***Cognitive Pluralism***

Precis and Chapter Descriptions

(Draft in conjunction with manuscript preparation, August 2014)

Steven Horst

Department of Philosophy, Wesleyan University

[shorst@wesleyan.edu](mailto:shorst@wesleyan.edu)

Cognitive Pluralism is the thesis that we understand the world through a variety of mental models of different content domains. Each mental model provides a framework for thinking and reasoning about a particular domain, employing a distinctive representational system that is suboptimized for thinking about that domain. A model involves an internal ontology of kinds, properties, relations, processes, and transformations relevant to phenomena within its domain, and different models have distinctive internal ontologies and constitutive inference patterns distinctive to their domains. Models, moreover, are *idealized*, both in that (a) each model speaks only to its own domain, bracketing everything else about the world and that (b) a model may represent its domain only imperfectly, though generally well enough for particular theoretical and practical purposes. To the extent that a model is good enough for such purposes, it is *apt*. We can often think about the same objects, events or states of affairs by triangulating them through multiple models. But the facts that each model is idealized, and that different models are idealized in different ways, can produce mismatches between models, which can lead to inconsistencies, puzzles and paradoxes.

This book introduces Cognitive Pluralism and explores its possible implications for epistemology, semantics, and metaphysics. It is written for a mixed audience of philosophers, cognitive scientists, and the educated public.

**Part I – Unities and Disunities of Mind and Understanding**

Part I of this book presents as a foil for Cognitive Pluralism a “standard view” that is assumed in mainstream epistemology, semantics, truth theory and theory of reasoning, which does not accord mental models any special role. Instead, the standard view treats the basic types of units found in thinking as analogous to the differently-sized units found in language and logic: units that are word-sized, sentence-sized, and argument-sized. It is sentence-sized judgments or beliefs (or their propositional contents) that are candidates for truth. Sentence-sized units have semantic values that are largely functions of the meanings of their constituent concepts plus their compositional syntax (or, if forms of judgment are not literally syntax, some structural feature analogous to syntax). Atomists and holists differ over whether semantic value is constructed from the ground up, beginning with atomic concepts, or whether it is tied to the entire network of concepts, beliefs and inferential commitments; but they agree upon the basic *units* of analysis, with the sole exception that holists count the entire network as itself the governing unit. Sentence-sized units are the basic candidates for warrant as well as truth, and hence the principal units of epistemology, though foundationalists and coherentists differ over whether warrant attaches to some beliefs individually and others through their being the conclusions of valid arguments from other warranted beliefs (foundationalism) or because of a global coherence between beliefs (coherentism). At least one type of model – the scientific theory – is often treated as a paradigm case of knowledge. But at least in pre-Kuhnian philosophy of science (and to a surprising degree even today), theories tended to be treated either as *propositions* (e.g., laws interpreted as universally quantified claims) or as a set of propositions and inferential rules.

**Chapter 1, “Unities of Mind and Understanding”,** introduces a set of issues about “unities” and “disunities” of mind and understanding. Questions about the unity of mind have arisen, at least since Plato, in the form of a tension between the intuition that the mind is in some sense a simple unity (an indivisible substance, a single representational medium, or the a set of conscious experiences linked by the transcendental unity of apperception) and the acknowledgment of multiple psychological *faculties*. Traditional lists of faculties included intellect, imagination, the passions, sensation, and the will, and contemporary cognitive sciences have greatly expanded the list. The presence of multiple and possibly dissociable faculties not only calls into question the possibility of viewing the mind (or in older works, soul) as a simple unity, but also raises questions about whether semantic and epistemological issues can be confined to “the intellect” (the faculty that thinks and reasons in language-like propositions). However, questions about the unity of understanding – unity of knowledge, unity of science – are generally not approached as questions about the mind and its faculties, but as questions about how (or whether) the *objects* of belief hang together in particular ways: say, as a single theory of everything, a comprehensive worldview, or at very least as a set of mutually consistent beliefs. Such questions are generally not posed as questions about understanding, belief, or even science *as we find it*, as our minds generally contain any number of inconsistent beliefs or commitments, but as questions about questions about whether our beliefs can, in principle, be uni*fied* in some particular way, such as a single theory of everything or by inter-theoretic reduction.

**Chapter 2, “Modular and Central Cognition”,** examines a familiar contemporary strategy for walling off what we might call “the space of meaning and reasons” – our concepts, beliefs, and inferential dispositions – from concerns about the disunity of psychological faculties. Following Fodor (1983), many philosophers and cognitive scientists distinguish “central cognition” comprised of intentional states from a variety of “mental modules” that perform special tasks in very different ways. Advocates of modularity differ in how stringently they define the criteria for modularity, and consequently for how many modular systems there might be, but agree that modules are special-purpose mechanisms designed for particular tasks, which represent and process information in distinctive ways, often producing conceptually-laden beliefs only as their outputs into central cognition. Fodor’s list of features treats modules as also being characterized by informational encapsulation, cognitive impenetrability, fast processing, typical nativistic developmental sequences, and characteristic neural localizations. Advocates of massive modularity typically add that modules must be adaptive products of natural selection, though they may relax some of Fodor’s other criteria. Modules and central cognition can work together, but what philosophers have traditionally thought of as “thinking” and “reasoning” are all located within central cognition, with modules lying largely at the perceptual and motor “periphery”, supplying inputs to central cognition and (in the case of any motor modules) acting upon the outputs of central cognition. Modules may work as tightly-interdefined units, but they are non-conceptual and non-intentional. Central cognition, by contrast, is domain-neutral, allowing us to think and combine thoughts about any domain.

**Chapter 3, “Beyond Modularity and Central Cognition”,** presents a number of lines of research that suggest that this bifurcation of central versus modular cognition is not a useful division. Core Knowledge Systems, Folk Theories, and scientific theories are all constituted by tightly-interdefined domains that require their own representational systems and inferential patterns. Core Knowledge Systems and Folk Theories are developmentally canalized, and the former are evidenced as early as we can test infants, while scientific theories are among the most rarefied forms of learned understanding. Moreover, attempts to model human understanding in artificial intelligence, such as Minsky’s frame-based approach, succeeded only by treating everyday understanding in terms of domain-specific “frames” or “mental models”. Domain-specific models seem to the basis for much of our understanding and reasoning about a wide variety of things, all of which are examples of paradigmatically intentional states and processes, yet they are strikingly different from Fodorian modules and those of massive modularists in that they run the gamut of options when it comes to how they are acquired, whether they are species-typical, and whether they are products of natural selection. My conclusion from this is *not* that all of cognition is modular, but that there is a highly significant type of cognitive unit – a mental model of a content domain – that is found *quite generally* as a way of structuring understanding.

**Part II – Cognitive Pluralism and the Architecture of Understanding**

**Chapter 4, “Cognitive Pluralism”,** gives an initial presentation of a view of understanding that places domain-centered mental models at its core: Cognitive Pluralism. Our cognitive architecture is characterized by what we might think of as a basic design strategy to understand the world through a proliferation of mental models that do good-enough jobs of providing epistemic and practical grip upon different aspects of the world for a variety of theoretical and practical purposes. Each model is idealized, both in the sense that it brackets a great deal about the world to cast light upon a particular set of patterns found within it, and in the sense that it may provide only an imperfect grasp of even those phenomena of which it treats, and may distort them. (For example, a theory that treats bodies as point-masses or persons as ideally informed perfect reasoners.) A model is called *apt* to the extent, and in the contexts, that it does a good job of doing so. A model has a characteristic internal ontology, representational system, and inferential patterns, and the tacit rules and representational system define a possibility space of representations that can be framed in terms of the model. The concepts and inferential patterns found within a model are constitutively interlinked with one another, but not in a similar way with things outside the model.

**Chapter 5, “The Plausibility of a Pluralistic Cognitive Architecture”,** explores why a pluralistic cognitive architecture is a *good* approach to building understanding into organisms with finite minds. Part of this is an exploration of well-covered territory about the progressive evolution from very simple to more complex minds through an accumulation of new abilities and adaptive representational capacities. Even before the theory of evolution, philosophers had often been happy to view the minds of non-human animals as a “grab bag of good tricks”, but even philosophers who accept evolution have often been disposed to minimize the role of these in human cognition. This seems a wrong-headed approach, as legacy capacities are often still quite adaptive (sometimes more so than the use of sophisticated theories) and evolution is conservative in eliminating older features. The more original part of the chapter argues that a proliferation of special-purpose models that are *learned* – perhaps a consequence of a “modeling engine” – is also a good approach to allowing more sophisticated minds to form an open-ended number of new models of relevant content domains. Cognitive Pluralism is thus argued to be plausible on general theoretical grounds: it is what we should expect on evolutionary principles, and is what seems to be needed to confer flexible open-ended understanding upon finite minds.

**Chapters 6 (“Models”) and 7 (“Mental Models”)** develop one of the key notions of Cognitive Pluralism: that of a mental model. Chapter 6 explores a number of paradigm cases of *external* models – scale models, maps, architectural blueprints, computer models – as cognitive prostheses, noting characteristics of models, and also distinguishing some significantly different uses of the word ‘model’, such as that employed in model theory. Chapter 7 then explores more specifically the idea of a *mental* model, looking at a variety of paradigms: visual/spatial models that can be used in imagination, a mental model of chess, and models of social, moral, and cognitive systems.

**Chapter 8, “Models and Intuition”,** claims that a model-based account can help to explain intuitive reasoning and intuitive judgment. The first half of the chapter sorts through the somewhat confusing variety of characterizations of “intuition” in philosophy and psychology. The second half builds a case that we can explain why “intuitive” judgments have certain characteristics – immediate production, the absence of explicit conscious reasoning, and a sense of necessity – if we view them as products of processes in which the implications of the constitutive rules of models cause particular judgments to “pop out” as necessitated by the model. Conversely, a judgment can seem “counterintuitive” because of cognitive dissonance caused by detection of a violation of the rules of a model or the impossibility of accommodating it within a model’s representational system.

**Chapter 9, “Relations Between Mental Models”,** bundles together discussions of five ways models can be related to one another. 1) Models can be related by degrees of abstractness: a model provides a framework for representations, but something that is a representation in the system provided by one model can also serve as a model itself. 2) One model can be created as a variant upon another, through a process of alteration of the original model. 3) New models can be created through metaphorical transposition, in which the structure of one model is used as scaffolding for constructing a model of a distinct content domain. 4) We can *triangulate* objects and states of affairs in the world by viewing them through multiple models. 5) Two models can be dissonant with one another, by being formally incommensurable or inconsistent, or by producing conflicting interpretations of or predictions about a single state of affairs to which they are both applied.

**Chapter 10, “Models and Language”,** examines relationships between model-based understanding and language-like cognition. Model-based cognition is found quite widely in the animal kingdom, whereas language-like cognition is a specialized ability, perhaps found only in the human species. The semantic and inferential properties that we employ when reasoning in a language (or in language-like thinking) are in many cases actually derived from the ways concepts and inferential patterns are constitutively inter-related within a model. And language itself is not a representational system and is not itself a model. But linguistic thinking provides resources for reasoning, communication, public and private scrutiny of representations, inferences and assumptions, and hence possibilities for calling both beliefs and model-based understanding into question (and revising them) that are not available without it. Language can also provide an additional means through which models can be learned. However, the very fact that language and language-like thinking are available for explicit scrutiny in ways that model-based reasoning often is not can obscure the role of models in human cognition, and lead to a view of human cognition (such as the standard view) that overestimates the parallels between language and thought.

**Chapter 11, “A Speculative Cognitive Phylogeny”,** draws a broader and somewhat speculative picture of the role of model-based cognition, both in human thought and in an evolutionary context. Model-based cognition is distinguished from other basic strategies for conferring intelligence (specialized native abilities, conditioning, redundant systems, social distribution of cognition, incorporating the environment into the extended phenotype). The transition from having only special-purpose models that are evolutionary adaptations to being able to acquire new models through learning via a *modeling engine* is proposed as a major advance in intelligence. This is then supplemented by interactions with additional cognitive advances: language and language-like cognition, exploratory play, external information storage, and offline cognition.

**Chapter 12, “Issues About the Nature and Status of Models”,** concludes Part II with more philosophical discussion of issues about the nature and status of models. One of these concerns the identity conditions for a model: for example, should an alteration to a model be regarded as resulting in the production of a distinct model or as a revision to a model that persists through the change? I argue that there are distinct ways of “typing” models – semantically and psychologically – that seem to require opposite answers. A second question is what kinds of psychological systems should be counted as models: for example, should a completely “hard-wired” system, like the fly’s nervous system, in which perception directly drives motor response, be counted as involving models? I adopt a very broad construal of ‘model’ that counts something as a model if it affords a space of possible representations, even if these are not plausibly construed as involving *concepts* that can be separated from perception and motor control.

**Part III – Semantics, Epistemology, Disunities of Understanding**

**Chapter [13], “Epistemology, Semantics, and Truth”,** turns to implications of Cognitive Pluralism for epistemology, semantics, and truth. Models are not propositions, and hence are not the sorts of things that can be true or false, but aptness may be regarded as a separate alethetic attribute appropriate to model-sized units. The aptness of a model is also itself an important externalist criterion for epistemic warrant, which fits better with coherentist than foundationalist theories, but the unit in which coherence takes places is model-sized rather than global. But more broadly, neither foundationalism nor coherentism can be an adequate epistemological theory. Some beliefs are derived from other beliefs, and whether the derivation is valid surely affects warrant. But beliefs are not individually properly basic to the extent that they derive their warrant from being grounded in apt models. On the other hand, the ways different models are idealized may prevent them from being *globally* coherent or even consistent. I suggest that different epistemological accounts not be viewed as competitors providing necessary and sufficient conditions for warrant, but as exploring distinct dimensions of epistemic goodness and fault. Similarly, Cognitive Pluralism suggests a “regionalist” semantics located somewhere between atomism and holism. The semantic values of concepts are constitutively interlinked with other things within a model, but not outside of a model. Moreover, a model-based account provides a way of avoiding the problem of an “implicationist” semantics leading to holism because mental models are *psychological* units that are of intermediate size, and in which constitutive interrelations operate.

**Chapter [14], “The Multiple Lives of Concepts”,** takes a critical look at something that might be suggested by a model-based approach to semantics: namely, that concepts reside (only) within individual models. While much of the semantics of concepts is derived from models in which the play a role, there are two reasons to doubt this stronger “residentialist” view. On the one hand, it seems possible to have concepts (or something very like concepts) that are not embedded in models at all. On the other hand, it often seems to be the case that a single concept plays a role in multiple models. To accommodate these insights, I propose that concepts are in fact complex, multi-factor cognitive structures, which can have both model-based and more “lexical” components.

**Chapter [15], “Disunities of Understanding”,** returns to the topic of disunities of knowledge and understanding, including disunities of the sciences. I argue that, if understanding is grounded in a variety of idealized models optimized for different contexts, various types of dissonance between models can naturally arise: incommensurabilities, formal contradictions, and different predictions and anticipations. On the one hand, this suggests a reassuringly deflationary analysis of disunities of the sciences: that they may to a large extent be artifacts of the ways we model phenomena in the world, and need not point to a disturbing “disunity of the world”. On the other hand, it is quite possible that our ways of understanding the world make us *incapable* of forming a way of understanding things that is at once comprehensive and self-consistent – a principled barrier to certain types of unification of knowledge and understanding.

**Chapter [16], “Cognitive Illusion”,** presents, from an epistemological standpoint, the idea that our cognitive architecture may produce certain types of *illusion*. This is “illusion” in more or less the Kantian sense of the production of beliefs that seem to be knowledge but in fact are products of the operation of the mind’s own mechanisms in a fashion that could not produce knowledge. This chapter picks up on the idea, developed in Chapter 8, that many intuitions are based in the features of particular mental models. But such intuitions are no better than the models from which they spring. Inapt application of models can produce illusory intuitions. And as Kant argued, the mind also has special drives to unify its understanding which can produce illusory intuitions that seem to have the character of necessity, a theme that will be returned to in Chapter 19.

**Part IV: Cognitive Pluralism and Metaphysics**

**Chapter [17], “Cognitive Pluralist Metaphysics,”** explores the fact that, like Kant’s Transcendental Idealism (perhaps the best-known cognitivist philosophy), Cognitive Pluralism can also be adopted as a view with metaphysical implications. Like Kantian Idealism, Cognitive Pluralism is opposed the kind of “naïve” realism that assumes that the world comes pre-divided in a single, canonical and mind-independent way into objects, kinds, properties, relations and processes, which it is the mind’s job to reflect faithfully. But whereas Kant posited a single system of Understanding grounded in the Categories (and a single system of Sensibility closely united to it), Cognitive Pluralism posits a number of mental models of different content domains, challenging Kant’s assurance of the “Unity of Reason”. Cognitive Pluralism also suggests that there might be a variety of types of cognitive illusion produced by features of the mind, though these are not limited to the projection of Kantian Ideas of Reason through a series of prosyllogisms, but have potentially much broader scope. Cognitive Pluralism also has an important feature that helps it to avoid a significant limitation of Kantian Idealism. In Kant’s system, whatever is assured to be true of all possible experience because it arises from features of Pure Sensibility or Pure Understanding is thereby deemed necessary and a possible subject of synthetic *a priori* knowledge, and indeed one cannot even frame the question of the aptness of the Categories or Forms of Sensibility coherently. Cognitive Pluralism, by contrast, allows for the possibility of using one model as a corrective to another, a feature that is found precisely because it employs multiple models. In assessing the aptness of one model, we *treat* another *as though* it were an unidealized canonical representation – adopting a kind of “realist stance”.

**Chapter [18], “Intuitions and Necessitarian Metaphysics”** engages two possible consequences of Cognitive Pluralism for analytic metaphysics. The first of these follows upon earlier discussions of intuitions. If intuitions are based in the rules of mental models, they are only as good as the aptness of the models in which they are based. It seems unlikely on a variety of grounds that models tooled towards more mundane ends should prove trustworthy in producing intuitions about metaphysical necessity and possibility. Moreover, models are idealized, and there is reason to think that intuitions about necessity and possibility would have to be unidealized and completely general in their aptness. Second, the standard framework for talking about necessity and possibility – possible worlds semantics – seems to require that we posit things like “the set of all propositions” and to assume that these could all be assigned truth values in a consistent fashion. The first seems to assume the very sort of realism which the Cognitive Pluralist metaphysician, qua cognitivist, rejects. The second is threatened by the possibility that incompatibilities between our models could prove principled and abiding.

**Chapter [19], “Metaphysical Illusions”,** returns to the topic of cognitive illusion. Following Kant’s pioneering work in the Transcendental Dialectic, I consider how a number of long-standing philosophical ideas might in fact be illusory products (not necessarily *false*, but not *known* or *necessary* either, in spite of intuitions to the contrary) of either the application of particular models or of drives to unify our understanding to an ideal conclusion. Some of the “case studies” involve appeals to recent work in cognitive science, but the list bears some resemblance to Kant’s: dualism, personal immortality, reductionism, determinism, the causal closure of the physical world, and promiscuous teleology.

**Chapter [20], “Objects and Objectivity”,** returns to positive accounts Cognitive Pluralism might provide of difficult philosophical notions: objects and objectivity. The treatment of the notion of an object in the most general sense draws upon earlier discussions of triangulation. In order for triangulation – or the ability to think of something as the selfsame object through changing states and changing beliefs about it – the mind must possess a special resource for thinking about *individuals*, as opposed to kind- or property-instances. This is a beguiling notion, as in a certain sense it has no content beyond a pure “that-ness”, and can give rise to an illusory idea of “bare particulars”. But it also provides a way of fixing the intended reference of our thoughts in a fashion that allows for changes in our ways of conceiving it. This, in turn, also provides the foundation for a way of thinking about objec*tivity*. Even if all models are idealized, we can attain successively more “objective” understandings of it by forming models that bear fewer traces of the *peculiarities* of our relationships to them.

**Chapter [21], “Self and Mind as Subject and Object”,** at last returns to questions raised at the outset about unities and disunities of mind. Like the Transcendental Idealisms of Kant and Husserl, Cognitive Pluralism gives the mind’s own operations a distinctive place in epistemology and metaphysics. However, I do not share Husserl’s ambition of making it a new *foundation* for all other knowledge, and in particular, do not think that even a pure phenomenology produces indubitable or unidealized understanding of its subject matter. We have multiple models of the mind, which do not line up exactly. Some, like those produced by the sciences, are explicitly theoretical. Others, like those of Folk Psychology, are like theories in having their own (often tacit and unquestioned) internal ontologies. Moreover,these, and even the conscious mental states that are available for phenomenological reflection, are all *interpreted* in some particular model-based terms, which can potentially be apt in some circumstances but not in others. In particular, dispositional notions of belief modeled upon explicit judgments and/or upon language are potentially suspect in how faithfully they represent what they report. Dennett’s interpretationist view has something right: much of our thinking *about* thinking involves taking an interpretive stance, though there is probably more than a single “intentional stance” used as an interpretive framework. Eliminativism also has something right: even when intentional explanation works well, it misrepresent what it reports. But the misrepresentation does not consist in positing things that are nonexistent, but thinking about them through a model that is not fully apt to its domain.

**Chapter [22], “Noesis”,** takes up, in a quite speculative mode, a final theme about unities and disunities of mind and understanding. Some philosophers (notably Platonists and Neoplatonists) have claimed that our minds are capable of a kind of nondiscursive understanding that is very different from ordinary, scientific, and philosophical thinking. In the metaphor of the divided line, Plato refers to this as *noesis*. A few philosophers and a larger number of spiritual practitioners have claimed to experience this mode of awareness, and claim that in it, understanding is not divided by concepts and representations. If such a mode of understanding is really possible, it *would* confer a kind of unified cognition very different from those that most of us experience. However, as it is something that most of us have never experienced, it is something for which we can only *hope*, and can neither confirm nor describe.