

Intuitive Ontologies, Critical Metaphysics, and Cognitive Pluralism

For the December 2011 Conference “Psycho-Ontology” hosted by the Shalem Institute in Jerusalem

Steven Horst, Ph.D.
Department of Philosophy, Wesleyan University
Middletown, CT 06459 USA
shorst@wesleyan.edu
www.stevenhorst.com

Preface

When I saw that I was to speak on the last day of the conference, I experienced a moment of panic. On the one hand, there was the very real risk that I would end up re-hashing territory already covered, and that long expositions of those things would prove redundant, and others left under-developed.

I decided, in the end, to scrap the parts of my talk in which I tried to articulate the case for a cognitivist metaphysics – that is, a metaphysics that does not treat the kinds of objects, properties, relations and processes found in the world as ultimate metaphysical bedrock, but emphasizes the mind’s own role in constituting a world of objects. And, indeed, anyone participating in a conference like this probably already has an understanding of the appeal of such a metaphysics, even if they wish to resist it in favor of a more commonsense realism or go beyond it by grounding cognition in practices or in bodily engagement with the world. So, I shall proceed on the assumption that we are taking cognitive metaphysics seriously

as a general philosophical programme. Towards that end, I shall do four things in my talk:

- First, I shall briefly distinguish several senses of the word ‘ontology’ (and more broadly, ‘metaphysics’) as used within philosophy and the cognitive sciences.
- Second, I shall outline several lines of research in the cognitive sciences that suggest that human understanding does not come in the form of a single, consistent and comprehensive world-view, but rather comes in the form of mental models of particular content domains. This is the first component of a view I wish to advocate which I call Cognitive Pluralism.
- Third, I shall explore the epistemological implications of Cognitive Pluralism: in particular, how it might explain why our understanding, from commonsense knowledge domains to scientific theories, resist wholesale unification.
- And finally, I shall develop some issues that emerge when one considers Cognitive Pluralism as a view in critical metaphysics.

1. Notions of “Ontology” in Philosophy

In the first edition of *Encyclopaedia of Philosophy* (Edwards, 1967), the entry on ‘ontology’ (written, somewhat surprisingly, by Alisdair Macintyre), distinguishes two uses of the word. A modern usage, which Macintyre associates with Quine, is

essentially concerned with the *inventory* of the world: “ontology”, in this sense, is concerned with the question of *what exists*. I shall refer to this as *inventory ontology*. Now, there is a very uninteresting sort of inventory ontology that amounts to nothing more than a laundry list of what sorts of things there are: cabbages and kings exist, unicorns and perpetual motion machines do not. The philosophically serious form of inventory ontology, corresponding to what Frank Jackson (Jackson, 1997) calls “serious metaphysics”, is concerned with deeper questions about the *fundamental kinds* of things: e.g., are material objects the only things that exist, or are there also souls, properties, states of affairs, mathematical objects like numbers and sets, etc.? Materialism, dualism, and mathematical Platonism are perhaps the best-known inventory ontologies.

One of my misgivings about the term “psycho-ontology” is that even “serious” inventory ontologies that are relatively deflationary – say, those that do not count properties, states of affairs, or numbers among *things that exist* – are generally part and parcel with some kind of account of these other categories. Ontology is seldom completely divorced from the broader realm of *metaphysics*, and so “psycho-metaphysics” might be a more apt characterization.

The second usage of the word ‘ontology’ flagged by Macintyre is the older notion of “the study of being *qua* being”. This sort of project – which is concerned, not with what kinds of things exist, but with *what it is to be* an object, a property, and so on – I shall call *critical ontology*, and its broader context *critical metaphysics*. Familiar positions in critical metaphysics would include various forms of idealism, pragmatism, social constructionism, and instrumentalism.

Of particular interest to us at this conference are critical metaphysics of a sort that I shall label as *cognitivist*. Cognitivist metaphysics are those that attempt to ground metaphysical categories like *substance* and *property* in something about our own cognitive architecture. Idealisms, particularly the Transcendental Idealisms of Kant and Husserl, stand as paradigms of this sort of project. And we might be inclined to include as cognitivist (at least in a broad sense) views which take things a step further by trying to ground cognition itself in embodied practices: for example, American Pragmatism, Existential Phenomenology, and Embodied Cognition.

Cognitivist critical metaphysics, and indeed critical metaphysical projects in general, tend to have few implications for inventory ontology, save perhaps for the assumption that there are such things as *minds*. Philosophers like Kant and Peirce, for example, were perfectly content to view the question of what particular kinds of things exist as an empirical question, to be settled by experience and particularly by the sciences. Their philosophical concern is, rather, that even when we have settled such empirical questions, we have not yet reached metaphysical bedrock, as the most general categories like *substance* and *property* still need to be cashed out in cognitivist and/or pragmatic terms.

2. “Ontology” in Cognitive and Developmental Psychology

Psychologists have also begun to use the word ‘ontology’. Their interest, however, is not in putting forward philosophical theories about what exists or the

nature of being, but in describing *how people think* – their concepts, categories, and the inferential patterns and data structures that go along with these. Since the time of Piaget, developmental psychologists have suggested various ways in which children at different developmental stages have “theories” or “schemas” that represent the world differently from adult cognition. And cognitive psychologists have explored the “implicit” or “intuitive ontologies” involved in both juvenile and adult cognition.

Such studies are most directly parallel to inventory ontology in philosophy, as they are concerned with how the mind represents the world by dividing it into particular kinds, properties, etc. But the serious proponent of any form of critical ontology that cashes out the nature of object-hood in terms of the mind and its concepts ought to take note of the best contemporary research in cognitive psychology. And thus psychological theories about implicit ontology have – or at least ought to have – implications for critical ontology as well.

There are, of course, a number of ideas in cognitive psychology that could be explored here, but I shall concentrate on several distinct lines of research that share a common theme: that human understanding comes in the form of mental models of individual content domains, rather than smaller units such as conceptual atoms and individual propositions, or larger units, like the holistic Quinean web of meaning and belief. This is, I think, a theme that is under-explored by philosophers, whether their area of interest be metaphysics, epistemology, semantics or philosophy of mind. It has, however, been a growing interest of mine over about the past fifteen years. (Cf. Horst 2007, 2011.)

2.1 Core Knowledge Systems

One line of research that has attracted a great deal of attention recently is the Core Knowledge Systems (CKS) hypothesis. (For an excellent overview of this research, see Spelke and Kinzler 2007, which I shall quote from below.) According to CKS, infants already seem to think differently about “Agents” (things that behave without obvious outside impetus) and inanimate “Objects” (basically, contiguous solids that move only when pushed or pulled). They are also able to distinguish small exact ordinalities (1, 2 and 3) and to make inexact relative comparisons of larger groupings (say, a group of 20 dots versus a group of 5), and to possess some level of geometric understanding of the world around them. According to CKS, these abilities require the postulation of several distinct mental models – Core Knowledge Systems for Agents, Objects, (two for) Number, and Spatial Geometry.

Studies of human infants and non-human animals, focused on the ontogenetic and phylogenetic origins of knowledge, provide evidence for four core knowledge systems (Spelke, 2004). These systems serve to represent inanimate objects and their mechanical interactions, agents and their goal-directed actions, sets and their numerical relationships of ordering, addition and subtraction, and places in the spatial layout and their geometric relationships. Each system centers on a set of principles that serves to individuate the entities in its domain and to support inferences about the entities’ behavior. Each system, moreover, is characterized by a set of signature limits that allow investigators to identify the system across tasks, ages, species, and human cultures. (Spelke and Kinzler 2007, 89)

The CKSs for Agents and Objects each seem to have an implicit ontology (Agents and Objects, respectively, where I have capitalized the words to indicate their proprietary usage within the theory). And the child’s understanding of these two domains of things can be characterized by distinct sets of rules.

The core system of object representation centers on the spatio-temporal principles of cohesion (objects move as connected and bounded wholes), continuity (objects move on connected, unobstructed paths), and contact (objects do not interact at a distance) (Aguiar & Baillargeon, 1999; Leslie & Keeble, 1987; Spelke, 1990). These principles allow human infants as well as other animals to perceive object boundaries, to represent the complete shapes of objects that move partly or fully out of view, and to predict when objects will move and where they will come to rest. Some of these abilities are observed in the absence of any visual experience, in newborn human infants or newly hatched chicks (Valenza, Leo, Gava & Simion, in press; Regolin & Vallortigara, 1995; Lea, Slater & Ryan, 1996). [Spelke and Kinzler 2007, 89]

It is important to note that the notion of “Object” here is inextricably linked with the ways things are constituted *as* Objects. Philosophers, for example, are used to applying the word ‘object’ much more broadly – say, to abstract objects, to persons, to Cartesian souls, to sets. In this quote, however, the “Object System” is used in explicit contrast with systems oriented towards intentional agents or sets. The relevant notion of “Object” is further contrasted with things to which it does not apply, like liquids and heaps. (ibid, 90) In other words, the claim is more precisely that there is a psychological system that is applied to spatio-temporal things that are perceived as being cohesive and continuous. Indeed, even this is not quite right, as it suggests that *first* there is some *independent* test for properties like cohesion and continuity, and *then*, if these are met, the Object System is applied. But in fact the application of principles of cohesion and continuity are supposed to be themselves part of the Object System – indeed, one might venture to say that, if there is such a core system, *this is where our ideas of cohesion and continuity, as well as Object-hood come from.*

From early on, children treat people, animals and other stimuli that display signs of agency differently from inanimate Objects, liquids and things like piles of sand. The behaviors of the former are treated as goal-directed, and infants tend to

mirror them in their own behavior. Some such mirroring, such as mimicry of four stereotypical facial gestures, has been observed very early indeed, perhaps as early as hours after birth.

Spatio-temporal principles do not govern infants' representations of agents, who need not be cohesive (Vishton, Stulac & Calhoun, 1998), continuous in their paths of motion (Kuhlmeier, Bloom & Wynn, 2004), or subject to contact in their interactions with other agents (Spelke, Phillips & Woodward, 1995). Instead, the intentional actions of agents are directed to goals (Woodward, 1999), and agents achieve their goals through means that are efficient (Gergely & Csibra, 2003). Agents also interact contingently (Johnson, Booth & O'Hearn, 2001; Watson, 1972) and reciprocally (Meltzoff & Moore, 1977). Agents do not need to have perceptible faces (Johnson, Slaughter & Carey 1998; Gergely & Csibra, 2003). When they do, however, infants use their direction of gaze to interpret their social and non-social actions (Hood, Willen & Driver, 1998; Johnson *et al.*, 1998), even as newborns (Farroni, Massaccesi, Pividori & Johnson, 2004). In contrast, infants do not interpret the motions of inanimate objects as goal-directed (Woodward, 1998), and they do not attempt to mirror such actions (Meltzoff, 1995).

Interpreting Agents (again capitalized to denote a stipulative theoretical usage) differently from inanimate Objects is clearly a feature of adult cognition as well. Such an Agency-sensitive module could be a uniquely human development, in which case it would differ in a crucial respect from the other hypothesized core systems. However, a similar profile of Agent-directed cognition has been evidenced by studies of non-human animals.

Goal-directedness, efficiency, contingency, reciprocity, and gaze direction provide signatures of agent representations that allow for their study in non-human animals and in human adults. Newly hatched chicks, rhesus monkeys, and chimpanzees are sensitive to what their predators or competitors can and cannot see (Agrillo, Regolin & Vallortigara, 2004; Flombaum & Santos, 2005; Hare, Call & Tomasello, 2001). These studies accord well with the physiological signatures of 'mirror neurons', observed in captive monkeys, which selectively respond to specific actions performed by the self and others (see Rizzolatti, Fogassi & Gallese, 2002, for a review). Mirroring behavior and neural activity occurs in human adults as well (Iacoboni, Woods, Brass, Bekkering, Mazziotta & Rizzolatti, 1999), and representations of goal-directed action guide adults' intuitive moral reasoning (Cushman, Young & Hauser, in press). Together, these findings provide evidence for a core system of agent representation that is evolutionarily ancient and that persists over human development. (*ibid.*, 90)

As empirical claims, each of the hypotheses about particular Core Systems should be taken individually. However, the fact that they share some common

features is of both scientific and philosophical interest. (Indeed, with such issues it is hard to find a clear dividing line between philosophically-oriented science and scientifically-oriented philosophy.) There are a number of such features that are attributed to all four core systems:

- 1) *Species-typicality* found in both children and adults
- 2) *Nativism* – the systems appear too early to be products of learning
- 3) *Analogs* in closely-related species
- 4) *Domain-specificity* – the systems are applied to particular types of stimuli and not to other types
- 5) *Proprietary representational system* – each system represents its subject-matter in a particular fashion, affording particular types of information and expectations about it
- 6) *Characteristic breakdown patterns*
- 7) *Signature limits*

The first three features – species typicality, nativism, and analogs in other related species – are closely related. The purely empirical claims of species-typicality and the existence of animal analogs serve as evidence for the stronger claim of nativism. (The term ‘nativism’ is both ambiguous and controversial. I shall charitably read ‘innate’ as meaning something like “species-typical, developmentally canalized, and early-appearing”.) The earlier an ability appears, the less likely that it is a product of learning. Similarly, its presence in related species that are incapable of

distinctively human types of learning (such as linguistically-based instruction and acculturation) makes it more likely that the presence of the ability in the each species is a result of a common evolutionary heritage. In some cases, the argument for nativism can be additionally strengthened with a poverty of the stimulus argument.

The remaining four features – domain-specificity, proprietary representational systems, characteristic breakdown patterns and signature limits – are likewise closely linked. With individual innate systems, the notion of “domain-specificity” cannot be understood as being given independently of the system itself. It is not as though the mind *first* determines that something is an “Object” in the relevant sense, and *then* applies the Core Object System to it. Rather, we must suppose that the fact that we constitute a particular class of stimuli in this way is part and parcel with the system. Likewise, we must suppose that, if there is a single Core Agency System, it is the Agency System itself that is involved in constituting stimuli as Agents, rather than depending on some other system to pre-classify them. Indeed, what we *mean* by ‘Agent’ is in large measure determined by the features, such as goal-directed action, that are ascribed by the system. On the other hand, the opposite inference does not follow. A representational system can be domain-specific without being nativistic. (Analogously, a domain-specific system in a computer can equally well be implemented through hardware design or software.)

Domain-specificity is also closely related to the presence of a proprietary representational system. There is, of course, a more generic reason for this as well: any representational system must be some *particular sort* of representational

system. Representational systems are typified by their elements, relations and transformations. There is no such thing as a *generic* representational system.¹ But the Core Systems are, additionally, optimized for particular types of tasks. The sort of representations *needed* for Object-tracking in space are different from those needed for interpreting Agents.

“Optimized”, however, is not really the right word, as it suggests that the system indeed performs optimally. But this is not quite correct. An object-tracking system that affords tracking of two or three Objects is not *better* than one that affords tracking of five or ten, and one that is applied only to contiguous solids is not better than one that is also applied to liquids or piles of sand. The “optimization” here is really more like *satisficing*, or hitting upon a *good-enough* solution to an ecologically-defined task. This is, in general, the sort of thing we see in evolved systems, however, and is in no way peculiar to psychology.

The use of any particular representational system, moreover, inherently involves breakdowns in cases where that system does not represent its domain aptly, and cases where it faces performance limitations, such as limited representational space or computational power. We would expect such limitations to be more pronounced in nativistic systems because they are not flexible – not subject to on-the-fly revisions, as learned representational systems at least sometimes are.

¹ One might think of a natural language as an obvious exception to this claim. However, languages are not really “representational systems” in the sense that implies a *model* at all. Rather, they provide a flexible system for expressing ideas whose semantic basis is in models.

To these features explicitly mentioned by Spelke and Kinzler, I should add two more that are at least implicit in their account.

- 8) *Partial autonomy* – core systems operate largely independently of one another (are functionally autonomous, or nearly so)
- 9) *Pragmatic framing* – core systems are tied to a limited set of pragmatic goals and ways of interacting with the environment

Core systems cross-classify the things in their domains. The same stimulus – say, an organism – may be interpreted both as a contiguous Object and as an intentional Agent. But the connection is not necessary. Both adults and children can interpret non-contiguous stimuli as Agents, and indeed from an early age we can make sense of the idea of human selves and other entities continuing to exist after the death of their bodies. (It is because of this dissociability, argues Paul Bloom, that we are “natural-born dualists”. (Bloom, 2004)) Moreover, the conditions that trigger each core model are independent, and the content of the models allow for little in the way of cross-domain inference. But the autonomy is not complete. We can, as stated, identify the selfsame stimulus as both a contiguous Object and an Agent, and indeed we can be aware that we apply both categories to the selfsame object. And, on the basis of this, we can combine insights gained through multiple systems. We can, for example, see that a mouse has a bit of its tail snipped off and infer that it is in pain and wants to run away, or ascribe to the cat a perception of the mouse as a contiguous solid. The Core Systems may be autonomous, or largely so, in their

internal processing, but their “outputs” are available for combination, at least in minds like our own.

I should also stress that these core systems seem to be shaped, not merely *cognitively*, but also (and perhaps primarily) *pragmatically*. They are evolutionary solutions to ecological problems faced by organisms in dealing with their environment. They are essentially involved in tasks like *tracking* Objects (rather than simply representing them abstractly), *predicting the behavior* of Agents (rather than simply imputing agency to them in the abstract) and *orienting adaptively* in an environment (rather than simply representing it *as spatial*).

2.2 Folk Theories (aka Intuitive Theories)

A line of research distinct from, but closely related to, the CKS hypothesis is guided by the view that our commonsense understanding of people, animals, artifacts and the physical world can be seen as the product of a handful of “Folk Theories” that are arrived at through some combination of development, learning and culture. (Adam Morton has dubbed this view the “theory-theory”.) These are referred to by such names as “Folk Psychology”, “Folk Biology” (or perhaps separate Folk Theories for zoology and botany), and “Folk Physics”. Some psychologists who are advocates of this idea, such as Alison Gopnik, take the “theory” part of the description quite seriously, positing that the child arrives at these understandings of the world through processes of hypothesis-formation and –testing.

Folk Theories are also domain-specific, and there is some similarity between their domains and those of CKSs. But Folk Theories are generally understood to include more sophisticated understanding of the world than is posited for CKSs, and to be acquired over a developmental timeframe reaching into early childhood. One test for a mature Folk Psychology (or “Theory of Mind”), for example, is the *false belief attribution test* – that is, whether the child understands that other people do not necessarily have the same beliefs and desires as she does. This ability is thought to generally be in place around age four (Boyer, 2002). Likewise, Folk Biology might include an understanding that animals come in species, that they are born and die, have parents of the same species, possess species-typical physiognomies, diets and means of locomotion, etc.

One thing that I should very much like to see the developmentalists address and come to some consensus on is the relationship between CKSs and Folk Theories. It seems plausible that the former are used to bootstrap up to the latter, which are in turn used in the formulation of still more adequate models of aspects of the world, such as scientific theories. But there are interesting questions that are not, to the best of my knowledge, resolved, such as whether CKSs are replaced by Folk Theories of related domains, or continue to exist alongside them in the cognitive toolkit, and if the latter, whether Core and Folk systems of roughly the same domains ever have conflicting expectations.

Some theory-theorists have attempted to trace out the implicit ontologies and metaphysics of some of the Folk Theories. I shall use as an example Pascal

Boyer's discussion of the concepts of Folk Biology (or, more exactly, Folk Zoology).

Boyer makes repeated and explicit use of a psychological notion of ontology:

An ontology postulates a certain number of kinds of objects in the world as well as the fundamental properties that differentiate them from other kinds. Ordinary ontological categories denote such broad domains as events, physical objects, animate beings, abstract objects, persons, plants, animals, etc. Obviously, such categories are not just labels for classes of objects thrown together for no good reason. All ontological categories are based on (generally tacit) assumptions about the specific characteristics of the different classes that are distinguished. [Boyer CASBS]

Intuitive ontologies, moreover, play an important role in organizing and facilitating cognition:

Ontological assumptions are part and parcel of ordinary, everyday cognitive processes. Even the most trivial activities require a set of fast, automatic decisions about the identification objects in the environment, as well as the inferences made about them. Subjects may not, and in fact need not be aware of these ontological categories and principles. This is why their content and organisation can best be revealed by experimental studies, and by developmental studies in particular. These show how intuitive ontologies develop early, before subjects can formulate explicit, integrated conceptions of the types of objects that can be found in the environment. Moreover, such studies demonstrate the constraining function of intuitive ontologies. They make conceptual acquisition possible by restricting the range of inferences that can be made about given objects, by orienting subjects towards certain preferred inferences.

Like CKS theorists, Boyer holds that intuitive ontologies are domain-specific and underdetermined by experience:

Recent experimental studies of children's early intuitive ontologies have shown that children's inferences about particular aspects of the experienced world have two salient features. First, they are to a certain degree under-determined by experience, and appear much earlier than the child's actual interaction with the world would lead us to imagine. Second, they are domain-specific, involving structurally different principles and expectations according to the aspect of reality they focus on. Generating physical expectations from the experienced physical behaviour of objects is not done in the same way as generating biological expectations on the basis of observed biological features. Experimental studies allow us to describe a variety of domain-specific conceptual structures (for a general survey, see Hirschfeld and Gelman 1994). Here I will only mention three structures of particular importance, namely:

(1) an intuitive physical module, which produces explanations and predictions concerned with the physical properties of solid objects;

(2) an intuitive biological inference module, which specifies principles to do with the essential nature of certain biological properties, and their connection to membership of

particular, mutually exclusive and jointly exhaustive, classes of beings;

(3) an intuitive intentional explanation module, which infers speculative explanations of behaviour in terms of the existence and causal properties of hidden entities such as beliefs, intentions, perceptions.

About the biological inference module he writes:

From the earliest age, there is a categorical distinction between "animate" and "inanimate" objects. This distinction is then enriched with a variety of specific intuitive principles which describe biological aspects of live beings. There are several different kinds of principles at work in the child's "intuitive biology". First and foremost, there is a taxonomic principle, following which the space of all species is divided in a number of mutually exclusive, jointly exhaustive natural classes, arranged in a taxonomy (Atran 1991). Second, the identification of living kinds also activates an "essentialist principle" following which an undefined internal principle, which is exclusive to the species, causes the external features of living objects as well as certain aspects of their behaviour. For instance, even pre-schoolers assume that membership of a kind is more important than observable features, in predicting the typical behaviour of an unknown animal (Gelman 1986), (Massey & Gelman 1988; Becker & Ward 1991)³. These theoretical assumptions are also manifest in children's reactions to putative scenarios of artificial transformation from one kind to another. Such changes are judged more plausible between types of artefacts than from one living kind to another, even in cases where an animal is described as having gradually acquired the other's outside appearance or behaviour (Keil 1986). Third, this essentialist understanding of living things leads children to evaluate a number of properties as directly caused by "inner" mechanisms.

These "natural classes" have special highest-order kinds, like ANIMAL, which Boyer calls *Ontological Categories*. Ontological Categories like ANIMAL differ from ordinary sortal concepts like DOG or OSTRICH, not only in being (logically) more abstract superordinal concepts, but also (psychologically) as serving as a kind of *template* used in the coining of new concepts. All such concepts are, in Boyer's view, rich information structures, with "slots" for things like a species' mode of procreation, physiognomy, habitat, etc. The Ontological Category, however, leaves many of these slots blank. When one encounters a new type of animal, one creates a new concept essentially by copying the structure of the Ontological Category and filling in the blanks. This not only builds general knowledge of animality into all species

concepts, but also guides the concept-learner in what further information to seek about the species.

The Ontological Category also encodes knowledge (or at least assumptions, as we are not concerned with the accuracy of concepts or categories here) that does not need to be filled in on a case-by-case basis: for example, that animals are born and die, and that they have parents of the same species. Having a blank slot for species-particular information also encodes general assumptions, such as that all animals eat something and are able to move about of their own accord, even though different species each have their own characteristic diets and modes of locomotion.

This explanation provides at least a partial account of *intuitive understanding*. (In calling a judgment “intuitive”, I mean merely that it was not reached by explicit reasoning using propositions. In particular, calling a judgment “intuitive” is not meant to imply that it is, or is the product of, something innate.) We have intuitive understanding that animals eat and move, not by dint of some propositional reasoning, but because this can, as it were, be *read off* the structure of the Ontological Category. By contrast, *counterintuitive* concepts and beliefs arise when a value is written into a slot that is contrary to the (default) rules of the Category: say, an animal that is immortal or a bird whose father was an elephant.

One way of looking at Boyer’s account would be to say that Folk Zoology includes, and perhaps is centered around, a mental model of animality that is encoded into the data structure of the Ontological Category ANIMAL and (hence) into all of its subordinate species concepts. It is a model of a *specific domain*. This does not mean that none of the concepts involved in Folk Zoology appear in other

mental models – for example, specifying an animal’s habitat invokes some understanding of physical environment, and physiognomy is partially geometric in character. But the learning and use of the model are largely independent of other models (say those of Folk Psychology or Folk Physics), and adjustments to Folk Zoology (say, when one learns that there are sessile animals that do not move about, or that there are hitherto-unanticipated forms of reproduction such as cell division) need not ramify to cause changes in the others, or vice-versa.

2.3 Semantic Networks and Frame Theory

Both the CKS hypothesis and the theory-theory postulate a small number of knowledge domains that are more or less species-typical and which appear before (sometimes long before) adulthood. But the idea of domain-specific models with proprietary data types, inference rules, and representational systems has also been applied in a much broader fashion by Knowledge Representation (KR) theorists in artificial intelligence (AI), beginning in the 1970s with the so-called “second generation” of AI research. Whereas the first generation of AI research attempted to model deductive inferences in sentential or predicate calculus form, this research hit a roadblock when it tried to account for semantically-based and situation-specific understanding, such as that dogs have teeth or that dessert is served after the entrée. Second-generation AI researchers thus set out to try to determine how the mind might encode semantic relations and situational understanding.

One idea that emerged out of this research was the *semantic network*. The basic idea here is that semantic relationships between concepts can be understood by thinking of concepts as nodes in a network, and the semantic relations between them as links. The idea of a “web” of belief and meaning had also been suggested by philosophers like Quine and Davidson. However, Quine and Davidson held that each person’s concepts and beliefs make up a single and comprehensive network, so that changes in any one part of the network amount to changes in *every* concept. The notion of a semantic network in AI, however, leaves open the question of how “local” or “global” such networks might be. And, intuitively, this seems like a good thing. Changes in one concept in a scientific theory may require adjustments of the other concepts employed in that theory. But it seems unlikely that, whenever a new theory replaces an old one, we all need to rethink our concept GRANDMOTHER or our understanding of chess as well.

A second idea born in second-generation AI is what Marvin Minsky calls a *frame*. Minsky characterizes frames in the following fashion:

A frame is a data-structure for representing a stereotyped situation, like being in a certain kind of living room, or going to a child's birthday party. Attached to each frame are several kinds of information. Some of this information is about how to use the frame. Some is about what one can expect to happen next. Some is about what to do if these expectations are not confirmed.

We can think of a frame as a network of nodes and relations. The “top levels” of a frame are fixed, and represent things that are always true about the supposed situation. The lower levels have many *terminals*–“slots” that must be filled by specific instances or data. Each terminal can specify conditions its assignments must meet. (The assignments themselves are usually smaller “sub-frames.”) Simple conditions are specified by markers that might require a terminal assignment to be a person, an object of sufficient value, or a

pointer to a sub-frame of a certain type. More complex conditions can specify relations among the things assigned to several terminals.

Collections of related frames are linked together into *frame-systems*. The effects of important actions are mirrored by transformations between the frames of a system. These are used to make certain kinds of calculations economical, to represent changes of emphasis and attention, and to account for the effectiveness of "imagery."

The most straightforward example of a frame might be one's understanding of a well-defined game such as chess. To understand chess, one must be able to represent a certain stock of piece-types, the board as a system of possible positions for pieces, the initial state, the movement types for pieces, the capture rules, the goal or end state, and a few special rules such as those for castling and queening. These component parts of this understanding are not independent, but tightly interwoven: changing the movement rules for knights so that they move like bishops could equally well be described as inventing a different game in which there were no knights, but twice as many bishops. On the other hand, the rules and representational types allow for an indefinite number of different sequences of play and heuristic strategies, and because the typology is not directly linked to physical type, the game can be played with an indefinite variety of "pieces", including the odd sorts of "pieces" that must be spoken of in the cases of computer chess and mental chess.

Understanding of chess involves understanding of a domain with a proprietary ontology (the piece types), states (positions), relations (threat, capture, checkmate), and dynamic possibilities (the rules of movement). The "kinds" involved are not "natural" kinds – they are products of invention and "exist" only

insofar as there are people who know and observe particular conventions. And the typology is largely orthogonal to other ways of carving up the world, such as physical ways. Virtually any physical thing (even some meriological monstrosity) can, in principle, count as a chess piece, and in mental chess, there are no physical pieces at all.

The typology of other frames can have stronger connections to other domains. Take Roger Schank's (1975) example of the understanding we need in order to understand dining at a restaurant. This includes understanding particular kinds of agent (customers, waiter, chef, perhaps hostess and cashier), types of objects (table, menu, meal courses, utensils, check), and operations (seating the customer, bringing the menu, ordering, paying the bill). But there are also strong default assumptions that, say, the agents will be human beings, the meal courses things that people eat, etc. (Though the idea of an alien or ghostly restaurant with different particulars is far from unintelligible, and a comedy skit could be built around a restaurant whose waitstaff bring things that people are incapable of eating.) The kinds, states, relations and activities that are assumed by the restaurant frame are thus somewhat abstract (at least some are *functional* in nature), and some of them are proprietary to the frame. A waiter brings food, but not everyone who brings food is a waiter, as you will quickly discover if you treat your spouse or mother as one at the dinner table.

Clearly, the number of frames an adult human mind can deal with is very large, and most such frames are neither species-typical nor developmentally canalized. And if one counts "domains" frame by frame, there are a surprisingly

large number of these as well. Each frame involves its own implicit ontology, states, relations, and changes. In some cases, these are closely related to those of other frames; but in other cases they are not. Likewise, some have close relations with CKSs and Folk Theories, while others do not: A waiter is an Agent, a thinking being, an animal, and a physical particular. A chess piece can be any of these, or none at all. Of course, a particular thing may fall into the classifications of any number of frames, each of which licenses distinct inferences and expectations; and sometimes these may lead to conflicting conclusions.

2.4 Scientific Theories and Mental Models

Philosophers will no doubt see a connection here with the idea that scientific theories are tightly-integrated units, in which the meanings of the individual concepts are inseparable from their relations to one another and hence to the theory as a whole. The notion of 'mass', for example, changes with changes in gravitational theory, but other physical notions not directly involved in the theory, such as 'charge', may remain untouched. Some psychologists have indeed taken the view that the child's acquisition of understanding of a domain should be viewed on the model of hypothesis-testing and theory acquisition in science. There are, I think, important continuities, but I prefer to view scientific theories as a special case of a more ubiquitous phenomenon also found in Core Knowledge, Folk Theories and frames, which I shall call *mental modeling*.

3. Cognitive Pluralism as Psychology

The idea that understanding comes in the form of multiple domain-specific models is importantly different from the prevailing philosophical accounts of semantics and epistemology. The fundamental semantic units (mental models of domains) are neither atomistic nor holistic, but somewhere in-between. And the warrant for a model is not foundationalist, nor is it globally coherentist, though coherence plays an important role within models of individual domains. We need a general name for the thesis that our cognitive architecture achieves understanding through a proliferation of domain-specific models. Over the past decade, I have begun to call this view *Cognitive Pluralism*.

I shall shortly suggest that Cognitive Pluralism can serve as a thesis in epistemology and metaphysics as well, but it is, first and foremost, a psychological claim about cognitive architecture. Its major theses are as follows:

1. The mind achieves understanding through models of specific content domains.
2. Each model employs a particular representational system that has its own stock of kind-concepts for things, properties, relations, etc., and implicit inference rules that give rise to intuitive judgments about the model's domain.
3. The representational system of each model is well-suited (though not necessarily perfectly suited) to representing its domain well enough for particular tasks of understanding, interaction and prediction.
4. Models can be idealized in at least three ways: (a) they *bracket* or abstract away from other features of the world (e.g., a model of gravity says nothing

about charge), (b) they may *distort* their subject-matter (e.g., by treating bodies as point-masses or collisions as ideally elastic), and (c) quantitative models that can be used in calculation often involve finite approximations of the values of constants.

5. The (internal) semantic values of concepts are at least partially fixed by their embeddedness in mental models; and because the truth values of propositions cannot be determined until the concepts are fixed, the truth or falsity of a proposition cannot be determined until we have pinned down what model is in play. We thus need a way of talking about the fidelity of models themselves that is distinct from truth/falsity. I shall call this feature *aptness*.

One might be tempted to see Cognitive Pluralism, thus described, as a variant upon the Massive Modularity Hypothesis. There are certainly similarities, in that both postulate a large number of special-purpose models of different domains. But beyond this, the views are really quite different. ‘Modularity’ is, of course, a word that is used in many different ways. Indeed, I am increasingly inclined to avoid it altogether. But the Massive Modularists assume certain features in common with Fodor’s (1983) influential characterization, such as nativism (another word I have come to avoid whenever possible), automaticity, and standard neural localization, to which they add the criterion of being a product of natural selection. It should be clear that most Minskian frames, not to mention scientific theories, do not meet these conditions. It may be that CKSs do, in which case modules are either a subset of the Cognitive Pluralist’s mental models, or (if some modules are not models) an intersecting class.

Cognitive Pluralism is thus after something that is orthogonal to many debates about modularity. It is neither about the etiology of understanding, nor is it about its neural basis. Its claims are intended to apply equally to developmentally

canalized species-typical modules that are products of selection, and to scientific theories and games like chess that are learned only by a few individuals in particular cultural contexts. The proliferation of special-purpose models may well be a reasonable strategy for the evolution of minds, but it seems also to be a more general design strategy that has been retained and transposed into a different timeframe for acquisition of learned understanding. Another way of seeing the difference would be to say that Cognitive Pluralism sees the boundaries between models in terms of epistemic and semantic relations rather than biological ones. It is a thesis about the architecture of *cognition* – how a mind like ours is composed – but not about the architecture of the biological systems that underlie it.

4. Cognitive Pluralism as Semantics and Epistemology: Disunities of Understanding

One of the reasons I became interested in this pluralistic view of cognitive architecture is that it can explain disunities of understanding. Often, our understandings of two domains either fail to intersect at all or, worse, generate conflicting predictions and paradoxes. Relativistic Gravitation and Quantum Mechanics, for example, are widely viewed as formally incompatible, and the same is true in less striking ways for other pairs of models employed in the sciences. And there are philosophical paradoxes of long standing involving matters such as taking ourselves to have both a physical and a mental nature, or in combining a law-governed view of nature with the assumption of free will that seems necessary if ethical categories are to be intelligible.

In addressing such disunities, there are two principal temptations: to chalk them up to present-day misunderstandings that will eventually be overcome, or to take them as indicating that at least one of the models we employ is radically mistaken – say, that there is no *real* quantum indeterminism (only hidden variables), or that free will and hence ethical responsibility are illusions. Cognitive Pluralism opens up a third type of explanation: that at least some such disunities are *artifacts of the mismatches between the representational systems of the models involved*.

Every model must employ some *particular* representational system, and *which* one it employs is constrained on the one hand by the pragmatic interests in how the model is used successfully, and on the other hand by deep features of cognitive architecture that set limits on which types of representational systems human minds are capable of employing. Models are shaped by their individual utility, and not by a presumption of formal unity. But two representational systems can be incompatible with one another – Euclidean and Lorenzian geometries, for example, have incompatible axioms. And the representational system that a human mind finds most apt for understanding of, and interaction with, one domain may similarly turn out to be formally incompatible with that needed for another. Thus, disunities of understanding may often turn out to be predictable consequences of Cognitive Pluralist assumptions about cognitive architecture.

This may initially strike you as a variation on the view that disunities are a consequence of present-day misunderstandings, and that all we need to do is retool our models in forms that are mutually consistent. But note that the idea that this can in fact be done involves some heavy-duty tacit empirical assumptions about the

human mind. And, on reflection, these assumptions are at best not obviously true, and indeed seem somewhat implausible when scrutinized. First, the models that are developmentally canalized probably cannot be revised or erased. Second, we do not know what limits the deep architecture of the mind may impose upon *just what* ways of representing things are really available to us. And hence, third, it is a real empirical possibility that there may be pairs of models such that we cannot attain to a single super-model that is self-consistent without it losing some of the explanatory insight or practical grip on the world afforded by two incommensurate models employed in tandem.

This possibility – of a *principled and abiding epistemic and semantic pluralism* – is one that I think we should take seriously. I do not mean that we should give up trying to integrate our various models of the world. I mean, rather, that we should not *assume*, as either a metaphysical or an epistemological axiom, that our understanding of the world can be integrated into a single, comprehensive, self-consistent system. Indeed, from an evolutionary standpoint, the ability to do so – to have minds optimized for unified theoretical understanding – would seem to be a wonderful accident. Moreover, a flexible pluralistic architecture has real advantages. A being with a single way of representing things would be held captive to whatever forced errors were implicit in its representational system, while a mind with multiple models and flexibility can use one model to test and compensate for another, and also use in tandem to triangulate a reality that cannot be captured from a single representational perspective.

While I think that a principled and abiding epistemic and semantic pluralism is a reasonable empirical thesis, I should emphasize that it is not entailed by *de facto* psychological pluralism. That is, one can accept Cognitive Pluralism's psychological thesis as a description of the default state of human understanding as we in fact find it without drawing the pessimistic conclusions about the prospects of unifying knowledge.

5. Cognitive Pluralist Metaphysics

Similarly, one can accept the psychological and epistemological claims without taking the leap into cognitivist critical metaphysics. Indeed, some realists might actually take a good deal of solace in the thought that the apparent disunities of understanding (and particularly scientific disunities) may be artifacts of our cognitive architecture, and hence need not imply that the world itself is in some way incoherent or nonsensical. In the spirit of this conference, I shall lay aside that dispute here and instead explore what happens if one *does* combine a cognitivist metaphysics with a psychological and epistemological pluralism. This question is, to the best of my knowledge, largely unexplored within analytic philosophy, and it gives rise to several puzzles and surprises.

5.1 Multiple Ontologies

The first puzzle is this. The cognitive metaphysician holds that the categories of metaphysical ontology are a kind of projection of the intuitive ontology employed by the human mind. But the psychological pluralist holds that there is not just *one* such intuitive ontology, but several, perhaps a great many, and that these do not always play well together. Each model has implicit constitutive rules for the sortalkinds it employs, and these cross-classify individual things. Sometimes two models have little or no connection with one another except for the fact that they can be applied to the selfsame object. But sometimes they generate contradictory expectations about how a thing will, should, or even *must*, behave.

There are some famous philosophical examples that illustrate the point. Kant had to retreat to the distinction between the phenomenal and noumenal worlds to reconcile his deterministic interpretation of theoretical reason with his view that we must posit freedom to make sense of moral discourse.² Aristotle objected to Plato's hypostatization of the Forms with the Third Man argument, which reveals a potential problem for an ontology that includes both abstract and concrete objects. And every philosophy of mind textbook begins with a rehearsal of the difficulties that result from positing separate realms of *res extensa* and *res cogitans*. But similar problems have emerged from pairs of scientific models, such as wave and particle models in classical optics or Relativity and Quantum Mechanics in contemporary physics. What we would *like* to do in such cases is to find a technical solution that

² I argue in *Laws, Mind, and Free Will* that a proper understanding of scientific laws in no way implies a commitment to determinism.

shows that the apparent problem was due to ignorance or misunderstanding. But if we take seriously the proposal that it may be due to mismatches between the best models available to the mind for different aspects of experience, it may be that the reason for the longevity and the beguiling character of such problems lies deeper, and that such solutions are beyond our reach. It may be, for example, that there are aspects of our deeply-rooted conceptions of physical objects, agency, and number, that we cannot shake off, and which lead to paradoxes when applied to particular problems.

Moreover, while these elevated problems are the ones that have vexed philosophers and scientists, similar problems may occur with pairs of much more mundane Minskian frames. Normativity is not confined to ethics – we find it in things like the game of chess as well – and if, say, we play chess with pieces that are large concrete statues or human actors, there may indeed be problems analogous to those of whether ought implies can or those involving agency and freedom. (The rules say I must move out of check, but the king is too heavy for me to lift. I try to advance the queen, but the actress playing the queen will not go where I tell her to go.) Likewise, when two normative social models require contrary actions – is Antigone to bury her brothers as a dutiful sister or obey the *archon* as a dutiful citizen? – we find that there is not one “ethical realm” but many, and when these worlds collide, we have the makings of classical tragedy.

So the fundamental problem facing a pluralist critical metaphysics is that it does not seem to be able to provide a single ontology that is broad enough to encompass everything. There is no global ontology, but only a patchwork of

intersecting local or regional ontologies. We seem, in short, to be at risk of losing the World in favor of either many worlds or, better, many realms that fall short of the kind of comprehensiveness required by the notion of a 'world'. Even for those of us comfortable with the epistemological prospect that we cannot attain anything as comprehensive and consistent as a world-view, the idea that there may, in some sense, be no such thing as "the World" seems deeply puzzling and at least vaguely threatening.

5.2 What is "The World"?

The most obvious big issues to be faced concern the metaphysical implications of the idea that our understanding comes neither in atomistic concepts nor in a holistic world-view, but in domain-sized chunks. The epistemological implications of this may indeed seem radical enough in their own right: if there are abiding incommensurabilities between models that we cannot do as well without, it may be that no one has ever had (or ever will have) such a thing as a world-view, at least if a world-view is something that is supposed to be both comprehensive and self-consistent. But if we try to move from epistemology to metaphysics, the very notion of "the world" seems called into question – at least, again, if by a world we mean something comprehensive and self-consistent.

Perhaps it is easiest to approach this subject by clarifying first some things that I do *not* mean. First, I am not talking about the idea that different people and different cultures have "different world-views". The issues I am interested in are not about whether the ways your beliefs and concepts are different from mine, or

even from those of other species, but in how, among my own modes of understanding (and likewise among yours) there may be mental models that are either totally independent of one another or, worse still, are in deep conflict with one another.

Second, we sometimes speak loosely of “the natural world”, “the moral world”, “the business world”, and so on. If one were to take such talk naively, it might suggest a commitment to many “worlds” that we each inhabit. But in fact common sense is not naïve in this way: the papers, copy machines, and so on that one deals with in “the business world” are also regarded as objects that exist in nature. The metaphysical problem that arises for the Cognitive Pluralist is rather something like this: a “world” or *cosmos* has generally not been understood by philosophers and scientists simply as the aggregate of all that there is (a very large set of individual things), but in some kind of *system* with an overarching order that unites it. (“World”, after all, is a translation of *cosmos*, which means “order”.) But if the order that we find comes in domain-sized units, whose inter-relations are seldom better than partial, sometimes conflict, and sometimes are non-existent, we need to find a way of thinking of “the world” that goes beyond domain boundaries *without* assuming that there must be a super theory of everything that would unite all that we understanding into a single, comprehensive, self-consistent *model*. And for purposes of *my* project, it would be useful to have a cognitivist approach to this question as well – which is to say, some sort of account of how our cognitive abilities are not strictly limited to conceiving of propositions and making inferences *within* the confines of a single domain.

One familiar proposal that some might find promising is the idea that *language* is a domain-general representational system. This idea is quite familiar in discussions of mental modularity, where it tends to go by the name *central cognition*. Some advocates of such views see acquisition of *natural* language as the key step, while others posit a Language of Thought that already has many or all of the necessary features. While I am somewhat skeptical of the LOT hypothesis, it has an important advantage over the natural language hypothesis, in that it seems clear that non-linguistic animals and pre-linguistic children are *already* able to coordinate understanding that is derived from different sensory modules (a cat can identify the mouse she sees as what she first heard scurrying about) and from different mental models (e.g., different Core Knowledge Systems). But of course this may not suffice to shed light on the metaphysical problem of “the World”, as babies and non-human animals may have no such concept.

Language *does* seem to provide a medium that allows for cross-domain *inferences*, and for other things such as taking problems from one domain and trying to apply the resources of another towards their solution. And this is indeed a huge cognitive advance over any sorts of minds that cannot do so. And yet language itself does not really add any systematic understanding, except perhaps insofar as it canalizes our emerging understanding of the world into the grammatical categories of languages. And it is not clear that this clarifies more than it obscures. Much of our non-linguistic experience – say, our experiences of sounds and our understanding of kinesthetic processes – translates very poorly into a subject-verb structure. And philosophers are all too familiar with pseudo-problems that can

arise from the ease with which we do such things as nominalize verbs or adjectives, and thereby risk the illusion of thinking of them as *things* in their own right.

The approach to this problem that I find most promising is derived from Husserl's distinction between immanent and transcendent aspects of cognition. Incompatibilities between mental models take place within the immanent aspects of cognition – with how aspects of the world are modeled domain by domain. (This, I think, corresponds more or less to what analytic philosophers call “internalist” components of semantic content.) And if this were the whole story about human understanding, there might indeed be a troubling problem for metaphysics. But there would seem to be important ways in which model-based understanding is *not* the whole story. For one thing, even when models are incommensurable, we are perfectly capable of thinking of thoughts cast in the terms of two models as being about the selfsame thing – persons are both Objects and Agents, light is both wave and particle. The ostensive transcendent component of cognition that allows us to track objects and properties even through changes in our beliefs and concepts (corresponding roughly to “externalist” components of content) also allows us to triangulate them through multiple models, even when these models cannot be neatly combined. At the level of individual objects and kinds of phenomena, this is perhaps an alternative account of what Kant was after in speaking of the noumena. But what about “the world”?

Kant, of course, spoke of “the world” as a dialectical illusion. But his reasons for saying this were epistemological: the world as a comprehensive nexus of cause and effect cannot ever be *experienced in an empirical intuition*. Cognitive Pluralists

need not accept this Kantian conception of knowledge, and are free to embrace alternative standards of warrant upon which we can, for example, have knowledge of unobservable theoretical entities, normative claims, and even supernatural claims. Nor need we accept the Kantian distinction between phenomena and noumena. (Indeed, my suggestion in the previous paragraph suggests an alternative way of getting at the idea that objects transcend our immanent or phenomenal ways of conceiving of them.)

It is, I think, less daunting to try to think first about the notion of “object” or “thing”, in the sense of a concrete individual. When I think about a particular thing – about Socrates, or about my living room sofa – I may think about it in many ways. I think of Socrates as a human being, a philosopher, a martyr, an Athenian, the founder of a school, the teacher of Aristotle and Speusippus, etc. Each of these ways of conceiving him is tied to its own characteristic types of inferences. In Minsky’s terms, each is embedded in one or more frames. Together, assuming everything is working right, these various ways of thinking *triangulate* a concrete individual. And when I think of Socrates under any of these aspects, there is an aspect of the thought that aims at tracking a concrete individual in a fashion independent of whether the concepts I am wont to apply to him are correctly applied.

There are, of course, several ways in which such cognitive triangulations might need to be readjusted. I might at one time have been unaware that, say, my representation of the Founder of the Academy in fact tracked the same person as my representation of the teacher of Aristotle. And I might have mistakenly believed that Socrates wrote comedies under the penname of “Aristophanes”. I think that the

most likely account of how this comes about is that there is some component of thought that is, as it were, *name-like* rather than *predicate-like*. That is, something like Husserl's "pure x" whose function is purely to track an individual. This is a crucial component of thought because it allows us to keep (putative) reference pinned down while sense is adjusted. In some extreme cases, there may be no concrete sense assigned at all, as when, in the dark of night, one suddenly is struck with the thought "There is something out there." It is also crucial to our ability to juxtapose our different frame-based understandings of a single thing. Indeed, without such a capacity, it is not how we could think of the targets of thoughts with different contents as being the selfsame individual.

So here we have a view on which there are (at least) two complementary semantic and epistemic relations to objects of cognition. One corresponds roughly to the idea of *sense* and is encoded in concepts and expressed in predicates. (Though with the twist that concepts are not atomistic, definitional, or holistic, but embedded in models, which are the fundamental units of sense.) We might thus expect the other to correspond to *reference*. But (at least as this idea has been developed in extensional semantics) that would not be exactly correct. It would be better to say that this element – Husserl's "pure x" – is more like an *ostensive* or perhaps *name-like* element, whose function is to track particulars through changes in assumptions about what concepts are rightly applied to them. This allows two degrees of freedom. On the one hand, I can hold the sense fixed: I can, say, hold constant my understanding of CHIMP through changing beliefs about which individuals rightly fall under that concept (Bonzo, Noam Chimpsky, but not King

Kong). On the other hand, I can continue to track an individual through changing beliefs about how to categorize it: I first mistake Kong for a chimp, and then change my mind and decide that he is a gorilla.

I should point out that, without these complementary aspects of cognition, we would be semantically and epistemically impoverished. If understanding really worked like predicate calculus, we would have no way of tracking an individual through changing assumptions about what predicates apply to it. And if we had only name-like ways of thinking about things, we would be unable to reason about them.

There is, however, a twist to this analysis. For the same kind of duality or complementarity is found at the level of the relationships between concepts and properties. Concepts (or at least some concepts, such as those of so-called “natural kinds”) have both a sense-like component and a name-like component as well. That is, if we take concepts as *psychological continuants*, so that we may talk of conceptual *change* as changes *to a concept* over time, we must suppose that concepts have (at least) two components. Consider a paradigm case of conceptual change: the transition from the pre-Aristotelian assumption that whales are fish to the recognition that they are mammals. From a psychological standpoint, we want to be able to say that a single thing – a single psychological continuant – is somehow changed or edited. In this case, the sense and inferential patterns associated with it must change while something remains anchored so it is still the same concept. What remains anchored? It is a name-like or ostensive element that aims at tracking a sortal kind.

The claims I am making here are, first and foremost, *psychological* (or, in Husserl's sense, though not Nagel's, phenomenological). They are analyses of what our minds are doing as we track objects and properties through changing assumptions about them. But if we are interested in cognitivist *metaphysics*, what might be the metaphysical payoff? With respect to individual things, Husserl's observation, while not put forward as metaphysics, is reminiscent of Aristotle's claim that particular things are always a "this-such" (*tode ti*) – that is, they have an element of the sort that can be expressed by a predicate (the "suchness") and an element that is more ostensive in character. Performing a cognitivist inversion, we might say that the *notion* of "thing" or "object" (in the generic, not the CKS sense of the word) results from the intersection of two distinct semantic resources: one which grasps objects through the application of concepts, and another which tracks them as individuals even through changing assumptions about what concepts should be applied to them.

But what about the notion of "the World"? From the perspective of a thinker at any point in time, the range of possible variations in combinations of the predicate-like and name-like features through which we understand things present themselves as an inchoate and indefinitely vast possibility space. (This may, I think, what Heidegger called "the background".) We intuitively understand that there are processes through which the triangulation of reality can be refined, and in so doing, to fix our understandings of particulars and properties with increasing fidelity. We then project that there might be an ideal conclusion of this process, in which our concepts of kinds, properties, and individuals (if indeed our cognitive relations to

individuals are rightly called “conceptual”) come to a stable equilibrium. This anticipated equilibrated understanding, when projected onto the objects of cognition, we call “the World”.

At this point, philosophers will be reminded of Kant’s discussion of “the World” as a posited Idea of Pure Reason, and his claim that it is a dialectical illusion. For Kant, the issue was that the world as a comprehensive system of cause and effect could never be a possible object of an empirical intuition. For the Cognitive Pluralist, the concerns are different. They are concerns about whether such an “equilibrium” (a) is attainable, given the empirical facts (whatever they may be) about our cognitive architecture, and (b) whether “equilibrium” requires a single, comprehensive, and self-consistent system. If it does require this, and the pessimistic epistemological conclusion (that we cannot reach a single self-consistent system that has all of the explanatory power of multiple but inconsistent models) holds good, then the Idea of “the World” would indeed seem to prove a kind of dialectical illusion. If it does not require this, and the equilibrium is attainable – well, I am not sure how to think about this, to be quite honest! But is there another alternative?

I think that the alternative may be to treat “the world” as a kind of *normative* or *ideal* notion, which might be expressed in a kind of epistemological maxim: “Treat your beliefs, concepts, and reference to individuals as though they could be integrated, if not by you then by a mind that thought in the same terms but without limits of performance, into a comprehensive and self-consistent system.” In pragmatic terms, this normative notion of “the world” would undergird our efforts

to reach ever more exacting and comprehensive understanding. But, because it is a normative notion, it does not imply that its ideal conclusion can actually be met. This solution is, of course, quite in the Kantian tradition, as Kant viewed the posits of theoretical and practical reasoning quite charitably, even if he did not think them productive of knowledge. His treatment of “the world” is perhaps an important exception to this, as he regarded it as illusory in ways that the Ideas of God, the Soul and Freedom were not. But if my treatment gives “the world” a better status...well...I am happy to take credit for “saving the world”!

5.3 Problems for Necessitarian Metaphysics

Cognitive Pluralism may raise problems for necessitarian metaphysics – or, more specifically, for a modal realist version of possible world semantics (PWS) generally, and for an integration of PWS with scientific theories in particular.

Here is a familiar, middle-of-the-road way of developing the notion of possible worlds. Take the set of all propositions. To each of these there corresponds a state of affairs--an abstract entity corresponding to a way the world might be, locally. Now define a function mapping the set of propositions onto truth values. Each such mapping picks out a global state of affairs, which we may designate a “world”. (Again understood abstractly as a way things might be, or might have been.) Not all assignments of truth-values are consistent. If we assign TRUE to “there is at least one dog” and FALSE to “there is at least one animal”, we have a contradiction. Worlds corresponding to such an assignment of values have

incompatible states of affairs, and are not really possible. Not even God could (*pace* Cartesian voluntarism) bring such a world into being. So let us designate worlds as *possible* worlds just in case they consist in jointly compatible states of affairs, and correspond to a consistent assignment of truth values to the set of propositions. A proposition P is necessarily true if it is true at all possible worlds, necessarily false if it is false at all possible worlds, and possibly true if it is true at one or more worlds. A state of affairs S is necessary if it obtains in all possible worlds, possible if it obtains in one or more possible worlds, and impossible if there are no worlds in which it obtains. Particular modal logics (such as T, S4 and S5) are distinguished by axioms regarding necessity, possibility and counterfactuals.

Can a cognitive pluralist embrace modal metaphysics, thus cashed out? It is a tricky question. At one level, the cognitive pluralist is bound to treat modal metaphysics and possible worlds semantics like any other models. If there are problems to which they are aptly suited, more power to them. But at another level, PWS is problematic if it is taken as revealing the deep, fundamental and mind-independent structure of metaphysical reality, and as a canonical tool for revealing metaphysical truths. It is problematic both from the cognitivist perspective and from the pluralist perspective.

Cognitivist suspicions are aroused by the too-easy use of expressions like “the set of all propositions” or “the set of all states of affairs”. This smacks of the very type of realism that the cognitivist, *qua* cognitivist, rejects. If this machinery requires us to assume that the world divides itself in a canonical and mind-independent way into states of affairs, the cognitivist ought to reject this. But

perhaps this problem is not insuperable. Perhaps the talk of “all propositions” can be cashed out in terms acceptable to the cognitivist: say, as “the contents of all possible judgments (for minds like ours)”. And likewise “all states of affairs” as “the states of affairs corresponding to all possible judgments (for minds like ours).” This would render PWS acceptable to the cognitivist. However, it is not clear that it gives the modal metaphysician all that she is looking for. If we restrict the space of propositions and states of affairs to things human minds are capable of conceiving, we are in danger of leaving out things that our minds are *not* suited to conceiving. And unless one takes the (unlikely) view that there is nothing that is beyond our ken, this falls short of what one needs for genuine metaphysical necessity and possibility. (One cannot help but hear Kant’s ghost complaining of attempts to speak of propositions or states of affairs outside of the realm of possible experience, and the conflation of phenomena with noumena, even if advocates of PWS do not take themselves to be talking about noumena, or even reject that Kantian machinery altogether.)

Qua *pluralist*, the cognitive pluralist finds PWS problematic on other grounds. For if it should indeed turn out that we cannot understand the world through a single self-consistent system that retains all of the explanatory value of our current mishmash of more local models, then it would seem that any enterprise predicated on the assumption that the ways the world might be are captured by maximally large consistent sets of propositions is bound to fail. Indeed, it could turn out that the *actual* world is not “possible” in this proprietary sense. We might be able to find “worlds” that meet this description by taking away some of the properties that come

into conflict in our best models – say, a world with relativistic gravitation, but no quantum strong/weak/electromagnetic force. But all of the worlds that are sufficiently like our own to be interesting for counterfactual purposes would likely be ruled out, and so necessitarian metaphysics would be deprived of most of its traditional applications.

5.4 Metaphysical Intuitions

Finally, the psychological and epistemological sides of Cognitive Pluralism would seem to give us reason to question the fidelity of our metaphysical intuitions. First, intuitive judgments are those that can be “read off” the implicit rules of a model. Their psychological “obviousness” is arguably simply an artifact of this, and no sign of any deep necessities lying in the world, as is illustrated when two sets of intuitions come into conflict in a single person’s mind. The persistence of many paradoxes over millennia of our intellectual history suggests at very least that the psychological force of an intuition is no guarantee of its fidelity, as often these involve equally strong, yet conflicting, intuitions.

Whether the intuitions associated with a model are dependable for a given purpose depend in turn on the aptness of the model itself for that purpose. We may well make a default assumption that most of our models are *good enough*, judged on pragmatic criteria that do not stray too far from the context in which the model was produced. But consider what kinds of assumptions we would have to make about the origins of human minds to warrant the assumption that models developed by

evolution or learning for mundane purposes would be exacting enough to prove dependable for questions of metaphysical necessity and possibility. One would, I think, need something on the order of Descartes' assumption that a benevolent deity created us to be theorists capable of avoiding all metaphysical error and grasping the real natures of created things. And even many theists (myself included) have not seen theism as implying Rationalism.

Now one might think that this should all lead me into the school of philosophers who say we should give up analytic metaphysics (and perhaps all intuition-based inference) altogether. But actually, this is not where I end up. For it is not as though we could step outside of intuitive reasoning and understand the world in some other way. Any model we adopt produces its own intuitive inferences. Moreover, the fact that a model has persisted over generations suggests that it must at least be *pretty good*, and not something we should cast aside unless we have something to put in its place that does all of the same work, only better. What I think is called for are two levels of caution. First, when we have models that are apt in their own contexts, yet conflict, and we cannot definitively show one better than the other in the conflict cases, we should be wary of intuitions arising from either model, at least in the known conflict cases. Second, even when we do not have conflicting models, our attitude of assent to the intuitions produced by those we do have should be hedged by a background awareness that the understanding afforded by those models is partial, perspectival, and consequently that there may arise cases where they do not hold good that we are currently blind to.

There is a more vexing problem with models that are so deeply canalized as to be species-typical and not unlearnable. Consider the progressively richer understandings of persons and agency, from the Core Agency System, to Folk Psychology, to the assumption of freedom supposed by many of the most sublime ethical systems. Some people have believed that they have good reason to think that some of these assumptions, such as that of freedom, are false, because they think that the sciences entail determinism. I think this deterministic view of science is unwarranted (see my 2011), but my point here is different: I suspect that Kant was correct in saying that we cannot really do without the assumption of freedom in practical reasoning – and that in daily life even, say, B.F. Skinner reasoned for the most part about ethical matters as the rest of us do. Similarly, some have taken developments in physics to suggest that our deeply-entrenched bias towards thinking of a world of persistent objects with determinate properties and relations such as exact location is quite inapt as an understanding of the physical world. Yet even they do not cease to think in the ordinary object-oriented ways in daily life.

It is with these models – those that are so canalized into our cognitive architecture that we cannot shake off their attendant intuitions – that we are perhaps led most compellingly to the concern that they must be *either* synthetic *a priori* truths or forced errors, and we cannot step outside of them to settle which it is. To this, I shall say two things. First, I think that Cognitive Pluralism fares better than most views in accommodating the problem, as it allows us to test one model against another, even if we cannot rid ourselves of either, and because it views all models as necessarily idealized, and hence not to be held to a standard that only a

non-idealized understanding could meet. Second, our minds are not built only, or even primarily, for theoretical understanding, but to live a fully human life. And if our greatest theoretical efforts, be they laudable or misguided, cannot shake off the intuitions needed to lead a fully human life, this only shows that our minds are indeed well-designed.

Works Cited

Bloom, P. (2004). *Descartes' Baby: How the Science of Child Development Explains What Makes Us Human*. New York: Basic Books.

Boyer, P. (2002). *Religion Explained*. New York.

Edwards, P. (Ed.). (1967). *Encyclopedia of Philosophy*. Macmillan.

Horst, S. (2007). *Beyond Reduction*. New York: Oxford University Press.

Horst, S. (2011). *Laws, Mind, and Free Will*. Cambridge, MA: MIT Press.

Jackson, F. (1997). *From Metaphysics to Ethics: A Defense of Conceptual Analysis*.
Oxford University Press.