
- Horgan, T. and Potrč, M. In press. Addressing Questions for Blobjectivism, *Facta Philosophica*.
- Horgan, T. and Tienson, J., eds. 1988a. *Spindel Conference 1987: Connectionism and the Philosophy of Mind*. *Southern Journal of Philosophy* 26, Supplement.
- Horgan, T. and Tienson, J. 1988b. Settling into a New Paradigm, *Southern Journal of Philosophy* 26, Supplement, 97-114. Reprinted in T. Horgan and J. Tienson, eds., *Connectionism and the Philosophy of Mind* (Kluwer, 1991).
- Horgan, T. and Tienson, J. 1989. Representations without Rules, *Philosophical Topics* 17, 27-43.
- Horgan, T. and Tienson, J. 1990a. Connectionism and the Kuhnian Crisis in Cognitive Science, *Acta Analytica* 6, 5-17.
- Horgan, T. and Tienson, J. 1990b. Soft Laws, *Midwest Studies in Philosophy* 15, 256-79.
- Horgan, T. and Tienson, J., eds. 1991. *Connectionism and the Philosophy of Mind*. Kluwer.
- Horgan, T. and Tienson, J. 1992a. Cognitive Systems as Dynamical Systems, *Topoi* 11, 27-43.
- Horgan, T. and Tienson, J. 1992b. Structured Representations in Connectionist Systems?, In S. Davis., ed., *Connectionism: Theory and Practice* (Oxford), 195-228.

- Horgan, T. and Tienson, J. 1993. Levels of Description in Nonclassical Cognitive Science, *Philosophy* 34, Royal Institute of Philosophy Supplement, 159-88.
- Horgan, T. and Tienson, J. 1994a. Representations Don't Need Rules: Reply to James Garson, *Mind and Language* 9, 38-55.
- Horgan, T. and Tienson, J. 1994b. A Nonclassical Framework for Cognitive Science, *Synthese* 101, 305-345.
- Horgan, T. and Tienson, J. 1995. Connectionism and the Commitments of Folk Psychology, *Philosophical Perspectives* 9, 127-52.
- Horgan, T. and Tienson, J. 1996. *Connectionism and the Philosophy of Psychology*. MIT Press.
- Horgan, T. and Tienson, J. 1997. Précis of *Connectionism and the Philosophy of Psychology*, *Philosophical Psychology* 10, 337-56.
- Horgan, T. and Tienson, J. 1998. Resisting the Tyranny of Terminology: The General Dynamical Hypothesis in Cognitive Science, *Behavioral and Brain Sciences* 21, 643.
- Horgan, T. and Tienson, J. 1999a. Authors' Replies, *Acta Analytica* 22, 275-87.
- Horgan, T. and Tienson, J. 1999b. Short Précis of *Connectionism and the Philosophy of Psychology*, *Acta Analytica* 22, 9-21.
- Horgan, T. and Tienson, J. 2001. Deconstructing New Wave Materialism. In C. Gillett and B. Loewer, eds., *Physicalism and Its Discontents* (Cambridge), 307-18.
- Horgan, T. and Tienson, J. In press. The Intentionality of Phenomenology and the Phenomenology of Intentionality. In D. Chalmers, ed., *Philosophy of Mind: Classical and Contemporary Readings* (Oxford).
- Horgan, T. and Timmons, M. 1991. New Wave Moral Realism Meets Moral Twin Earth, *Journal of Philosophical Research* 16, 447-65; and in J. Heil, ed., *Rationality, Morality, and Self-Interest: Essays Honoring Mark C. Overvold* (Rowman and Littlefield, 1993).
- Horgan, T. and Timmons, M. 1992a. Troubles on Moral Twin Earth: Moral Queerness Revived, *Synthese* 92, 221-60.
- Horgan, T. and Timmons, M. 1992b. Troubles for New Wave Moral Semantics: The 'Open Question Argument' Revived, *Philosophical Papers* 21, 153-75.



- Horgan, T. and Timmons, M. 1993. Metaphysical Naturalism, Semantic Normativity, and Meta-Semantic Irrealism, *Philosophical Issues* 4, 180-203.
- Horgan, T. and Timmons, M. 1996a. Troubles for Michael Smith's Metaethical Rationalism, *Philosophical Papers* 25, 203-231.
- Horgan, T. and Timmons, M. 1996b. From Moral Realism to Moral Relativism in One Easy Step, *Critica* 28, 3-39.
- Horgan, T. and Timmons, M. 2000a. Copping Out on Moral Twin Earth, *Synthese* 124, 139-52.
- Horgan, T. and Timmons, M. 2000b. Nondescriptivist Cognitivism: Framework for a New Metaethic, *Philosophical Papers* 29, 121-53.
- Horgan, T. and Timmons, M. In press. Conceptual Relativity and Metaphysical Realism, *Philosophical Issues*.
- Horgan, T. and Tye, M. 1985. Against the Token Identity Theory. In E. LePore and B. McLaughlin, eds., *Act and Event: The Philosophy of Donald Davidson* (Basil Blackwell), 427-43.
- Horgan, T. and Tye, M. 1988. Braving the Perils of an Uneventful World, *Grazer Philosophische Studien* 31, 179-186.
- Horgan, T. and Woodward, J. 1985. Folk Psychology is Here to Stay, *Philosophical Review* 94, 197-226.
- Levine, J. 1983. Materialism and Qualia: The Explanatory Gap, *Pacific Philosophical Quarterly* 64, 354-61.
- Potrč, M. 2000. Justification Having and Morphological Content, *Acta Analytica* 24, 151-73.
- Quine, W. V. O. 1966. The Ways of Paradox. In his *The Ways of Paradox and Other Essays* (Random House).
- Stark, H. 1993. Rationality is Not Algorithmic, *Kinesis* 20, 16-30.
- Stark, H. 1994. Connectionism and the form of Rational Norms, *Acta Analytica* 12, 39-53.
- Stark, H. 1999. What the Dynamical Cognitive Scientist Said to the Epistemologist, *Acta Analytica* 22, 241-60.
- Timmons, M. 1999. *Morality without Foundations: A Defense of Ethical Contextualism*. Oxford University Press.
- Tienson, J. 1997. What the Differences Are: Reply to Hardcastle, *Philo-*

*sophical Psychology* 10, 385-9.

Tienson, J. In press. Questions for Blobjectivism, *Facta Philosophia*.

Tye, M. 1995. *Ten Problems of Consciousness: A Representationalist Theory of the Phenomenal Mind*. MIT Press.