Science and Mind: How theory change illuminates ordinary thought

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor in Philosophy in the Graduate School of The Ohio State University

By

Timothy James Fuller

Graduate Program in Philosophy

The Ohio State University

2012

Dissertation Committee:

Richard Samuels, Advisor

Louise Antony

Declan Smithies

William Taschek
Copyright by
Timothy James Fuller
2012
Abstract

This dissertation imports insights from the philosophy of science into the philosophy of mind. In order to do so, I first distinguish importantly different types of theories of scientific inference, including normative, population-level, competence, and performance theories. Then, I trace the unique implications that each type of theory has (or fails to have) for theories of everyday adult belief revision, cognitive development, and cognitive architecture. In particular, I argue that scientific theory change and ordinary adult belief revision are underwritten by the same kinds of cognitive processes, those that can be characterized by information flow and the principles governing it. On this view, ordinary belief revision in adults and scientific theory change are importantly different from childhood learning, because the latter is underwritten by distinct kinds of cognitive processes. Cognitive development is thus needed to bridge the gap between childhood learning and adult scientific inference. I construct a model of this development in terms of Bayesian networks and related advancements in machine learning. Finally, I argue that the failure of “The Extended Mind” to characterize scientific inference strongly suggests that “The Extended Mind” has no legitimate applications to contexts of ordinary cognition either.
Acknowledgments

I owe a special debt of gratitude to Richard Samuels, who patiently and generously taught me how to research and write a publishable paper. I believe this skill will be useful in the future. I would also like to warmly thank Louise Antony, who first introduced me to the joys of thinking rigorously about empirical issues in the philosophy of mind. All of the following deserve my thanks for providing valuable feedback on drafts of this dissertation or its central ideas: William Bechtel, David Danks, Justin D’Arms, Mike Fereira, William Friedman, Alison Gopnik, Joe Levine, Amy Masnick, Daniel Pearlberg, Harini Raghupathi, Abe Roth, Tim Schroeder, Laura Schulz, Declan Smithies, Shannon Spaulding, William Taschek, Neil Tennant, Kelly Trogdon, Carl Voss, and Jiji Zhang.
Vita

1998.......................................................... B.A., Philosophy, Cornell University

2002- present........................................... Graduate Teaching Associate/Lecturer Ohio State University

Fields of Study

Major Field: Philosophy
# Table of Contents

Abstract ........................................................................................................ iii

Acknowledgements ....................................................................................... iv

Vita ................................................................................................................ v

List of Figures ................................................................................................ ix

Chapter 1: Brief Introduction ...................................................................... 1

Chapter 2: Scientific Inference and Ordinary Cognition: Fodor on Holism and
Cognitive Architecture (with Richard Samuels) ........................................ 6

1. Fodor on Cognitive Architecture and Central Processing .................... 8
   1.1. Stage-Setting .................................................................................. 8
   1.2 Assembling Fodor’s Critique .............................................................13

2. Varieties of Scientific Globalism .............................................................17

3. Does Any Version of Scientific Globalism Have Implications For Theories of
Cognitive Architecture? ...........................................................................21
   3.1. Normative Globalism .....................................................................22
   3.2 Population-level Globalism .............................................................27
   3.3 Competence Globalism ..................................................................39
   3.4 Performance Globalism ..................................................................45
      3.4.1 *The Argument From Confirmation Holism* ............................46
      3.4.2 *The Teleological Argument* ..................................................49
      3.4.3. *The Argument from Remoteness* ........................................52
List of Figures

Figure 1. DAG
Chapter 1: Brief Introduction

Can the way science works tell us anything about the way minds work in ordinary circumstances? According to many influential philosophers, psychologists, and cognitive scientists, there is much that can be learned by comparing ordinary thought to processes of scientific theory formation and revision. Thus, over the last half century, this powerful comparison has shaped prominent views—indeed entire research programs—across a wide variety of areas in the cognitive sciences. But what reasons are there to look to science for insight into the nature of ordinary thought? And if there are any insights to be gained, what kinds of ordinary thought might be illuminated? Further, what in particular might the nature of science tell us about these kinds of thought? This dissertation seeks to address these questions and improve their answers, and thereby shed light on the relationship between science and everyday cognition.

Overall, my investigations chart the following course: I begin by distinguishing importantly different types of theories of scientific inference, including normative, population-level, competence, and performance theories. Then I trace the unique implications that each type of theory has (or fails to have) for theories of everyday adult belief revision, cognitive development, and cognitive architecture. In particular, I argue that ordinary adult belief revision and scientific theory change should be understood as instances of the same kinds of cognitive processes, those that can be characterized by
information flow and the principles governing it. On this view, ordinary belief revision in adults and scientific theory change are unlike childhood learning, which is underwritten by distinct kinds of cognitive processes. Cognitive development is thus needed to bridge the gap between childhood learning and adult scientific inference. I construct a model for this development in terms of Bayesian networks and related advancements in machine learning. Finally, I argue that the failure of “The Extended Mind” to characterize scientific inference strongly suggests that “The Extended Mind” has no legitimate applications to contexts of ordinary cognition either. In what follows, I briefly outline how each chapter in this dissertation contributes to establishing the above claims.

Chapter two (with Richard Samuels) examines a well-known view on the relationship between science and everyday cognition that is found in the work of Jerry Fodor. The focus of Chapter two is on Fodor’s highly influential claims about the global or holistic nature of scientific inference and the putative implications that this has for extant theories of cognitive architecture. This chapter introduces a general complication with drawing conclusions about the nature of ordinary cognition from claims about the nature of scientific hypothesis formation and confirmation. The challenge is to distinguish importantly different types of theories of scientific inference, including normative, competence, population-level, and performance theories. Failing to do so, it is argued, has led to fundamental misunderstandings of the relationship between science and everyday cognition, including the misunderstandings contained in Fodor’s critiques of theories of cognitive architecture. Having so argued, this chapter outlines more promising relations that hold between theories of scientific inference and theories of
ordinary cognition. In particular, the chapter defends two important similarities: the first between normative theories of scientific inference and normative theories of everyday inductive inference; the second between scientific and everyday reasoning competence.

Chapter three argues that Fodor is not alone in likely conflating different types of theories of scientific inference and mistaking their implications for cognitive science. I argue that a similar mistake plagues one prominent version of the “theory theory,” advocated by Alison Gopnik and others, according to which early cognitive development is highly similar to scientific theory change. In the first part of this chapter, I criticize the “strong” thesis that scientific inference and early cognitive development are subserved by the very same mental mechanisms. I contend that Gopnik has failed to distinguish among population-level, normative, competence, and performance theories, and that the strong thesis fails to hold for any of them. In the second part, I argue that even a weaker analogy between childhood cognitive development and scientific inference has severe limitations. In particular, I maintain that scientific inference is centrally subserved by cross-domain processing, while early cognitive development is not. The hope is that the inadequacies in Gopnik’s position are instructive. For one, they suggest a tighter analogy, which I briefly defend, between adult quotidian cognition and scientific theory formation and confirmation.

In Chapter four, I consider in detail recent empirical support for the “theory theory” that might appear to come from research on Bayes nets as models of human learning. I agree with prominent proponents of Bayes nets modeling that the associated formalism accommodates some significant aspects of both early childhood and scientific
learning. However, I disagree with such proponents that Bayes nets research supports the view that there is a high degree of general similarity between both types of cognition. Instead, reflections on the limitations of a standard Bayes net formalism reveal important differences between scientific and early learning. In turn, these differences suggest what I call *The Process Development Hypothesis*, according to which the cognitive development of adult scientific inference includes the advent of new kinds of learning process for forming and revising theories. In addition, I defend an accompanying set of proposals for characterizing this development that are drawn from recent advancements in machine learning. I maintain that the development of adult scientific inferential capacities can be understood, in several central cases, in terms of significant generalizations of a standard Bayes net formalism, or in terms of cognitive processes coming on-line that are appropriately modeled by distinct formalisms.

The topic of Chapter five is *Externalism* and its potential implications for theories of scientific cognition. The core idea behind Externalism is that cognition, at least some of the time, is underwritten by cognitive mechanisms, representational vehicles, or other cognitive structures, the boundaries of which are located beyond the biological boundaries of individuals. Thus, according to externalist theories of scientific inference, scientific theory formation and revision are sometimes underwritten by cognitive processes that occur beyond the biological boundaries of individual scientists. I argue that two kinds of externalist theories of scientific inference deserve to be sharply distinguished. The first I call *The Extended Scientific Mind*, which is an application of The Extended Mind to contexts of scientific inference. The second I call *population-level*
theories of scientific inference, which characterize the dynamics of groups or communities of scientists at the population-level. I argue that population-level theories of scientific inference are more plausible than The Extended Scientific Mind. If I am right, a similar moral likely holds for extended and population-level accounts of ordinary cognition.
Chapter 2: Scientific Inference and Ordinary Cognition: Fodor on Holism and Cognitive Architecture (with Richard Samuels)

Abstract: Do accounts of scientific theory formation and revision have implications for theories of everyday cognition? We maintain that failing to distinguish between importantly different types of theories of scientific inference has led to fundamental misunderstandings of the relationship between science and everyday cognition. In this paper, we focus on one influential manifestation of this phenomenon which is found in Fodor’s well-known critique of theories of cognitive architecture. We argue that in developing his critique, Fodor confounds a variety of distinct claims about the holistic nature of scientific inference. Having done so, we outline more promising relations that hold between theories of scientific inference and ordinary cognition.

Introduction

What implications do accounts of scientific inference have for theories of quotidian cognition? According to many prominent theorists, the implications are wide-ranging and profound. Thus over the past half century or so, influential work on perception, cognitive development, and reasoning have all been heavily informed by —and modeled on—
conceptions of the processes by which scientific hypotheses are formed, tested and revised.

This paper has three main aims. First, we consider and reject one important manifestation of the influence that conceptions of scientific inference have had on theories in cognitive science: Jerry Fodor’s highly influential claims about the *globality* of scientific inference and the putative implications that this has for theories of cognitive architecture. Second, we articulate a range of more plausible respects in which accounts of scientific inference can inform theories of ordinary cognition. Finally, in developing the above points, we provide a taxonomy of kinds of theoretical relations that plausibly hold between claims about scientific inference and claims about ordinary cognition, and thereby provide a framework for further research.

Here is how we proceed. In section 1, we sketch Fodor’s well-known picture of everyday non-demonstrative reasoning and explicate how his conception of the global nature of scientific inference informs this picture. In section 2, we articulate four importantly different kinds of claim regarding the globality of science that appear to be run together in Fodor’s work—what we call normative, population-level, competence, and performance versions of globalism. In section 3, we first identify two conditions—*Plausibility* and *Connectability*—that any claim about the globality of scientific inference would need to satisfy in order to have implications for theories of cognitive architecture. We then consider each globality claim in turn, and argue that none satisfy both Plausibility and Connectability. We therefore conclude that Fodor’s claims about the globality of scientific inference do not have interesting implications for theories of
cognitive architecture. Nonetheless, we do not wish to claim that there are no ways in which theories of scientific inference can substantively inform theories of ordinary cognition. With this in mind, in section 4, we conclude by articulating a range of more plausible proposals about the significance of theories of scientific inference for theories of everyday thought.

1. Fodor on Cognitive Architecture and Central Processing

Over the past four decades or so, Fodor’s view of central processes has exerted an enormous influence on debates within cognitive science. In capsule form, his view is that the central processes responsible for belief revision and decision-making are, in crucial respects, ‘global’ or ‘holistic’. Moreover, he argues that this fact is one that cannot be readily accommodating by any extant account of cognitive architecture, including connectionism, massive modularity, classical computationalism, and their associated hybrids (Fodor 1975, 1983, 2000, and 2008). Whether the holistic character of thought has any such negative implications is a disputed matter that one of us has discussed in detail elsewhere (Samuels 1998, 2005, 2010). But for present purposes, this is not the issue that concerns us. Rather, we are concerned with the reasons Fodor provides for his globalism in the first place —reasons that, as we will soon see, turn crucially on claims about the holistic character of scientific inference.

1.1. Stage-Setting
Before we set out Fodor’s critique of theories of cognitive architecture, some stage-setting is in order. Specifically, we need to explain some key pieces of terminology, and provide a brief overview of some of Fodor’s commitments. Since much of this will be familiar to many, we endeavor to keep things brief.

1. **Cognitive Architecture.** Though the term has been used in a variety of ways, the notion most relevant to our present discussion—one adopted by Fodor, Pylyshyn and many others—is that a theory of our mental architecture describes the invariant physical structures that underwrite the capacities and achievements of cognition (Fodor, 1975; Pylyshyn, 1984). In Pylyshyn’s words, the invariant features of a cognitive architecture are ‘cognitively impenetrable’: they remain constant across changes to an agent’s beliefs, goals, desires and other intentional states, and remain unchanged by the cognitive tasks that the agent performs (1984). Importantly, what falls out of an account of the cognitively impenetrable features of a cognitive system are the ‘set of possible cognitive processes that are allowed by the structure of the brain’ (Pylyshyn, 1996, p.49). Prominent proposals that seek to characterize our mental architecture include various forms of classicism and connectionism, hybrid classical-connectionist architectures, and a range of modularity hypotheses. Our purpose here is not to advocate or assess any of these views. Rather, we are concerned primarily with the general relationships that obtain between cognitive architecture as such and claims about the nature of scientific inference.

2. **Fodor’s taxonomy of cognitive mechanisms.** To a first approximation, Fodor (1983) maintains that human cognitive architecture can be divided into three broad
categories of cognitive mechanism: input systems for low level perceptual processes, output systems for motor control, and central systems, which are paradigmatically responsible for various kinds of ‘higher’ cognitive process, including decision-making and belief revision.\footnote{Fodor, following Peirce, uses the expression ‘belief fixation’ (1983, 2001, 2008) to describe a core aspect of belief revision: the process of generating new beliefs on the basis of current perceptual inputs and prior beliefs. This is a ‘central process’ according to Fodor because it occurs downstream from the processing of perceptual input systems and evaluates their outputs in light of, for example, beliefs stored in memory. Instead of ‘belief fixation,’ we will use the broader term ‘belief revision’ in what follows.} Though it will be useful for heuristic purposes to adopt this taxonomy, as far as we can tell nothing we say here turns on its acceptance. What does matter is Fodor’s claim that those systems responsible for belief revision underwrite a kind of non-demonstrative inference. On Fodor’s view, everyday beliefs are formed and revised in the light of perceptual evidence that underdetermines those beliefs. As a consequence, Fodor maintains, the central processes responsible for belief formation and revision can be characterized in terms of the projection of hypotheses that are then either accepted or rejected in light of background information (Fodor 1975, 1983, 2000, 2008).

3. The Analogy Between Scientific Inference and Quotidian Reasoning. Suppose that a central function of ordinary reasoning is to provide non-demonstrative assessments of empirical hypotheses. Then it may well seem natural to turn to philosophical accounts of scientific inference for insight into the nature of everyday belief revision, since both centrally involve the formation and confirmation of hypotheses. With little argument, Fodor appears to make precisely this turn:
Central systems are … ‘largely unconscious’ and very little is known about their operation… However, it seems reasonable enough that something can be inferred about them from what we know about explicit processes of non-demonstrative inference – viz. from what we know about empirical inference in science (Fodor, 1983, p.104).

Twenty-five years later, Fodor’s thinking is essentially unchanged:

…the psychological fixation of empirical beliefs is closely analogous to the scientific confirmation of empirical hypotheses (individual psychology = science writ small). I suppose, for example, that the fixation of perceptual beliefs is initiated by the (subdoxastic) registration of sensory information, from which beliefs about the distal percepts are (subdoxastically, nondemonstratively) inferred; roughly, the inference runs from sensory states to their distal causes (Fodor, 2008, p.113).

In short: There is, according to Fodor, a striking similarity between the function of scientific hypothesis formation and confirmation on the one hand, and central processes of belief revision on the other. But if this is so, then it may seem reasonable to draw conclusions about the nature of central processes from claims about scientific inference. And this is precisely what Fodor does. Specifically, he argues that central
processing is ‘global’ because scientific inference is ‘global’ in two crucial respects, which he calls isotropy and Quineanism.

4. Two kinds of Globality. In Fodor’s terminology, a rational inference, paradigmatically the confirmation of a hypothesis, is isotropic just in case the facts relevant to that hypothesis’s confirmation ‘might be drawn from anywhere in the field of previously established truths’ (Fodor, 1983, p.105). In the first instance, Fodor defines isotropy in terms of epistemic relations that hold between a given hypothesis and a set of background facts. That is, Fodor holds that (dis)confirmation of a hypothesis, in the general case, might come from any domain, however individuated, of one’s background theory. In addition, however, Fodor also characterizes isotropy as a property of cognitive processes, in particular those processes that can identify and appreciate these epistemic relations (this distinction will be important for our discussion in Section 3.2.). In this second sense of isotropy, Fodor claims that successful confirmation of a hypothesis must be subserved by isotropic cognitive processes that are capable ‘in principle’ of bringing any relevant component of a background theory to bear on that hypothesis’s confirmation. Isotropic cognitive processes might be characterized by contrasting them with the informationally encapsulated processes that Fodor believes perceptual input systems implement, and which have restricted access only to a proprietary database of information. Isotropic processes, in contrast, are highly unencapsulated and essentially have access to all of the available background theory that might be relevant to a hypothesis’s confirmation.
An inference, paradigmatically the confirmation of a hypothesis, might also be global if it is *Quinean*: that is, if ‘the degree of confirmation assigned to [the] hypothesis is sensitive to properties of [an] entire belief system.’ Fodor claims that hypothesis confirmation is Quinean when ‘… [a]s it were, the shape of our whole [theory] bears on the epistemic status of … the hypothesis’ (Fodor, 1983, p.107). In maintaining that scientific inference is Quinean, Fodor appears to believe that he is doing little more than drawing out an obvious consequence of the familiar claim that confirmation is a species of abduction. As ordinarily construed, abduction—or inference to the best explanation—involves the assessment of hypotheses not merely in terms of their empirical adequacy, but also their possession of such theoretical virtues as simplicity, conservativeness, coherence and consistency (Harman, 1965). According to Fodor, however, these virtues are defined over entire theories or belief systems. And intuitively it may seem that a hypothesis’s simplicity, coherence, consistency, etc. is not defined relative to some proper subset of a background theory or belief system. Rather, a scientific hypothesis might appear to enjoy such theoretical virtues only when it is coherent, consistent, etc. with *all* of one’s background commitments. In sum, if Fodor is correct, then scientific inference is in general global because it is, to a significant degree, both Quinean and isotropic.

1.2 Assembling Fodor’s Critique

We are now in a position to assemble the pieces of Fodor’s argument from the global character of scientific inference to conclusions about the nature of central processes and
the cognitive architecture on which they depend. Broadly speaking, the argument can be broken down into two stages.

*Stage 1. The argument for the globality of central processes.*

As we have already seen, Fodor maintains that there is a high degree of functional similarity between scientific inference and belief revision:

(1) Scientific inference and belief revision are both centrally involved in the non-demonstrative assessment of empirical hypotheses.

Further, we have already seen that, according to Fodor:

(2) Scientific inference is Quinean and isotropic.

But if this is so, then given the high degree of functional similarity between everyday belief revision and scientific inference, we may have some reason to suppose that:

(3) Central processes for everyday belief revision are also Quinean and isotropic.
Finally, since theories of cognitive architecture purport to characterize all and only those cognitive processes that are allowed by the structure of the human brain, it follows from (3) that:

(4) The cognitive architecture on which central processes depend must be able to subserve Quinean and isotropic processes.

For Fodor, (4) is seldom treated as the final conclusion of an argument. Rather, it is used as a basis for criticizing extant accounts of cognitive architecture.

Stage 2: Drawing out implications for cognitive architecture.

Since this stage of the argument isn’t our central concern, our characterization will be correspondingly brief. Given (4), however, it is clear that:

(5) Any adequate theory of cognitive architecture should posit structures that can subserve Quinean and isotropic processes.

Notoriously, Fodor has argued at great length that computationalists, modularity theorists and connectionists cannot account for such global processes (Fodor 1983, 2000, 2008). There is a very long story regarding why he thinks that this is so but we must leave that story untold here. That said, Fodor maintains (with characteristic pessimism!) that:
(6) No extant theory of cognitive architecture permits Quinean or isotropic processing.

In which case, we have reason to draw the following bleak conclusion:

(7) All extant accounts of the cognitive architecture on which reasoning and belief revision depend are fundamentally unsatisfactory.

That is: massive modularity, connectionism, classicism, and their associated hybrids are all implausible as theories of the invariant and impenetrable features of central systems. This line of argument is not the only critique that Fodor has made of theories of cognitive architecture, but we believe it is the centerpiece of Fodor’s decades-long skepticism toward every major theory so far proposed.

Fodor’s overall critique has been challenged on a variety of grounds. Pinker and Carruthers, for example, challenge the inference from (1) and (2) to (3) on the grounds that everyday reasoning is crucially disanalogous from scientific inference. They argue that the latter is typically secured by many individuals working over long periods of time, with sophisticated instruments and other resources, while the former is typically secured by individuals over relatively brief time periods (Pinker, 2005; Carruthers, 2003). Others have challenged premise (6) by arguing that Fodor’s pessimism regarding extant theories of cognitive architecture is unfounded (Samuels, 2005, 2010; Pinker, 2005). What has
received rather less attention, though it is equally central to Fodor’s overall argument, is
the commitment to the globality of scientific inference in the first place (i.e., premise (2)).
We propose to remedy this oversight by focusing on what we call Fodor’s Scientific
Globalism: roughly, the thesis that scientific inference, including hypothesis formation
and confirmation, is Quinean or isotropic.

2. Varieties of Scientific Globalism

Clearly, Fodor’s inference from Scientific Globalism to a conclusion about cognitive
architecture is not a demonstrative one. Rather, it is intended as a plausibility argument:
one that renders its conclusion prima facie plausible. Nonetheless, we maintain that
Fodor’s argument fails to do even this much. Whatever implications Scientific Globalism
may seem to have for theories of cognitive architecture derives from conflating
importantly different versions of the thesis. Or so we argue. In this section we set out
what we take to be four main kinds of globalism about scientific inference. In subsequent
sections, we evaluate each version with respect to both its plausibility and its suitability
for supporting claims about our cognitive architecture.

The first kind of Scientific Globalism we consider is what might be called
Normative Globalism. This is the thesis that:

2 In terms of the argument sketched in section 1.2, our rebuttal amounts to the following: Premise (2) is
systematically ambiguous and no plausible reading supports the conclusion (4) that theories of cognitive
architecture must permit Quinean or isotropic processes.
**Normative Globalism**: Scientific inference, including hypothesis formation and confirmation, *ought* to be Quinean or isotropic.³

Though it is unclear whether Fodor intends to endorse Normative Globalism, he sometimes writes as if he does. For example, in one passage he maintains that confirmation is isotropic because any background information might ‘in principle … be relevant to determining what else [one] *ought* to believe’ (Fodor, 1983, p.105, emphasis ours). Thus here (and elsewhere) in Fodor’s discussion of scientific inference, there is the suggestion that Quineanism and isotropy have the status of normative demands.

The second kind of Scientific Globalism that we detect in Fodor’s work is a *population-level* claim about scientific inference. To say that Fodor’s Scientific Globalism is a population-level claim means that the thesis does not purport to characterize the inferential activities of individual scientists, but instead purports to characterize the dynamics of scientific theory construction and confirmation at the level of institutions and communities of researchers. *Population-level Globalism* is the thesis that:

*Population-level Globalism*: The population-level dynamics of the scientific communities that construct and confirm theories are Quinean or isotropic.

³ We opt for a disjunctive formulation of each version of Scientific Globalism. A disjunctive formulation is intended to forestall the possibility that Fodor’s critiques might go through if either a claim about isotropy or a claim about Quineanism were, in isolation, both plausible and suitable for informing theories of cognitive architecture. Nothing much turns on opting for a disjunctive formulation of Scientific Globalism except that it makes our position more difficult to argue for.
Once again, it is unclear whether Population-level Globalism is a thesis that Fodor endorses. Nevertheless, there are places where he seems to do just that. So, for example, he claims:

Confirmational isotropy is a reasonable property for nondemonstrative inference to have because the goal of nondemonstrative inference is to determine the truth about a causal mechanism – the world – of whose workings we are arbitrarily ignorant. That is why our institution of scientific confirmation is isotropic … (Fodor, 1983, p.107, emphasis ours).

Part of the reason it is uncertain whether Fodor really intends to endorse either Population-level or Normative Globalism is that he does not typically differentiate between different types of theories of scientific inference when discussing the apparent global character of scientific confirmation. In the only passage we know of where he explicitly addresses the topic of what a theory of scientific confirmation is supposed to do, he appears to endorse an account on which such theories provide a characterization of a scientific reasoner’s competence:

…A theory of scientific confirmation is basically just a moral of what a scientist would do if he was doing what he is taught to do. It’s just a model of belief fixation in a reasonably idealized scientist, just as a grammar is a model of
grammatical intuitions of a reasonably idealized speaker (Fodor, 1987, MOM discussion).

This suggests a third kind of Globalism, viz.:

*Competence Globalism:* Quineanism or isotropy are properties of a scientist’s inferential competence with respect to confirming and constructing scientific theories.

Though the competence/performance distinction is an exceedingly familiar one (Chomsky, 1965), there are a number of crucial issues about cognitive competences that we will need to consider later in greater detail. Nevertheless, the rough idea is that a competence theory —in the present context for scientific inference— is one that abstracts from the details of the actual inferences that scientists perform, and instead provides a suitably idealized characterization of the inferences that scientists would make under conditions that are, in some appropriate sense, optimal. Specifically, accounts of this sort abstract from such ‘performance’ factors as time, memory, speed of processing and energy limitations.

This leads us to the last version of Scientific Globalism that we discuss here, what we call Performance Globalism. If a competence construal of the global character of scientific inference is one that abstracts from the details of the processes that scientists actually engage in, then a performance reading of Scientific Globalism is one that
purports to capture facts about the cognitive processes that scientists actually have. That is:

*Performance Globalism*: The actual inferential processes exhibited by individual scientists in the course of hypothesis formation and confirmation are Quinean or isotropic.

As with Competence Globalism, this claim is intended to capture some fact about individual scientists. But in contrast to Competence Globalism, it does not abstract away from performance limitations. Indeed, precisely the purpose of a performance theory is to capture facts about cognitive processes under conditions in which they are subject to such constraints. This point will be important to our discussion later on.

3. Does Any Version of Scientific Globalism Have Implications For Theories of Cognitive Architecture?

So far we have argued that there are four main readings of Scientific Globalism found in Fodor’s writings. In this section, we argue that each fails to do the work required of it. More precisely, we argue that no version of Scientific Globalism meets both of the following obvious adequacy conditions on having implications for theories of cognitive architecture:
Plausibility: We have reason to suppose that scientific inference satisfies the relevant globality thesis.

Connectability: On the supposition that scientific inference satisfies the relevant globality thesis, there is reason to believe that our cognitive architecture can subserve Quinean or isotropic inferential processes.

Clearly, any version of Scientific Globalism must satisfy both of these conditions in order to support Fodor’s contention that our cognitive architecture subserves Quinean or isotropic inferential processes. If a given globality thesis is implausible, then nothing is warranted by the mere fact that it follows from the thesis in question. Moreover, even if a given globality thesis is plausible, this alone would fail to support Fodor’s criticisms unless the relevant thesis were also to provide some reason to accept the claim that our cognitive architecture can subserve Quinean or isotropic inferential processes. What we now propose to argue is that there is no reason whatsoever to believe that any of the versions of Scientific Globalism we have identified satisfy both Plausibility and Connectability. In which case, there is no reason to believe that any version of Scientific Globalism supports Fodor’s criticisms of extant theories of cognitive architecture.

3.1. Normative Globalism

According to Normative Globalism, scientific inference ought to be Quinean or isotropic. However, even on the assumption that global inferences are desirable as a normative
standard for scientific inquiry, the claim that such a thesis has important implications for theories of cognitive architecture is *prima facie* implausible. The problem should be obvious enough: Normative Globalism is supposed to be a thesis about how science *ought* to be. But descriptive claims about what is in fact the case seldom follow from such normative claims. Thus, even on the assumption that Normative Globalism is plausible, this likely implies nothing about how, as a matter of fact, scientific inference proceeds. And if Normative Globalism doesn’t imply any descriptive claims about scientific inference, it surely doesn’t imply any descriptive claims about the globality of quotidian cognition either. On the face of it, then, Normative Globalism doesn’t begin to satisfy Connectability.

But perhaps this summary denial that Normative Globalism satisfies Connectability is too quick. Though there are no doubt many possible lines of response, by far the most plausible relies on the adoption of some form of an epistemic *ought implies can* (OIC) principle. The idea behind such OIC principles is that just as in matters ethical one has no obligation to ø unless one can ø, so too in matters epistemic our capacities constrain our obligations. According to many, normative accounts of reasoning are subject to such a principle (Goldman, 1986; Kitcher, 1992; Kornblith, 1992; Harman, 1986; Stein, 1998; Samuels et al., 2004). To take one well-known example of the implications of an OIC principle, if human belief systems are too large to permit the preservation of truth-functional consistency, then it is not the case that we *ought* to preserve the truth-functional consistency of all our beliefs (Cherniak, 1986). Now suppose that Normative Globalism is correct. Then, given an additional commitment to
OIC, it would appear to follow that scientists have the capacity for Quinean or isotropic reasoning. And on the plausible assumption —of which more in Section 4— that the minds of scientists and non-scientists are much alike, it would seem to follow that our cognitive architecture licenses Quinean or isotropic inferences. Thus, if one endorses an OIC principle and grants Normative Globalism’s plausibility, then Normative Globalism would appear to satisfy Connectability.

Unfortunately, the present strategy for defending Normative Globalism’s implications for theories of cognitive architecture is rather less plausible than it may initially seem. First, epistemic OIC principles are far from uncontentious and many theorists interested in normative principles of reasoning roundly reject them. Second, even waiving this objection, in the present context the endorsement of both Normative Globalism and OIC constitutes an unstable position since ought-implies-can considerations in fact provide reason to reject Normative Globalism. In order to see the point, consider theories of inductive inference, such as Carnapian and Bayesian theories, that have commitments similar to Normative Globalism. In particular, both kinds of normative theory are committed to a version of isotropy about non-demonstrative inference because they embrace a ‘requirement of total evidence’:\footnote{As Fitelson (2008) points out, a requirement of total evidence for Bayesians, unlike for Carnap, is often an implicit rather than explicit commitment.}
… in the application of inductive logic to a given knowledge situation, the total evidence available must be taken as the basis for determining the degree of confirmation (Carnap, 1950, p.211).

In fact, however, advocates of epistemic OIC principles, such as Harman, Goldman and Gigerenzer, have prominently criticized assumptions like total evidence precisely because they apparently require cognizers to exceed their capacities (e.g., Harman, 1986; Gigerenzer 2001; Goldman 1986; Kitcher, 1993). More specifically, they reject such normative requirements because the computational complexity of algorithms that conform to them are likely to vastly outstrip the resources of any individual, scientist or otherwise (Harman 1986). In a Bayesian framework, for example, a requirement of total evidence entails that scientists, when updating in the light of new evidence, should conditionalize on every combination of statements in their available background theory, perhaps including total science. But notoriously, updating in this fashion leads to combinatorial explosion (Osherson, 1995). As the number of statements (or beliefs) increases, the number of operations that the reasoner must perform increases exponentially, which is, for computational complexity theorists, the \textit{sin qua non} of a computationally intractable process.\(^6\) On the face of it, then, such considerations support

\phantomsection

\footnote{To give some idea of just how expensive updating can be, on standard Bayesian accounts, the equations for assessing the impact of new evidence on our current beliefs are such that if one’s system of beliefs has \(n\) elements, then computing the new probability of a single belief, \(B\), will require \(2^n\) additions (Harman, 1986). Thus on the hyperconservative assumption that we possess just 100 beliefs, calculating the probability assignment of a belief \(B\) on the basis of new information would require the performance of}
the implausibility of simultaneously endorsing both Normative Globalism and OIC principles.

In the present context, however, such considerations concerning Normative Globalism’s putative tension with OIC principles need to be handled with caution. Since Fodor is highly skeptical of computational accounts of central processes, it would be question-begging to assume that computational intractability \textit{precisely} delimits the boundaries of our reasoning capacities. After all, one natural reading of the Fodorian position is that some alternative to classical computational proposals is required in order to make sense of everyday belief revision. Nevertheless, even for non-computationalists, there is a weaker assumption about the status of computational complexity analyses that should be common ground, namely: evidence of combinatorial explosion is at least a \textit{prima facie} indicator of debilitating levels of cognitive resource-taxation. Thus if one has grounds to think that a task is combinatorially explosive, but no good reason to suppose that we can nevertheless perform it, then one has grounds (albeit defeasible ones) for thinking that we do not perform that task. To reject this weaker assumption is tantamount to claiming that computational complexity analyses are irrelevant to establishing architectural claims. But if the only way to establish that Normative Globalism satifies Connectability is to \textit{assume} the irrelevance of analyses of computational complexity, surely one cannot legitimately invoke Normative Globalism —as Fodor invokes Scientific Globalism— in the rejection of computationalism. After all, if more than $10^{30}$ additions, which is considerably more than the number of microseconds that have elapsed since the Big Bang!
computationalism is true, then such analyses are clearly relevant; and to suppose otherwise would already imply the covert denial of what one explicitly purports to refute.

Given the plausible assumption that combinatorial explosion is a *prima facie* indicator of debilitating levels of cognitive resource taxation, there really is no good reason to suppose that scientists routinely solve combinatorially explosive tasks when confirming hypotheses. Thus, there is reason to doubt that scientists can live up to the normative standards of Bayesian confirmation theory or the similar standards set by Normative Globalism. In which case, there is substantial, though defeasible, reason to doubt that Normative Globalism respects OIC. In short, a strategy that appeals to an OIC principle for establishing that Normative Globalism satisfies Connectability likely fails. Since this was the most plausible defense of Normative Globalism satisfying Connectability, there is no reason to believe that the thesis has any implications for theories of cognitive architecture.

### 3.2 Population-level Globalism

Next, consider Population-level Globalism. We propose to grant, at least for the moment, that it is plausible that the population-level dynamics of scientific communities are Quinean and isotropic. Even so, there is an obvious *prima facie* reason to doubt that Population-level Globalism satisfies Connectability, and thus to doubt that the thesis has any implications for theories of cognitive architecture. The reason is that it is far from clear that the properties of a given population are invariably, or even typically, shared by members of that population. In which case, even if the inferential processes of the
scientific community were isotropic and Quinean, this would do little to support the claim that the inferential processes of individuals, including individual scientists, exhibited these same global properties.

The worry above turns crucially on the empirical claim that properties that are instantiated at a population-level are not invariably or even typically instantiated at the corresponding individual-level. However, overwhelming evidence for this claim has been amassed in recent decades; and though we cannot hope to survey much of the evidence here, we take it that the point is by now a familiar one.⁷ So, for example, the development of harvester ant colonies appears to exhibit sensitivity to various kinds of information about the efficient allocation of colony resources, such as when to gather and store food, and when to work on colony infrastructure. But it turns out that no individual ant instantiates a decision-making procedure for appropriately directing tasks in the light of colony needs. Instead, individual ants are guided by algorithms defined over the frequency of smells so that population-level informational sensitivity results from lots of very local operations (Greene & Gordon, 2007). Much the same is true in the case of mound construction in termites fish foraging, bird flocking, human crowds, traffic jams, and many other things besides (Resnik, 1997).⁸ For that matter, if we do not make the assumption that populations must be comprised of biological organisms—as opposed to

⁷ The sorts of examples that we are discussing here are often associated with claims about the existence of properties that are, in some sense, emergent. For the record, we are inclined to think that much recent discussion of emergence has been rather confused. But in any case, our present point has little to do with a proper characterization of emergence. Our point is simply that population-level properties and relations are very typically not mimicked at the level of the individuals that comprise the populations.
⁸ Many similar examples can be found in: Holldobler & Wilson (2008).
merely particulars of some sort—then pretty much all of science stands testament to the fact that the properties of populations are very typically not shared by the particulars that comprise them. Thus many of an organism’s properties are not possessed by the cells from which they are composed; many of a cell’s properties are not possessed by the molecules that comprise it; many of an organic molecule’s properties are not possessed by its constituent atoms, and so on. From this vantage, the claim that individual entities lack many of the properties possessed by the populations to which they belong should be little more than a banal truism.

Is there any reason to suppose that the population-level dynamics of scientific communities are an exception to this claim? If there is, it certainly is not because complex human activities are somehow exempt from the generalization that population-level properties are not invariably mimicked at the level of individuals. Economic activity, for example, is an exemplar of an enormously complex population-level human activity. Moreover, it is one that seems akin to scientific practice in that principles of rationality are routinely invoked to describe and explain it. Yet the failure of individual members to invariably share properties exhibited at a population-level is implicit in several hundred years of ‘invisible hand’ explanations and predictions of market activity. To take just one example, Plott (2000) catalogued some of the Iowa Electronic Futures Market’s successes in predicting election outcomes and other future events. This is one example of a population-level description of market activity that makes no predictions concerning individual market participants. In particular, treating futures markets as assigning probabilities to future events does not entail that any individual participant
assigns the same probabilities to these events. Indeed, a central idea behind futures markets is that they can, and in some cases do, out-perform individual experts. Of course none of the above excludes the possibility that scientific confirmation is an exception to the generalization that properties of populations need not be ‘mirrored’ at the level of individuals. In particular, it might be that the Quineanism or isotropy of scientific communities in some way demands that individual scientists also exhibit these properties. Dominic Murphy has recently argued for such a conclusion with respect to isotropy (Murphy, 2006). Murphy readily endorses a socially distributed conception of science on which confirmation and other processes routinely involve an array of complex interactions and collaborations between scientists. Moreover, he allows that socially distributed processes of confirmation are isotropic. Nevertheless, pace Carruthers, Pinker and others, Murphy denies that a socially distributed conception of science is reason to doubt that individual scientists have isotropic minds. On the contrary, he maintains that the isotropic character of individual cognition is a precondition for the existence of the sort of science that makes more than local progress on small-scale problems (ibid.). Murphy’s considerations take the form of what might loosely be termed a Transcendental Argument for Isotropy. Specifically, he maintains that results from one branch of science could not figure in the (dis)confirmation of hypotheses in another unless individual scientists were able to recognize the mutual, confirmational significance of information derived from disparate regions of science. Thus, to take a well-worn example, Lord Kelvin’s claim that Darwin wasn’t entitled to his assumptions about the antiquity of the Earth could not have exerted an influence on the (dis)confirmation of Darwin’s theory.
unless Kelvin, Darwin and others had been able to recognize that astrophysical findings were, in this instance, relevant to biology (Murphy, 2006, p. 562; see also Antony 2003 for a striking example that illustrates a similar phenomenon). For Murphy, this sort of case both illustrates the isotropic character of individual scientific belief, and highlights an important fact about the preconditions for the sort of science we have: ‘if individual scientists did not have isotropic minds, the upshot of all that distributed inquiry would be nugatory’ (ibid., p.562).

To a first approximation, the conclusion of Murphy’s Transcendental Argument for Isotropy, and what he believes examples such as the “Kelvin/Darwin” case illustrate, is that individual scientists have the capacity to recognize the relevance of any item of information to the relevance of any hypothesis. This is how we understand Murphy’s claim that scientists have isotropic minds. In contrast, we maintain that Kelvin/Darwin examples, as well as Murphy’s related transcendental considerations, though not uninteresting, show far less than this. That is, we hold that all that is required for isotropy to obtain at the population-level, and all that Kelvin/Darwin cases clearly support, are far weaker claims. Specifically, we maintain that Kelvin/Darwin cases most clearly support the following two conclusions:

*The Capacity to Identify Cross-domain Confirmational Relations:* There are scientists who have identified some of their background beliefs from one domain that are relevant to the confirmation of some hypotheses from other domains, where domains are individuated along disciplinary boundaries.
The Capacity to Appreciate Cross-Domain Confirmational Relations: There are scientists who, when presented with information from one domain, have appreciated the relevance of that information to the confirmation of some hypotheses from other domains, where domains are individuated along disciplinary boundaries.

As we see it, Kelvin’s role in the Kelvin/Darwin case illustrates the first conclusion, whilst Darwin’s role illustrates the second. In contrast to Murphy’s claim that individual scientists have isotropic minds, the above claims are not—and should not be—contentious. As far as we can tell they have no implications for debates over extant theories of cognitive architecture since pretty much every party to these debates assumes that minds exhibit these capacities. Moreover, even if Murphy’s transcendental considerations and Kelvin/Darwin cases do have implications for theories of cognitive architecture, they would do so not because they establish any claims about isotropy, but because they support the weaker conclusions above. Or so we propose to argue. To see why we reject Murphy’s Transcendental Argument for Isotropy and parse Kelvin/Darwin examples in this way, we need to draw two crucial distinctions. Our contention is that whatever plausibility Murphy’s arguments for individual isotropic minds may appear to have derives from a tendency to collapse these distinctions. If we are right, then there is no reason to believe that isotropy at the individual level is a requirement for isotropy at the population-level, and thus no reason to believe that Population-level Globalism
satisfies Connectability.

**Distinction 1: Cross-Domain versus Isotropic Capacities**

Murphy’s contention, recall, is that in order for science to be global at the population-level, individual scientists must have isotropic minds, i.e., they must have the capacity to recognize the relevance of *any* item of information to the relevance of *any* hypothesis. Quite obviously, Kelvin/Darwin cases do not directly demonstrate this capacity. Rather, what such cases most obviously show is that some scientists, some of the time, are able to recognize the mutual significance of beliefs in *some* particular domain or disciplinary field with hypotheses in *another*. This weaker capacity we will refer to as a *cross-domain* inferential capacity. We readily acknowledge that there are lots of examples in which scientists exhibit such capacities. After all, this is precisely what Kelvin, Darwin, and all those who follow the debate succeed in doing. At the risk of stressing the crushingly obvious, however, there is an enormous difference between the claim that scientists have cross-domain inferential capacities and the claim that their minds are isotropic— that they can recognize the relevance of *any* belief to *any* hypothesis. Nor is there any obvious inferential connection from claims about the former to claims about the latter. That is, directly inferring isotropic minds from even a multitude of Kelvin/Darwin cases would be a gross error (It would be like inferring from the facts that John loves Jane, Eric loves Sue, etc … to the conclusion that everyone loves everyone as opposed to the conclusion that everyone loves someone!). In order to see why we don’t believe that Kelvin/Darwin
cases or Murphy’s Transcendental Argument for Isotropy provide reason to believe that scientists have isotropic minds, however, we need to enforce a second distinction.

*Distinction 2: Appreciation versus Identification*

In addition to the distinction between isotropic and cross-domain capacities, consider the following quite different kinds of cognitive achievement:

1. *Confirmational Identification*: The *identification* of confirmational relations between one’s background beliefs and a given hypothesis; and

2. *Confirmational Appreciation*: The *appreciation* of confirmationally significant information after this information has already been successfully sought out and identified.

This distinction is systematically elided in Murphy’s discussion, in large measure because he uses the term ‘recognition’ in such a way as to cover both sorts of activity. We maintain, however, that ‘recognition’ is ambiguous between identification and appreciation, and getting clear on this fact helps us see why isotropic scientific minds is not a precondition for the sort of science that we have, and are not supported by Kelvin/Darwin cases.
Let us start with the second cognitive achievement on display in the Kelvin/Darwin case: the appreciation of the relevance of information from one domain (e.g., astrophysics) for the (dis)confirmation of a hypothesis in another (e.g., biology). This task was accomplished by Darwin when Kelvin and his collaborators presented a set of salient propositions about the earth’s age (and an argument for its significance). Under these circumstances, Darwin inspected and duly appreciated the significance of information for his own theory that others had already identified. Does the fact that Darwin and other scientists appreciated such cross-domain confirmational relations support isotropic scientific minds? Such a claim of isotropy would entail that Darwin and other scientists’ cognitive processes were not architecturally restricted from accessing any potentially relevant information for the confirmation of a given hypothesis. And this would require processes that are highly unencapsulated— for example, processes that can exhaustively search memory and in principle bring any belief to bear on (dis)confirmation tasks. Contrary to what many appear to suppose, however, the phenomenon of confirmational appreciation has almost nothing to do with informational encapsulation and therefore provides little to no support for isotropic minds. Rather, the phenomenon bears on the extent to which scientific cognition is domain-general. More precisely, it concerns the extent to which processes of confirmational assessment can take inputs from disparate domains of inquiry. In our view, then, the phenomenon provides evidence for the existence of some domain-general cognitive system(s) that can integrate information across multiple domains. But it is important to be clear that this assumption is uncontentious. Indeed, every major theory of cognitive architecture of which we are
aware readily acknowledges this sort of domain-generality. For example, almost all theories allow for a domain-general working memory of the sort that might be implicated in solving the task of assessing an argument (Evans, 2010). Indeed, even the most radical advocates of modularity readily accept the existence of such systems (Sperber, 2004, p.14). Moreover, although working memory is domain-general in a way that would allow for the simultaneous consideration of information from remote domains of inquiry, it is not an isotropic system. On the contrary, all accounts of working memory assume that it is subject to heavy architectural restrictions (Baddeley, 2007), typically on the number of items whose activation is maintained for such purposes as conscious reasoning.

The present considerations undermine Murphy’s Transcendental Argument for Isotropy and his interpretation of Kelvin/Darwin cases in the following sense: In developing his argument for the claim that non-nugatory science presupposes isotropic minds, Murphy relies in part on the observation that scientists, such as Darwin, are able to recognize the significance of Kelvin’s claims for the theory of natural selection. But if “recognize” here means appreciate, then it is apparent that his observation does not militate in favor of isotropic minds. Instead, it supports the commonplace claim that our mental architecture contains at least one domain-general system, such as one involved in working memory. Thus, even though capacities for cross-domain appreciation may be a
precondition for the kind of science that humans have, such capacities are not at all contentious among theorists interested in human cognition. ⁹

**Confirmational Identification.** Let’s turn our attention to the second kind of cognitive task highlighted by the Kelvin/Darwin example, namely, the phenomenon of confirmational identification. Even if the *appreciation* of confirmational relations between information drawn from remote domains doesn’t require isotropic minds, it may still seem that the initial *identification* of such relations does. That is, one might believe that unencapsulated cognitive systems underwrote Kelvin’s identification of relevant information on solar cooling when he made confirmation judgments on evolutionary hypotheses about terrestrial organisms. Certainly, Kelvin/Darwin cases and the existence of non-nugatory science strongly suggests that any restrictions on a scientist’s capacities of identification, especially on her access to background theory, will not trace traditional disciplinary boundaries in the sciences. But this fact falls far short of providing support for isotropic scientific minds. Instead, all that appears to be required to account for the relevant phenomena are capacities of *cross-domain* identification—viz., the capacity to identify confirmational relations between *some* background beliefs from one domain for hypotheses in another. What sorts of restrictions on processes of identification might be consistent with such cross-domain capacities yet inconsistent with full-blown isotropy? In

---

⁹ Indeed, the sole kinds of cognitive architectures that appear to be incompatible with the appreciation of cross-domain relevance relations of the sort highlighted by Kelvin/Darwin cases are architectures that consist entirely of a variety of cognitive mechanisms, all of which are highly encapsulated and have restricted access to a proprietary database. While this resembles the late United States Senator from Alaska Ted Steven’s notorious description of the Internet as ‘a series of tubes’, we know of no theorist who has endorsed such a conception of our mental architecture.
short, many kinds. There are lots of possible architectural constraints on cognition that would allow for cross-domain identification without permitting isotropy. For example, processes of search and retrieval might be subject to constraints related to the recency of the information being accessed, the structure and organization of memory, the character of recall, and so on. These factors may constrain access to background beliefs, but they are not restrictions that necessarily mirror traditional disciplinary boundaries in the sciences, and they may not render any particular domain of inquiry inaccessible to any other. In which case, they may well permit cross-domain identification but fall far short of licensing isotropy.

Now we are in a position to make our response to Murphy’s Transcendental Argument for Isotropy and his invocation of Kelvin/Darwin cases. Our central claim is that the non-nugatory nature of science does not require individual isotropic scientific minds. All it requires is that there be a sufficiently large number of individual scientists who are capable of identifying and appreciating cross-domain confirmational relations. The collaborations and complex interactions between scientists who have cross-domain inferential capacities may, as Murphy claims, give rise to isotropy at the population-level. But we cannot in this context presume to aggregate over such achievements by the community of science, and infer that each individual scientific mind is in principle capable of attaining all of them. That would be to assume precisely the conclusion of what The Transcendental Argument For Isotropy was supposed to show, namely, that isotropy must obtain at the individual-level in order to obtain at the population-level. Since we should not assume this claim in order to argue for it, the upshot is that there
isn’t any convincing reason to believe that scientific minds are isotropic or that Population-level Globalism satisfies Connectability. The upshot for Population-level Globalism is that even assuming its plausibility, the thesis provides no basis for rejecting prominent classicalist, connectionist, or modularity theories of our mental architecture.

The reader may be concerned that we have not given full consideration to the evidence that examples such as the Kelvin/Darwin case may provide for every version of Globalism. We agree. And we discuss in more detail the evidence that performances of scientific inferences, like those illustrated by the Kelvin/Darwin case, may seem to provide for Performance Globalism in section 3.4. Our purpose in this section has been to rebut Murphy’s Transcendental Argument for Isotropy and argue that Kelvin/Darwin cases do not illustrate that individual isotropic minds are a precondition for having the kind of science that we have.

### 3.3. Competence Globalism

Let’s turn now to Competence Globalism – the thesis that Quineanism or isotropy are properties of a scientist’s inferential competence. Again, we will for the sake of argument grant that the thesis satisfies Plausibility and focus instead on assessing whether the thesis satisfies Connectability. As we see it, Competence Globalism, even if true, provides us with no reason to endorse claims about our cognitive architecture. Nonetheless, in order to make our case, we need first to consider an argument that appears to show precisely the opposite. In what follows, we first set out this argument. We then highlight a problem with the argument, and explain why this problem suggests a general difficulty with
inferences from claims about cognitive competences to claims about cognitive architecture. Before spelling out the argument it is important to be clear on the notion of a competence that is relevant here. Notoriously, the notion of competence has been used in a variety of non-equivalent ways by cognitive scientists (Stein, 1996). In many contexts, competences are construed as bodies of mentally represented information. For example, a speaker’s competence for a natural language is sometimes characterized as an internally represented grammar of the language – an integrated set of generative rules and principles that entail an infinite number of claims about the language (Chomsky, 1965; 1986). But this notion of a competence is not the one that should be adopted for the purposes of assessing Competence Globalism. If it were, then Competence Globalism would very obviously fail to show that our cognitive architecture allows for Quinean or isotropic reasoning; and this is because these are properties of inferences as opposed to bodies of information. Rather, in the context of assessing whether Competence Globalism satisfies Connectability, the most plausible construal of ‘competence’ is something like ‘performance under idealized circumstances’ where such circumstances are taken to abstract from a wide range of factors, such as memory limitations, distractions, shifts of attention and interest, and errors (random or characteristic) that routinely affect actual performance.

With this clarification in hand, consider the following argument for the claim that Competence Globalism satisfies Connectability:

(1) Theories of scientific reasoning competence characterize, at least in part, what a cognitive system can do under idealized conditions.
(2) In which case, if Competence Globalism is true, then the cognitive systems that underwrite scientific reasoning, can, under idealized conditions, engage in global inferences.

(3) But a theory of cognitive architecture just is an account of what a cognitive system can do. It describes the possible processes allowed by the structure of the brain.

(4) Therefore, if Competence Globalism is true, then there is reason to believe that our cognitive architecture can subserve Quinean or isotropic inferential processes.

The above argument may appear to establish that Competence Globalism satisfies Connectability. After all, premise 1 is true more-or-less by definition; 2 clearly follows from 1; and as noted earlier, a central function of a theory of cognitive architecture is to delimit the ‘set of possible cognitive processes that are allowed by the structure of the brain’ (Pylyshyn, 1996). In which case, if Competence Globalism is true —and hence captures performance under idealized conditions— then our cognitive architecture must apparently allow for Quinean and isotropic inferences; which is precisely the conclusion that Fodor requires.

For all that, the argument is invalid because it turns on an equivocation. Competence theories and theories of cognitive architecture both, in a sense, concern what we can do —what our cognitive capacities are. But the modality in question —the relevant sense of ‘can’— is not the same in the two cases. Specifically, the modality
relevant to competence theories is significantly less restricted in that it abstracts away from constraints that are central to theories of cognitive architecture. This, we maintain, gives rise to an equivocation in premises (2) and (3) such that the conclusion in (4) fails to follow. One way to see the point is to notice that the performance-competence distinction cross-classifies with the distinction between aspects of cognition that are architectural and those that are not. In particular, on standard ways of drawing the competence-performance distinction much of what comprises a theory of cognitive architecture is centrally concerned with performance issues. As a consequence, the idealizations associated with cognitive competences need not be accommodated by theories of cognitive architecture.

Some well-known examples from linguistics illustrate these different idealizations. It has long been known that there are sentences that speakers judge as unacceptable even though linguists frequently conclude that they are sanctioned by our linguistic competence. One such phenomenon involves so-called Garden Path sentences, for example:

The horse raced past the barn fell.

Another involves the phenomenon of center-embedding:

The president whom the prime minister who the mob that the FBI threatened owned insulted cried.
Though not typically judged acceptable by English speakers, such sentences are plausibly well-formed according to theories of a speaker’s grammatical competence. But plausible explanations for performance failures in recognizing their well-formedness routinely appeal to limitations on the speaker’s short-term memory-buffer or to limitations on the attention mechanisms subserving intuitive linguistic judgments (Fitzgerald, 2010). Importantly, such explanations of these failures turn on claims about cognitive architecture. That is, the explanations posit invariant and cognitively impenetrable features of our minds in order to explain the fact that typical speakers fail to recognize the grammatical properties of these sentences. But if this is so, then theories of cognitive architecture cannot be making the same idealizations as competence theories do. Indeed, rather than specifying our capacities in such a way as to abstract away from performance limitations, theories of cognitive architecture seem centrally dedicated to specifying and explaining these limitations. To put the point another way, just because center embedded and Garden Path sentences are grammatical —sanctioned by our competence— does not mean that a theory of cognitive architecture should incorporate the sorts of memory and attentional capacities that would be required to parse such sentences. On the contrary, in relation to such cases, a central function of a theory of cognitive architecture is typically to explain why speakers often do not, and in some cases cannot, parse such sentences as grammatical.

Our discussion of linguistic competence implies a corresponding moral for theories of scientific reasoning competence, and for Competence Globalism’s prospects.
for satisfying Connectability. The moral is this: Even if global inferences were part of the inferential competence of scientists, it does not thereby follow that the cognitive architecture of a scientist should allow her to make such inferences. On the contrary, the idealization under which a scientist were able to perform Quinean or isotropic inferences may well be one in which she had an entirely different cognitive architecture. Thus, even if Quineanism or isotropy are features of a scientific reasoner’s competence, that fact alone does not provide reason to believe that the cognitive architecture on which scientific reasoning depends permits Quineanism or isotropy. For all Fodor has argued, Quinean and isotropic inferences might violate architectural constraints on memory, attention, or other cognitively impenetrable limitations that restrict performances of scientific reasoning. Indeed, we have given some reason in our discussion of Normative Globalism to suppose that global reasoning might well exceed the limitations imposed by a scientist’s cognitive architecture.

Let us be clear, however, about what we take the above considerations to show. The above argument does not, of course, establish that global inferences, in Fodor’s sense, necessarily conflict with the invariant and cognitively impenetrable features of a scientist’s cognitive architecture. This, we take it, remains a largely open empirical question. Rather, our point is that Competence Globalism fails to satisfy Connectability: Its truth would not provide good reason to believe that our cognitive architecture allows for Quinean or isotropic processing. In view of this, we conclude that claims about the globality of scientific reasoning competence are not an appropriate basis from which to criticize extant theories of our cognitive architecture.
3.4. Performance Globalism

The final version of Globalism we consider here is Performance Globalism: the thesis that the actual inferential processes of individual scientists are Quinean or isotropic. In contrast to the three forms of Globalism discussed above, it is not implausible that Performance Globalism satisfies Connectability. To deny that Performance Globalism satisfies Connectability would in fact require believing that the minds of scientists are discontinuous with the rest of human minds in quite fundamental respects. Specifically, it would require that the cognitive architecture of scientists allows for cognitive acts that are impossible for everyday reasoners. Now there is, to be sure, significant room for debate concerning the overall similarity between everyday cognition and scientific reasoning (see e.g., Knobe, 2010). But so far as we know, there is as yet no evidence whatsoever that the minds of scientists have a different architecture from the minds of quotidian reasoners. Indeed, much of the evidence—which we discuss briefly in section 4—points in the opposite direction. Thus, we will suppose that the balance of evidence is quite compatible with the view that the minds of scientists and non-scientists are, in fundamental respects, much alike. In view of this, we propose to grant, at least for the sake of argument, that Performance Globalism satisfies Connectability.

But does Performance Globalism satisfy our other constraint – Plausibility? We maintain that the answer is ‘No’. In order for Performance Globalism to satisfy Plausibility, it would need to be plausible that individual scientists in fact engage in
genuinely global inferences. But there is no good reason, we maintain, to believe that this is so. Notice that it follows from our previous discussions in sections 3.1 – 3.3 that Performance Globalism does not enjoy support from the mere observation that theoretical virtues such as simplicity, conservativeness, and the like apparently guide non-demonstrative scientific inference. Such an observation does not by itself render Performance Globalism plausible because we cannot assume that non-idealized accounts of non-demonstrative inferences will define, as Fodor does, simplicity, conservativeness, etc. over entire belief systems. What are instead required to establish Performance Globalism’s plausibility are empirical considerations that support the globalism of actual inferences by individual scientists. With this point in mind, we turn to rebutting what we take to be the three main arguments that Fodor offers for Scientific Globalism (Fodor 1983, 2000, and 2008).\textsuperscript{10} The upshot is that even though Performance Globalism is the only version of Scientific Globalism that satisfies Connectability, Fodor’s central arguments for Scientific Globalism do not begin to establish that Performance Globalism satisfies Plausibility.

3.4.1 \textit{The Argument From Confirmation Holism.} Fodor’s first argument rests on the claim that confirmation holism (alternatively, the ‘Quine-Duhem’ or ‘Q-D’ Thesis) enjoys broad support among philosophers of science:

\footnote{Space prevents us from considering here every argument Fodor has made that might appear to support Performance Globalism. In particular we leave aside Fodor’s arguments that borrow from Goodman’s New Riddle of Induction (1983; 1987) and that invoke analogical capacities (1983). We note that these arguments, though important to earlier work, have not persisted into Fodor’s later treatments, in contrast to the arguments we address here.}
Recent discussion of confirmation (from, say, Duhem forward) have increasingly emphasized the holism of nondemonstrative inferences by claiming, in the limiting case, that whole theories are the proper units for the evaluation of such inferences … (Fodor, 2008, p.123).\footnote{See also Fodor, 1983, p.111.}

But if the Q-D Thesis is true as a description of actual scientific inferences, then, so the argument continues, performances of scientific hypothesis confirmation are either Quinean, isotropic, or both. In which case, we would seem to have some sort of evidence for the plausibility of Performance Globalism.

In fact, however, Performance Globalism does not enjoy support from recent work on confirmation theory. First, in order for endorsements of confirmation holism to support Performance Globalism, the putative consensus would need to concern the inferential practices of individual scientists. But even supposing a widespread endorsement of confirmational holism by philosophers of science, there is not the slightest reason to suppose that the claim being endorsed should be construed as a thesis about the performances of individual scientists as opposed to a normative, population-level, or competence claim. In which case, any consensus in favor of Q-D would not provide reason to believe that Performance Globalism satisfies Plausibility.

Second, on our reading of the literature, Fodor grossly exaggerates any consensus in favor of the Q-D Thesis, especially any readings that would support Performance
Globalism. In order to appreciate the concern, consider the following formulation of the Q-D Thesis:

Strong Q-D: Only entire belief systems stand in the confirmation relation to evidence E.

Suppose that this were a standard formulation of the Q-D Thesis. Moreover, suppose that it was intended to capture the confirmational practices of individual scientists. Under such an assumption, there would be a plausible inference from Strong Q-D to Performance Globalism. Nevertheless, we maintain that insofar as there is any consensus regarding confirmation holism it is certainly not for Strong Q-D, but for weaker versions of the thesis. For example, we suspect that the following view instead has quite widespread support:

Weak Q-D: Only a hypothesis conjoined with auxiliary statements stands in the confirmation relation to evidence E.

In fact, on our reading, Weak Q-D, or some other substantially weaker version of Q-D, is far more likely than Strong Q-D to be what contemporary philosophers of science typically have in mind when discussing confirmation holism (at least in the realist literature). So, for example, Sober defines the Q-D Thesis thus:
… it isn't the hypothesis $H$ alone that predicts whether [observation] $O$ will be true; rather, it is the conjunction $H \& [\text{auxiliary assumption}] A$ that has this implication. The controversial … epistemological point that Quine (1953) famously defended in ‘Two Dogmas of Empiricism’ is that what gets confirmed and disconfirmed by observations is not $H$ taken by itself, but the conjunction $H \& A$ (Sober, 1999, p.73).12

Weak Q-D, however, does not support any version of Scientific Globalism and certainly not Performance Globalism in particular. This is because Weak Q-D is compatible with only a small subset of a background theory ever informing a given confirmation judgment. Thus, neither Quineanism nor isotropy need be instantiated by cognitive processes that respect the form of confirmational holism with the most currency in the philosophy of science. We conclude that recent views on the putatively holistic nature of scientific confirmation provide no reason to suppose that Performance Globalism satisfies Plausibility.

3.4.2 The Teleological Argument. The second of Fodor’s arguments for Scientific Globalism that we will consider is teleological in character and rests on assumptions

12 Sober in fact denies even Weak Q-D and claims that hypotheses can be tested singly in many cases (1999). For similar formulations of the Q-D thesis that are much closer to a weak as opposed to a strong reading of Q-D, see Strevens, 2001, p.44 and Ariew, 1984, p.319. Indeed, even Quine is plausibly read as having eventually rejected Strong Q-D.
concerning the proper function of science. The argument proceeds from a claim about the nature of the world we inhabit:

(1) The structure of the world is a connected, causal system (Fodor, 1983, p.106).

Next, Fodor maintains that our epistemic predicament with respect to this causal structure is precarious:

(2) ‘[W]e don’t know how the connections are arranged’ (ibid, p.106, emphasis Fodor).

This predicament, Fodor holds, is precisely what science was designed to cope with, since:

(3) Science, ‘broadly construed,’ is an attempt to track the causal structure of the world (1983, pp. 105-6).

Fodor appears to conclude from these three premises that isotropy is a property of scientific hypothesis confirmation:

Confirmational isotropy is a reasonable property for nondemonstrative inference to have because the goal of nondemonstrative inference is to determine the truth
about a causal mechanism – the world – of whose workings we are arbitrarily ignorant. That is why our institution of scientific confirmation is isotropic ...
(Fodor, 1983, p.105).

In general, we are not fans of teleological arguments (neither, for that matter, is Fodor in most contexts). If, however, one is going to run an argument whose premises appeal to the proper function of a cognitive process, the account of this functioning should not be biased solely in favor of accuracy. In the general case, an account of the proper function of a cognitive process should incorporate additional and competing desiderata relating to, for example, speed, energy, time, and the use of other cognitive resources. And there is no reason to believe that a more comprehensive teleological account — for example, one that factors in feasibility and efficiency — will also incorporate Quineanism or isotropy.

Given the bias toward optimality in Fodor’s remarks on the function of non-demonstrative scientific inference, Fodor’s teleological considerations are simply too weak to meet a reasonable burden of proof for positing global processes.

An analogous and plainly fallacious line of reasoning regarding the proper function of vision should make this clear:

The purpose of initial visual processing is to detect edges and light contrasts in a world of whose shapes and shadows we are arbitrarily ignorant. A flexible and unencapsulated visual system would be optimally suited for such detection, given that we don’t know in advance how any given distal layout is arranged.
Should we conclude from this argument that initial visual processing is likely isotropic and therefore does not rigidly follow heuristics? Plainly not, and certainly not by Fodor’s own lights, as he has argued elsewhere that visual illusions strongly support the encapsulation and inflexibility of early vision (Fodor 1983). The general moral is that any account of a cognitive system’s proper function that solely incorporates optimality does not legitimate conclusions about the character of that cognitive system. This holds in particular for Fodor’s remarks on the proper function of the cognitive systems underwriting scientific reasoning. Since his Teleological Argument omits desiderata other than accuracy, it fails to provide reason to believe that Performance Globalism satisfies Plausibility.

3.4.3. The Argument from Remoteness. The last argument that we consider is based on considerations we have already encountered in our discussion of the Kelvin/Darwin case in section 3.2. As we have remarked, such an example most obviously shows that scientists have cross-domain inferential capacities. What is less clear is whether such examples could also support isotropic scientific minds. There is an argument, however, suggested in remarks by Fodor that might appear to take us from claims about cross-domain inferential capacities of the kind illustrated by Kelvin/Darwin cases to conclusions about isotropy. We call this The Argument from Remoteness and find indications that Fodor endorses it, for example, when he writes: ‘In principle, our botany constrains our astronomy, if only we could think of ways to make them connect’ (Fodor,
In fact, the Kelvin/Darwin case suggests that some scientists can make such connections, since Kelvin connected up claims about solar radiation with claims about the evolution of terrestrial organisms, including those studied by botanists. The following characterization of The Argument from Remoteness might appear to provide reason to believe in isotropic scientific minds given that scientists can draw particular kinds of cross-domain connections:

_The Argument from Remoteness._ In cases such as the Kelvin/Darwin example it would seem that the pieces of information that are brought to bear on the (dis)confirmation of Darwin’s theory are recruited from a domain — astrophysics — that is exceedingly _remote_ from terrestrial biology. Surely (so the line of thought goes) if a scientist like Kelvin can connect up items of information that are _this_ remote from each other, then it is unlikely that there are _any_ architectural constraints on what items of information Kelvin could connect. Further, since Darwin, Huxley and everyone else was able to _recognize_ the problem — and moreover that there are many other similar cases — it cannot be that Kelvin’s achievement was an improbable and isolated occurrence. Rather, it would seem to be a pervasive capacity of scientists to draw such connections between arbitrarily ‘remote’ items of information. In which case, it might seem that some performances of scientific inference support the claim that scientists have isotropic inferential capacities.
We begin our response to The Argument from Remoteness, and the support it may appear to provide for Performance Globalism, by disposing of a red herring. We suspect that part of the intuitive attraction behind some claims of isotropy is that, when confronted with Kelvin/Darwin cases, most scientists and laypeople are able to recognize the possibility that epistemic relations may hold between propositions from different domains. And with some extrapolation it might be tempting to maintain that:

Propositions from any domain (however individuated) can, in principle, bear epistemic relations —paradigmatically confirmational relations— to any other proposition, including those in vastly different domains.

Call this propositional isotropy. The question of whether propositional isotropy characterizes science as a body of knowledge is, we think, an interesting matter. But for present purposes —and this is the crucial point— it is largely irrelevant. What’s at issue here does not concern the potential range of confirmational relations that can hold between propositions. This is an issue for epistemologists and confirmation theorists. Rather, we are focused on causal-explanatory issues in empirical psychology, and, as such, we are concerned with the potential range of causal relations that can hold between mental states. In view of this, the relevant construal of isotropy is what we call cognitive isotropy:
Individual scientists are able (where relevant) to bring any background belief to bear on the confirmation of any hypothesis.

In other words, cognitive isotropy does not merely assert, as propositional isotropy does, that any proposition can in principle bear epistemic relations to any other. It further maintains that these epistemic relations between propositions are “mirrored” by the causal relations between the actual mental states of scientists, and thus that the cognitive processes subserving judgments of confirmation are radically unencapsulated. As an illustration of this distinction, consider that it is one thing to accept that all propositions bear relations of consistency and inconsistency to each other. But it is quite another to claim that humans can, for any proposition, determine whether it is (in)consistent with any other. Conflating the two claims would entail that no human has ever unknowingly held contradictory beliefs. Our present point is that even if one endorses a claim about the propositional isotropy of science, a claim about the cognitive isotropy of individual scientists remains an almost entirely open matter.

With this distinction in mind, what do Kelvin/Darwin Cases tell us about the casual relations that can hold between the mental states of scientists? In order to answer this question, we will again invoke the distinction that we made in section 3.2 between appreciation and identification. As with Murphy’s Transcendental Argument for Isotropy, we believe that this distinction will aid in removing the force that The Argument from Remoteness may appear to martial. First, consider cross-domain
appreciation: The fact that information from arbitrary domains of inquiry, no matter how remote, can be appreciated by the minds of individual scientists should not lead us to posit isotropic processes. Rather, what appreciation of this sort supports is the existence of domain-general cognitive systems, such as those involved in working memory tasks. But as we have already remarked, none of this is at all contentious amongst theorists interested in human cognitive architecture and none of it supports the claim that the minds of scientists, or any other individuals, are cognitively isotropic.

What, however, does The Argument from Remoteness and accompanying Kelvin/Darwin case say about processes of confirmational identification? Here we believe the issue is subtler. Certainly the phenomenon of identifying, in the first place, remotely related pieces of information when confirming hypotheses is consistent with cognitive isotropy. However, this phenomenon is also apparently consistent with the existence of cognitive processes that have the weaker capacity for cross-domain identification but that are not cognitively isotropic. We suspect that there is at present no way to conclusively adjudicate between these competing accounts. Nevertheless, we offer a prima facie consideration in favor of merely positing cross-domain inferential capacities as opposed to cognitively isotropic scientific minds.

On a natural reading, accounts that posit capacities of cross-domain identification and accounts that posit cognitive isotropy yield different predictions regarding how commonplace cross-domain confirmation from relatively isolated areas of inquiry should be. If scientist’s minds were cognitively isotropic, one should expect it to be quite commonplace for scientists to identify ‘remote’ connections between evidence and
hypotheses. In contrast, if scientists merely have cross-domain inferential capacities, one should not expect scientists to invariably or routinely identify ‘remote’ connections across seemingly isolated domains. In our view, the evidence appears to conform to the prediction made by accounts that posit merely cross-domain identificational capacities. As an example of the kind of evidence that we believe supports our view, consider the discovery of penicillin, which highlights that scientists do not invariably identify available background information that is relevant to the formation and confirmation of hypotheses. Cherniak’s (1986) discussion makes precisely this point:

At least a decade before Fleming’s discovery of penicillin, many microbiologists were aware that molds cause clear spots in bacteria cultures, and they knew that such a bare spot indicates no bacterial growth. Yet they did not consider the possibility that molds release an anti-bacterial agent (Cherniak, 1986, p.50).

Cherniak reasonably speculates that the belief that mold causes bare spots in Petri dishes was ‘filed’ in the memory of individual scientists ‘under the category of practical laboratory lore’ on undesirable contamination (1986, p.51). On this plausible assumption, relevant information related to the practical use of maintaining culture-growth in a laboratory was not identified and recruited for the purpose of forming and confirming hypotheses in microbiological theory. Rather, in this case, practical information regarding molds was not ‘connected up’ with microbiological hypotheses. Famously, the connections were instead drawn as a result of causal interactions in the world rather than
causal transitions among Fleming’s mental states. That is, Fleming mistakenly left open a Petri dish containing a *Staphylococcus* plate culture, the culture became contaminated, and only afterwards did Fleming pursue anti-bacterial medicine derived from the *Penicillium* fungi.\(^{13}\) In this case, many scientists, including the inventor of penicillin, did not identify background information that was relevant to the projection and acceptance of a scientific hypothesis.

Naturally, such an example fails to conclusively demonstrate that cognitive isotropy is not a property of a typical scientific mind, since the mere fact that Fleming and others did not trace relevant epistemic relations between propositions does not show that they were architecturally restricted from doing so. Our real point is this: If the present considerations don’t count against cognitive isotropy, then neither does the history of science count in its favor either. In particular, if the failure by Fleming and others to identify relevant information does not speak against cognitive isotropy, then the Kelvin/Darwin case and accompanying Argument from Remoteness does not support it. Of course our denial of cognitive isotropy is an empirical claim and as such might be defeated by the production and analysis of compelling evidence regarding the inferential capacities of scientists. By our lights, however, such evidence has simply not been provided. The balance of evidence appears to provide no reason whatsoever to believe that the mental architecture of individual scientists licenses cognitively isotropic processes; instead, it appears to merely the support the weaker claim that scientists have

---

\(^{13}\) Cherniak draws an anti-globality conclusion from this and similar examples, namely, that ‘the web of belief is not merely tangled; its fabric … is ‘quilted’ into a patchwork of relatively independent subsystems’ (1986, p.51). For our purposes, only the weaker claim that Kelvin/Darwin examples do not support cognitive isotropy is necessary and that is all we maintain here.
capacities for cross-domain appreciation and identification. If our read of the evidence is warranted, then The Argument From Remoteness provides no reason to accept that Performance Globalism satisfies Plausibility.

We conclude, then, that three central arguments found in Fodor’s work which might appear to support Performance Globalism in fact fail to do so. In which case, Performance Globalism fails to satisfy Plausibility and therefore has no implications for theories of cognitive architecture. Our overall assessment of the prospects for any version of Scientific Globalism to inform theories of cognitive architecture is similarly pessimistic. Despite playing a central role over the past several decades in Fodor’s criticisms of classicism, modularity theory, connectionism and their associated hybrids, there does not seem to be any reason whatsoever to accept that any version of Scientific Globalism—including Normative, Population-level, Competence, and Performance Globalism—undermines extant theories of cognitive architecture. As we see it, Fodor has simply failed to make a range of appropriate distinctions among different theories of scientific inference and failed to acceptably track their implications for cognitive science.

4. On the Proper Treatment of the Relationship Between Scientific Inference and Ordinary Cognition

In this section, we explore more plausible connections between theories of scientific inference and theories of ordinary cognition. Though we earlier criticized Fodor’s
specific attempts to draw conclusions about ordinary cognition from claims about the
globality of scientific inference, we suspect that Fodor is broadly correct in thinking that
explicit inductive practices in science can illuminate aspects of ordinary cognition that
are inaccessible to conscious reflection. In fact, we believe our previous discussion points
in the direction of such connections, though we will restrict ourselves to a few brief
comments here that enjoy only prima facie plausibility.

First, there are relatively uncontroversial relations that likely hold between
theories of scientific inference and theories of ordinary cognition. For example, extant
research suggests that accounts of scientific inference can provide valuable formal
resources as well as analogical source material for the construction of substantive
hypotheses regarding the nature of ordinary learning and everyday reasoning. These types
of relatively uncontroversial connections might be illustrated by recent fruitful work on
interventionist theories of causal inference and related research on Bayesian networks (of
which more later), or by accounts in developmental psychology of the bodies of
information that children can learn that have drawn inspiration from structural
descriptions of scientific theories.

Second, however, there are more controversial connections— for example
plausibility arguments that purport to draw conclusions for the nature of ordinary
cognition from claims about the nature of scientific inference. We believe a general
framework for assessing such plausibility arguments can be generated from a more
precise diagnosis of where, precisely, Fodor’s plausibility argument concerning the
globality of central processing went wrong. In Section 3 we argued that there is no reason
to believe that any version of Scientific Globalism has implications for theories of
cognitive architecture. But notice that the majority of our objections did not turn on
particular claims about the *globality* of scientific inference per se; rather, our objections
turned largely on general difficulties associated with drawing conclusions about one type
of inference from claims about inferences of another type. Thus, Normative Globalism’s
failure to satisfy Connectability turned largely on general obstacles to drawing
descriptive conclusions from claims about inferential norms. Similarly, Population-level
Globalism’s failure to satisfy Connectability turned largely on the general failure of
population-level theories to have implications for individual-level claims, including
claims about individuals’ inferential capacities. And Competence Globalism failed to
plausibly entail conclusions for theories of cognitive architecture because, in general,
competence theories of reasoning invoke unique kinds of idealizations in comparison to
other kinds of theories of cognition. Our criticisms of Performance Globalism, in
contrast, did not follow this pattern precisely because theories of cognitive architecture
and performance theories are both dedicated to descriptively characterizing, with similar
levels of idealizations, the actual cognitive achievements of individuals. Therefore, no
general complications with drawing implications across types of theories of cognition
arose.

This pattern suggests a general negative moral or rule of thumb for making
plausibility arguments concerning the nature of scientific inference and ordinary
reasoning:
The Negative Levels Principle: In the general case, one cannot render plausible inferential theories of one type (e.g., concerning ordinary cognition) on the basis of inferential theories of another type (e.g., concerning scientific inference). According to this principle, population-level theories of inference, as a rule, won’t render claims about individuals’ inferential processes plausible; normative claims about the nature of inference won’t, as a rule, render claims about population-level inference plausible; and so on. We strongly suspect that The Negative Levels Principle has important implications that go well beyond Fodor’s work. For example, one of us has argued elsewhere that a similar conflation between different types of theories of scientific inference and theories of ordinary cognition plagues a prominent version of the ‘theory theory’ in developmental psychology, according to which early cognitive development and scientific theory change are underwritten by most or all of the same learning processes. On this view, proponents of this version of the theory theory have, in effect, flouted The Negative Levels Principle by drawing inferences from claims about population-level processes of scientific theory change to claims about the processes that underwrite cognitive development in individual children (Fuller, forthcoming).

This general negative moral, however, also suggests a positive framework for appreciating the most plausible and fruitful connections between theories of scientific inference and other kinds of theories of cognition. In particular it suggests:

The Positive Levels Principle: As a rule, to the extent that one can render plausible an inferential theory with reference to another theory, the two theories will be of the same type.
According to this principle, the plausibility arguments that are most likely to be successful are those that trace connections between, for example: normative theories of scientific inference and normative theories of everyday inference; performance theories of scientific inference and performance theories of commonsense inference; and so on. Is The Positive Levels Principle itself plausible? The framework it suggests rests contingent on many difficult and unresolved issues in epistemology and empirical psychology. Consequently, our support for it will rest on no more than defeasible claims that enjoy prima facie reasonableness. However, we do believe there are examples that provide inductive support for the idea that research which traces connections between different types of theories of scientific inference —normative, population-level, competence, and performance— and corresponding types of theories of ordinary cognition is likely to be fruitful.

As one such example, consider recent work on models of causal learning, including research on Bayesian networks. The associated formalism was developed partially from underpinnings in the philosophy of science (Glymour, Spirtes Scheines 1993) and for some time has plausibly been taken to model central aspects of inductive scientific inference concerning causal systems, observed correlations, and experimentation (Pearl 2000). Even their most enthusiastic proponents do not claim that Bayes nets are fully general models of scientific causal reasoning. However, Bayes nets are reasonably regarded as having had successes as models of core aspects of scientific inference, and in particular there are some reasons for viewing Bayesian networks as contributing to theories of scientific reasoning competence. First, Bayes nets have been
treated by their advocates as computation-level descriptions (Gopnik & Glymour 2006) in Marr’s sense, where Marr modeled this level of description after Chomsky’s competence theories in linguistics (Marr 1982). Second, in many contexts, Bayesian networks efficiently process massive data sets in ways that far outstrip the performance capabilities of human reasoners. This is in fact one of their central purposes as expert systems (Glymour & Cooper 1999). Thus, Bayes nets appear to be associated with idealizations that are most connected to theories of cognitive competencies. On the assumption that Bayes nets models aspects of scientific reasoning competence, The Positive Levels Principle implies that Bayes nets are also likely to have applications for competence theories of ordinary reasoning. And in fact this prediction appears to have been born out, given that the Bayes net formalism has been plausibly employed to model features of childhood and adult causal reasoning, in particular reasoning that is sensitive to (conditional) dependencies and observed interventions (Gopnik & Schulz 2007; Gopnik et al. 2004). The success of this recent research suggests that one and the same formalism likely characterizes central features of idealized inferential capacities associated with both scientific and ordinary reasoning competences. Further, the successes of Bayes nets as psychological models speaks to the general productivity of exploring connections between competence theories for both ordinary and scientific cognition. And we believe it provides (albeit weak) inductive support of the general framework suggested by The Positive Levels Principle.14

14 In fact, one of us has argued that Bayes nets likely do not model as many aspects of scientific reasoning as they do of childhood learning (Fuller, under review). However, this fact does not undermine the claim
Our final reflections concern what we view as two related issues:

a) Why have many theorists, perhaps including Fodor, found some form of Globalism about both scientific inference and everyday cognition plausible?

b) Are there in fact any plausible versions of Scientific Globalism that have correspondingly plausible analogs with respect to ordinary cognition?

As we see it, these issues are related because an answer to the second provides an answer to the first; and moreover, the fact that many have found plausible some form of Globalism about both scientific and ordinary cognition provides evidence relevant to addressing the second. Let us first be clear about what we are not arguing. Though in sections 3.1 – 3.3 we granted, for the sake of argument, that Scientific Globalism was plausible on normative, population-level, and competence readings, we are as a matter of fact quite uncertain whether any of these theses is in fact correct. However, we do believe that Normative Globalism and Competence Globalism enjoy at least prima facie plausibility; and further that if plausible, corresponding normative and competence claims about the globality of quotidian cognition are plausible to roughly the same extent. Though we won’t argue for these conclusions in any detail, we propose to close by considering some considerations that militate in favor of both claims.

that Bayes nets have significant applications as models for both types of cognition. In addition, with respect to everyday causal reasoning competence in adults, there is reason to believe that Bayes nets are equally plausible as models of this kind ordinary cognitive competence as they are of scientific reasoning competence.
Let us start by considering why we take Normative Globalism to enjoy at least some *prima facie* plausibility. As noted in section 3.1, it is exceedingly commonplace for well-articulated normative theories of inductive inference to incorporate global constraints of some sort. So, for example, both Carnapian and Bayesian approaches to probabilistic inference incorporate a requirement of total evidence; and many discussions of normative abductive inference presuppose that theoretical virtues—such as simplicity and consistency—are defined with respect to entire belief systems. Again, let us stress that we are quite unsure that Normative Globalism is in fact the right normative commitment to adopt for characterizing of scientific inference. But it is not hard to discern from where the plausibility of such commitments derives. Normative theorists who cleave to such commitments are tracking a strong intuition regarding conditions on good scientific inference—an intuition that seems to apply to general principles of inference, but is also supported by reflection on specific actual and counterfactual cases. So for example, on the face of it, the fact that \( H \) is probable given some subset of my evidence, \( E \), seems insufficient for endorsing \( H \), if I possess additional relevant evidence that renders \( H \) improbable. Hence the attraction of total evidence. Analogously, the fact that \( H \) is consistent with some proper subset of my beliefs seems inconsequential if I also possess other beliefs inconsistent with \( H \). *Mutatis mutandis* for simplicity. Our point is not that these (putative) global normative demands should be adopted but that the face plausibility of normative constraints are strongly dependent on intuitions about the quality of inferences, and that, as a matter of fact, many theorists concerned with normative theories of scientific inference have tended to feel the intuitive pull of some
form of Normative Globalism. Whether or not the intuitive plausibility of Normative Globalism should be resisted is, we think, a complex issue; and neither of us is inclined to maintain that intuition alone resolves the matter. Among other factors, we suspect that issues about the status of OIC principles must be resolved before any great confidence can be placed on the intuitions that support Normative Globalism. But for present purposes our point is simply that Normative Globalism has a \textit{prima facie} plausibility in virtue of its intuitive support that warrants taking it seriously.

If Normative Globalism is \textit{prima facie} plausible, then in our view so too is Competence Globalism. In fact, we believe that Competence Globalism’s \textit{prima facie} plausibility is supported by the very set of intuitions about (good) scientific inference that also lent Normative Globalism its face plausibility. Why believe that one and the same collection of intuitive judgments supports both Normative and Competence Globalism? This view is in fact supported by the widespread acceptance—at least since the inception of contemporary linguistic research—that intuitive judgments constitute a central source of evidence not just for normative theories of inference, but also for theories of cognitive competences (Chomsky 1965, Devitt, 1997; Stich, 1990). (Indeed, according to some theorists, intuitive judgments are the \textit{only} relevant evidence for both kinds of theories (Cohen, 1981).) But if this is so, then intuitive judgments about the quality of particular (scientific) inferences lend credence to claims about the character of the inferential competence that underlies these intuitive judgments. In our view, then, Competence Globalism is \textit{prima facie} plausible for much the same reasons that Normative Globalism is.
Note that the above considerations, if correct, might well explain why so many — Fodor included — have found some form of Globalism about scientific inference attractive. The reason is that some form of Scientific Globalism is at least *prima facie* plausible — viz. the normative and competence versions. Further, we speculate that this might go some way toward explaining the widespread attractiveness of claims about the globality of scientific inferential *performance* as well. After all, intuitive judgments in general don’t come clearly labeled! In particular, we suspect that intuitive judgments regarding the quality of an inference — intuitions that derive from an inferential *competence* — and unreflective empirical judgments regarding a *process* of inference are easily confounded. Indeed, one central task of this paper has been to focus attention on precisely these distinctions, and on what we believe is a pervasive tendency to elide them.

Let us return to the issue of whether any forms of Scientific Globalism have a place in the framework suggested by The Positive Levels Principle; that is, whether Normative and Competence Globalism have similarly plausible analogs for ordinary cognition. According to the Positive Levels Principle this should be so; and as it happens, we maintain that it is. For it seems that precisely analogous considerations to the ones that supported Competence and Normative Globalism also support corresponding normative and competence claims about everyday reasoning. Consider first normative theories of everyday inference. As with theorists of scientific inference, those who have reflected on normative accounts of ordinary inductive inference appear to find global normative demands supported by their intuitions. Indeed, even those normative theorists who are principally interested in scientific inference have not claimed that the boundary
between science and non-science marks the limits of their theory. To the extent that science *per se* is taken to mark an interesting boundary, it is typically because (reasonably enough) normative theorists often suppose that scientific contexts tend to be amongst those contexts in which inference is at its best. Thus, normative theories that incorporate global demands very frequently purport to apply not merely to science but to good inductive inference quite broadly. And precisely the same intuitive judgments that support, in scientific contexts, the normative appropriateness of total evidence or standards of global consistency appear to be operative in non-scientific contexts as well. In which case, there is reason to believe that an analog to Normative Globalism for ordinary thought is, like Normative Globalism itself, *prima facie* plausible.

Similar points hold for an analog to Competence Globalism in accounts of everyday reasoning competence. As we pointed out with respect to normative and competence theories of scientific inference, intuitive judgments about the quality of inference which support normative claims about reasoning will also provide an evidential basis for corresponding claims about reasoning competence. Thus, since a normative version of Globalism is plausible about ordinary inductive inference, claims about the globality of everyday reasoning competence are also plausible. In other words, there is reason to believe that an analog to Competence Globalism for ordinary cognition is plausible to more or less the same extent as Competence Globalism itself. Since this is what The Positive Levels Principle predicts, we believe that this connection between theories again provides (albeit weak) inductive support for the general framework suggested by the principle. In addition, this plausible connection between theories of
scientific inference and theories of ordinary cognition may well help to explain why so many have found globality claims about both scientific inference and everyday reasoning attractive. If we are correct, then normative and competence readings of Globalism for both types of cognition are plausible, and are plausible to more or less the same extent. We hasten to add, however, that none of the roles we have explored for Scientific Globalism in the framework suggested by The Positive Levels Principle support Fodor’s criticisms of theories of cognitive architecture. Rather, this exploration is another attempt to learn from Fodor’s mistakes when tracing implications between theories of scientific inference and theories of ordinary cognition.

Conclusion

The preceding has been an attempt to characterize a range of respects in which accounts of scientific inference might plausibly inform theories of ordinary cognition. We formulated our characterization by reflecting on Fodor’s mistaken attempts to draw morals for theories of cognitive architecture from claims about the nature of scientific inference. We hope to have established in particular that claims about the globality of scientific hypothesis formation and confirmation do not in fact undermine extant proposals about our mental architecture. In so doing, our aim has been to lay some important groundwork for future investigations into the relationship(s) between scientific inference and everyday thought.
Chapter 3: Is scientific theory change similar to early cognitive development?

Gopnik on science and childhood

Abstract: I offer a sustained and two-part critique of Alison Gopnik’s position that early cognitive development is highly similar to scientific theory change. In the first part, I criticize the “strong” thesis that scientific inference and early cognitive development are subserved by the very same mental mechanisms. I contend that Gopnik has failed to distinguish among several different types of theories of scientific inference – including population-level, normative, competence, and performance theories – and that the strong thesis fails to hold for any of them. In the second part, I argue that even a weaker analogy between childhood cognitive development and scientific inference has severe limitations. In particular, I maintain that scientific inference is centrally subserved by cross-domain processing while early cognitive development is not. The hope is that the inadequacies in Gopnik’s position are instructive. For one, they suggest a tighter analogy, which I briefly defend, between adult quotidian cognition and scientific theory formation and confirmation.

1. Background on “Theory Theory” and Gopnik
Gopnik endorses a version of the “theory theory,” which denotes a cluster of views in developmental psychology all of which maintain that central aspects of early cognitive development are similar to scientific theorizing. My focus here is on Gopnik’s version of the theory theory in particular, which holds that information acquired by infants and children has “structural, functional, and dynamical” features similar to scientific theories (Gopnik & Meltzoff 1997, pp. 35-36). With respect to structural features, Gopnik claims that childhood knowledge in several domains is, roughly, a system of lawlike generalizations that underwrite counterfactuals, contain ontological commitments, and classify entities according to their underlying causal structure (as opposed to their superficial characteristics). The functional features of childhood knowledge, according to Gopnik, are similar to scientific theories in that children use their knowledge to generate predictions, interpret evidence, and provide explanations. Further, Gopnik gives a realist interpretation of these structural and functional features: scientific theories and childhood knowledge are approximately veridical models of the world and often give rise to accurate predictions. In turn, Gopnik’s realism is coupled with an adaptationist explanatory framework. She argues that the ability to form approximately accurate theories and predictions gave past children a differential fitness advantage, and thus that the “cognitive devices” underlying childhood development were selected for (Gopnik 1996, p. 496).

---

15 “Theory theory” is sometimes used to refer solely to the view that attributions of mental states and predictions of behavior are underwritten by an internal folk psychological theory (e.g., Morton 1980). I will not follow such a restrictive usage in this paper.  
Most theory theorists (as well as some modularity theorists)\(^\text{17}\) endorse many or all of the above claims about the structural and functional features of childhood knowledge. The focus of our discussion will not be on static properties of bodies of information, however, but rather on claims concerning dynamical features of childhood knowledge acquisition. Theory theorists are unique in holding that infant and childhood knowledge is generated and modified, in many cases, by the same types of process scientists use to form and confirm theories. Gopnik’s distinctive account of these dynamics is based on a distillation of a variety of developmental data (much of it collected by Gopnik herself) regarding children’s improving knowledge in various domains: including knowledge of a lexicon, object classification, movement and spatial relations between objects, and belief and desire attributions that help predict actions. In particular, Gopnik claims that early cognitive development in each of these domains undergoes the following stages, which are transitional between an old and new theory (Gopnik & Meltzoff, pp. 78-9):\(^\text{18}\)

STAGE 1: A set of beliefs or theory is in place and if potentially recalcitrant or countervailing evidence is encountered it is ignored and essentially treated as noise.

---

\(^{17}\) Modularity theory is usually understood as a competitor to the theory theory. There are a variety of ways of understanding the notion of a module, but typically modules are construed as mental organs dedicated to performing inferences about a proprietary domain. The basic point of contrast is that modules grow or mature and their increased capacities are the product not of learning but of innately specified ontogenesis.

\(^{18}\) One might wonder why Gopnik does not provide formation stages of a child or infant’s initial theory, rather than trace transitional stages between theories. Gopnik claims, however, that with early cognition the appropriate analysis is that “it’s theories all the way down!” (Gopnik & Meltzoff 1997, p.72).
STAGE 2: As additional countervailing evidence is encountered, modifications to the theory are made by adding ad-hoc auxiliary hypotheses. These additions are intended to accommodate apparently recalcitrant data, but often reduce a theory’s overall coherence or simplicity.

STAGE 3: A new theory or model is formulated, often in terms of ideas that were peripheral to the old theory, that better satisfies general criteria of coherence and simplicity.

STAGE 4: Before adopting a new theory, a period of intense experimentation and observation occurs with the purpose of selecting between old and new alternatives.

These stages can be plausibly ascribed to scientific theory change, and principally on the basis of these putative shared dynamical features, Gopnik argues for the central thesis of her version of the theory theory, namely: “the cognitive processes that underlie science are similar to, or indeed identical with, the cognitive processes that underlie much of cognitive development” (Gopnik & Meltzoff 1997, p.32). This thesis admits of several readings. I consider below two readings that are importantly different and especially deserving of separate treatment:

\[
\text{\underline{\text{\footnotesize{19}}} Similarly, Gopnik claims: “Scientists and children both employ the same particularly powerful and flexible set of cognitive devices” (1996, p. 496).}
\]
*The Strong Theory Theory Thesis*: The theory-forming mechanisms underlying early cognitive development and scientific inference are numerically identical.

*The Weak Theory Theory Thesis*: The processes subserving early cognitive development and scientific inference are highly similar, though numerically distinct.

1.1. Strong vs. Weak

1.1.1. The strong reading

Gopnik in places appears to endorse the Strong Theory Theory Thesis, and there are various reasons for doing so given her explanatory project. For instance, Gopnik advertises that her version of the theory theory can answer the following question about the existence of science:

Where did the particularly powerful and flexible devices of science come from? After all, we have only been doing science in an organized way for the last 500 years or so; presumably they didn’t evolve so that we could do that (Gopnik 1996, p. 496).

Gopnik’s answer is that scientific inference is underwritten by the very same “powerful and flexible” mental mechanisms used in early cognitive development. Since Gopnik believes she has an explanation for the existence of the “cognitive devices” underlying
childhood development in terms of natural selection, she believes she also has an explanation for the existence of science: “Science is a kind of spandrel, an epiphenomenon of childhood” (Gopnik 1996, p. 490). This explanation is available only on a strong reading of Gopnik’s central thesis. If the cognitive processes subserving science were merely similar to the processes subserving childhood development, Gopnik’s thesis would not explain how or why such similar but distinct processes arose. Further, only the Strong Theory Theory Thesis can answer, in Gopnik’s preferred fashion, the related “1492 Problem”,20 viz. “Why did science arise circa 1492 in the West but not before or elsewhere?” The explanatory framework offered by Gopnik is that in most adults, theory-generating mechanisms are present but relatively dormant, at least in the sense that they are not operating at their full capacity to produce and modify theories about one’s surroundings. According to Gopnik, however, if we recreate the conditions of childhood for scientists – leisure time, exploration and collection of new types of evidence, social support, etc. – a scientist might again use those very same theory-forming cognitive mechanisms (Gopnik & Meltzoff 1997, p.122). Gopnik offers an account according to which such conditions were not present at any time until circa 500 years ago in Europe.

Relatedly, on this account everyday adult cognition is not similar to scientific inference or childhood development. Scientific inquiry and childhood development, but not adult cognition, occur during periods when the requirements for survival are suspended, and when the relevant parties concentrate on acquiring roughly veridical

conceptions of their physical and social surroundings. Though each child is like a “little scientist,” “not much” of this cognitive activity, according to Gopnik, “survives into adulthood” (Gopnik 1996, p. 490). Rather, the success of adult scientists “literally comes from their childishness” (Gopnik 1996, p. 561). This feature of Gopnik’s view – whereby childhood development but not adult reasoning is similar to scientific inference – will become important later. Presently I am emphasizing that only the Strong Theory Theory Thesis is consistent with Gopnik’s explanatory framework for addressing the 1492 Problem.

Finally, Gopnik appears to prefer the Strong Theory Theory Thesis because it avoids a vacuity worry about weaker readings of her central claim:

It is all very well to suggest that children’s learning mechanisms are analogous to scientific theory-formation. However, what we would really like is a more precise specification of the mechanisms that underlie learning in both scientists and children (Gopnik & Schulz 2004, p. 371).

Though this specification is not central to Gopnik’s original formulation of her version of the theory theory, she has subsequently settled on particular mechanisms that underwrite much of childhood development and scientific inference, namely Bayes nets.²¹

The above three reasons, then, are roughly what is at stake for Gopnik if she does not endorse the strong version of her central thesis. Only a strong reading is consonant

with her framework for explaining the existence of scientific theory-forming devices and their employment circa 1492. Further, a strong reading straightforwardly avoids the worry that her central thesis lacks substance.

1.1.2. The weak reading

In contrast, I believe the Weak Theory Theory Thesis might achieve some explanatory aims that are apparently vital to Gopnik’s project. For example, Gopnik adopts an adamantly realist view of science as well as an account of childhood theories according to which they are roughly veridical.\textsuperscript{22} The Weak Theory Theory Thesis could provide a partial account of the putative veridicality of both scientific theories and childhood knowledge, on the assumption that the types of process generating both bodies of information produce approximately true accounts of one’s surroundings. Further, a weak reading might also explain the convergence that Gopnik highlights in both science and childhood development. Gopnik claims that scientists tend to arrive, at least over time, at a consensus regarding the truth of theories (Gopnik 1996, p. 488).\textsuperscript{23} Similarly, most children converge to a remarkable degree on highly similar theories in a variety of domains. Again, we might suppose that a weak reading of Gopnik’s central thesis could provide an account of this shared convergence, on the assumption that processes with the dynamical properties that Gopnik describes tend to converge over time on the same sorts

\textsuperscript{22} In fact she calls veridicality the “sine qua non” (Gopnik 1996, p. 497) of childhood development: “The crucial fact about cognitive development, and cognition in general, is that it is veridical, it gives us a better understanding of the world outside ourselves” (Gopnik & Meltzoff 1997, p. 72).

\textsuperscript{23} Gopnik also highlights some convergence at the stage of scientific theory formation. She points out that breakthroughs often happen nearly simultaneously – e.g. evolutionary theory, calculus, the structure of DNA, etc. (Gopnik & Meltzoff 1997, p. 189)
of bodies of information. Finally, the Weak Theory Theory Thesis might well provide a useful organizational principle for understanding central aspects of infant and childhood cognitive development.

In the critical discussion that follows, I give separate consideration to the strong and weak reading of Gopnik’s central thesis. I will not dispute any portion of the bulk of developmental data that Gopnik offers in favor of either reading. That is, I will not argue that her claims about the dynamics of childhood learning are unpersuasive (though some have). Instead, I claim her version of the theory theory faces a stumbling block that is at least as crucial: its claims about science.

2. Criticisms of The Strong Theory Theory Thesis

In this section I argue that Gopnik’s account of the dynamics of scientific theory change are ambiguous among several different types of theories of scientific inference, none of which have the relations to early cognitive development asserted by the strong reading of Gopnik’s central thesis. I examine Gopnik’s claims about theory change dynamics as a component of population-level, normative, competence, and performance theories of scientific inference. My central criticism is that on any disambiguated claim about theory dynamics, the Strong Theory Theory Thesis is implausible given what I call the “Disjoint Problem.”

24 E.g., Carey & Spelke (1996) or Carey (2009).
2.1. Population-level Theories of Scientific Theory Dynamics

Recall Gopnik’s four stages of putatively shared theory development. In the case of science, it is natural to wonder whether Gopnik’s stages are intended to describe the dynamics of individual scientists or of the scientific community. If her description of the dynamics of scientific theory change is construed as a component of a population-level theory of scientific inference, it might be that such features are emergent properties, distributed solely over the coordinated activities of the scientific community. In that case, the Strong Theory Theory Thesis would enjoy no support whatsoever. In fact, however, Gopnik is resistant to the idea that her stages of scientific theory change should be thought of as applying only to populations of scientists, not individuals: “Ultimately, the sociology of science must consist of a set of individual decisions by individual humans to produce or accept theories” (Gopnik 1996, p. 487). I propose to grant, for the sake of argument, that each of Gopnik’s stages is effected by the cognition of individual scientists, such that the subvenience base of the relevant cognition is circumscribed inside the heads of individual scientists. Unfortunately, Gopnik’s view still faces:

**The Disjoint Problem:** The stages of scientific theory change are often, and perhaps typically, instantiated by distinct subsets of the scientific community with few or no common members.

The Disjoint Problem poses a challenge to the Strong Theory Theory Thesis because it highlights different cognitive dynamics in a typical scientist and a typical child. Gopnik’s central support for postulating identical mechanisms underlying both childhood development and scientific theorizing is developmental data purportedly showing that
each infant or child undergoes all of Gopnik’s dynamical stages. The Disjoint Problem suggests the same is not true of individual scientists; they typically do not go through all of Gopnik’s stages. It follows that the Strong Theory Theory Thesis is not only unsupported but implausible. The Disjoint Problem requires Gopnik to explain why the very same theory-forming mechanisms produce different cognitive dynamics in individual scientists and individual children.

2.1.1. The Disjoint Problem: an example

Gopnik herself does not present evidence of a single scientist who has undergone all of the stages she ascribes to scientific theory change. Nor does she outline how her stages characterize any particular period of scientific theory formation and confirmation. Plausibly, however, her stages are present in a central strand of the development of evolutionary theory. Below I briefly reconstruct some relevant theoretical dynamics in terms of Gopnik’s stages in order to highlight their disjoint nature. In particular, I focus on attempts to account for the geographical distribution of diverse flora and fauna, which Darwin regarded as a (if not the) central explanatory task of evolutionary theory.25

STAGE 1: Initially, natural theologian biogeographers dismissed or ignored evidence relating to the geographical distribution of diverse fauna, as such evidence might have been taken to conflict with the reigning orthodoxy of a

25 Kitcher (1993) places this at the heart of Darwin’s project; in what follows I borrow heavily from the second chapter, “Darwin’s Achievement.”
single creation event of immutable species. Their project instead was to locate a single creation event and unearth natural principles of subsequent geographical dispersion.

STAGE 2: Ad-hoc auxiliaries were added to accommodate apparently recalcitrant evidence of locally similar plants and animals in a region, but relatively dissimilar plants and animals in otherwise similar, but non-local, environments. (For example, there were marsupials in Australia, but not in other continents at similar latitudes.) Thus, natural theologians proposed and came to accept that there were several “different centers of creation,” and sought to locate and describe principles of distribution flowing from these distinct centers.

STAGE 3: Under the pressure of increasing evidence of local similarity and non-local diversity, of puzzling patterns of historical and current geographical distribution, etc., Darwin and others (e.g., Baden Powell) proposed an alternative model that included species transmutation. Darwin of course is notable, in addition, for proposing natural selection as the central mechanism driving descent with modification.

STAGE 4: Following the publication of On the Origin of Species, a period of intense gathering of evidence, debate, and testing ensued. Eventually, the scientific community converged in its assessment of the truth of Darwin’s theory.
If Gopnik is correct about the dynamics of early cognitive development, such dynamics are shared by an important and plausibly representative example of scientific theory change. Even on this assumption, however, the Disjoint Problem undermines the Strong Theory Theory Thesis. In what follows, I show that for each pairwise combination of the above stages there may be few, and perhaps no, associated common members.

Stages 1 and 2 are Disjoint: The ad-hoc auxiliaries of “different centers of creation” were originally proposed in the late 1700s and came to be orthodoxy among natural theologians only several generations later, in the mid 19th century (Kitcher 1993, p. 44). So, many who accepted the auxiliaries from Stage 2 never dismissed or ignored apparently recalcitrant data associated with Stage 1 (and few who accepted the auxiliaries likely proposed them). It follows that there are likely few members common to Stages 1 and 2.

Stages 2 and 3 are Disjoint: Few who were working in the old paradigm of immutable species developed an alternative theory of descent with modification propelled by natural selection. Further, there is no evidence that Darwin, or anyone who proposed alternative evolutionary models, ignored or dismissed evidence that originally gave rise to the addition of “different centers of creation” auxiliaries.26 Thus it is entirely plausible

26 See Carey (2009) for discussion. For decades prior to formulating his theory of evolution, Darwin worked within a different, and also idiosyncratic, “Monad Theory,” which largely consisted of an amalgamation of existing theories that held that species were transmutable. There is no evidence Darwin developed “different centers of creation” auxiliaries, or dismissed evidence that undermined single or multiple creation centers. Further, after recording the central components to his theory of evolution in a notebook, Darwin did not engage in an intense period of testing in order to select between his old and new
that there is little to no overlap between Stages 2 and 3 and there is almost certainly no overlap between Stages 1 and 3.

Stages 3 and 4 are Disjoint: Few who eventually accepted evolutionary theory proposed it and thus there is likely little overlap between Stages 3 and 4. Further, many who accepted the theory of evolution did not convert to it from another theory, but were instead new to the field. Evolutionary theory came to be accepted only after decades of controversy. This means that the scientific community’s acceptance came about in large part because of members passing in and out of the community, rather than the community changing its collective mind at any one time. Thus there is likely little overlap between Stages 2 and 4 and almost certainly no overlap between Stages 1 and 4. The sets of scientists associated with these stages are disjoint.

If the example of evolutionary theory is reasonably representative, the Disjoint Problem provides general grounds for doubting that the very same cognitive devices give rise, on an individual level, to radically different developmental dynamics in a typical scientist and a typical child.\(^{27}\) And there is reason to suppose that, in the general case, only a minority of scientists propose significant modifications to a given reigning theory, 

\(^{27}\)Gopnik anticipates that one might give a population-level reading to her claims about science and then object, as I am doing, that her account of childhood development is apparently individual-level. That is, she attempts to argue that any sort of population-level account of science would be equally plausible for describing children: “children are less isolated than the term ‘little scientist’ is likely to imply. They live in a rich social structure with much opportunity for contradiction, instruction, and the linguistic transmission of information. We are not dealing with a contrast between a non-social process and a social one, but between two different types of social organization” (Gopnik & Meltzoff, 1997, p. 24). But this doesn’t solve the Disjoint Problem. It is not the case that disjoint subsets of the community in which children are embedded undergo some developmental stages but not others. Instead, each child, according to Gopnik, undergoes every stage.
with the majority normally working within that theory (Kuhn 1972). Further, it is reasonable to believe of the general case that those who accept new theories are often new to a field and typically converge on extant proposals rather than generating them (Collins 2001), and so forth.

The Strong Theory Theory Thesis’s implausibility exemplifies the difficulties associated with drawing implications from population-level accounts of scientific inference for models of individual cognition (and vice versa). Such inferences are appropriate only if we assume that individual members of the scientific community invariably share the properties of the community as a whole. The Disjoint Problem shows one significant difficulty with making such an assumption. None of what has been said, however, need undermine the Weak Theory Theory Thesis. Useful analogies may be found in many places, and perhaps the processes subserving scientific theory change at the population level are interestingly similar to the processes subserving early cognitive development. I reserve my criticisms of the Weak Theory Theory Thesis for Section 3.

2.2. Normative Theories of Scientific Theory Dynamics

Gopnik might defend the Strong Theory Theory Thesis against the Disjoint Problem by insisting that her claims concerning the stages of scientific theory change are about individual scientists, but were never intended to be purely descriptive. Rather, perhaps Gopnik’s claims about theory dynamics concern what individual scientists should or would do under certain idealizations. Gopnik in places writes as if her characterization of
theory dynamics is intended as a component (or perhaps a complement) to a *normative* theory of scientific inference:

Science … proceeds in various and haphazard ways. But it nonetheless manifests a kind of logic, and converges on a truthful account of the world. A cognitive view of science might provide an important bridge between normative or logical and sociological views of science. It might indeed be true that there were particular kinds of abstract, logical structures that characterized the important cognitive achievements of science. And it might also be true that looking at the practice of actual science, we might see rather little of that abstract logical structure. …

Quite typically, an abstract structure underlies some human cognitive activity that is not at all apparent in superficial phenomenology or practice. Often, that structure is related in interesting ways to the structures we would invent if we constructed an ideal machine to perform that cognitive activity (Gopnik & Meltzoff 1997, p. 88).

This is a difficult line to pursue, however, given that Gopnik’s central thesis is an apparently descriptive claim about the identity or similarity of the processes underlying scientific theory change and early cognitive development. Gopnik attempts to undercut general worries that her approach runs roughshod over a normative/descriptive gap, but
one might nonetheless regard such an approach as cavalier. Here I will emphasize a particular complication for Gopnik given the Disjoint Problem.

Recall that the Disjoint Problem highlights that the actual cognitive dynamics of individual scientists and individual children typically diverge. It follows that any idealizations that bring these dynamics into accord would have to invoke much greater levels of idealization when characterizing scientific reasoning, as opposed to early cognitive development. The problem in a nutshell is that the Strong Theory Theory Thesis cannot survive if Gopnik is characterizing highly idealized scientists on the one hand and more or less actual children on the other. It is barely coherent, let alone plausible, to suppose that the cognitive mechanisms of ideal scientific agents are numerically identical to the mechanisms found in actual child learners.

The upshot for population-level and normative theories of scientific inference is similar: drawing implications for models of individual cognition from normative theories of scientific inference (and vice versa) is fraught with peril at best and ill-formed at worst. Again, however, it might be the case that a normative model of scientific inference could characterize processes that are highly similar to the processes that subserve childhood development. The Weak Theory Theory Thesis is still in play.

2.3. Competence Theories of Scientific Theory Dynamics

We should also consider, however, whether Gopnik intends to give an account of the dynamics of scientific theory change that enjoys idealizations of a different sort, and that
might preserve the Strong Theory Theory Thesis. Perhaps instead of characterizing the
cognitive structures of ideal scientific agents, Gopnik is characterizing actual scientists
under certain idealized conditions. In particular, we can consider that Gopnik intends her
claims about theory dynamics as a component of a competence theory of scientific
reasoning. Much might be said on how best to understand the notion of a theory of
scientific reasoning competence, but here I will just briefly outline, with the aid of an
analogy to linguistic competence, how a competence/performance distinction could serve
in a defense against the Disjoint Problem.

In parlance Chomsky has made popular, the set of actual grammaticality
judgments made by a speaker is that speaker’s linguistic performance. By contrast,
formulating a theory of a speaker's linguistic competence (as opposed to her
performance) involves abstracting away from linguistically irrelevant “performance
conditions” that might compromise grammaticality judgments. Paradigmatically, such
performance-related conditions include “memory limitations, distractions, shifts of
attention and interest, and errors (random or characteristic) in applying knowledge of …
language in actual performance” (Chomsky 1965, p.3). If we adopt something like this
Chomskian model in characterizing scientific reasoning competence, we might suppose
that the set of actual validity judgments made by a scientific reasoner is that reasoner’s
performance. In contrast, we might seek to characterize her competence by abstracting
away from “irrelevant” performance conditions that compromise judgments of successful
scientific inference, also including memory limitations, attention shifts, etc. The
theoretical motivation behind invoking a competence/performance distinction in
understanding scientific reasoning would be to characterize the cognitive structures principally dedicated to underwriting successful scientific inference, while designating as “performance errors” failures of those structures to successfully interact in practice with distinct cognitive structures.

How might a competence/performance distinction serve to support the Strong Theory Theory Thesis? First, we might modestly reinterpret the thesis to assert that the cognitive structures underlying scientific reasoning competence and the structures underlying early cognitive developmental competence are identical. In line with this reinterpretation, we might next attempt to explain the apparent divergence in practice highlighted by the Disjoint Problem as solely a divergence of performance, not of underlying competence. In short, the suggestion is that it is because an individual scientist makes more performance errors than an infant or child typically makes during cognitive development that the scientist ordinarily does not go through all of Gopnik’s stages.

Even with this cursory characterization of such a defense of the Strong Theory Theory Thesis, one would be justified in taking a dim view of its prospects. In the absence of a proposal from Gopnik about what performance errors are rife among typical

---

28 Again, there are some passages where Gopnik might be interpreted as providing a competence theory of both scientific inference and childhood cognitive development. For example, she claims that intermediate stages in acquiring a folk physical theory should not be explained in terms of “information processing or memory limitations” (Gopnik & Meltzoff 1996, p. 97). Since these are plausibly thought of as performance conditions, one might interpret Gopnik as seeking instead to characterize an infant or child’s competence in forming theories. Similarly, Gopnik appears to abstract away from “performance errors” by scientists, including “motivational impurities” that compromise successful scientific inference (Gopnik & Meltzoff 1997, p. 91). Again, one might interpret such an approach as attempting to provide a theory of scientific reasoning competence.
scientists but not typical children, there is little reason to regard this reinterpretation of the Strong Theory Theory Thesis as plausible. Indeed, most of what are traditionally thought of as performance conditions that could compromise a given cognitive competence are shared by children and scientists: memory limitations, shifts of attention and interest, distractions, etc. It follows that we have no reason to expect an explanation of the dynamical divergence between individual scientists and children in terms of performance errors. Thus there is no reason to accept that the cognitive structures underlying an individual scientist’s reasoning competence and a child or infant’s learning competence are identical. In sum, the Strong Theory Theory Thesis is not lent greater plausibility if we read Gopnik’s claims about the dynamics of scientific inference as a component of a competence theory of scientific reasoning.

We have reviewed the prospects for supporting the Strong Theory Theory Thesis via Gopnik’s claims about scientific theory change as components of population-level, normative, and competence theories of scientific inference. All of these readings have been found wanting: either implausible, unsuitable for childhood development, or both. It seems, however, that a performance theory of scientific inference would be uniquely suited to support descriptive claims about the nature of the mechanisms underlying childhood cognitive development and scientific reasoning. Unfortunately, the Disjoint Problem highlights that a performance theory of scientific theory dynamics does not appear to support the Strong Theory Theory Thesis either, since the performances of scientists and children radically diverge.
The upshot is that no reading of Gopnik’s dynamical claims about scientific
theory change supports the Strong Theory Theory Thesis. In turn, Gopnik has no account
of the existence of science and its appearance circa 1492. I maintain that the failure to
appropriately trace implications between different types of theories of scientific inference
and cognitive science is not unique to Gopnik;\textsuperscript{29} rather, the failure of the Strong Theory
Theory Thesis serves as a stark and hopefully instructive illustration.

3. Limitations of the Weak Theory Theory Thesis

In this section I argue that even the Weak Theory Theory Thesis has severe limitations.
Recall that the Weak Theory Theory Thesis asserts that the processes underlying both
eyearly cognitive development and scientific inference are “highly similar.” In fact, Gopnik
claims early cognitive development is more similar to scientific inference than everyday
adult belief-formation, reasoning, or planning. In contrast, I argue that central features of
scientific inference, namely cross-domain analogical reasoning and the formation of
cross-domain unifying explanations, are insignificant to early cognitive development.
This important dissimilarity between childhood cognition and scientific inference,
however, is in fact an area where performances of both adult quotidian cognition and
scientific inference exhibit significant similarities. If my argument is correct, there is
reason to both doubt the Weak Theory Theory Thesis and look favorably on an analogy

\textsuperscript{29} Cf. Fuller and Samuels (unpublished ms) on Fodor.
between everyday adult cognition and scientific inference.30 I begin by describing why childhood cognitive development typically does not exhibit cross-domain processing, then proceed to characterize cross-domain processing in scientific inference, and finally highlight similar patterns in adult cognition.

3.1. Domain-Specificity in Early Cognitive Development

Though Gopnik flirts with the idea that domain-general processes might account for childhood development in a variety of areas, she also wants her claims to be consistent with variegated ontogenetic pacing among cognitive systems. That is, the cognitive structures responsible for infants’ and children’s success in a variety of domains, including those Gopnik compares to scientific theories, do not develop uniformly or obviously interact. Rather, acquired bodies of knowledge appear to develop over independent timeframes and without (much) information transmission. Gopnik & Meltzoff write:

… we will argue that these domains may be relatively independent of each other in development. That is, though there are changes in all of these areas at around 18 months on average, the changes are not highly correlated and do not constitute a cognitive stage in Piaget's sense... In this sense our argument will be in line with many recent suggestions that

30 Others have endorsed such a view, including Fodor 1983, 2000, and 2008; Newell and Simon 1976; and, perhaps, Einstein: “The whole of science is nothing more than a refinement of everyday thinking” (1936, p. 59).
development is domain-specific. Indeed, we have called this idea "the specificity hypothesis" … (Gopnik & Meltzoff 1997, p. 74).

The “specificity” of early knowledge acquisition has two components. First, the relevant cognitive systems do not significantly interact with each other during development; in general, folk biological, physical, and psychological theories mature independently. Second, despite mastering complex systems of lawlike and interrelated generalizations in all of these domains, such capacities are not deployed more generally. For example, though young children can learn folk psychology, they have trouble with traffic signs (Segal 1996).

This suggests that two important properties, often invoked to taxonomize cognitive processes, are present to a significant degree in childhood processing. First, the processes underwriting early cognitive development are to a significant degree “domain-specific” – i.e., their class of inputs is significantly restricted. Second, the processes are “informationally encapsulated” – i.e., any stored information available to early cognitive processes is significantly restricted and not accessible to other cognitive systems.\footnote{See Fodor (1983) for the introduction of such notions into cognitive science. There he claimed these properties are instantiated “in degrees,” which I am assuming above. Indeed, the critique of the Weak Theory Theory Thesis pursued here is Fodorian in many respects (Fodor 2000, 2008).} I suggest that these two central features of cognitive processes underwriting childhood development are not present in scientific inference. That is, the processes subserving scientific inference enjoy significant levels of apparent domain-generality and
informational unencapsulation, and this striking dissimilarity undermines the Weak Theory Theory Thesis.

3.2. Cross-Domain Processing in Scientific Inference

Cross-domain processing is central to scientific inference, particularly in the area of hypothesis formation and theory construction. Examples of performances of cross-domain analogical reasoning in the history of science are legion: The Rutherford atom model was analogically based on the solar system, Pasteur analogized disease to fermentation (Thagard 1997), Huygens and Young modeled theories of light waves on sound waves (Thagard 1978), and Kepler seemed to think about the solar system in any and every form of analogy -- in terms of lodestones, sailors, and orators at a crowd, to name a few. Perhaps Kepler was unique in his placement of analogical reasoning at the very heart of scientific inference, but we would be remiss not to also count cross-domain analogical reasoning as central to hypothesis and theory formation. Further, analogical reasoning is present in other stages of scientific inference, including periods of so-called “normal” science, when theories are being assumed rather than challenged.33

32 “I especially love analogies, my most faithful masters, acquainted with all the secrets of nature…” (quoted in Gentner and Brem et al. 1997).
33 Here is an example of cross-domain reasoning in medical diagnosis: “The CD4 lymphocyte depletion seen in advanced HIV-1 infection may be likened to a sink containing a low water level, with the tap and drain both equally wide open. As the regenerative capacity of the immune system is not infinite, it is not difficult to see why the sink eventually empties. It is also evident from this analogy that our primary strategy to reverse the immunodeficiency ought to be to target virally mediated destruction (plug the drain) rather than to emphasize lymphocyte reconstitution (put in a second tap)” (quoted in Thagard 1997).
Also central to scientific hypothesis formation is generating explanations that “unify,” especially in the sense of accounting for information from multiple domains. We might call this “cross-domain unification”, of which examples from the history of science are also legion. Darwin and Newton, for example, certainly provided unifying explanations that accounted for phenomena from multiple domains (however individuated). And as Friedman points out, the kinetic theory of gases was praised and accepted in large part because it integrated with previous theories of planetary motion and of the motion of falling bodies near the earth’s surface. As with analogical reasoning, we again have reason to believe that cross-domain processing plays a central role in scientific inference.

In contrast, there is little evidence that cross-domain unification or cross-domain analogical reasoning plays a central role, or indeed any significant role, in early cognitive development. Gopnik describes babies, infants, and young children as scientists, but does

---

34The following working definition of “unification” should suffice for our purposes: formulating and explaining “a maximum number of facts in terms of a minimum of theoretical concepts and assumptions” (Feigl 1970, p. 12); or “reducing the total number of independent phenomena that we have to accept as ultimate or given” (Friedman 1974, p. 15). Some have identified unification as not just central, but the *sine qua non* of scientific explanation. For example, Feigl claims the “aim of scientific explanation throughout the ages has been *unification*” (1970, p. 12). Friedman calls unification the “essence” of scientific explanation (1974, p. 15). Kitcher: “the search for scientific explanation is governed by a maxim once succinctly formulated by E. M. Forster. Only connect” (1981, p. 530). We can work with weaker assumptions here, but the extent to which unification is central to scientific inference should not be minimized.

35 Darwin regarded unification from multiple domains as strong evidence for his theory: “It can hardly be supposed that a false theory would explain, in so satisfactory a manner as does the theory of natural selection, the several large classes of facts above specified.” Notice Darwin’s comparison of this facet of scientific inference to quotidian adult cognition: “It has recently been objected that this is an unsafe method of arguing: but it is a method used in judging of the common events of life, and has often been used by the greatest natural philosophers” (*Origin of Species*, p.476, emphasis added).

36 Newton’s theory of dynamics and gravitation ties together the law of falling bodies, the law of the pendulum, Kepler’s laws, and many diverse theorems of celestial mechanics (Glymour 1980).

37 Friedman 1974, p. 11-12.
not provide any evidence that they engage in any type of analogical reasoning, let alone cross-domain analogical reasoning. Further, there is little data showing that babies, infants, or young children provide explanations, to any extent comparable to scientific inference, that unify diverse classes of facts. This disparity is reason to criticize the Weak Theory Theory Thesis for asserting a high degree of similarity between the processes subserving scientific inference and early cognitive development. The criticism might be framed as a failure to respect a central goal of theories of cognitive processing, namely to provide a model of the flow of cognitive information and principles governing that flow. In this sense, the fact that scientific inference apparently enjoys significant levels of domain-generality and informational unencapsulation, while early cognitive development does not, undermines the Weak Theory Theory Thesis. In fact, this criticism applies whether one is construing scientific inference as a component of a population-level, normative, competence, or performance theory. Perhaps such a significant limitation on the analogy between scientific inference and early cognitive development raises the question of whether there are higher levels of similarity between scientific inference and other types of cognition.

3.3. Cross-Domain Processing in Adult Quotidian Cognition

---

38 Susan Carey does argue that children engage in analogical reasoning in the case of forming concepts of integers at the rather late age of between 4 and 6 (2009). The particulars of the case of number concepts, as well as Carey’s distinct version of the theory theory, deserve treatment elsewhere. I remark only that it is not obvious that Carey has a plausible example of analogical inference in this case, and does not claim to be providing evidence that young children engage in cross-domain analogical reasoning in particular. In addition, she has remarked in person that she does not believe that babies and infants make analogical inferences, or any type of cross-domain processing, similar to scientific reasoning.
As it happens, everyday quotidian cognition apparently exhibits both cross-domain analogical inference and the formation of cross-domain unifying explanations. That is, there is significant evidence that a tighter analogy holds between adult quotidian cognition and scientific theory construction and confirmation than between scientific inference and early cognitive development. And indeed, studies of cross-domain analogical inference show that it is routine in both adult problem-solving and adult decision-making. In a famous study by Gick & Holyoak (1983), subjects were asked to solve a problem concerning a doctor’s treatment of cancer. Many more subjects were successful when instructed about a structurally similar source problem regarding generals and a fortress.\footnote{Gick & Holyoak (1983).} Such studies support the presence of a capacity for cross-domain processing in adults that is not, as far as we know, shared by babies, infants or young children. Further, the capacity is often exercised by adults, as argued in a study concerning decision-making and reasoning about Quebec’s proposed 1995 secession.\footnote{Blanchett & Dunbar (2001) looked at magazine articles, op-eds, and political speeches about the proposed 1995 secession of Quebec from Canada and found, among 400 examples, 200 analogies where “the range of source categories was very diverse, ranging from agriculture, the family, sport, magic, to religion” (Dunbar 2002, p. 161).} The above considerations constitute an initial case for believing that cross-domain analogical reasoning is rife both in performances of scientific inference and adult quotidian cognitive tasks.

Similarly, explanations that provide unified accounts of facts from multiple domains are also generated by everyday adult cognition, even though everyday adult reasoning does not seek to unearth laws with the same generality of scope sought by
many branches of science. Some mundane examples should suffice to make the point that
unifying hypotheses are routinely generated in everyday contexts. If the average adult
knows that a man has knowledge of Latin, professes to celibacy, has a clean-cut
appearance and so on, she might abductively infer that the man is a Catholic priest
(Peirce 1958). Similarly, one might generate a common cause hypothesis41 for two events
from different domains that are correlated but don’t cause each other. For example,
suppose one notices, roughly at the same time, cuts on a pet and a hole underneath the
porch. Quotidian cognition routinely generates a unifying common cause hypothesis
(e.g., the arrival of a wild animal) rather than explaining these events separately. I submit
that the very mundane nature of the above examples speaks to their representativeness of
everyday adult cognition. They are dull examples precisely because they are routine
cognitive achievements. There is little evidence, however, that babies, infants, or young
children exhibit similar patterns of inference to any significant degree.

What this suggests is that we might expect a tighter analogy to hold between
scientific inference and everyday quotidian cognition than between early cognitive
development and science. Perhaps such an analogy will assist in understanding the
cognitive flow of information and the principles governing that flow in everyday adult
thought and scientific inference. There is reason to suppose that we might apply lessons
from the philosophy of science to models of everyday quotidian cognition, and vice
versa.

41 Informally, the principle of the common cause applies when two events A and B are correlated with each
other conditional on a third event C. C “screens off” both A and B from being causes of each other and C is
a common cause of both A and B (Reichenbach 1956).
4. Conclusion

The preceding is a sustained critique of one prominent version of the theory theory, Gopnik’s account of the dynamics of scientific theory change and early cognitive development. The thesis that scientific theory change and early cognitive development are subserved by the very same mechanisms, I have argued, is implausible given the Disjoint Problem. Once one separates out different types of theories of scientific inference – including population-level, normative, competence, and performance theories – it is apparent that Gopnik’s thesis does not hold for any of them. This supports a cautionary moral: that we should carefully consider what type of theory of scientific inference is at issue when drawing conclusions from the philosophy of science for cognitive science. Next, even a tight analogy between early cognitive development and scientific inference has severe inadequacies, since the latter is significantly subserved by cross-domain processing while the former isn’t. This insight inspired a second moral: a tighter analogy likely holds between adult quotidian cognition and scientific inference. There is reason to find this plausible given that quotidian cognition also uses significant cross-domain processes, including cross-domain analogical inference and the formation of cross-domain unifying explanations. The hope has been that the inadequacies in Gopnik’s position are instructive and point the way toward more fruitful implicational relations between the philosophy of science and cognitive science.
Chapter 4: Scientific Inference, Cognitive Development and Bayes Nets

Abstract: This paper criticizes an influential version of a research paradigm in developmental psychology – the “theory theory” – which holds that early childhood and scientific learning are both underwritten by the same kinds of cognitive processes. I argue instead that there are important kinds of learning processes that underwrite scientific theory formation and revision that are distinct from those that underwrite early cognitive development. Further, I advance a range of novel hypotheses concerning the cognitive development of adult scientific inferential capacities. In particular, I characterize the emergence of unique learning processes that underwrite scientific reasoning in terms of significant generalizations of the standard causal Bayes net formalism and in terms of cognitive processes coming on-line that are appropriately modeled with distinct formalisms.

1. Introduction

Arguably, the central purpose of cognitive science is to provide a mechanistic account of cognitive processes. According to a highly influential mechanistic account of learning processes, learning mechanisms are appropriately modeled by Bayesian networks. Recent
research on Bayes nets has shown that they have important applications not only as expert machine learning systems (Pearl 1988, 2000; Spirtes, Scheines, and Glymour 2000), but also as models of human learning, especially early cognitive development (Gopnik and Schulz 2004; Gopnik et al. 2004; Gopnik and Glymour 2006; Schulz et al. 2007) and scientific inference (Glymour 2000, 2010; Pearl 2000). The aim of this paper is to assess what Bayes nets research tells us about the general nature of both types of cognition.

My assessment has two components. First, I challenge a common assumption among proponents of psychological modeling with Bayes nets, namely, that Bayes nets research is happily wedded with an influential research paradigm in developmental psychology known as the “theory theory”. According to the version of the theory theory I criticize, early cognitive development and scientific inference are, in general, highly similar and perhaps underwritten by the very same learning mechanisms.42 I maintain that this version of the theory theory is likely false because there are kinds of learning in science that are dissimilar to learning in childhood. If I am correct, the moral to draw from Bayes net research is that, in spite of its success in modeling some aspects of both scientific inference and early cognitive development, we should not expect that, in general, both types of learning will be accommodated by the same kinds of models.

Second, I explicate and defend a hypothesis about the cognitive development of adult scientific inferential capacities that arises from challenging the theory theory.

According to The Process Development Hypothesis I defend, the development of adult

---

42 This view is defended in Gopnik (1996); Gopnik and Meltzoff (1997); Gopnik and Wellman (1994).
scientific inferential capacities includes the advent of new kinds of learning process for acquiring and confirming theories. In addition, I advocate a set of accompanying proposals for characterizing this development. These proposals are inspired by the observation that apparently unique learning processes in science are likely not accommodated by the standard Bayes net formalism. Thus, I pursue a characterization of these novel kinds of learning in terms of significant generalizations of the standard Bayes net formalism or in terms of cognitive processes coming on-line that are appropriately modeled by distinct formalisms.

The structure of this paper is as follows: In the next section, I briefly introduce some aspects of Bayes nets research and an animating paradigm for much of this research – the so-called “theory theory”. In Section 2, I discuss in more detail features of the standard Bayes net formalism that are relevant to understanding its successes as well as some of its limitations. In Section 3, I highlight aspects of scientific inference that are likely not accommodated by the standard Bayes net formalism and that are likely dissimilar to early cognitive development. In addition, I defend The Process Development Hypothesis and provide ways of characterizing the development of these aspects of adult scientific cognition. In Section 4, I elaborate further on The Process Development Hypothesis; in Section 5, I conclude.

2. Bayes Nets and the Theory Theory
For present purposes, we can think of Bayes nets research as composed of a standardized formalism on the one hand, and a deployment of that formalism for psychological modeling on the other. On the formalistic side, Bayes nets are a class of graphical structures that encode relations of probabilistic (in)dependence between variables. The associated formalism contains a variety of algorithms for generating these graphical structures from sets of variables and for assigning a distribution of probabilities to the variables on the graph. On the psychological modeling side, the graphical structures are treated as models of structured mental representations of causal systems in the environment. And the algorithms for generating Bayes nets and for assigning them a joint probability distribution are treated as models of human learning processes. Proponents of (causal) Bayes nets have principally invoked the formalism to model central aspects of both early childhood learning and scientific hypothesis formation and confirmation.

In the background of this research is the “theory theory,” which denotes a cluster of views in developmental psychology, all of which maintain that cognitive development and scientific inference are in some respects importantly similar. The bold version of this view that will be our focus was defended some fifteen to twenty years ago by some of the same proponents of more recent psychological modeling with Bayes nets, including Alison Gopnik and her collaborators. According to this version of the theory theory, “[s]cientists and children both employ the same particularly powerful and flexible set of

---

43 Some proponents of psychological modeling with Bayes nets view the associated formalism as providing a representational framework for characterizing learning processes at the “computational level” (in Marr’s sense). Others, such as Gopnik and Glymour, assert in addition that the Bayes net formalism provides an “algorithmic level” description of learning processes (2006). Thanks to David Danks for discussion on this point.
cognitive devices” (Gopnik 1996, 496). This “expansive” version of the theory theory, as I will call it, is distinct from weaker views that share the same name. For example, some weaker versions of the theory theory in developmental psychology merely posit that bodies of information in children are structurally similar to scientific theories (e.g., Carey and Spelke 1996; Wellman and Gellman 1998). And some merely treat the acquisition of new concepts in childhood as similar to the formulation of theoretical concepts in science (Carey 2009). In contrast, the expansive version of the theory theory advocated by Gopnik and collaborators incorporates these weaker views and asserts, in addition, that early cognitive development and scientific theory change exhibit the same dynamical properties and in general are underwritten by identical learning mechanisms.

Importantly, this expansive version of the theory theory views cognitive development solely in terms of the acquisition of new concepts and knowledge, not in terms of the acquisition of new types of learning process. Childhood and scientific learning, on this view, are both underwritten by the very same versatile and powerful learning mechanisms as they are fed more and better information. Indeed, according to Gopnik and collaborators, all of the learning processes underwriting scientific theory change were already operative in the earliest stages of cognitive development. Thus, Gopnik and Meltzoff claim that “[s]cientific progress is possible because scientists employ cognitive processes that are first seen in very young children” (1997, 32).

Though the expansive version of the theory theory is at first glance a provocative view, the history of the psychology of learning contains many theories that characterize learning from its earliest appearance in childhood to its latest application in adulthood as
underwritten by repeated applications of the same processes. For example, many
behaviorists held that operant and classical conditioning underwrite learning in the
general case; Piaget held that learning in general is underwritten by processes of
“accommodation” and “association” (1952); and many connectionists have argued that
learning is principally underwritten by associationist processes that are implemented by a
connectionist architecture (e.g., McClelland and Thompson 2007). The expansive version
of the theory theory represents a continuation of research paradigms within this tradition.
It claims in particular that the stages of cognitive development are *rationally* related and
that cognitive change is invariably the result of repeated applications of learning
mechanisms that acquire, confirm, and revise theories.

In their early formulations of the expansive version of the theory theory, Gopnik
and other developmental psychologists had no definite proposals as to the specific
cognitive mechanisms that implement theory change in both science and childhood.
However, the subsequent turn to Bayes nets as a candidate learning device for
underwriting both kinds of cognition was a natural choice for at least two reasons. First,
expert machine learning systems had already shown the Bayes nets formalism to be both
fruitful and mechanically implementable. Thus, Bayes nets enjoyed some general
plausibility as psychological models. Second, the plausibility of Bayes nets as models of
scientific inference had already been ably defended (e.g., Glymour and Cooper 1999;
Pearl 2000) prior to their adoption and employment in developmental psychology.

It seems likely that many developmental psychologists view recent modeling of
early cognitive development with Bayes nets as a continuation of research within the
older paradigm of an expansive version of the theory theory. For example, in a paper summarizing some successes of Bayes nets modeling, Gopnik and Schulz claim to be providing a “more precise specification of the mechanisms that underlie learning in both scientists and children” (2004, 371).\(^{44}\)

In what follows, I examine this older research paradigm in light of more recent research on Bayes nets and related advancements in machine learning. I believe that proponents of psychological modeling with Bayes nets have shown that Bayes nets model some important aspects of both early cognitive development and scientific inference. But granting these successes leaves open the question of whether the aspects modeled by Bayes nets are representative of both types of cognition generally. I provide reason to believe that they are not. My intent is not to deny merely that Bayes nets in particular model every aspect of both early cognitive development and scientific inference. This denial is relatively uncontroversial even for enthusiastic proponents of psychological modeling with Bayes nets.\(^{45}\) Rather, I also provide reason to doubt that a common collection of distinct models will accommodate both scientific inference and early cognitive development. That is, I maintain that there is reason to believe that both types of cognition are not subserved, in general, by a variety of common learning processes.

I should emphasize that the putative differences I will highlight between scientific inference and early cognitive development are, at least with respect to evaluating an expansive version of the theory theory, intended to be special. Not just any difference

\(^{44}\) See also Gopnik et al. (2004, 4).

\(^{45}\) See Gopnik and Glymour (2006, 42) and Glymour (2004, 784) for recognition that Bayes nets likely do not model every aspect of early childhood and scientific learning.
between science and childhood can undermine an expansive version of the theory theory. In fact, proponents of this view can accommodate many apparently distinctive features of scientific inference – for example, greater mathematical rigor, reliance on sophisticated instruments, the positing of unobservables, and so on. A Gopnik-style theory theorist can claim that such distinctive features are the result of the very flexible and powerful learning mechanisms she posits. In other words, it is compatible with an expansive version of the theory theory that different theoretical outputs should result when more and better evidential inputs and background information are fed into the proposed versatile learning mechanisms underlying theory change.

The features of scientific inference I will call attention to, however, are not so readily accommodated by an expansive version of the theory theory. This is because the apparently unique features of scientific inference I highlight (in Section 3) are not accommodated by the standard Bayes net formalism. Since this formalism characterizes the central learning mechanism so far proposed by proponents of the expansive theory theory, then such differences between scientific inference and early cognitive development speak to differences in the types of learning process underwriting both kinds of cognition, and not merely to differences in the information fed into such learning processes.

In order to better understand the implications of Bayes nets research on the nature of both early cognitive development and scientific inference, I turn to characterizing the associated formalism. This formalism is well known, and my intent is not to give a thorough characterization of it here. Instead, I will explicate some of its features that are
central to our discussion of both its successes and limitations in modeling human learning.

3. What Is a (Causal) Bayes Net?

Bayes nets are Directed Acyclic Graphs (DAGs) that encode conditional probabilistic dependencies between nodes or variables on the graph. The graph below is both directed (the arrows are unidirectional) and acyclic (there are no paths that trace the direction of the arrows which begin and end with the same variable):

![Bayes Net Diagram]

Figure 1.

Not all DAGs are Bayes nets, however. Bayes nets are DAGs with a joint probability distribution\(^46\) on the variables in the graph that respects certain constraints, and with an interpretation – or “parameterization” – of its variables and arrows that falls within a

\(^{46}\) A “joint probability distribution” is a specification of the probability of all possible combinations of values for the variables on a DAG.
specified range. A parameterization of a causal Bayes net, which has been the focus of psychological modeling, assigns some form of causal influence to the graph’s arrows, including positive or inhibitory, probabilistic or deterministic. In addition, parameterizations of causal Bayes nets typically treat their variables as propositional.

The language of familial relations is ordinarily used to describe connections between variables on a Bayes net, and a little of this terminology will assist in understanding how these variables are assigned probabilities. In the graph depicted above, A \(\rightarrow\) B, and thus we can speak of A as the parent of B, and B as A’s child. In addition, we may wish to refer to the set of a variable’s children and its children’s children. Call this set for a given variable its descendants. A central condition governing the joint probability distribution to all the variables on a Bayes net is:

*The Causal Markov Condition (CMC)*: For any variable X, X is independent of all its non-descendants (i.e., its non-effects), conditional on X’s parents (i.e., its direct causes). The CMC says that X’s parents (direct causes) in a Bayes net “screen off” every other variable except X’s descendants (effects) from influencing X’s value. That is, according to the CMC, given a probability assignment to X’s parents, the values of X’s non-descendants do not impact the value of X.

The data from which the graphical structure of a Bayes net is “learned” are sets of variables that contain information about their probabilistic dependencies. I will mention two features of the learning methods that have been important for modeling early cognitive development and scientific inference. The first is that the CMC, along with
other conditions on a Bayes net’s joint probability distribution, constitute constraints on
the type of possible graphical structures that can be generated out of sets of variables. In
addition, these constraints facilitate inferences between probabilistic (in)dependencies
and causal structure. The CMC in particular facilitates inferences from causal relations to
probabilistic independencies, while other constraints facilitate inferences from causal
relations to probabilistic dependencies. Second, so-called “intervention” learning
methods have been developed that can be represented on causal graphs by eliminating
every arrow leading in to a particular variable while leaving every other relation intact.
Pearl calls this “graph surgery” (2000, 52). A family of such techniques allows one to
isolate dependencies and to distinguish between different types of causal structure. In
general, causal information can be garnered from Bayes nets because they facilitate
inferences from observed patterns of probabilistic dependence to predictions about the
effects of interventions and vice versa.

Proponents of psychological modeling with Bayes nets plausibly claim that both
young children and scientists construct causal maps of their environments from observed
contingencies and the effects of causal interventions. For childhood learning, a growing
body of evidence and research suggests that important aspects of the early acquisition of
intuitive biological, psychological, and physical theories can be modeled by causal Bayes
nets. Similarly, the causal Bayes net formalism has been plausibly applied to model
experimentation and causal reasoning in science. In short, Bayes nets research provides
significant reason to believe that both types of learning are, in some respects, importantly
similar. A central question this paper addresses, however, is whether such results support
the expansive view endorsed by Gopnik and her collaborators that childhood and scientific learning are in general both underwritten by a common array of learning processes.

In the next section, I challenge this view by highlighting putatively unique learning processes underlying scientific hypothesis formation and confirmation. For each type of presumptively unique learning in science, I argue that an adequate model likely requires one or more of the following: new methods for learning or assigning probability distributions to graphs, new kinds of graphical structures, or distinct non-graphical formalisms altogether. Thus, on the assumption that the standard Bayes net formalism models mental representations and learning processes in both scientists and children, there is reason to classify such departures from the standard Bayes net formalism as models of new kinds of learning process.

I claim these new kinds of learning in science are plausibly involved in scientific reasoning about cyclic causal systems, analogical hypothesis formation, and theoretical identifications and reductions. The enumeration of such differences between childhood and adult scientific learning undermines expansive versions of the theory theory and instead supports:

*The Process Development Hypothesis:* The development of adult scientific inferential capacities from childhood learning includes the advent of new kinds of learning process for acquiring and confirming theories.
I will say much more about this hypothesis in Section 4. As a first approximation, the central idea is that cognitive development – and the development of adult scientific inferential capacities in particular – involves not just the acquisition of new concepts and knowledge but also the emergence of genuinely new ways of learning. In addition, I will defend an accompanying set of proposals for characterizing this development. All of these proposals posit cognitive processes that can be modeled by departures of varying degrees from the standard Bayes net formalism, up to and including learning processes coming on-line that are appropriately modeled by distinct formalisms.

I should emphasize that the considerations offered in the next section for The Process Development Hypothesis and its accompanying proposals are hardly dispositive. As one might expect in the exploration of a range of novel proposals in developmental psychology, the proposals are tentative, and the evidence in their favor is thoroughly defeasible. My primary purpose is to render a neglected family of hypotheses plausible candidates for future empirical inquiry.

4. Limitations of Bayes Nets and the Theory Theory

The considerations in this section follow a common structure. First, I emphasize a central feature of scientific inference that I provide reason to believe is not a feature of early cognitive development. Next, I argue that this aspect of scientific inference is also likely not accommodated by the standard Bayes net formalism. Taken together, these two points support the claim that scientific inference is underwritten by a kind of learning process.
that does not underwrite early cognitive development. Finally, I offer remarks about how one might understand the development of the putatively novel learning process, either in terms of significant generalizations of the standard Bayes net formalism or in terms of distinct models. The cumulative effect of these considerations, I maintain, casts doubt on expansive versions of the theory theory and supports The Process Development Hypothesis.

4.1 Cyclic systems

Learning and reasoning about cyclic causal relations is an important component of scientific inference. Cyclic systems are pervasive in nature and have been successfully modeled by scientific theories in a wide variety of theoretical domains. For example, in the medical sciences, understanding disease entails understanding that it weakens the immune system, which can lead to a stronger disease, and so on. Similarly, appreciating the biochemical pathways associated with disease transmission requires understanding the nature of feedback loops in these pathways (Thagard 2003). Cyclic modeling of ecosystems and the climate are commonplace. Melting snow causes less sunlight refraction, which causes more melting snow, etc. In general, appreciating the greenhouse effect involves understanding cyclic causal relations between heat absorbed by the Earth’s surface and heat transmitted to the atmosphere. In addition, the social sciences utilize a class of “non-recursive” models, in which a variable X is a function of Y and vice versa. This model class is valuable for understanding bidirectional relations between,
for example, supply and demand, poverty and crime, and so on (Glymour 2010). In sum, there is every reason to believe that appreciating and representing cyclic systems is a central aspect of scientific hypothesis formation and confirmation.

Is reasoning about cyclic systems also a central feature of early cognitive development? There are no studies, so far as I have been able to discover, establishing that infants and young children definitively lack a capacity to reason about cyclic causal relations. So the evidence for or against such a capacity is hardly conclusive. It is suggestive, however, that whereas reasoning about cyclic causal systems is nearly everywhere one cares to look in science, it is nowhere to be found in standard texts on early causal reasoning (e.g., Koslowski and Masnick 2010) or in the body of research produced by theory theorists.

In addition, there is indirect evidence that infants and young children generally lack sophisticated capacities for representing and reasoning about cyclic systems, including cyclic causal systems. This evidence comes from deficits in children’s understanding of *temporal* cyclic structures, especially before age five. In particular, children under five do not exhibit a hallmark of cyclic representation in reasoning about the seasons or cyclic daily routines (breakfast, lunch, dinner, bedtime, breakfast, etc.), namely, facility in reasoning forward and backward from arbitrary points in a cycle (Friedman 1990). Further, though young children can grasp sequential temporal structure, they do not appear to differentiate non-cyclic from cyclic ordering when

---

47 Reverse-order reasoning tasks are taken, in part, to measure reasoning with cyclic as opposed to linear representations. The assumption is that if there is linear representation, then there are distinctive difficulties associated with tasks that require tracing the linear sequence backwards (hence the backward alphabet test as a measure of sobriety).
presented with permutations that preserve cyclic structure (Friedman 1989, 1990). In contrast to adult scientific inference, the above considerations provide reason to believe that a domain-general capacity for learning, representing, and reasoning about cyclic systems is not a significant component of early cognition.

Does the standard Bayes net formalism have the resources to model learning and reasoning about cyclic causal systems? Since Bayes nets are a species of Directed Acyclic Graphs, there is some reason to believe that representing cyclic relations requires at least a modest departure from the standard formalism and perhaps accommodation by distinct models. In fact, dynamic Bayesian networks were developed relatively quickly in the Bayes net literature as a generalization of the standard formalism precisely in order to model reasoning about cyclic systems (Ghahramani 1998). Dynamic Bayes nets involve a sequence of Bayes nets, each of which might be interpreted as a time-slice of a cyclic system, and all of whose variables are temporally indexed. A transition matrix then characterizes conditional probabilities that hold between variables in different time slices.

More substantial departures from the standard formalism include Neil et al. (2000)’s generalization to computing cyclic graphs. They claim that aspects of the standard Bayes net formalism can be applied to cyclic graphs if the formalism is first applied to the graph’s smaller acyclic components. Then, with the assistance of new estimation techniques, a joint probability distribution can be assigned for the larger cyclic structure. Another generalization of the standard Bayes net formalism to cyclic structures is the extension by Spirtes et al. (2000) of intervention learning methods to cyclic graphs. Some trends in machine learning, however, attempt to characterize learning and
reasoning about cyclic systems with formalisms that are distinct from Bayes nets. For example, Neville and Jensen (2007) argue that relational dependency networks have advantages over dynamic Bayes nets and other generalizations of the standard formalism, including greater efficiency and the capacity to recover cyclic structure when temporal information is unknown.\(^{48}\)

My purpose here is not to advocate any particular approach to modeling learning and reasoning about cyclic systems. Rather, I am highlighting a first example of an important component of scientific inference that (a) likely falls outside the standard (causal) Bayes nets formalism and (b) is likely dissimilar to early cognitive development. I maintain that such a feature of scientific learning provides initial grounds for doubting an expansive version of the theory theory. Following the evidence where it currently points instead provides reason to believe that scientific reasoning about cyclic causal systems is not underwritten by a kind of learning process that underwrites early cognitive development. There are, however, a variety of approaches in the machine learning literature for characterizing the adult scientific inferential capacity to reason about cyclic systems. Particular proposals of interest concern modeling this development with some type of generalization of the standard Bayes net formalism, though alternate formalisms might also be considered. Either approach provides support for The Process Development Hypothesis and casts doubt on the view that cognitive development is invariably underwritten by repeated applications of the same learning processes.

\(^{48}\) For example, Neville and Jensen (2007) discuss purportedly superior applications of relational dependency networks in scientific contexts, including reasoning about co-located proteins in a cell and tuberculosis infection among subjects in close contact.
Reasoning about cyclic causal systems, then, is our first example that advances these broad points. In the next sections I discuss two additional components of scientific inference – analogical hypothesis formation and theoretical identifications and reductions – that lend these claims further support.

4.2 Analogical Inference

Examples of analogical reasoning in the history of science are legion: the Rutherford atom model was analogically based on the solar system; Darwin modeled natural selection on artificial selection; Pasteur analogized disease to fermentation (Thagard 1997); Huygens and Young modeled theories of light waves on sound waves (Hesse 1966); and Kepler seemed to think about the solar system in any and every form of analogy – in terms of lodestones, sailors, and orators at a crowd, to name a few (Gentner et al. 1997). Perhaps Kepler was unique in his placement of analogical reasoning at the very heart of scientific inference, but any account of scientific theory construction that did not also classify analogical reasoning as central to hypothesis and theory formation would be inadequate. For the purpose of comparing scientific inference with early cognitive development, note that the above examples of analogical reasoning in scientific contexts plausibly involve appreciating abstract relations between sources and targets in different cognitive domains.\(^{49}\) In addition, it is doubtful that such reasoning could be

\(^{49}\) There is substantial dispute over how to individuate *domains* in cognitive science. Nothing turns on the outcome of such disputes here, however, since our examples are cases of prima facie information flow.
characterized as a species of enumerative induction. The same is not necessarily true of
the less sophisticated similarity-based judgments in early cognitive development. Or so I
argue below.

Following Gentner and Ratterman (1991) and Gentner and Namy (1999), it
appears that, to the extent that analogical inference is present in early cognition, there are
importantly different levels of sophistication at different ages and even importantly
different levels of sophistication with respect to different cognitive domains. Around four
months is the earliest age for which there is evidence that infant inference might be
characterized in terms of the canonical analogical form A:B::C:D. Such inferences,
however, involve expectations for future objects based on observations of past objects
that are highly similar with respect to properties that are directly perceptually detectable,
such as shape and color (Baillargeon 1991). Indeed, such “similarity-based” judgments
might also be characterized as a species of enumerative induction where the same is not
the case for the previous examples of sophisticated analogical hypothesis formation in
science. Thus, there is some reason to differentiate such similarity-based cognition at
these early stages of development from the kind of analogical reasoning in scientific
discovery.

Later, at about two years of age, children typically recognize similarity and
difference relations that are also centered around properties that are directly perceptually
detectable, but with more latitude. For example, similarity-based predictions and

across cognitive domains however one might nontrivially individuate them for the purposes of cognitive
psychology.
expectations for red round apples and red round balls might be the same. Around three years, children first begin to classify objects for the purpose of similarity-based judgments according to properties other than shape, color, etc., including especially their function. For example, children at this age show the ability to make A:B::C:D inferences about objects that share a capacity to soak up water but that might be of quite dissimilar shapes or colors (Gentner and Namy 1999, 489-90). This new ability to appreciate functional and more abstract relations is classified by Gentner and Ratterman not as a mere difference in degree with respect to earlier analogical capacities, but as a difference in kind. This “relational shift,” as Gentner and Ratterman call it, however, is not apparently a feature of reasoning in every domain about which early cognizers make similarity-based judgments. For example, three-year-olds do not typically classify biological entities according to function; e.g., they do not typically group predators or carnivores together (Genter and Ratterman 1991). This suggests that the relevant cognition is domain-specific to an extent that is also distinct from the cross-domain processing associated with analogical scientific hypothesis formation.

Some time in their fourth year of life and into their fifth, children’s capacities for analogical reasoning begin to resemble more substantially the capacities of adult reasoners, including scientists. I will not trace further developmental stages that Gentner and collaborators attribute to children over the age of four that culminate in the capacity, present in most children by age nine, to appreciate higher-order relations between relations and to map across cognitive domains. The above is sufficient to illustrate the contrastive point between the sophisticated, cross-domain analogical inferences made by
adult scientists and what are likely lower-level, similarity-based judgments and domain-specific analogical inferences made during early stages of cognitive development. This contrastive point suggests that some of the cognitive processes underlying scientific hypothesis formation are different in kind from the processes underlying early cognitive development.

There has been very little work on applying the Bayes net formalism to model analogical inference. A recent exception is Holyoak et al. (2010), who develop a generalization of the standard formalism for implementing analogical mapping between sections of Bayes nets, and in particular attempt to model Darwin’s comparison between artificial and natural selection. There are a few reasons, however, for thinking that both the standard Bayes net formalism as well as generalizations of this formalism are not particularly suited, in the general case, to model analogical inference. First, the most plausible models of analogical inference characterize it as a form of mapping between structured representations (Hummel and Holyoak 1997). The variables in Bayes nets, however, are unstructured despite sometimes denoting complex propositions. In this sense, the standard Bayes net formalism appears to be too “coarse-grained” to capture many cases of analogical inference. Sections of causal Bayes nets, however, are structured representations and generalizations of the standard formalism that accommodate mapping between these sections, such as those pursued by Holyoak et al. (2010), perhaps have a role to play in understanding some cases of analogical inference.

But there is a second worry that raises concerns about the capability of even generalizations of the Bayes net formalism to provide an inclusive model of analogical
inference. The worry is that the relations between nodes in a Bayes net are restricted to causal or probabilistic relations; yet it is not plausible that such restrictions are generally operative in analogical inference. Instead, non-causal and non-probabilistic relations – such as spatial, part/whole, and ordering relations – appear to be involved in successful mapping of source representations onto target problems. For example, Kepler’s attempt to understand the solar system in terms of an orator at a crowd suggests that structured representations of spatial relations between a centered body and accompanying satellites were operative, and a similar point applies to the development of the Rutherford model of the atom.

In short, there is reason to believe both that analogical hypothesis formation is central to scientific inference and that many such inferences resist accommodation by standard Bayes nets models. As with reasoning about cyclic causal systems, the most promising accompanying proposals for modeling sophisticated scientific analogical hypothesis formation might well include some significant generalizations of the standard Bayes net formalism. In contrast, however, it appears that the relevant and plausible proposals for modeling scientific inference in this case are more likely to incorporate formalisms that are distinct from the family of graphical structures associated with Bayes nets. Thus, we have reason to believe that sophisticated analogical scientific hypothesis formation is dissimilar in important respects to early cognitive development. Further, it appears that the development of sophisticated, cross-domain analogical inference by adult scientists might be characterized, at least in part, as the development (after the age of five) of new and distinct kinds of learning process. In this way, we have again motivated
The Process Development Hypothesis and undermined the view that learning in children and adults is underwritten by the same collection of learning processes.

4.3 Theoretical Identifications and Reductions

Providing an account of theoretical identifications and reductions should be a component of any adequate model of scientific theory formation and revision. Important historical examples of theoretical identifications include substances being identified with the chemical compositions of materials, the identification of contagions with germs in the formulation of the Germ Theory of Disease, and the identification of sound with a mechanical wave transmitted through a molecular medium. Indeed, proponents of psychological modeling with Bayes nets are themselves proposing a kind of theoretical identification, namely, that central components of theory construction in both science and early childhood should be identified with learning causal Bayes nets. In addition to theoretical identifications, we might also consider the closely related phenomenon of theoretical reductions, such as the reduction of classical genetics to molecular biology or the reduction of a cluster of thermodynamical concepts concerning temperature to central portions of statistical mechanics. As with analogical hypothesis formation and reasoning about cyclic causal systems, theoretical identifications and reductions are not de minimus components of scientific inference. Rather, they are a central and important way in which science advances.
Are theoretical identifications or reductions features of early cognitive development? There are no studies establishing that babies and young children lack the capacity to engage in theoretical identification or reduction, so the question cannot be answered with full certainty. As with reasoning about cyclic systems, however, it is striking that theoretical identification and reduction are highly visible aspects of scientific inference, yet wholly absent from extant accounts of early cognitive development. This is especially surprising given the popularity of theory-theorist approaches in developmental psychology. For if expansive versions of the theory theory were true, we would expect the existence of widespread inter-theoretic reduction and identification in early childhood. Yet theory theorists have not provided the slightest evidence that earlier and later stages of intuitive theory acquisition are connected via bridge laws or that early knowledge acquisition can be characterized in terms of identification or reduction. This provides reason to believe that intuitive childhood theories, in stark contrast to scientific theories, do not change via theoretical identification and reduction.

On the assumption that this difference between science and childhood obtains, how might we characterize the development of an adult scientific inferential capacity for theoretical identification and reduction? The standard (causal) Bayes net formalism does not appear to have much promise for accommodating such inferences. The main reason for this pessimism is that theoretical identifications and reductions involve relations of noncausal probabilistic dependence; that is, the variables denoting the parties being identified or reduced plausibly bear probabilistic, but not causal, relations. If such relations are represented on a directed acyclic graph, however, they violate the Causal
Markov Condition (CMC), which, as we have seen in Section 2, is a central component of the standard Bayes net formalism. Since the CMC says that a variable should be independent of all its non-effects given its direct causes, any noncausal probabilistic dependencies between propositional variables violate the CMC. Theoretical identifications and reductions are in fact one type of noncausal probabilistic dependency that belongs to a larger class of mathematical, logical, and lawful relations, none of which are readily represented by causal Bayes nets.

There is therefore reason for psychological modeling of theoretical identification and reduction in science to look beyond the standard (causal) Bayes net formalism. As with reasoning about cyclic systems and analogical inference, generalizations have been developed to accommodate reasoning with variables that bear noncausal but probabilistic relations. For example, Williamson describes an application of a Bayes net expert system to a context of scientific inference – attempting to assist in colonoscopies – where the Bayes net system encountered difficulties with data sets that bore mathematical, rather than causal, relations (2004, 63). In that instance, positing hidden or latent variables that screened off mathematically related variables helped to preserve causal relations of interest while still recovering relations of noncausal statistical dependence. This generalization of the formalism arguably compromises standard parameterizations and interpretations of Bayes nets as causal maps, however. My point is not that this particular strategy will necessarily illuminate models of scientific reasoning about theoretical identifications and reductions. Rather, I maintain that there is no reason to believe that there are principled limitations on generalizations of the Bayes net formalism for
successfully modeling reasoning about noncausal probabilistic dependencies in general and theoretical identifications and reductions in particular. Indeed, Glymour (2007) has claimed that dynamical Bayesian networks can accommodate many cases of identification and reduction in science (though Glymour invokes an idiosyncratic understanding of “local” relations of identity and reduction between micro and macro states).

Our reflections on theoretical identification and reduction again suggest that a central feature of scientific inference (a) falls outside the standard Bayes net formalism and (b) is likely dissimilar to early cognitive development. In turn, we have some reason to suppose that scientific reasoning about theoretical identifications and reductions is subserved by a unique type of learning process that does not subserve early cognitive development. Moreover, our reflections suggest that we might profitably characterize the development of the capacities underlying theoretical identifications and reductions either in terms of generalizations of the standard Bayes net formalism or as the advent of a new kind of cognitive process that is accommodated by a distinct formalism. If this approach to theoretical identifications and reductions is apposite, then there is a third central aspect of scientific inference – in addition to reasoning about cyclic causal systems and analogical hypothesis formation – that at once undermines the expansive version of the theory theory embraced by many developmental psychologists and instead supports The Process Development Hypothesis.

5. The Process Development Hypothesis
The Process Development Hypothesis (PDH) is not a single, well-defined hypothesis but a family of proposals for characterizing the cognitive development of adult scientific cognition. What the proposals have in common is that they characterize this development not merely in terms of the acquisition of new concepts and information, but also in terms of the emergence of genuinely new forms of learning.

This family of proposals might be thought of as an explanation for a striking phenomenon with respect to modeling human learning with Bayes nets. The phenomenon is this. Causal Bayes nets appear to be a plausible model for much of the acquisition, before the age of five, of intuitive theories especially in the physical, biological, and psychological domains. Indeed, Bayes nets models are arguably the best extant proposals for characterizing these central features of early cognitive development. When it comes to modeling scientific inference, however, the standard Bayes nets formalism seems to have unique and pronounced limitations. Thus, the formalism appears to accommodate a significantly greater part of early childhood learning than adult scientific learning.

The PDH is an explanation and characterization of this gap. It suggests that the features peculiar to scientific inference that fall outside the standard Bayes net formalism and that are dissimilar to early cognitive development should be understood as the emergence of new kinds of learning. The particular support for the PDH given here implies that there are at least three new kinds of learning process that underwrite scientific theory change: learning and reasoning about cyclic systems, sophisticated analogical hypothesis formation, and reasoning about theoretical identifications and
reductions. Further, I have argued that these proposed new kinds of learning can be characterized via models with departures of varying degrees from the standard Bayes nets formalism, up to and including distinct formalisms.

Why characterize these developmental models of novel scientific learning processes in terms of their relation to the standard Bayes nets formalism? On the assumption that Bayes nets model central features of early cognitive development, general considerations of parsimony favor such characterizations. That is, simplicity considerations recommend, ceteris paribus, explaining subsequent cognitive development in terms of modifications to existing cognitive structures rather than in terms of entirely distinct cognitive structures. Where appropriate, that is just what the PDH recommends. In addition, I have argued that the development of generalizations of the standard Bayes nets formalism in the machine learning literature gives us reason to be optimistic that generalizations of some form might accommodate important features of scientific inference, especially including reasoning about cyclic causal systems and theoretical identifications and reductions. Naturally, we should also be open to the possibility that positing learning processes coming on-line that are modeled by distinct formalisms might better accommodate the relevant psychological data. Of the three kinds of new learning I have argued for here, developmental hypotheses of this kind look most appropriate for characterizing analogical hypothesis formation.

A final and brief elaboration of the PDH does not concern how the putative new learning processes underwriting scientific inference should be modeled, but rather during what period(s) in the cognitive development of individuals such learning processes
emerge. My earlier criticisms of an expansive theory theory, if correct, entail that the PDH likely does not characterize cognitive development before the age of five. However, this leaves open a wide range of time periods during which these novel learning processes might come on-line. The particulars may vary for each kind of learning process, but I would like to address, in general, whether the PDH should be thought of as characterizing a non-scientist-to-scientist transition or a child-to-adult transition.

Consider first a non-scientist-to-scientist interpretation of the PDH. Someone attracted to this interpretation might hold that scientists come to acquire proprietary learning capacities in virtue of undergoing scientific training. Perhaps one might hold that the transition from everyday adult reasoning to scientific cognition is effected by culturally scaffolded processes that occur in virtue of participating in scientific institutions, substantially collaborating with members of the scientific community, and becoming instructed on the methods and history of scientific inquiry. This interpretation of the PDH predicts significant disparities between normal adult reasoning capacities and scientific inferential capacities. In particular, it predicts such disparities with respect to reasoning about cyclic systems, engaging in sophisticated analogical hypothesis formation, and formulating theoretical identifications and reductions.

In contrast, an interpretation of the PDH as characterizing a child-to-adult transition does not make these predictions. Rather, it predicts that everyday adult reasoning should also exhibit facility with learning about cyclic causal systems, sophisticated analogical hypothesis formation, and theoretical identifications and reductions. Note that if a child-to-adult interpretation of the PDH is correct, it provides
partial support for the position that there is a fundamental similarity between adult
quotidian cognition and scientific hypothesis formation and confirmation. This is a
position that many have held,\(^50\) including the view that a stronger analogy likely holds
between scientific inference and quotidian adult cognition than between scientific
inference and early cognitive development. I will briefly say why I think this framework
is plausible here and thus supports a child-to-adult interpretation of the PDH.

The central reason to prefer a child-to-adult interpretation of the PDH is that typical
adults do not appear to exhibit the relevant deficits predicted by the non-scientist-to-
scientist interpretation. For example, adults do not have deficits in understanding cyclic
structures generally, as deficits for temporal cyclic reasoning do not exist much beyond
age six (Friedman 1990). With respect to reasoning about *causal cycles*, such a capacity
would appear to be required for adults to navigate many simple feedback systems
encountered in everyday life. For example, there is a feedback loop between most home
thermostats and furnaces, where the states of the one are typically determined by the
states of the other. It is prima facie reasonable that adults who are responsible for
maintaining these systems comprehend this basic feature of their structure. With respect
to scientific analogical hypothesis formation, there is significant evidence that adults
possess similarly sophisticated capacities and may utilize them in problem-solving tasks
(Gick and Holyoak 1983). So again, there is reason to believe that everyday adult
reasoning and scientific inference are similar in important respects. The development of

\(^{50}\) For example, Fodor (1983, 2001, 2008); Newell and Simon (1976); and, perhaps, Einstein: “The whole
of science is nothing more than a refinement of everyday thinking” (1936, 59).
reasoning about theoretical identifications and reductions is not as straightforwardly placed inside a child-to-adult transition period, though we may believe that the recognition of Frege cases by typical adults could provide evidence for some of these same sorts of cognitive capacities exhibited by scientists. For these reasons, an interpretation of the PDH as characterizing a child-to-adult transition, as opposed to a non-scientist-to-scientist transition, appears most plausible. I will leave further elaborations on the development of adult scientific inference for future empirical inquiry, however.

6. Conclusion

In conclusion, I offer some remarks on both what I have attempted to show and what I have not attempted to argue. First, the foregoing discussion has not undermined every version of the theory theory in developmental psychology, only the expansive version advocated by some of the same developmental psychologists who have pursued research on psychological modeling with causal Bayes nets. Second, I have not established that every learning mechanism proposed by theory theorists, or that might be proposed by theory theorists, necessarily has the same limitations I have claimed restrict standard causal Bayes nets.\(^{51}\) It is of course open to empirical inquiry whether theory theorists might identify learning mechanisms that plausibly underwrite early cognitive

\(^{51}\) For example, I have not addressed the distinct proposals of learning process by Tenenbaum et al. (2006) or by prominent connectionists (e.g., McClelland and Rogers 2007), both of which also posit a high degree of similarity between early childhood and scientific learning.
development and that also plausibly underwrite sophisticated analogical hypothesis formation, reasoning about cyclic causal systems, and reasoning about theoretical identifications and reductions. The preceding might be taken as a challenge for theory theorists to produce a plausible model of such learning mechanisms.

What have I attempted to show? First, the most developed and significant learning mechanism so far proposed by proponents of an expansive version of the theory theory has several important and apparently unique limitations with respect to modeling scientific inference. Second, these limitations provide reason to believe, in spite of the success of the Bayes nets formalism in modeling some aspects of both childhood and scientific learning, that unique learning processes underlie scientific theory formation and revision. Third, the conjunction of the previous points constitutes a substantial challenge to an influential research paradigm in developmental psychology that views childhood and scientific learning as underwritten by repeated applications of the same learning processes. Fourth, Bayes nets research and related advancements in machine learning instead support the PDH and accompanying proposals for modeling scientific learning processes in terms of departures of varying degrees from the standard Bayes nets formalism. Finally, I hope to have rendered plausible the idea that working with a different paradigm in developmental psychology might lead to a richer understanding of the cognitive development of scientific inference.
Chapter 5: Extended Scientific Minds and Population-level Theory Change

Abstract: According to externalist theories of scientific inference, scientific theory formation and revision are sometimes underwritten by cognitive processes that occur beyond the biological boundaries of individual scientists. Two kinds of externalist theories deserve to be sharply distinguished. The first I call *The Extended Scientific Mind*, which is an application of The Extended Mind to contexts of scientific inference. The second I call *population-level* theories of scientific inference, which characterize the dynamics of communities of scientists at the population-level. I argue that population-level theories of scientific inference are more plausible than The Extended Scientific Mind. If I am right, a similar moral likely holds for extended and population-level accounts of ordinary cognition.

1.0 Introduction
This paper addresses two foundational issues in the cognitive science of scientific inference.\(^5^2\) The first concerns the extent to which *externalist* frameworks are appropriate for characterizing scientific theory formation and revision; the second concerns the relationship between internalist and externalist theories. Addressing both issues, I will argue, also has important implications for understanding the nature of ordinary cognition.

According to *Internalism*, cognitive mechanisms, representational vehicles, or other cognitive structures are located within the head or skin of individual cognitive agents. An internalist framework for characterizing scientific inference is captured by:

*Scientific Internalism*: The cognitive systems underwriting scientific theory formation and revision are circumscribed inside the head or skin of individual scientists.

*Externalism* has been characterized by a collection of concepts and terms, including “embodied,” “distributed,” “situated,” “extended,” “collective,” “vehicle externalism” and “locational externalism,” to name the most prominent.\(^5^3\) The core idea of Externalism is that cognition, at least some of the time, is underwritten by cognitive mechanisms, representational vehicles, or other cognitive structures whose boundaries are located beyond the biological boundaries of individuals. I am centrally concerned in this article

---

\(^{5^2}\) I am here concerned with *descriptive* theories of scientific inference as opposed to, for example, normative or ideal agent accounts of scientific inference that are often associated with inductive logics or Bayesian confirmation theory. On the distinction between different types of theories of scientific inference, see Fuller & Samuels (in preparation).

\(^{5^3}\) Neither Internalism nor Externalism is intended to characterize the *content* of cognitive states. Instead, they are theses about the boundaries of the mechanisms, representational vehicles, or other physical structures that underwrite cognition.
with the extent to which externalist theories appropriately characterize scientific inference. Thus, consider an externalist framework captured by:

*Scientific Externalism*: Scientific theory formation and revision are underwritten by cognitive systems whose boundaries are located beyond the head or skin of individual scientists.

Though I will assess a cluster of externalist proposals simultaneously, two types of externalist frameworks for characterizing scientific inference ought to be sharply distinguished despite often being run together. The first I call *The Extended Scientific Mind*. According to a theory of this type, the cognitive systems underwriting an individual scientist’s inferences can “extend” out onto her extra-cranial environment and incorporate elements with she is intimately intertwined. As I construe The Extended Scientific Mind, it characterizes the cognition of individual scientists and thus is an *individual-level* theory. That is, as I am understanding extended scientific minds here, they do not extend onto other scientists. In contrast, *population-level* theories of scientific inference characterize the dynamics of groups of scientists and communities of researchers. Theories of this type posit cognitive features that are instantiated by these groups or populations and not by the individual scientists who comprise them.

Proponents of externalist accounts of cognition in general often simultaneously endorse both extended and population-level theories. For example, Wilson (2010) claims: “The extended mind thesis identifies cognitive systems themselves as reaching beyond individuals into their physical and social environments” (p.171). Similarly, proponents of Scientific Externalism in particular have simultaneously endorsed both The Extended
Scientific Mind and population-level theories of scientific inference. Giere (2002) claims: “…parts of the cognitive process [underwriting scientific reasoning] take place not in anyone’s head … The cognitive process is distributed among humans and material artifacts” (p.287). A taxonomical approach suggested by such remarks is one that treats population-level theories as a special case of The Extended Mind; that is, it treats population-level theories as characterizing particular cases where an individual mind extends out into a community in which it is embedded. This taxonomy implies that if population-level theories are viable, then they support the more general extended accounts of which they are an instance. Such a taxonomy, I will argue, is mistaken. As I seek to establish in what follows, extended and population-level accounts of scientific inference are not equally plausible and therefore there is reason to sharply distinguish them.

This paper has four central aims: first, to undermine the viability of The Extended Scientific Mind; second, to provide support for population-level accounts of scientific inference.
theory change; third, to advocate for relations of *autonomy*, as opposed to *competition*, between internalist and externalist theories of scientific inference; and fourth, to provide reason to believe that similar claims hold for theories of ordinary, non-scientific cognition.

In order to accomplish these aims, I begin in Section 2 by considering three broad views on the relationship between internalist and externalist theories of scientific inference. I then introduce The Extended Mind in general and The Extended Scientific Mind in particular. In Section 3, I object to extended scientific minds on the grounds that they are either explanatorily superfluous or explanatorily destructive. I suggest that this objection also undermines extended mind accounts of “higher” cognition generally. Finally, in Section 4 I defend population-level theories of scientific inference that bear relations of autonomy to internalist theories.

**2.0 Internalist and Externalist Theories**

Three approaches to the relationship between internalist and externalist theories of scientific inference sufficiently carve the relevant terrain for our purposes. The first is:

*Hard Internalism*: Externalist accounts of scientific inference should be reduced to, or replaced by, internalist theories.

According to Hard Internalism, theories of scientific inference with internalist commitments will satisfy the central goals of explanatory accounts of scientific reasoning. What central explanatory goals should theories of scientific inference seek to accomplish? I will assume that among these goals are the following: (a) to explain the powerful and flexible capacities of scientific learning processes to generate
approximately true accounts of the world in the face of evidence that under-determines these accounts; and, relatedly, (b) to explain the capacity to successfully assess hypotheses not merely in terms of their empirical adequacy, but also in terms of such theoretical virtues as simplicity, consistency, conservativeness, coherence, and explanatory power. We can understand Hard Internalism as the view that only theories within an internalist framework are likely to accomplish these goals.\textsuperscript{55} Hard Internalism is a view with some currency in cognitive science; for example, it is an implicit commitment in Chomsky’s (2000) discussion of a “science-forming-faculty”\textsuperscript{56} and is explicitly endorsed in Fodor’s work.\textsuperscript{57}

A second approach to the relationship between internalist and externalist theories of scientific inference is:

\textit{Autonomous Externalism}: Neither internalist nor externalist theories of scientific inference legitimately reduce or eliminate each other. Instead, they bear relations of autonomy.

\textsuperscript{55} A useful analogy for characterizing Hard Internalism might be found in the debate over whether natural selection should solely be understood as a causal process at the individual-level (Bouchard & Rosenberg 2004), or whether it is a statistical concept that most appropriately applies at the level of populations of organisms (Walsh, Lewens, & Ariew 2002). Hard Internalists, like Bouchard & Rosenberg on natural selection, do not see a legitimate role for non-individualistic accounts. Instead, Hard Internalists maintain that scientific hypothesis formation and confirmation is solely addressed by individualistic theories.

\textsuperscript{56} Chomsky (2000) posits a “science-forming faculty” (SFF), viz.: “certain components of the mind … that enter into naturalistic inquiry, much as the language faculty … enters into the acquisition and use of language” (2000, p.22). According to Chomsky, the SFF is a “biological system” with a proprietary “scope and limit” (2000, p.83). Chomsky’s implicit commitment to Hard Internalism lies in his appeal to the SFF’s biological boundaries to explain why science has failed to make progress in so-called “mysterious” domains, where the SFF not only cannot solve certain problems but cannot formulate them correctly. Without a commitment to Hard Internalism, the SFF’s biological limitations would not explain why externalist scientific cognition could not make progress in mysterious domains.

\textsuperscript{57} See Fodor (2005, p.27) for an explicit rejection of population-level theories and (2007) for a rejection of The Extended Mind in general. See Fuller & Samuels (under review) on Fodor’s characterization of individualistic scientific reasoning competence. See Gopnik & Meltzoff (1997) for a rejection of population-level theories of scientific inference (p.15).
“Autonomous” should here be understood as “resists reduction or elimination” in a sense familiar from discussions over whether the special sciences—for example, biology or psychology—resist being reduced by chemistry, physics, or other “hard” sciences (cf. e.g., Fodor 1977). “Autonomy” in this sense is not meant to connote “unrelated.” Instead, an Autonomous Externalist might compare the relationship between internalist and externalist theories of scientific inference to the relationship between intimately related subfields—such as, for example, between population-level and molecular genetics, or between microeconomics and macroeconomics.

The last type of approach to the relationship between internalist and externalist theories of scientific inference I consider is:

**Competitive Externalism**: Externalist theories of scientific inference are successful competitors with internalist accounts. In central cases, internalist theories should be replaced by, or subsumed under, theories with externalist commitments.

Among these three broad views—Hard Internalism, Autonomous Externalism, and Competitive Externalism—I argue that Autonomous Externalism is the most likely to be correct. Specifically, I argue for the viability of population-level theories of scientific inference and for the claim that they plausibly bear relations of autonomy to internalist theories of scientific inference. If I am right, it follows that Hard Internalists are mistaken. So too, I maintain, are proponents of The Extended Scientific Mind. I argue that extended scientific minds are implausible precisely because they are invariably wedded with an ambitious version of Competitive Externalism. Establishing this claim is
the central purpose of the next sections. First, however, I will characterize The Extended Mind in general and then turn to characterizing extended scientific minds in particular.

2.1 The Extended Mind

The Extended Mind purports to characterize a cognitive agent’s systematic interactions with her environment in terms of cognitive systems that contain both the agent and any environmental elements with which she is intimately intertwined. On this externalist approach, when a cognitive agent is appropriately coupled with extra-cranial elements in her environment, her cognitive vehicles, mechanisms, or other cognitive structures can “extend” beyond their biological boundaries and integrate with these environmental elements. The extended cognitive systems that form in such circumstances, according to proponents of The Extended Mind, do not supervene on the intrinsic properties of a cognitive agent despite implementing her cognitive achievements. Rather, extended mind theorists hold that the borders of cognitive systems do not invariably coincide with the biological boundaries of individuals. Indeed, on many extended mind accounts of cognition, a cognitive agent’s cranial boundaries are treated as more or less “arbitrary” in terms of tracing the perimeters of cognitive mechanisms, representational vehicles, or other cognitive structures (Wilson & Clark 2006). The tendency to adopt internalist frameworks for characterizing cognition is often regarded as a kind prejudice, perhaps based on assumptions inherited from Descartes.

Inspiration for The Extended Mind has been drawn from a variety of fields that describe organism-environment relations, especially including niche construction in
biology and ecology. Rob Wilson and Andy Clark, for example, draw on work in these areas when characterizing extended cognitive systems:

Thinking is a kind of building, a kind of intellectual niche construction that appropriates and integrates material resources around one into pre-existing cognitive structures. In cognition, agents modify or augment the capacities that those pre-existing structures enable (Wilson & Clark 2006, p.142).

According to Wilson and Clark’s version of The Extended Mind, relations to the environment that systematically augment and modify internal cognitive structures can generate extended cognitive systems. In particular, they hold that brains may “functionally integrate” with their environment so as to form extended information processors whose decomposable parts include both intra and extra-cranial components. One result, according to this version of The Extended Mind, is that an extended subvenience base that incorporates extra-cranial elements —or “wideware” (Wilson 2004)— allows for cognitive demands to be “offloaded” from internal onto external resources (Wilson & Clark 2006, pp.143-144). It is part of the promise of The Extended Mind that we might better understand our ability to intelligently interact with the environment, and to understand the nature of our cognitive achievements in general. In what follows, I will focus on Wilson and Clark’s particular version of The Extended Mind, though I believe the discussion has implications for a broader set of views.58 My

58 There are many versions of The Extended Mind on offer. My selection of Wilson and Clark’s version is partially for convenience’s sake, and another version would serve equally well to fix ideas. In addition,
specific interest is in evaluating a class of proposals according to which extended cognitive systems underwrite “higher” cognition including, paradigmatically, conscious planning, decision-making and (scientific) reasoning. I will not comment on proposals of extended “lower” cognition, for example concerning aspects of early perceptual processing or sensorimotor control (e.g., Noe and Reagan 2001; Chemero 2008).

A canonical example of how a cognitive agent may use aspects of her extracranial environment to augment and modify existing cognitive structures is the case of Otto:

Otto suffers from Alzheimer's disease, and like many Alzheimer's patients, he relies on information in the environment to help structure his life. Otto carries a notebook around with him everywhere he goes. When he learns new information, he writes it down. When he needs some old information, he looks it up. For Otto, his notebook plays the role usually played by a biological memory (Clark and Chalmers 1998 p.14).

In this example, extended mind theorists claim that Otto and his notebook, given their systematic relations, together form the relevant subvenience base of the cognitive systems that underwrite the completion of Otto’s memory tasks. In addition to notebooks, extended mind theorists have claimed that the cognition of agents may extend onto a variety of inanimate and animate objects in their environment, including pencil and paper, navigational instruments, computers, waiters at restaurants, and many other

however, their version allows for the decomposition of extended cognitive systems into internal and external components, whereas some other versions do not (e.g., Chemero 2008). As I argue in what follows, this is a desirable feature of Wilson and Clark’s view.
environmental elements besides. As I am understanding The Extended Mind, however, it is an *individual-level* theory and cases where externalist cognitive systems are putatively constituted by multiple individuals should be considered separately under the auspices of a population-level theory.

There have been a variety of attempts by extended mind theorists to formulate general principles, the satisfaction of which would warrant positing extended cognitive systems. My intent here is not to exhaustively survey these proposals, nor the putative counterexamples offered by critics of The Extended Mind; instead I will briefly characterize the kinds of situations in which extended mind proponents typically claim that extended cognitive systems underwrite cognitive achievements. A “Parity Principle” has notoriously played a central role in the characterization of such situations, which Clark & Chalmers (1998) describe thus:

> If . . . a part of the world functions as a process which, were it to go on in the head, we would have no hesitation in accepting as part of the cognitive process, then that part of the world is (for that time) part of the cognitive process (1998).^{59}

Even enthusiastic proponents of The Extended Mind, however, do not typically regard the Parity Principle as a strict sufficient condition whose satisfaction guarantees the existence of extended cognitive systems. Rather, this principle, and principles like it, are more likely to be treated as a rough guide to the kinds of situations where cognitive

---

^{59} Giere has endorsed a similar principle in arguing for his version of Scientific Externalism:

> What makes a distributed cognitive system “cognitive” is that it produces an output that, attributed to an individual person, would clearly be a cognitive achievement ... (2007, p.23).
models with extended systems are likely to be fruitful. For example, Clark has recently referred to the Parity Principle as a “rule of thumb” or heuristic (Clark 2010, p.44).

Additional guides to positing extended cognitive systems that are offered by extended mind theorists include: when an individual has secure, stable, and unfettered access to environmental elements (so extended systems aren’t merely fleeting); when agent-environmental interactions can be characterized in terms of a consistent feedback loop; and when environmental elements systematically “augment” or “modify” internal cognitive structures. Clark, for example, describes extended cognitive systems as forming in “the presence of continuous mutually modulatory influences linking brain, body and world” (1997, p. 163). With the goal in mind of assessing The Extended Mind’s applicability to contexts of scientific inference, I will briefly draw some morals from previous discussions over the Parity Principle and related proposals for positing extended cognitive systems.

The extensive discussion over extended accounts of memory and the case of Otto in particular has yielded results that are relevant to evaluating extended scientific minds. Adams and Azaia (2001, 2010) point out that processes of memory storage and retrieval that rely on information encoded in the brain are quite different from interactions with the environment that rely on external records and prompts. In particular, internal processes that access and store internally encoded information typically exhibit a variety of unique properties: they are subject to Weber’s law, negative transfer, have associated recency,
priming, chunking effects, and so forth. In contrast, there is no evidence that processes which rely on external prompts and records, as illustrated by the case of Otto, exhibit these same features. For some internalists, such differences provide reason to doubt that there is any significant parity between internal memory and Otto’s procedures, and thereby provide reason to doubt the propriety of positing extended cognitive systems.

In response, externalists such as Clark have claimed that the functional similarities invoked by the Parity Principle are supposed to capture “the potential for some form of higher-level unification despite mechanistic dissimilarities” (Clark 2010, p.51). According to this response, internalists who object to the case of Otto are individuating functional roles too finely. Extended mind proponents argue, instead, that a Martian might have a memory that does not exhibit the uniquely human features highlighted in discussions over the differences between Otto’s procedures and normal memory. Thus, according to extended mind proponents, we should not exhibit a kind of chauvinism in our taxonomy of cognitive systems which excludes either Otto or Martians.

The morals I draw from this discussion are three-fold. First, there is reason to posit, for some purposes at least, purely internal cognitive processes. The case of memory shows that cognitive science legitimately characterizes cognitive phenomena that supervene on internal cognitive mechanisms. Second, it follows that there is a kind of explanatory asymmetry between the internalist and extended mind proponent. The special explanatory burden of an extended mind proponent is to justify positing extended

---

60 See also Rupert (2009).
cognitive systems in addition to internal cognitive structures. In the absence of such justification, skepticism toward extended cognitive systems is the default reasonable position. Third and finally, if extended cognitive systems exist, their processing will count as cognitive according to a relatively coarse-grained individuation of functional roles. If one adopts more fine-grained individuations—perhaps ones that invoke characteristics associated with unique structural features of the brain—an internalist framework is more likely to be appropriate. With these morals in mind, I turn to evaluating extended scientific minds.

2.2 The Extended Scientific Mind

Given a typical scientist’s apparent heavy and systematic reliance on various aspects of her environment, it would appear that scientific reasoning is a prime candidate, if any “higher” cognition is, for being underwritten by extended cognitive systems. Many scientists have built a kind of intelligence niche in their environments—they’ve constructed laboratories, journals, conferences, educational institutions, and sophisticated scientific instruments.

Scientific instruments in particular would appear to be environmental elements with which scientists are intimately intertwined and with which scientists are plausibly functionally integrated. Scientific instruments appear to augment and modify scientific cognition, at least to whatever extent notebooks, navigational equipment, or pencil and paper augment and modify ordinary cognition. In addition, scientists often have stable, unfettered, and reliable access to scientific instruments, and appear to use these instruments to “offload” such cognitive tasks as sophisticated measurement, information
storage, data retrieval and analysis, and a variety of other tasks besides. In addition, feedback loops plausibly exist between individual scientists and their instruments, since, we may believe, instruments are often calibrated and recalibrated in response to their own outputs. For these reasons, scientists and scientific instruments plausibly instantiate precisely the kind of “mutually modulating influences” that extended mind proponents have claimed warrant positing extended cognitive systems in other contexts. It would appear, then, that extended cognitive systems whose components include scientific instruments should exist, on the assumption that extended cognitive systems of any kind subserve scientific hypothesis formation and confirmation. Indeed, it appears reasonable to believe that extended scientific minds should exist if The Extended Mind has application to any contexts of “higher” cognition at all. In which case, the prospects for extended scientific minds may be an important test case for The Extended Mind in general.

What kinds of relations might The Extended Scientific Mind plausibly bear to internalist theories of scientific inference? There is reason to believe that The Extended Scientific Mind is most naturally wedded with some version of Competitive Externalism. On the assumption that scientific hypothesis formation and confirmation is sometimes achieved by scientists who are functionally integrated with scientific instruments, a purely internalist account that omits augmenting and modifying aspects of an extended information processor might reasonably be regarded as fundamentally inadequate. This inadequacy might be illustrated by analogy. Suppose there are cases where extended cognitive systems implement a scientist’s cognitive achievements. In such cases, a purely
internalist account of scientific cognition would be like a description of a radio’s operations that focuses solely on the radio’s transformers or transistors but omits any mention of its antennae or amplifier. Such a narrow focus (so the extended scientific minds proponent might maintain) would be to mischaracterize the cyclical, augmenting, and modulating relations that can hold between a scientist and scientific instruments in her environment. A better alternative, according to a proponent of The Extended Scientific Mind, would be to broaden one’s approach to individuating cognitive systems and embrace systems whose decomposable parts include both a scientist’s internal cognitive structures and any scientific instruments with which she is functionally integrated. If such cases of functional integration give rise to extended cognitive systems, a proponent of extended scientific minds might appear to have every reason to either replace or subsume purely internalist alternatives with her extended account. In other words, Competitive Externalism most aptly characterizes the relationship that The Extended Scientific Mind plausibly bears to internalist theories of scientific inference.

Extended scientific minds have recently been endorsed by both Ronald Giere and Barton Moffat. Giere claims that “distributed cognitive systems,” especially in the area of experimental science, may contain an individual scientist together with any elements of the scientist’s environment that are capable of “external representation” (2007, p.314). According to Giere & Moffatt, this may include such elements and instruments as “Cartesian coordinates,” “telescopes,” “maps,” and a “pedocomparator” (2003, p.5). In evaluating the prospects of The Extended Scientific Mind, I will treat their proponents, including Giere and Moffat, as essentially engaging in a prediction. This general
approach to understanding Externalism has been advanced by Susan Hurley: “… [E]xternalism predicts that some good psychological explanations … will be externalist” (2010, p.107). Adopting Hurley’s reasonable approach, we can understand proponents of The Extended Scientific Mind as predicting that some amount of good psychological explanations of scientific reasoning will invoke extended cognitive systems.

Note that this prediction suggests the availability of very modest versions of The Extended Scientific Mind, according to which an extended account is appropriate in only a small number of contexts of scientific inference. In fact, I will argue that modest versions of The Extended Scientific Mind are essentially unavailable, and that if extended scientific minds have applications in some contexts then they are likely to have application in a wide variety of circumstances. Thus my strategy in undermining extended scientific minds is to undermine the prediction that extended accounts have wide applications across a variety of cases of scientific theory formation and revision. I maintain that there is simply no reason to suppose that positing extended scientific minds across a broad range of circumstances is explanatorily warranted; instead, there is good reason to believe that extended scientific minds are purchased only with explanatory costs that are highly likely to outweigh their benefits. This strategy for undermining extended scientific minds contrasts with strategies pursued in other contexts by critics of extended cognitive systems. In particular, it contrasts with attempts to identify a plausible necessary condition that all cognitive systems must satisfy, and then establish that extended cognitive systems fail to satisfy this condition. This is an approach that has been pursued by a number of others; I will not remark on their successes here, other than to
note that my strategy is different. In addition, I note that if my critique of extended scientific minds is cogent, then it may well have implications for The Extended Mind in general. Establishing these claims is the central aim of Sections 3.0 – 3.4.

3.0 The Objection to Extended Scientific Minds

An important explanatory cost that The Extended Scientific Mind faces concerns its prospects for adequately addressing:

*The Input/Component Problem*: The Extended Scientific Mind potentially treats environmental information that serves as *input* to information processors as functionally integrated *components* of those processors.

Wilson & Clark acknowledge that, for proponents of The Extended Mind in general, drawing a distinction between informational inputs and functionally integrated environmental components has been “surprisingly hard to pin down” (Wilson & Clark 2006, p.43). However, they claim that the distinction is “intuitive enough,” and that there are obvious cases where portions of the environment do *not* serve as mere input to an information processor but instead become functionally integrated components of that information processor.61 To illustrate the distinction, Wilson & Clark offer the example of a radio, whose primary function is:

… to receive, decode and play contents borne by radio signals …anything locally added that helps it do so (e.g. a better transistor at some key point,

61 The input/component distinction I am drawing has obvious affinities to the coupling/constitution distinction drawn by Adams and Azaia (2001). The main difference is that Adams and Azaia’s distinction is more general and thus raises a distracting question over whether this more general distinction has application in areas of science outside the study of cognitive mechanisms and information processing (see Ross and Ladyman (2010)). The input/component formulation of this distinction avoids such a distraction.
a signal amplifier, etc.) looks like an augmentation of the system rather than a mere input. This is so whether the additional transistor or amplifier falls inside the pre-existing boundaries of the radio or lies beyond them (2006, p.44).

We will have an opportunity to examine whether such a comparison to an exemplar case provides adequate resources for addressing The Input/Component Problem. My current purpose is to stress the potential explanatory costs for extended mind proponents of insufficiently addressing it. In the absence of a viable input/component distinction, extended mind theorists run the risk of treating all informational inputs from the environment as components of cognitive systems. Such a proliferation of extended cognitive systems in contexts of scientific inference would likely entail a multiplicity of explanatory superfluous cognitive systems—ones that form every time a scientist consults a book, makes an observation, or retrieves a print out of data. In addition to such a radical proliferation of extended cognitive systems, drawing an input/component distinction is plausibly required to characterize processes of scientific inference which take in evidential inputs from the environment and subsequently form, confirm, and revise theories. In short, the input/component distinction is not a mere metaphysical nicety, but a highly plausible requirement for characterizing the powerful and flexible cognitive systems that underwrite scientific inquiry. If, in the course of our investigations, we find that proponents of The Extended Scientific Mind are committed to a proliferation of extended cognitive systems without clear guidelines for distinguishing
inputs from components, it follows that their prospects for successfully addressing The Input/Component Problem are thereby reduced.

Proponents of population-level accounts of scientific inference, I will argue in section 4, do not face a similar challenge to their theory. Presently, we may appreciate that The Input/Component Problem poses a much less serious challenge for internalists. Certainly, an internalist must face the challenge of drawing some distinction between inputs and components. But a failure by internalists to draw such a distinction does not carry with it the potential cost of treating all informational inputs from the environment as components of cognitive systems, nor would it lead to a radical proliferation of cognitive systems. Rather, any proliferation or “enlargement” of cognitive systems for internalists would be circumscribed within the cranial boundaries of scientists. Thus, even in the absence of an internal input/component distinction for internal cognitive systems, an internalist may nevertheless draw an input/component distinction that distinguishes environmental informational inputs from components of cognitive information processors. Specifically, she may draw such a distinction by invoking the biological boundaries of perceptual systems — e.g., the surface of the retina or the surface of the outer ear’s external auditory meatus. Utilizing biological perimeters to draw an input/component distinction may gloss over desirable distinctions, for example, between perceptual input systems and the downstream neural structures that underwrite conscious reasoning. Nevertheless, even this coarse distinction would serve to distinguish environmental elements that function as inputs to an information processor as opposed to the processor’s components, and thereby facilitate explanations for the relative success of
generating approximately true accounts of the environment across a wide variety of
domains of scientific inquiry. In short, The Input/Component Problem is a uniquely acute
explanatory challenge for the proponent of The Extended Scientific Mind.

What is the import of The Input/Component Problem for The Extended Scientific
Mind? As I invoke The Input/Component Problem it is a central part of a larger challenge
to proponents of extended accounts of scientific cognition. The challenge is this: The
Extended Scientific Mind likely fails to accommodate the cognition that constructs and
modifies the very elements in the environment on which scientists heavily rely. In
particular, I will argue that proponents of The Extended Scientific Mind are likely unable
to accommodate the scientific cognition that underwrote, in the first place, the initial
invention, design and development of scientific instruments: Call this kind of scientific
cognition “instrument-developing” cognition. In contrast, call “instrument-aided”
cognition those inferential activities that are assisted by already existing scientific
instruments. I will argue that accounts which claim that extended scientific minds
underwrite instrument-aided cognition face insuperable difficulties in accounting for
instrument-developing cognition, difficulties that in fact expose The Extended Scientific
Mind as incapable of resolving the Input/Component Problem or as taking on other
undesirable explanatory costs. This challenge to the proponent of The Extended Scientific
Mind may be posed in the form of a dilemma, one that asks proponents of extended
cognitive systems whether instrument-aided cognition and instrument-developing
cognition are of the same or different kinds. That is, this dilemma challenges proponents
of extended scientific minds to determine whether instrument-developing cognition and
instrument-aided cognition perform the same or different kinds of function—whether they underwrite the same kinds of hypothesis formation, judgments of confirmation, insight into causal structure, or other kinds of cognitive achievement associated with scientific inference. I outline this dilemma briefly below and then pursue it at greater length in what follows:

The Dilemma: Either (i) instrument-developing cognition and instrument-aided cognition implement different kinds of function or (ii) they perform the same kinds of function.

The First Horn: Assuming (i), extended scientific mind proponents take on three explanatory costs:

Cost 1: (i) is prima facie implausible, since both instrument-developing cognition and instrument-aided cognition appear to engage in similar forms of non-demonstrative hypothesis formation and confirmation.

Cost 2: Extended accounts are likely to be less explanatorily powerful than internalist alternatives that classify instrument-developing cognition and instrument-aided cognition as performing similar kinds of function, in virtue of shared internal cognitive structures and capacities.

Cost 3: The proliferation of extended cognitive systems reduces the prospects for proponents of The Extended Scientific Mind to adequately address The Input/Component Problem.
**SUB-CONCLUSION 1:** Costs 1-3 are unlikely to be outweighed by the benefits of positing extended cognitive systems. So, on The Dilemma’s first horn, the best prediction is that extended scientific minds are either explanatorily superfluous or explanatorily destructive.

*The Second Horn:* On the assumption that (ii) obtains, two strategies for grasping the second horn are worthy of consideration:

*Strategy 1:* Instrument-aided cognition is often underwritten by *extended* cognitive systems, while instrument-developing cognition is in general underwritten by *internal* cognitive systems.

*Strategy 2:* Both instrument-aided and instrument-developing cognition are to the same extent underwritten by *extended* cognitive systems.

**SUB-CONCLUSION 2:** For both Strategies 1 and 2, positing extended systems is either explanatorily superfluous or explanatorily destructive.

**MAIN CONCLUSION:** On both horns of The Dilemma, extended scientific minds are either explanatorily superfluous or explanatorily destructive. Thus, the best prediction for the psychology of scientific reasoning is that extended scientific minds are explanatorily unwarranted.
I expand on The Dilemma in sections 3.1 – 3.4. The moral I draw from it is that extended scientific minds are undermined by reflection on the cognition that helps to design, construct, and develop the very elements in the environment on which scientists most heavily rely. Further, I argue that The Extended Scientific Mind is explanatorily suspect in large part because it is an ambitious competitor with internalist alternatives. That is, I argue that The Extended Scientific Mind is likely explanatorily undesirable precisely because it is wedded with a strong version of Competitive Externalism. The upshot is that population-level accounts of theory change that bear relations of autonomy to internalist theories are the most viable externalist framework for characterizing scientific cognition.

3.1 The Dilemma’s First Horn

On the assumption that there are extended scientific minds, it may seem plausible that instrument-developing and instrument-aided cognition are of different kinds. Perhaps a proponent of The Extended Scientific Mind might hold that instrument-developing cognition typically occurs prior to any scientific instrument becoming functionally integrated with any scientist. According to this approach to The Dilemma, when scientists do become functionally integrated with scientific instruments, the scientific instruments augment and modify existing internal cognitive structures such that new kinds of extended cognitive system form. Giere & Moffatt might be interpreted as grasping The Dilemma’s first horn:

The invention of new forms of external representation and of new instruments for producing various kinds of representations has played, and continues to play, a large role in the development of the sciences. From a cognitive science
perspective, both sorts of invention amount to the creation of new types of distributed cognitive system. So, for us, the notion of distributed cognition brings under one category such things as Cartesian coordinates and the telescope, both of which are widely cited as major contributions to the Scientific Revolution (2003: pp.4-5).

It is unclear from this passage whether Giere & Moffatt intend to merely propose that cognitive systems with new kinds of part are formed as the result of inventions of scientific instruments, or whether they claim that extended cognitive systems are formed with new kinds of function. The latter interpretation, on which Giere & Moffatt embrace The Dilemma’s first horn, asserts a view that faces an obvious prima facie objection: At first blush, both instrument-developing and instrument-aided cognition appear to be fundamentally alike. In particular, both appear to be constituted by similar forms of non-demonstrative hypothesis formation and confirmation.

As an illustration, consider the apparent similarity between instrument-developing and instrument-aided cognition with respect to the telescope. For example, Kepler was responsible for significant advances in the early stages of the development of the refraction telescope. He realized that combining a convex objective and a convex ocular lens would significantly increase magnification. He further realized that although this combination would produce an inverted image, the inversion could be reversed with the addition of a third convex lens. These design innovations increased magnification in the

62 I am classifying what Giere and Moffat here call “distributed cognitive systems” as a species of extended scientific minds.
short term from about 8x to 33x and in the longer term to nearly 100x (Anderson 2007). Later in the seventeenth century, Newton was responsible for designing and constructing the first working telescope that was centrally based on principles of reflection as opposed to refraction.

These kinds of telescope-developing cognitive processes likely included instances of hypothesis-formation, including the formation of design hypotheses that were informed by the best available optical theory. In addition, many of the cognitive processes underlying these advancements were likely also highly similar to processes of hypothesis confirmation, in particular to processes of optical hypothesis confirmation. There is every reason to believe that Kepler and Newton traced the implications of optical theory with respect to the refraction and reflection of light when improving the telescope. Thus, on the plausible assumption that hypothesis confirmation in general requires tracing a hypothesis’s entailments and predictions, Kepler and Newton’s cognitive achievements appear to include precisely the same kinds of cognitive achievement required for successful confirmation of optical hypotheses. In which case, there appears to be every reason to accept that Kepler and Newton’s telescope-developing cognition centrally involved processes that are of the same kind as processes of optical hypothesis formation and confirmation.

These glimpses of telescope-developing cognition suggest that it is markedly similar to the scientific hypothesis formation and confirmation that is assisted by telescopes. In the course of forming and confirming hypotheses, we may trust that astronomers who are aided by telescopes routinely gauge the refraction of light by the
Earth’s atmosphere. Further, there is every reason to suppose that the formation and confirmation of astronomical hypotheses routinely requires calculating the reflection of the sun’s light by non-solar bodies. In which case, telescope-aided cognition also centrally involves hypothesis formation and confirmation that is informed by the best available optical theory. Given these apparent and significant similarities, there is no evident reason to expect or desire that telescope-aided and telescope-developing cognition should be classified as performing different functions. Of course the possibility cannot be ruled out that proponents of extended cognitive systems might legitimately differentiate telescope-aided from telescope-developing cognition by invoking fine-grained distinctions between them. But pessimism about such a prospect is particularly warranted in light of the coarse-grained individuations of cognitive processes endorsed by proponents of extended cognitive systems in other contexts, including memory. In terms of addressing The Dilemma, this raises the following problem for proponents of The Extended Scientific Mind: Grasping The Dilemma’s first horn likely comes at a substantial explanatory cost if, as in the case of the telescope, instrument-developing and instrument-aided cognition apparently occupy the same kinds of functional roles.

Do the above reflections on telescope-developing and telescope-aided cognition generalize to typical acts of instrument-developing and instrument-aided cognition? There does not seem to be any reason to regard the telescope as unrepresentative of scientific instruments generally.\(^{63}\) Rather, telescope-developing and telescope-aided

\(^{63}\) Instead, the telescope appears to be a friendly example in some ways for the proponent of extended scientific minds who grasps The Dilemma’s first horn and claims that instrument-developing and
cognition are similar in ways we might expect instrument-aided and instrument-developing cognition to be similar in general. Instrument-developing and instrument-aided cognition likely both involve forms of hypothesis formation that is highly influenced by background scientific theories. Instrument-developing cognition in general likely generates design hypotheses and manifests insight into causal mechanisms of exactly the kind that scientific theories provide. Similarly, instrument-aided cognition often searches out new evidence about which to form hypotheses and facilitates insight into the world’s causal structure. In addition, there is every reason to expect that both instrument-developing and instrument-aided cognition often trace the predictions and consequences of scientific theories. In which case, both kinds of cognition appear to be underwritten by the kinds of cognitive processes that underwrite scientific hypothesis confirmation.

This plausible similarity between instrument-developing and instrument-aided cognition is further supported by a historical convergence on scientists who engage in both types of cognitive achievement. That is, scientists who have been in the business of forming, testing, and revising scientific theories have also often been in the business of inventing, designing and developing scientific instruments. This convergence provides additional reason to believe that fundamentally similar cognitive capacities underwrite instrument-aided cognition are distinct. The telescope in its earliest stages (3x magnification) wasn’t developed by scientists at all, but rather by craftsmen living in Europe in the sixteenth century about whom relatively little is known (Anderson 2007).

64 In addition to Kepler and Newton, prominent examples include: Huygens, who formulated scientific theories on the nature of sound and also designed the pendulum clock; von Neumann who made fundamental contributions to quantum mechanics and computer science as well as helping to design the first electronic computing machine; Tesla who contributed to electro-magnetics in both theory and application, and many more besides.
both instrument-developing and instrument-aided cognition. And if this is so, there is no apparent reason to expect or desire that these kinds of cognition should, in general, be classified as distinct in terms of their respective functional roles. Thus, grasping The Dilemma’s first horn appears to take on the significant explanatory burden of denying the prima facie plausible claim that instrument-developing and instrument-aided cognition are underwritten by cognitive systems that perform very similar kinds of functions.

In addition, a proponent of extended scientific minds who grasps The Dilemma’s first horn may well take on further explanatory costs. First, given a proliferation of extended cognitive systems that putatively underwrite instrument-aided cognition, she lowers her probability of satisfactorily addressing The Input/Component Problem. Second, this approach to The Extended Scientific Mind may be explanatorily destructive, given that an extended account of scientific inference whose generalizations classify instrument-developing and instrument-aided cognition as distinct would, ceteris paribus, be less powerful than internalist alternatives whose generalizations type them together. And there is reason to believe that a theory of scientific inference with internalist generalizations might well cover both instrument-developing and instrument-aided cognition since, plausibly, internal cognitive structures contributing to both might be shared. Thus, on the assumption that a theory’s explanatory power is partially determined by (a) the number of instances its generalizations cover, and (b) the resources with which it covers them, an internalist theory of scientific inference, would, ceteris paribus, be more powerful than the version of The Extended Scientific Mind under consideration. In short, positing extended scientific minds while grasping The Dilemma’s first horn
appears to be accompanied by substantial explanatory costs. As with any empirical theory, there is the possibility that proponents of The Extended Scientific Mind might overcome these substantial costs by offering a highly fruitful account of extended cognitive systems that underwrite instrument-aided cognition. Currently, however, the best prediction regarding proponents of extended scientific minds who grasp The Dilemma’s first horn is that they are, at best, positing explanatorily superfluous systems and, at worst, engaging in an explanatorily destructive enterprise.

3.2 The Dilemma’s Second Horn

A proponent of The Extended Scientific Mind who grasps The Dilemma’s second horn opts to treat instrument-developing and instrument-aided cognition as the same kind of cognition, as mental processes that perform the same kinds of function. Grasping The Dilemma’s second horn is prima facie more desirable than grasping its first, since, as we have remarked, instrument-developing and instrument-aided cognition are plausibly characterized as similar forms of non-demonstrative hypothesis formation and confirmation. In view of this, two importantly distinct strategies for grasping The Dilemma’s second horn are worthy of consideration. According to the first, instrument-aided cognition is often underwritten by extended cognitive systems while instrument-developing cognition is most often underwritten by internal cognitive systems. According to the second strategy, both instrument-aided and instrument-developing cognition are, to the same extent, underwritten by extended cognitive systems.

Strategy 1: An attraction to the first strategy might be motivated by the thought, which we have already encountered, that instrument-developing cognition should be
given an internalist characterization since it typically occurs prior to any scientist
becoming functionally integrated with any scientific instrument. This strategy, however,
may be less attractive than it initially appears for proponents of The Extended Scientific
Mind. Problematically, a commitment to an internalist characterization of instrument-
developing cognition, when conjoined with the claim that instrument-developing and
instrument-aided cognition perform the same kinds of function, implies that extended
scientific minds are likely explanatorily superfluous. Why? Because these two claims
jointly imply that scientific instruments do not significantly “augment” or “modify”
existing internal cognitive structures. That is, on the current strategy for grasping The
Dilemma’s second horn, extant internal cognitive structures already enjoy the capacity to
implement the same kinds of function as the cognitive systems that underwrite
instrument-aided cognition. In which case, it is unclear in what sense scientific
instruments could augment or modify existing cognitive structures and thus unclear that
extended cognitive systems play any legitimate explanatory role.

Recall that, in general, a central promise of positing extended cognitive systems
was their putative capacity to illuminate the “mutually modulating” relations that hold
between individuals and elements in their extra-cranial environments. Thus, in canonical
examples of extended minds —Otto and his notebook, long division with pencil and
paper— extra-cranial elements putatively served to augment and modify internal
cognition, which was by itself supposed to be insufficient to accomplish the relevant
cognitive tasks. In the absence of augmenting and modifying relations between extra-
cranial elements and internal cognitive structures, internalist theories would appear to
have all the requisite resources to characterize the relevant cognitive achievements. And so it is with extended scientific minds in particular. If the internal cognitive structures of scientists are sufficiently powerful to underwrite instrument-developing cognition, and instrument-developing cognition and instrument-aided cognition implement similar functions, then internalist theories would appear to have every resource to characterize both kinds of cognition. Thus, a proponent of extended scientific minds who grasps The Dilemma’s second horn appears at best to be positing explanatorily superfluous cognitive systems. More seriously, however, such a proponent runs the risk of failing to adequately address the Input/Component Problem. After all, the presence of augmenting and modulating relations—as the radio example was intended to illustrate—was supposed to provide the resources for drawing an input/component distinction. In the absence of such relations, positing extended scientific minds runs the risk not only of explanatory superfluity, but also explanatory destruction. Thus, the first strategy for grasping The Dilemma’s second horn appears to render extended scientific minds entirely explanatorily unwarranted.

Strategy 2. A proponent of extended scientific minds might grasp The Dilemma’s second horn from a different angle, however. According to a second strategy for grasping The Dilemma’s second horn, not only is instrument-aided cognition often underwritten by extended scientific minds, but the very scientific cognition that underwrote the invention, design, and development of scientific instruments in the first place was also often—and to the same extent—accomplished by extended cognitive systems. That is, a proponent of The Extended Scientific Mind who adopts this second strategy for grasping
The Dilemma’s second horn maintains that scientists who have engaged in instrument-developing cognition have often been functionally integrated with extra-cranial elements in their environments, perhaps including tools, non-tool artifacts, or other environmental aids.

If our appraisal of the options for addressing The Dilemma has so far been sound, adopting an extended account of both instrument-aided and instrument-developing cognition is the remaining viable strategy for proponents of The Extended Scientific Mind. What is the best prediction for this approach to The Extended Scientific Mind? This remaining strategy again appears to be accompanied by significant explanatory costs. For one, it entails that extended cognitive systems proliferate to a significantly greater extent than has heretofore been advertized by proponents of The Extended Scientific Mind. Thus, The Extended Scientific Mind requires, at a minimum, adopting a cluster of speculative empirical hypothesis about the functionally integrated nature of scientific instrument design and development, none of which have been defended by proponents of extended scientific minds, including Giere or Moffatt. Further, our appraisal implies that relatively modest versions of The Extended Scientific Mind, according to which extended accounts may be applicable in only a small number of cases, are essentially unavailable. Instead, the fundamental similarity between instrument-aided and instrument-developing cognition implies that The Extended Scientific Mind has vast implications for understanding the inferential activities of scientists if it has any implications at all. Finally, given that this second strategy for grasping The Dilemma’s second horn requires endorsing a significant proliferation of the number and types of
extended cognitive systems, it is accompanied by a reduced potential for successfully addressing The Input/Component Problem, and thus for securing a theory of scientific inference’s central goals.

The potential explanatory costs enumerated above are substantial and may well by themselves warrant skepticism towards this understanding of extended scientific minds. However, the remaining strategy for grasping The Dilemma’s second horn has explanatory costs that are likely greater still given that a dilemma with respect to tools and artifacts might be constructed that is structurally similar to the scientific instrument-based dilemma we have been exploring. That is, suppose a proponent of The Extended Scientific Mind entertains the version of the view currently under consideration, according to which instrument-developing cognition is underwritten by extended cognitive systems whose parts include tools, non-tool artifacts, or other environmental aids. Call this kind of cognition “tool-aided cognition”.65 As with The Dilemma above, we might now ask a proponent of extended cognitive systems whether this tool-aided cognition is of the same kind as tool-developing cognition, or whether both types of cognition implement different kinds of function.

Without running through all of the options for addressing this new tool-based dilemma, it is reasonable to believe that an extended mind proponent will, as she was pressured to in response to the scientific instrument-based dilemma above, end up adopting the same strategy and endorsing an extended account for both tool-aided cognition.

---

65 “Tool-aided cognition” is here intended as shorthand to capture not only cognition assisted by tools proper, but also cognition assisted by non-tool artifacts, or any other extra-cranial environmental elements constructed by cognizers that serve as aids for completing cognitive tasks. “Tool-developing cognition” should be interpreted to enjoy a similar scope.
cognition and tool-developing cognition. That is, similar reasoning would suggest that a proponent of extended cognitive systems should claim that such systems underwrite—to the same extent—both tool-aided and tool-developing cognition. In order to avoid explanatorily weak generalizations, a proponent of extended cognitive systems should maintain that both types of cognition are of the same type. Similarly, in order to avoid explanatorily superfluous posits, a proponent of extended cognitive systems should maintain that any internal cognitive structures are insufficient by themselves to accomplish the cognitive tasks associated with both types of cognition. Accomplishing these tasks, such a proponent should maintain, requires a cyclical, augmenting and modifying role to be played by extra-cranial environmental elements. Only by adopting the assumption that tools, non-tool artifacts, or other environmental aids occupy such roles, is there a prospect for drawing a viable input/component distinction. And only on this assumption is there sufficient warrant for positing extended cognitive systems rather than relying on the explanatory resources of internalist alternatives.

The explanatory insufficiency of purely internal cognitive structures to underwrite tool-aided and tool-developing cognition in the sense gestured at above is in fact an apparent commitment of views that many proponents of The Extended Mind have already endorsed concerning the history and development of sophisticated human intelligence. According to such views, tools, non-tool artifacts and other environmental aids have played an indispensible historical role in augmenting and modifying the development of cognition in our species. Tool-aided and tool-developing cognition, on this picture, are
both part of a long, symbiotic chain of feedback loops between cognitive agents on the one hand, and tools, non-tool artifacts and other aids in their environment on the other. Such expansive proposals are intimated by, for example, Clark & Chalmers in their early and influential characterization of extended cognition:

[T]he biological brain has … evolved and matured in … the presence of a manipulable external environment … Once we recognize the crucial role of the environment in constraining the evolution and development of cognition, we see that extended cognition is a core cognitive process, not an add-on extra (1998, p. 31-32).

More explicitly, in “Making Tools for Thinking,” Dennett (2000) proposes a jointly interactive role for the development of tools and intelligent cognition, according to which both mutually scaffold and augment each other, and which might lend support to an extended account of both tool-developing and tool-aided cognition. Dennett claims that tools and other artifacts may have been indelibly intertwined in the development of human intelligence in part by allowing for quicker and easier solutions to what would otherwise be resource-intensive and costly cognitive processing. Clark refers to views of this type as “Tools all the way down!” (Clark 2002, p.33).

3.3 Tools All the Way Down

Before assessing the plausibility of extended, “Tools all the way down!” views, notice that such expansive accounts are entirely inconsistent with modest versions of The Extended Scientific Mind (and, perhaps, The Extended Mind in general – see section 3.4). Instead, our reflections imply that proponents of extended scientific minds, in facing
a dilemma over how to account for the cognition that designed and developed scientific instruments on which scientists rely, have been pressured to embrace speculative and bold empirical hypotheses not just about the nature of the development of scientific instruments, but about the symbiotic development of extra-cranial elements and intelligence itself. It follows that The Extended Scientific Mind is not appropriately viewed as a modest prediction about one or more cases of possibly extended cognitive systems which underwrite scientific theory formation and revision; rather, the view is an ambitious and radical competitor to internalist theories of scientific inference and perhaps to internalist theories of cognition generally.

These bold and speculative empirical hypotheses, in addition to not being defended by proponents of The Extended Scientific Mind such as Giere or Moffatt, are again associated with potential explanatory costs in the form of a reduced capacity to address The Input/Component Problem. The sweeping historical coupling with the environment that is under consideration here would undermine the prospects for successfully drawing an input/component distinction, and thereby undermine any prospects for successfully characterizing the flexible and reliable cognitive processes that routinely cope with evidential under-determination in a variety of areas of scientific inquiry. Certainly, we cannot expect that an input/component distinction could simultaneously apply to a broad range of historical cases —including initial, intermediate, and mature stages in the development of both tools and human intelligence— and be successfully drawn by comparison to a few unproblematic exemplar cases of already highly functioning information processors such as a radio.
Instead, we are potentially losing any grasp whatsoever on the circumstances under which the environment serves to provide informational inputs to cognitive information processors, as opposed to functioning as components of those information processors.

These potential explanatory costs, even if substantial, might nevertheless be outweighed by the promise of contributing to a well-supported view on the symbiotic development of human intelligence on the one hand, and tools, non-tool artifacts, or other environmental aids on the other. I will argue, however, that the balance of evidence does not clearly favor such a view. It follows that the best prediction regarding extended (scientific) minds is that they are unlikely to be components of fruitful psychological accounts of (scientific) reasoning.

A central potential weakness with an extended “Tools all the Way Down!” account of cognitive development is its commitment to the inadequacy of internal cognitive resources to underwrite both tool-aided and tool-developing cognition. If this commitment is rejected, then, as I have already remarked, the explanatory role for extended cognitive systems is entirely unclear. However, a serious challenge to such a commitment might be made on the grounds that the design, development, and use of extra-cranial material resources presupposes sophisticated and high load internal cognitive contributions. In other words, one might reject the idea that coordination with environmental elements such as tools, non-tool artifacts and other environmental aids is a way of making dumber brains smarter.

Clark anticipates a similar worry —that he calls the “paradox of active stupidity”—
and ties the worry into extended accounts of tool-developing cognition:

[Perhaps]… making the moves that sculpt the environment so as to allow cheap problem solving itself requires expensive, advanced, design-oriented cogitation. The nasty upshot being that only clever brains could make their worlds smart so that they could be dumb in peace — a result which would deprive the tool-based scenario of its appealing role in explaining the origins of advanced, reflective thought and reason (Clark 2002, p.35).

Clark’s worry is that there is reason to reject a commitment to the inadequacy of internal cognitive resources for underwriting tool-developing cognition. In which case, an extended mind proponent’s prospects for addressing the tool-based dilemma would appear to be dim. Dennett, however, has floated a kind solution to Clark’s “paradox of active stupidity”. He claims that we may be able to trace the development of tools itself back to initial processes that don’t require expensive, higher-level intelligence, such as processes of “trial and error” or simply identifying suitable objects in the environment that are already tailored to aid with the completion of cognitive tasks (2000, p.59). Dennett claims to find some evidence for this solution in the famous studies by Wolfgang Kühler of non-human primates and their use of tools and other environmental props to secure their goals:
They [apes] had to have many hours of exposure to the relevant props — the boxes and sticks, for instance — and they engaged in much manipulation of these items. Those apes that discovered the solutions — some never did — accomplished it with the aid of many hours of trial and error manipulating (1996, p.121).

Jeffares (2010) advances a similar narrative concerning the origin and development of human intelligence whereby elementary cognition became intertwined with early tools and then began a cyclic (and therefore extended) process of mutual development. Jeffares claims, on the basis of anthropological findings, that very simple tools, such as stones with a few marks, were made use of along the hominid line and then were coupled with our hominid ancestors during the development of more sophisticated cognition:

Tool use, the ubiquitous need for tools, and tool carriage scaffolds cognition in ways that actually extend hominin cognitive capacities. At this point [early Homo], the extension is an increased ability to engage in tasks with a strategic payoff, as the tools act as a buffer to working memory and a cognitive aide memoir. This in turn may well feedback into a selection for increased working memory itself (Jeffares 2010, p.514).

If something like Dennett and Jeffares’s narrations are apt, perhaps we might believe that tool-aided and tool-developing cognition are both components of a longer historical chain of cyclic and functionally integrated interaction between cognitive agents and material resources in their environments. And perhaps we might believe that the explanatory
benefits of this account are likely to outweigh the potential explanatory costs associated with a speculative proliferation of extended cognitive systems.

Some evidence instead suggests, however, that antecedent and apparently internal intelligence predates the successful application of tools or non-tool artifacts. For example, studies of non-human primates, including gorillas and east African chimpanzees (Pan troglodytes schweinfurthii), suggest that goal-directed behavior without tools exhibits more sophistication than goal-directed behavior with tools. In particular, in manual food preparation by both species there is apparent planning involving multiple and highly distinct intermediate states (Corp & Byrne 2002). In contrast, this level of sophistication is not typically exhibited by non-human primate tool use, since, among other limiting factors, non-human primates in general do not use more than one tool to secure a goal (Santos et al. 2005). The study of tool use by our ancestors and contemporary non-human primates is an open and active area of research. But on the assumption that this research might shed light on the initial stages of human cognitive development, these studies suggest that sophisticated cognition in our species may well have developed prior to, and perhaps be required for, the initial development of tool use and non-tool artifacts.

There are additional considerations that shed doubt on any narration about the development of intelligence and extra-cranial environmental elements that includes long,

\[66\] Corp & Byrne also claim that these multiple intermediate states are “heirarchically organized,” though it is difficult to attach a concrete meaning to this description.\[67\] Santos et al. (2005) in fact discuss an exceptional case to this rule, namely, the activities of tamarins after extensive human training for using multiple tools to complete tasks that require two intermediate, “mean” states prior to securing an end or goal state.

172
historical chains of augmenting and modifying feedback loops between, on the one hand, tools, non-tool artifacts, or other environmental aids and, on the other, human brains that these extra-cranial elements putatively. Sterelny, for example, has plausibly argued that the use of tools in central cases requires sophisticated and highly intelligent internal cognitive contributions. He argues that this is particularly the case for tools and non-tool artifacts that are shared:

Jointly used epistemic artefacts are often less than optimal for any of their users: they need to be individualised at each use. Moreover, though human interactions are often co-operative, they are not exclusively so. The possibility of deception and the hidden agendas of others cannot be ignored. … Agents using common tools cannot afford to be dumb (2006, p.66).

Sterelny’s observations suggest that there may well be important gaps in any historical sequence of putatively extended feedback loops between cognitive agents and material resources in their environment. If correct, his claims undermine extended accounts of both tool-aided cognition and instrument-aided cognition when tools and scientific instruments are shared. With respect to scientific instruments there is reason to believe that such circumstances often obtain. That is, we may reasonably believe that scientific instruments are often shared by multiple researchers who are working on a joint project or by researchers whose separate projects rely on specialized and sophisticated equipment
in demand. Thus, we may wonder if sophisticated internal cognitive resources are in fact adequate and required for the task of managing shared scientific instrument use; and if so, there may be reason to believe that internal cognitive resources are also adequate in the general case to account for cognition that is assisted by tools and scientific instruments.

My present aim is not to conclusively establish Sterelny’s claims about the shared use of tools or scientific instruments. Nor is it my aim to conclusively establish that internal cognitive structures must have had the capacity to underwrite the development of tools and other artifacts by our species. Rather, the present point is that adopting a response to The Dilemma that incorporates a Dennettian “Tools all the way down!” view requires accepting a number of highly conjectural hypotheses about the origin and development of tool use that the balance of evidence does not appear to favor. In the absence of compelling evidence in the favor of these speculative hypotheses, there is no reason whatsoever to believe that such responses to both The Dilemma and its tool-based counterpart overcome the significant potential explanatory costs that accompany them. Instead, there is strong reason to doubt the prediction that grasping The Dilemma’s second horn by proliferating extended systems down a long historical chain of tool and human interaction will lead to plausible externalist hypotheses regarding the psychology of (scientific) reasoning.

An overall assessment of The Extended Scientific Mind in light of The Dilemma yields similarly pessimistic results. On either of The Dilemma’s horns, extended
scientific minds appear explanatorily unwarranted—either explanatorily superfluous or explanatorily destructive. Extended scientific minds are associated with significant risks of being unable to draw a viable distinction between inputs and components. Further, there is reason to believe that positing extended cognitive systems can lead to a mischaracterization of the cognition responsible for the invention, design, and development of scientific instruments and other environmental aids on which scientists rely. In short, there is more than substantial reason to doubt the prediction that extended scientific minds will be a productive component of a future cognitive science of scientific inference. Reflection on The Dilemma has in fact exposed The Extended Scientific Mind as a potentially radical competitor to alternative internalist accounts of scientific inference. In order to avoid being explanatorily superfluous, extended cognitive systems were required to incorporate environmental elements that augmented and modified existing internal cognitive structures, which by themselves were allegedly explanatorily insufficient to account for the cognitive achievements of individual scientists. In this sense, a central culprit in the likely failure of The Extended Scientific Mind as an explanatory hypothesis is its tight connection to strong forms of Competitive Externalism. In which case, we have apparent reason to reject Competitive Externalism along with The Extended Scientific Mind.

3.4 Scientific and Ordinary Minds

Does our discussion of The Extended Scientific Mind have implications for extended accounts of higher cognition generally? On the assumption that non-scientific, ordinary
cognition—for example, everyday conscious reasoning and belief revision—is fundamentally similar to scientific inference, then our previous discussion implies that The Extended Mind is likely explanatorily unwarranted in contexts of “higher” cognition in general.

One might attempt to block this inference, however, by denying that scientific cognition and ordinary cognition are, in the relevant respects, similar. According to this strategy, the models that are likely to enlighten scientific reasoning are not likely to be relevant to modeling everyday cognition. In which case, it would not follow that because models of extended cognitive systems fail to accommodate scientific inference that we should also expect models of extended cognitive systems to fail to accommodate ordinary cognition.

Such a strategy might appeal to general doubts that have been raised concerning the similarity between ordinary cognition and scientific hypothesis formation and confirmation. Pinker, for example, writes:

Consider that a given scientific inference is accomplished by a community of scientists, who work over long periods of time, use sophisticated mathematical and technological tools and pool their results in journals and conferences. … In contrast, a common-sense inference is typically accomplished by a single brain working over a time-span of a few seconds (2005, p.11).

Similarly, Carruthers claims:
What many different thinkers working collectively over the course of a lifetime can achieve – conducting painstaking searches of different aspects of the data and bringing to bear their different theories, heuristics and reasoning strategies – is a poor model for what individuals can achieve on their own in the space of a few seconds or minutes (2003, p.77).

While such considerations might lend general support for claiming that scientific inference and quotidian cognition are importantly dissimilar, they appear to provide cold comfort to a proponent of The Extended Mind who claims that extended cognitive systems underwrite ordinary, but not scientific, cognition. For if anything, Pinker and Carrthers’s reflections appear to suggest the reverse, namely, that scientific cognition is more likely to be accommodated by models with extended cognitive systems than ordinary, non-scientific cognition. There are separate grounds, however, on which a proponent of extended cognitive systems might dispute a fundamental similarity between scientific inference and everyday cognition that are not self-defeating in this way (e.g., Knobe 2010 outlines alleged differences between scientific and ordinary cognition with respect to the role that background moral considerations influence them). I have defended this general similarity elsewhere, however, (Fuller 2011; Fuller & Samuels, in preparation) and will not do so here. I merely note that on the assumption that scientific inference and everyday cognition are fundamentally similar, then our previous reflections constitute an important objection not just to The Extended Scientific Mind, but to The Extended Mind in general.

4.0 Population-level Theory Change
Not all versions of Scientific Externalism are plagued by the same difficulties that plague The Extended Scientific Mind. Population-level theories of scientific inference are not invariably wedded with ambitious versions of Competitive Externalism and thus not undermined by The Dilemma. For these reasons, the most promising and fruitful version of Scientific Externalism is a population-level account of scientific theory change. Or so I will argue in this section.

In fact, some population-level theories of scientific inference have been associated with Competitive Externalism. For example, some Marxist theorists from the 1930s held that the sociology of science was primary over individualistic accounts of scientific theory change and acceptance. These theorists hoped to replace the individual-level with the social (see Thagard 1994 for brief discussion). In contrast to The Extended Scientific Mind, however, population-level accounts are not invariably wedded with an ambitious form of Competitive Externalism. This point can be readily seen from comparisons to population-level accounts in other areas of inquiry, including the study of eusocial insects, bird flocking, traffic jams, population-level genetics, and a range of phenomena characterized by complexity theory. In these domains, properties of population-level dynamics are not invariably or even typically instantiated by individual members of these populations (See Fuller & Samuels for more extensive discussion). In turn, this suggests that population-level theories of scientific inference may well enjoy relations of autonomy, as opposed to competition, with internalist theories of scientific inference, where such autonomous relations to internalist theories were essentially
unavailable to The Extended Scientific Mind in part because it purported to characterize the cognitive achievements of individual scientists.

A comparison to population-level theories in other domains suggests that there is no more reason for population-level theories of scientific inference to compete with individual-level theories than there is for any theory at one “level” to compete with a theory at another. In this way, the co-existence of unreduced scientific disciplines speaks to the possibility of an autonomous, cohabitation between population-level and internalist theories of scientific inference. Further, there is reason to believe that cognitive features of population-level dynamics need not modulate or augment internalist cognitive structures in order to be explanatorily warranted. For The Extended Scientific Mind this was a central feature of its competitive nature and its struggle to avoid explanatory vacuity. In contrast, there is no reason to expect that properties of population-level dynamics causally interact in any sense with individual, internalist cognitive structures. *Pace* causal competition worries that have arisen in some discussions over non-reductive physicalism, there is every reason to believe that this type of causal competition is the exception rather than the rule.

With respect to The Dilemma, population-level theories of scientific inference are not committed to extended cognitive systems that incorporate scientific instruments and other inanimate environmental elements on which scientists rely. A proponent of population-level accounts of theory change may simply hold that interactions between a scientist and her environment are likely to be addressed at the individual-level. Thus, a population-level theory, unlike The Extended Scientific Mind, is not cast into doubt
because of its incapacity to accommodate the cognition involved in the invention, design, and development of scientific instruments and other environmental aids on which scientists rely. For similar reasons, population-level accounts do not face a substantial challenge of drawing an input/component distinction when characterizing cognitive systems. That is, they do not face the problem associated with proliferating cognitive systems such that the environment no longer provides informational inputs to information processors. Population-level cognitive systems, since they are composed solely of scientists and not other kinds of environmental elements may invoke the biological boundaries of scientists when drawing an input/component distinction.

The above differences between population-level theories of scientific inference and The Extended Scientific Mind are powerful reasons to sharply distinguish these versions of Scientific Externalism. Naturally, however, the mere fact that population-level accounts of scientific theory change avoid many of the explanatory costs associated with extended scientific minds does not by itself establish that there are proprietary cognitive properties instantiated by the dynamics of scientific communities. I do not have the space here to mount a significant defense of population-level theories of scientific inference and will not attempt to do so. Instead, I will sketch an example, owing to Kitcher, that speaks to the potential desirability of this kind of externalist framework, and to the kinds of relations that likely hold between population-level and internalist theories.

Kitcher (1990) and (1993) describes apparently proprietary cognitive features of the population-level dynamics of scientific theory formation and revision. For example, he claims that some processes of scientific hypothesis formation and confirmation are
rational and desirable allocations of cognitive resources only when considered at the population-level:

Only if we situate the individual in a society of other epistemic agents … does it begin to appear rational for someone to assign herself to the working out of ideas that she (and her colleagues) view as epistemically inferior (Kitcher 1990, p.8).

Kitcher reasonably maintains that this population-level perspective explains the rational pursuit and development of continental drift theory in the 1920s and 1930s, despite requiring what at the time were reasonably regarded as improbably large geophysical forces. Considered at the population-level, but not the individual-level, it was a desirable and rational division of cognitive labor for a minority of geologists to continue developing the theory of continental drift even when a majority of scientists had good reason to doubt its accuracy. This is a plausible example of cognitive properties that are instantiated solely by the dynamics of populations of scientists.

Note that Kitcher’s proposal assumes the viability of individual-level theories of scientific inference that are entirely compatible with an internalist framework. That is, his proposal first seeks to characterize individual variation among a population of scientists and then to assess the overall rational cognitive resource allocation amongst that population. In this sense, such a theory strongly appears to complement and supplement internalist theories rather than compete with them. On the assumption that population-level accounts of scientific inference of the kind that I have briefly sketched are representative of population-level accounts generally, there is reason to believe that,
unlike The Extended Scientific Mind, population-level accounts are likely to cohere with Autonomous Externalism as opposed to Competitive Externalism.

Further, if we accept that scientific inference and ordinary, non-scientific cognition are fundamentally similar, then this moral concerning population-level theories of scientific inference may have implications for the study of quotidian reasoning and belief revision. In particular, we should expect theorizing on “higher” cognition in general to divorce population-level theories —which have sometimes gone under the name of “social” or “collective” cognition— from The Extended Mind and assess both kinds of Externalism separately. If I am right, the former externalist framework is more plausible than the latter and thus there is reason to regard population-level accounts, not The Extended Mind, as viable candidates for future research. I will leave further speculation on Externalism as an account of ordinary cognition to future discussion, however.

5.0 Conclusion

We have examined a variety of frameworks for characterizing scientific inference. My review strongly suggests that research on externalist theories of scientific inference should be directed towards population-level as opposed to extended mind accounts. In addition, I have argued that we should not think of internalist and externalist theories of scientific inference as competitors, but rather as autonomous complements. On the assumption that scientific inference and quotidian cognition are fundamentally alike, a similar moral is also likely to hold for externalist accounts of ordinary cognition.
References


Collins, J. (2002). On the Very Idea of a Science Forming Faculty Dialectica, 56, 125-51


