

Alison Gopnik's Half-Truth: a Proposal for a More Nuanced Account of Two-Systems Reasoning in Particular Theoretical Domains

Submitted by

William Michael Kallfelz

Revised January 2, 2006

ABSTRACT

Total word count of body of paper (not including abstract and footnotes) : 7269

I sketch my proposal in the Introduction (I). I begin by critically reviewing Gopnik and Melzoff (1997) in (II.i) and Gopnik and Glymour (2002) in (II.ii), vis-à-vis aspects of Carruthers (2005). My central, negative claim in (II) is that Gopnik, et. al overlook two-systems' reasoning in their continuity claims rendering conceptual change in the sciences fundamentally analogous to cognitive processes taking place in early childhood learning. The latter chiefly involves System-1 dynamics, whereas the former involves interaction effects between System-2 and System-1 reasoning. My positive claims in (III) and (IV) are based on the argument that *any* account of scientific activity must be characterized in terms of the *dynamical interaction* between System-2 and System-1 reasoning. I advance a quantitatively-based cognitive historical program for investigating System-2/System-1 dynamics in (IV). I show how any given theoretical domain in a scientific research tradition can be partitioned, whether studied diachronically or synchronically, in terms of this two-systems' interaction dynamic. Specifically, the partitioning is in terms of equivalence classes indexed according to the System-2 dynamical weighing parameter α , which I introduce and define. The associated equivalence classes can be thought of as natural kinds, in a plural-aspects metaphysical scheme. I also discuss some of what I consider are the potentially unique methodological advantages of the program I suggest.

I. Introduction

Based on the present state of cognitive science and its fruitful relationship with contemporary philosophy of mind, naturalistic accounts have recently been offered, which attempt to answer questions concerning what, in our cognitive and social make-up,

makes science possible?¹ In their studies of learning in infancy and early childhood, Alison Gopnik, et. al (1997, 2002-5), for instance, argue that their hypothesized models of childhood learning constitute essential features of the cognitive patterns of activity of scientific theorizing shared by their adult counterparts. I argue, on the other hand, that Gopnik's claims best resemble half-truths. That is to say, Gopnik et. al have primarily focused on what may be essentially characterized by System-1 processes. I, on the other hand, argue that scientific theorizing essentially involves *interaction effects* between System-2 and System-1 processes.²

The above critique of Gopnik comprises my negative claim, i.e. that any account of the cognitive processes underlying scientific activity cannot overlook System-1/System-2 interaction effects. I devote the first section below to briefly critically reviewing Gopnik, in the light of my negative claim.

In the following sections, what I positively suggest deals with some of what I consider as consequential, when taking System-1/System-2 interactions seriously. System-1/System-2 interactions synchronically and diachronically partition the space of scientific theorizing activity, even in one particular theoretical domain, into equivalence classes: The associated representative elements reflect the *degree of influence* System-2 has over System-1 in this dynamical interaction. A difference in such a degree of influence translates into a difference *in kind* of cognitive activity, even in one particular theoretical domain in any field of science.

In other words, my thesis stems from a plural-aspects metaphysical conception of the domain of cognitive activity in science. Each natural kind, indexed accordingly by the degree of influence System-2 has over System-1, represents a unique aspect of the phenomenon of scientific activity. Abstracting away these aspects leaves one with an impoverished view of the phenomenon of science, misleading in its homogeneity.

¹ Carruthers, et. al. (2002), Carruthers (2005), ch. 6. Specific accounts are also offered in Petitot, et. al. (1999), as well as in Carrier, Massey, and Ruetsche (2000). Others, however, like Fodor (1983, 2000) give negative arguments: despite progress in the fields of cognitive science and philosophy of mind, Fodor argues that we still lack the conceptual resources necessary to account for phenomena like abductive reasoning.

² That an account of such interaction effects should form a crucial part of *any* cognitive and developmental psychological study of scientific activity, is perhaps best suggested by Peter Carruthers' characterization of science as a *whole-person activity*. (2005, ch. 6)

II. Critical review of Alison Gopnik's program

II.i) Introductory overview

I begin here by briefly summarizing salient aspects of Alison Gopnik's research program from 1997 to the present, which essentially bifurcates into her earlier "theory-theory" and her subsequent Causal-Bayesian net accounts. These two versions of Gopnik's research program, which *prima facie* seem to differ in kind, are nevertheless thematically interrelated. In Gopnik & Meltzoff (1997) a theor[izing]-theory account is presented in systematic detail,³ while in Gopnik & Glymour (2002), and in later writings, primary methodological emphasis is directed towards Causal-Bayesian nets. Such nets act as a possible means for characterizing how adults and children "recover causal facts by making implicit assumptions about the causal structure of the environment." (Gopnik & Glymour (2002), 119).

At the outset, Gopnik's separate accounts may strike the reader as having fundamentally separate research agendas. For instance, Gopnik and Meltzoff (1997) present their claims in a top-down fashion,⁴ in which a framework for a "robust theory" or research program shares a central and organizing motivation, best described by the following bold statement:

Processes of cognitive development in children are...*perhaps even identical with* the processes of cognitive development in scientists. [Scientific theory change is, after all]...one of the clearest examples we know of the derivation of genuinely new abstract and complex representations of the world from experience. (Gopnik & Meltzoff (1997), 3, italics added)

What makes the above notion a bold maneuver does *not* stem from its thesis of continuity. (The latter thesis is best understood as the claim that cognitive processes and

³ Gopnik & Meltzoff actually refer to their work in terms of a 'theory-theory' account of childhood conceptual development. However, I side with Carruthers' claim that Gopnik & Meltzoff's study should be more accurately understood as a '*theorizing*-theory,' since, "[t]he term 'theory-theory' is best reserved for...synchronic accounts of our abilities such as folk-physics and mind-reading, rather than for Gopnik and Meltzoff's diachronic version, according to which the theories in question are *arrived at by a process of theorizing* (as opposed, for example, to emerging through the maturation of a module.)" (Carruthers (2002a), n. 2, 76). Unless otherwise explicitly noted, I will refer to Gopnik & Meltzoff's account according to Carruthers' modification.

⁴ I am speaking informally here: the account offers, at the outset, some quasi-axiomatic general claims sharing both normative and descriptive content. "We will deliberately blur the distinction between the 'logic of discovery' and the 'logic of confirmation.'" (Gopnik & Meltzoff, (1997) 33)

reasoning necessary for science form *part* of the innate human endowment, such that they require only supplementation of changes in belief and desire, as well as changes in external socioeconomic resources for science to become *possible*.⁵ (Carruthers (2002a), 74-75)) Stated weakly, what distinguishes the theorizing-theory is the attempt to present a ‘middle way’ between domain-specific *modular* accounts and fully domain-general *connectionist* attempts to provide models of theorizing and learning.⁶ The strong point, on the other hand, rests on the supposition that such a middle-way gives the best account for the *developmental*, or functional features of theorizing.

Gopnik & Glymour (2002), in a “bottom-up” manner, hypothesize that Bayes nets *may* give the best account for modeling representations of the causal characteristics of local environments, which adults and children construct. The authors argue, by way of analogy, that humans construct such representations, constitutive of their theory-formation system, in a manner akin to properties of *visual maps* in computational theories of vision.⁷

[W]e want to propose that there is a separate cognitive system with other procedures devoted to uncovering causal structure, and that this system has some of the same abstract structure as the system of spatial mapmaking with which it must in many cases interact...Just as cognitive maps may be differentiated from other kinds of *spatial cognition*, causal maps may be differentiated from other kinds of *causal cognition*[.] [g]iven the adaptive importance of causal knowledge[.] (121-122, italics added)

⁵ Contrast this with the *non-continuity claim*, which maintains that our innate cognitive endowment is *insufficient* for scientific reasoning. The necessary cognitive materials have to be *socially constructed* and *learned* before scientific activity can commence. (ibid, 74-75) Such a distinction, of course, raises the rather non-trivial semantic chore of providing an adequate notion of “innateness.” Peter Carruthers (2002a), for example, suggests that a cognitive feature is describable as innate if (a) it’s not learned. However, such a notion “would...require us to know what learning is. I wish I did.” (n.1, 74). Alternatively one could state that a cognitive notion is describable as innate if (b) it follows channeled development, i.e., if it follows the same developmental trajectories and outcomes over and across a variety of different circumstances.

⁶ “Modularity theories [admittedly] may invoke auxiliary hypotheses about changes in information-processing ability to explain development...development is not the result of internal conceptual changes as it is in the theor[izing]-theory, but rather is the result of external non-conceptual changes in information processing.” ((1997),55) On the other hand “[c]onnectionist systems involve a (somewhat) more neurologically realistic computation than classical computation systems...From this perspective [of the theorizing-theory] connectionist systems are simply an alternative way of implementing particular models of the mind, [insofar as] connectionist systems are designed to detect complex statistical patterns of regularity in the relations between inputs.” (ibid., 60-61)

⁷ Gopnik & Glymour refer here to Marr’s (1982) ground-breaking work (119).

Extensive research abounds in the philosophy of science and in the decision sciences concerning how Bayes nets can employ computationally tractable algorithms to produce “reasonably accurate causal inferences.” (124) The authors draw from this body of research to base their claim on the notion that such nets and their associated algorithms may well depict the manner in which humans learn new causal relations by way of prediction and intervention.

As the theorizing-theory gives a ‘middle path’ between modular domain-specificity and maximal connectionist domain-generalality, so Bayes nets likewise hypothetically steer a middle course, possibly giving an optimal account for how humans may infer causal relations. (130)

[Modular] proposals do not seem to capture the wide range of explanations that develop in everyday life nor the way that cognitive development move[s] back and forth away from these domains [of folk biology, folk psychology, folk physics]. The idea of causal maps seems to capture both what is general and what is specific about everyday theories. (124)

In other words, the authors advance a specific claim concerning the *developmental* aspects of humans’ theorizing ability, in terms of how new causal relations may be learned.

A pluralist conception of reasoning underlies both accounts. In Gopnik & Melzoff (1997), for instance, active reasoning employs modular outputs, theory-formation, and empirical generalizations (in the form of scripts and narratives). In Gopnik & Glymour (2002), the causal representations that humans construct “[are] a *separate cognitive system* [working] with *other procedures* devoted to uncovering causal structure.” (121, italics added) I examine the finer points of Gopnik’s two accounts below. The primary object of my critical examination has to do with my claim that Gopnik excludes from her study System-2/System-1 interaction effects, focusing on System-1 processes alone.⁸

II.ii) A closer examination of Gopnik and Meltzoff

⁸ In addition, I also briefly discuss what I consider are other weaknesses and tensions, contrasting them with what in my opinion are more cogently dealt with in Carruthers (2002, 2005). The reader well familiar with Gopnik may skip to section III below.

Gopnik & Melzoff (1997) present the fourfold purposes of the study concerning their theorizing-theory (4-7):

- I.) To articulate and show how it can generate predictions not made by the other (modular and connectionist) theories.
- II.) Apply it to explain what is known about early infancy and childhood.
- III.) Cast the notions of cognitive development in terms of *theory – change*.
- IV.) Show that cognitive development *depends* on language development.⁹

The authors aim to show that the above four claims entail that young children learning to talk *experience*, and are *engaged* in, profound restructurings of their domains of knowledge, best understood in terms of scientific theory-change.(7) Moreover, they argue that cognitive scientists should devote more attention to the *dynamical processes* through which such rules and representations of the knowledge-domains are transformed from childhood to adulthood. In this respect, one could think of scientists as “big children,” but children should *not* be thought of as “little scientists.”¹⁰ “Scientific progress is possible because scientists employ cognitive processes [of development] that are first seen in very young children.”(32)

The concern for the lack of an adequate treatment of developmental and functional aspects of theories in other accounts (i.e., modular and connectionist) forms a central motivation of Gopnik & Melzoff’s project. Modularists, for instance, may give a reasonable account of *structural* features of theories (their abstractness, coherence, causality, ontological commitments¹¹) but fall short in the case of functional features such as prediction, interpretation, and explanation.(50) Furthermore:

[Though] both [theorizing-theory and modular] theories involve abstract entities and rules related to sensory input in only very indirect ways... *that there is some logic in the relation between input and representations is itself not enough to distinguish modular and theoretical structures.* (52, italics added)

Hence, it is the *experience* of dynamical theory-change, which distinguishes the theorizing-theory from modular accounts.(53) To further illustrate this point, a ‘parallel

⁹ “[C]hildren internalize the linguistic structures in organizing or explaining experience...adult language is another source of information, admittedly a particularly important one, that the child uses in theory construction...[w]hile the child is inferring the structure of the outside world, she is also untangling the structure of language itself.” (6)

¹⁰ In their critical review of Gopnik, Faucher, et. al’s (2002) title: “The baby in the lab-coat...” proves a little bit misleading here.

¹¹ Structural and functional features are discussed in Gopnik & Melzoff (1997) 34-48.

universe' *Gedankenargument* is advanced, in which a child is suddenly thrust into some universe U' whose laws and features are far different from those of our own universe (U). According to Gopnik & Meltzoff's account, the child will *learn* and hence *revise* her folk physics/biology/psychology, whereas the 'Kantian' modularist would argue that she would be at a loss to revise her innately (U -based) folk theories:

Modularity accounts have been proposed in a variety of cases that appear to involve...strong constraints on the possible thoughts we can think and the beliefs we can hold...Spelke et. al. describe their accounts as 'neo-Kantian' [for this reason.] *This seems like quite an accurate term for this trend in cognitive science in general.* Like Kant, these authors propose that certain conceptual structures are innately given and cannot be overturned by evidence. (58, italics added)

However, and admittedly from hindsight, the modularist could reply that the fallacies of question-begging and dichotomization are at work here. Dichotomization occurs because Gopnik equivocates 'modularity' with an especially strong and *innate* version (à la Fodor, Spelke). In the case of Fodor, for instance, the authors call attention to his distinction between 'peripheral' (modular) versus 'central' (amodular) processing systems:

Fodor advocates [peripheral processing] modules but contrasts them with... 'central process' knowledge, which includes scientific knowledge. In a way, our view is quite Fodorian. We also think there is a distinction...between peripheral modules and central processes, and that central processes include both ordinary everyday concepts *and* scientific concepts. (57, italics added)

One need only contrast Fodor's strong modular thesis with Carruthers' (2005) weaker one, to see that one need not assent to some 'neo-Kantian' (Spelke) or quasi-Platonist (Fodorian) rendition of modularity. Carruthers' (2005) massively modular thesis, in fact, offers a positive account for the flexible dynamics of conceptual change, both in the cases of scientific and practical reasoning.¹² But because Gopnik & Meltzoff *do* characterize modularity in this way, means that they cannot help but beg the question for their (theorizing) theory. This is evidenced, for instance, in the following remark:

Theory formation will turn out to involve some set of particular causal principles that get us from patterns of input to patterns of representation...These causal principles *must*, however, be *deeply and radically different* from the...principles that have been proposed for modular systems. (56, italics added)

¹² Chapters 6, 7. This of course also provides an alternative to Fodor's notion of a central processing system.

Carruthers' (2005) account, however, covers notions like theory-formation.¹³

Additionally, between the cognitively impenetrable extreme consisting of perceptual acts like judging an object's texture and distance, versus the other cognitively penetrable theoretical extreme of inferring an object's identity, Gopnik and Melzoff speculate in some 'intermediate' cognitive area. (65) The authors argue that any modular theory attempting to account for the contents and processes of such an intermediate zone could be easily reduced to their theorizing-theory.¹⁴ Again, however, Carruthers' (2005) weakly massive modular account does more than an adequate job of filling this intermediate region.¹⁵

Based on their aforementioned restricted notion of modularity, Gopnik & Melzoff's claims become problematic when they seek to extend their theorizing-theory into the cases of *empirical generalizations*.¹⁶ For instance, they claim that "a statement phrased in theoretical terms doesn't mean that it is a theoretical statement." (67) They invoke as an example discoveries made by the Hubble space telescope of objects in the universe older than what the Big Bang Theory would predict. The age of the objects is determined via an essential use of this cosmological theory, and yet, they claim, such observation-statements aren't theoretical, since they violate the very principles of the cosmological theory. But this claim is misleading, for it equivocates a theory's

¹³ In particular, in his discussion on the phenomenon of creativity (ch 5) and its relationship to abductive inference (ch 6).

¹⁴ "[T]he idea is that at some point the internal structure of the module, its internal representation and rules become subject to the same kinds of revision and restructuring as more theoretical kinds of representations and rules. In this way a module could be rewritten as an innate theory...we open up the module and look at what's inside, and turn it into a theory." (66)

¹⁵ "I present and defend...a notion of modularity much closer to the use of the term 'module' in biology and...even closer to the notion used by researchers in AI...[that of] a functionally distinct processing system of the mind, whose operations are at least partially independent of those of others, and whose existence and properties are partly dissociable from the others. Moreover, modular systems must be *frugal* in their use of information...hav[ing] internal operations that are widely *inaccessible* to other systems. [I] claim that the mind is composed of *many* functionally isolable processing systems, hav[ing] multiple input/output connections with others." (viii-ix) As the passage suggests, and later articulated in detail, both the *structural* and the *functional-developmental* features of such 'mid-level' cognitive activity, ranging from the bottom extreme of (quasi cognitively inaccessible) mate preference and cheater-detection to the uppermost (quasi reflective, deliberate cognitively transparent) level of folk psychology/biology/physics, and human practical reasoning are given nuanced and detailed accounts.

¹⁶ I.e. classes of statements lacking *all* the structural and functional features of theories, i.e. statements whose features are a proper subset of the following: abstractness, coherence, causality, ontological commitments, as well as predictive, interpretative, and explanatory power.

methodology with its ontology. Scientific theories get revised and refined all the time, by claims based on more accurate observations *within* the theory's paradigm.¹⁷

The authors argue in such a manner to maintain that their theorizing-theory is conceptually pluralist. “[M]odules, theories, and empirical generalizations are all in play literally from birth.”(68) On the other hand, they also claim that some deeply dynamical and tacit theorizing is working in the minds of children's conceptual developments. By way of analogy, for instance, they claim that tacit theory-development in childhood stands in the same relationship to updated information processes, as scientific theories do in the face of technological innovation. (70) One may respond, however, that by way of Occam's Razor, among other reasons, Carruthers (2005) offers up a more plausible alternative. Carruthers' weak massive modularity thesis gives the same account for the structural and functional aspects of theorizing without the additional assumptions associated with Gopnik & Melzoff's conceptual pluralism.

I have critically reviewed above in some greater detail what I consider are some of the weaker points in Gopnik & Metzoff's theorizing-theory, and aimed to show that Carruthers offers a more robust account. I have primarily done this to show that in principle a modular account *can* provide an adequate, if not a better, explanatory alternative for issues related to the dynamics of theory-change that Gopnik & Melzoff address. My primary reason, however, is not to argue in favor of weak massive modularity per se, but specifically because Carruthers depicts adult scientific and practical reasoning activity as essentially involving two-systems reasoning.¹⁸ Indeed, as I argue in section III below, one need not necessarily ascribe to Carruthers' modular account to agree with the latter two-systems claim.

¹⁷ This happens all the time in 'normal science' (Kuhn (1962)), so long as the theoretical framework, or paradigm, is preserved in the face of improvements or modifications of its sub-theories. Gopnik and Melzoff should have made the qualification that 'theory-laden' observation-statements aren't *entirely* theoretical. But doing so would undermine their distinction of empirical generalizations from theories here.

¹⁸ Carruthers (2005), ch 6, 7. I should also qualify that when bringing up Carruthers vis-à-vis Gopnik, it is important to keep in mind the distinction between Carruthers' philosophy *of psychology* versus Gopnik's *philosophical* psychology. Though their respectively empirical domains clearly overlap (Carruthers gives his own detailed response to Gopnik and Melzoff in ch.6, 13-17) Carruthers is not attempting to establish massive modularity as a foundation for a framework for psychology. Akin to Hume's quasi-empirical maneuvers, Carruthers is simply advancing an empirically grounded inference to the best explanation “to lay out the best case for it [massive modularity].” (x-xi)

But as my above critical review indicates, Gopnik and Melzoff do not bring up two-system reasoning in any meaningful fashion, when they discuss scientific theorizing activity. In my opinion, this is the source of the primary equivocation latent in the theorizing-theory: that cognitive processes taking place in early childhood learning essentially underwrite those governing the dynamics of scientific theory-change. (This has predictably spawned a host of critical responses.¹⁹) I also argue in section III that Gopnik’s research domain consisting of early childhood learning can be essentially characterized by System-1 processes alone, whereas her explanatory “target” of adult scientific activity is essentially comprised by interaction effects between System-2 and System-1 dynamics.²⁰ As in this section, in my more detailed review of Gopnik & Glymour (2002) (in II.iii below), I continue my critical line to reinforce my more general points of discussion I raise in the beginning of section III.

II.iii) A closer examination of Gopnik and Glymour

Harkening back to the theorizing-theory, Gopnik and Glymour (2002) open with the claim:

Cognitive psychologists have argued that much of our adult knowledge, particularly our knowledge of the physical, biological, and psychological world consists of ‘intuitive’ or ‘naïve’ or ‘folk’ theories...this idea, which we call the ‘theor[izing] theory’ rests on an analogy between everyday knowledge and scientific theories. (117)

In an effort perhaps to lay the theorizing-theory on more solidly empirical groundwork, the authors specifically claim that human cognition possesses a “non-egocentric cognitive map.” This map is comprised of patterns of causal relations inferred from, and abstracted out of the local environment, such that:

¹⁹ Carruthers (2002a), 83-84, Faucher, et. al. (2002), Carruthers (2005) ch. 6, 13-16. Carruthers calls into question their thesis along cognitive-psychological lines, by bringing up the issue that conceptual change in childhood learning occurs relatively efficiently and in a timely manner spanning across a variety of childhood cognitive abilities and styles. Whereas in science, even in one community of specialists, conceptual change occurs slowly and deliberately, and is rife with controversy. Faucher et. al question Gopnik and Melzoff along sociological lines, by arguing that the ‘richness of data’ children have access to is insufficient to explain the distinction between the complex and regimented degree of coordination exhibited by scientific communities versus the relatively informal manner in which children group together with their peers.

²⁰ I am borrowing Nersessian’s terms in her discussion on modeling and analogy-based reasoning in science. Analogies and abstractions thereof essentially involve source and target domains of knowledge. ((2002), 144)

- A.) Causal facts, “used and evidenced in prediction, intervention and control” are partially revealed in the correlation between the human agent and his or her immediate environment. (119)
- B.) Such non-egocentric causal maps compensate for the fact that “[causal] [d]ata available are limited in particular ways.” (ibid.) In other words, “[c]hildren and adults may observe correlations they cannot control or manipulate, [or] they may observe features they can control or manipulate only indirectly. [Whereas] other objects...may involve an enormous number of features, only some of which are causally related.” (ibid.)
- C.) Moreover, “[h]uman beings [may] have a theory-formation system, like the visual system...recover[ing] causal facts by making implicit assumptions about the causal structure about the environment.” (ibid.)

Such a cognitive theory, as suggested above in C.), advances the two-fold conjecture: a.) Humans construct causal maps in a manner analogous to the visual maps hypothesized as computed by the human visual system.²¹ b.) Perhaps *only* humans possess some cognitive system which computes “non-egocentric, abstract, coherent representations of causal relations among objects.”(124) Given the adaptive importance of visual and causal knowledge, just as visual-cognitive maps should be differentiated from more concrete and egocentric forms of *spatial* cognition, “causal maps” should be distinguished from a variety of more concrete and egocentric *causal* cognition, engaged in by a variety of animals.²²

As Gopnik & Melzoff (1997) advance general analogies between childhood development and scientific theory-change, the operative analogy between abstract visual and causal maps seems *prima facie* compelling. On the other hand, keeping the appropriate qualifications in mind (n.18 above), consider what Carruthers’ (2005) account has to say. While Gopnik and Glymour *stipulate* that only human beings may construct and compute causal maps, Carruthers (2005, chs. 4 & 5) *explains* how seemingly uniquely human cognitive activities emerge.²³ The latter comprises part of an evolutionary biological, psychological, anthropological, and sociological *explanatory account* Carruthers sustains throughout his book.

²¹ E.g., Marr’s (1982) “2¹/₂-dimensional” sketches, etc.

²² Gopnik & Glymour (2002, 121-122). For instance, the Garcia effect (i.e. the phenomenon of an animal avoiding a contaminated food source) would be an example of this concrete and egocentric form of causal cognition. (122)

²³ Activities, among others, like art and science, manifesting in the sophisticated, robust, flexible, meta-evolutionary matrix of language and culture.

In place of stipulated analogies, inference to the best explanation pervades the methodology throughout Carruthers' work. This is manifest in his most concrete and particular discussions ranging in topics in case-studies in entymology and ornithology (ch. 2) to his more broadly-sketched out terrain of human culture and creativity (ch. 5), science (ch. 6), and practical reasoning (ch. 7). The numerous (and inevitably weakly anthropomorphic²⁴) analogies²⁵ subserve as "gears" comprising Carruthers' explanatory engine. Analogies *regulate*, but do not *constitute* the numerous claims in Carruthers (2005), whereas the above visual/causal analogy seems to constitute the central claim in Gopnik & Glymour.

Another problematic notion resides in causal-map talk: "[C]ausal claims and inferences...play a central role in most scientific theories."(122) Again, *prima facie*, this seems plausible. However, the metaphysics and semantics inherent in the concept of causation is *deeply* problematic. It is not at all clear, for instance, what kind of a common reference the theoretical term shares across even *two* distinct theoretical domains, for example, like that of the Standard Model in particle physics versus the actuarial sciences.²⁶ Lest this worry I explore here appear as a red herring, let us restrict the concept of causality to just that shared by inductive-statistical and interventionist accounts.²⁷ (The authors clearly refer to this notion in their research on Causal Bayes Nets, or CBNs.) By doing so, however, the authors' claim that Bayesian nets model folk

²⁴ Consider, for instance, Hans Julius Schneider's (2000) late-Wittgensteinian critique of cognitive psychology as being an enterprise indiscriminantly reifying anthropomorphic metaphors. (The visual cortex 'weighing the evidence' of its inputs, is just one such example of a second-order metaphor, as it's based on the *legal* notion of 'weighing/considering' which in turn is based on the physical activity of weighing, (198).) To which the cognitive psychologist resoundingly responds: obviously terms referring to *information-processing* possess their irreducibly dispositional and metaphorical character, but this is just an example of the *strength* of model-based reasoning operating outside its domain of what is immediately observed. (Nersessian (2002)). It is not at all clear that Schneider's 'language game' account of mental terms adequately addresses this theoretical issue, much less the obviously empirical success of cognitive science has gained over behaviorism over the last half-century.

²⁵ Inherent, for instance, in functional notions of modules 'tagging' their outputs, etc.

²⁶ The particle physicist when discussing causation seems to defer to a wholly quantitative and fully *nomic-type* notion (e.g. couched in terms like locality, invariance, gauge group, etc.) whereas the actuarial scientist would invoke a *relational token* (sometimes statistical, when estimating health insurance rates for clients with risk factors, and other times brutally indexical, as in the case of an automobile accident investigation.) To argue that such latter folk-statistical notions reduce to the former physical notions raises the even more formidable problems of metaphysical and methodological reduction. Needless to say, since it's a well-known fact that quantum theory *cannot even account for molecular structure*, this renders any talk of 'reduction' from chemistry to physics a forlorn hope at best. From these considerations, as well as others, the concept of causality at best should be considered as some open-textured, cluster-concept.

²⁷ Salmon (1984), Spohn (2000), etc.

as well as scientific theories²⁸ is largely deflated. Such a restriction would shrink the ‘central role’ played by causation to the domain of folk-psychological interventionist accounts, and perhaps to *some* scientific theories (*not* inclusive of most contemporary physics, as n. 26 explains above).

Granted, this might be precisely what Gopnik and Glymour *want* to show: Metaphysical and semantic worries over causal talk aside, this still does *not* detract from their central point that the characteristic *patterns of inference* made by the scientists are a small step removed from those made by children themselves. Moreover, as their research indicates, Bayes nets may play a very central role here. By the same token, however, one should also keep in mind the cautionary point presented in the literature of the philosophy of science (and its sub-specialty, scientific explanation²⁹) concerning attempts to base scientific theorizing on causal claims and inferences.

Bayes nets are advertised as optimally capturing the dynamics of scientific learning, the way the theorizing-theory supposedly accounts for the dynamics of conceptual change.(130) “[C]ausal maps seem to capture both what is general and what is specific about everyday theories.”(124) Vis-à-vis claims (A), (B), (C) mentioned above, the authors advance the following proposal: Children and adults construct CBNs primarily because:

- (A’) CBNs show the agent how to infer causal features from other causal features of the system or object of interest.
- (B’) CBNs show the agent the consequences of his or her actions, allowing for appropriate intervention.
- (C’) CBNs show, from the agent’s observation and actions, how causal action can be learned. (124)

Since the 1970s philosophers of science have utilized CBNs as an aid in modeling how scientists may form predictions from observations, as well as how they predict effects from actions.(127, 130) (I should however emphasize that such a project in the philosophy of science involving the search for the representation of patterns of causal

²⁸ Gopnik & Glymour (2002), 123.

²⁹ Some, like Salmon (1989), argue that scientific explanations are fundamentally causal-mechanical, with their ideal limit being the Ideal Explanatory Text (IBE) which would provide a maximally complete detailed dynamical story for any phenomenon. Others, however, like Philip Kitcher (1989) argue that scientific explanations should *unify*, and hence don’t (or oftentimes *can’t*, if the unification account is synchronic) involve causation to any significant extent.

inference, is a far cry from the *constitutive* claims made by Gopnik and Glymour, concerning causation and scientific theorizing.)

The authors' treatment of CBNs striking a balance between domain-specificity and generality raises the question concerning encapsulation, or cognitive accessibility.³⁰ It is clear from the research cited that the subjects' patterns of inference, for both adults and children, describable via CBNs to a reasonable degree of accuracy, are tacit and non-conscious. For example, "[a]dults seem to unconsciously consider the relationships among possible causes, [i.e.,]...consider alternative causal graphs when they decide how much causal influence one event has on another event." (129) Moreover, "[w]e have *prima facie* evidence that children do, in fact, learn an incredible amount about the causal structure of the world around them. That is evidence that supports the theor[izing] theory in general." (130) Such learning is clearly unconscious. (131)

Cognitive inaccessibility of contents, however, implies non-conscious processes, but *not* the converse. Gopnik and Glymour leave the accessibility question unanswered. Their position seems ambiguous. Aside from the 'causal reasoning constituting scientific theorizing' claim I critiqued earlier, they also seem to equivocate adult capacities for learning new causal facts, "central for understanding scientific cognition," (130) with what is observed as tacit in causal schema-building in toddlers. This is warranted by their theorizing-theory: childhood patterns of inference and those of adults differ only in terms of degree, and in the contingencies of content and context.

However, concerning the question of accessibility, they may have run into an impasse. Based on their experimental findings on eighteen-month olds playing the "Blicket detector" game,³¹ for example, the speed and sophistication through which the Bayesian inferences are made, compared to the subjects' relatively simple 'mind-reading'

³⁰ Admittedly, though, domain generality/specificity and cognitive accessibility/ inaccessibility are separate dimensions of description. Nevertheless, I raise the issue here because the authors' renditions of modularity admit only the rather strong variety, i.e., Pinker's and Fodor's. These renditions depict modules as domain-specific *and* harboring cognitively inaccessible contents.

³¹ In the 'Blicket-detector' game, a machine lights up and plays music when certain objects, and not others, are dropped on it. Gopnik and Glymour summarize that: i) Among all the contingent properties of the objects, two-year-old children "swiftly and accurately" determine which of them have the "Blicket" causal property. ii) Moreover, these *very* young people understand the objects' causal powers in such a manner that respects the CBN formalization. For instance, "even very young children use a form of 'screening-off' reasoning to solve these problems." (131)

capability, obviously belies any notion that the children have cognitive access to these hypothesized non-trivial inference mechanisms.

Granted, regarding the question concerning cognitive accessibility, one must guard against conflating the *underlying form* of scientific theorizing with its associated *content*. Certainly, CBNs apply to the former. Nevertheless, the central motivation of Gopnik involves the phenomena of learning, development, and conceptual change. “Human capacities for learning new causal facts about the world may be marginal for understanding much everyday adult cognition, *but they are central for understanding scientific cognition.*” ((2002), 130, italics added) However one may push CBNs away from the arena of conscious reflection, or down to the cognitively inaccessible ‘assembly language’ of reasoning’s background computations, it is reasonable to think that there should be *some* degree of cognitive accessibility of our causal inference patterns.

Though children and scientists may find themselves in a similar predicament, insofar as they are continually challenged to learn new causal facts, nevertheless scientific learning and conceptual change certainly entails conscious deliberation.³² Though early childhood learning involves fast computations, which are clearly cognitively inaccessible, adult scientific learning and conceptual change, however sophisticated, is slow and deliberate. One would expect the adult scientist to have some cognitive access to the form (and certainly the content) of the dynamics of scientific conceptual change. I develop this point in more detail in section III.

Concerning the nature of the role played by CBNs, the authors qualify:

Whenever we apply computational work to psychological phenomena we have no guarantee that the human mind will behave in the same way as a computer. We...have important reasons to think that the two [indeed] will be different...The program we propose is therefore not to theorize that children or scientists are optimal data-miners, but rather to investigate in general how human minds [may] learn causal maps, and how much (and possibly how little) their learning processes accord with Bayes net computation and heuristics. (131)

³² Conscious deliberation shouldn’t necessarily be equated with just the phenomenology a scientist is engaged in, on the fly, when doing research. “[Such] reflective processes are largely the result of after-the-fact reconstructions of philosophers of science.” (Gopnik & Meltzoff (1997), 122) Certainly, for instance, when a scientific research group is engaged in theoretical fine-tuning, reflective or deliberative processes play some constitutive role. See n 19, above.

Children shouldn't be held up as standards for data-mining, just as much as it is erroneous to think of them as 'little scientists.' ((1997), 32)

In summary, the research on CBNs indicates a detailed investigation of System-1 processes. Even more than the theorizing-theory indicates, it is evident here that System-1 processes are the primary focus of Gopnik & Glymour's study. Though the authors do not address the issue of two-systems reasoning,³³ their renditions of conceptual change can be characterized by System-1 processes alone. This is inadequate for a proper account of scientific activity (which as I argue in the section below) must include interaction effects between System-2 and System-1 processes.

III.) Science involves interaction effects between System-2 with System-1 processes

One need not necessarily be a modularist to believe that human cognitive dynamical processes involve essentially two different systems. On the one hand, there are System-1 processes: These are fast, computationally cheap, and mostly tacit processing systems that function just 'well enough' modulo the constraints posed by the agent's immediate sensory environment. On the other hand, there are System-2 processes characterized as slow, deliberate, and computationally expensive. Carruthers (2005, ch.4, ch. 5) for instance, describes System-1 processes in terms of networks of modules and submodules delivering up their contents to a central-language processing module. The latter, in turn, is coupled to a mind-reading system, which in turn 'globally broadcasts' its outputs back down to certain modules in the System-1 architecture, whether in the form of inner speech, or mental imagery, 'Mentalese,' etc. System-2 is characterized by *cycles* of this activity. It is a slow process, due to the dynamical stability brought about by feedforward/feedback cycles.³⁴

³³ Gopnik, et. al (1997, 2002, 2003) take issues with innatist accounts, which might explain why the two-systems theory is omitted from discussion. Again, though, as I shall explain in section III, one need not advance some strong modularity claim (a la Fodor and Pinker, whom she does well to criticize) to advocate some kind of two-systems theory. Kahnemann (2002) advances no modularity claim at all.

³⁴ To clarify, in any dynamical system composed of an epi-systemic 'driver' and the responding system, the evolutionary equations admit of initially transient behavior before the system settles into some local dynamic equilibrium. For example, consider the dynamical system composed of a child on a swing and the adult pushing the child, i.e., the driver. Before the oscillations of the swinging child match the frequency of the driver, one observes briefly some unstable transient behavior: the swing wiggles and trembles (i.e. complex momentary sub-oscillations occur). In Carruthers' account discussed above, 'global broadcasting' is the driver and the System-1 is the responder.

Daniel Kahneman (2002) in his Nobel Prize address summarizes his empirical work in terms of a two-systems theory. He does not, however, offer any constructive theoretical account of how the two systems emerge from some theory of cognition, whether modular, connectionist, etc.:

An intuitive [System-1] judgment that violates a [System-2] rule...will be overridden, if the rule comes through early in the respondent's mind. This argument is not circular, because we have adequate scientific knowledge (as well as widely shared folk knowledge) about the conditions that facilitate or impede the accessibility of...[System-2] rules. (473)

Complex judgments and preferences are 'intuitive' (and therefore System-1) if they appear in the mind like percepts, i.e., effortlessly and quickly.(481) Moreover, such "[h]ighly accessible impressions produced by [System-1] control judgments and preferences, unless modified or overridden by the deliberative processes of [System-2]." (ibid.)

Consider, for example, the following question Kahnemann poses from his storehouse of carefully designed experimental test-questions:

- **Q1:** A ping-pong ball and a paddle cost a total of \$1.10. The paddle costs \$1.00 more than the ball. How much does the ball (and paddle) cost?

Reflecting on the phenomenology of answering the question, System-1 might reply with: "\$0.10 for the ball, \$1.00 for the paddle." System-2, on the other hand, restrains this initial judgment, perhaps in the form of some inner speech like: "Wait a second, the paddle costs a dollar *more*," etc. Perhaps the subject may sketch out a simple algebraic equation using pencil and paper (though algebra is not required to solve the problem), before arriving at the correct answer of \$0.05 for the ball, \$1.05 for the paddle.

Not only is the phenomenology of **Q1** exemplary of Carruthers' account, i.e. exemplary of cycles of 'global broadcasting' of System-2 and System-1 dynamics,³⁵ but above all this is exemplary of *scientific* cognition. When engaged in 'normal science' (Kuhn (1962)), i.e., problem-solving within a theoretical framework or paradigm, System-1 runs its heuristics, and delivers its contents using rough ampliative inferences, which often involve a host of quasi-perceptual 'models.' (Nersessian (2002))

³⁵ It may take several iterations of inner speech before one settles in on the right answer.

System-2, on the other hand with its arsenal of cognitive ‘head-and-hands’³⁶ protocols, methods, and formalisms, for the most part inherited publicly from one’s specialized education and training, monitors and constrains the ‘surface’ in the ‘solution space’ over which System-1’s dynamical trajectory evolves.³⁷ Sometimes, in the case of highly formal and abstract problems, System-2 presents a unique path or constrained trajectory in the solution space. For instance, when one attempts to solve a problem in mathematical physics, such problems are so highly constrained by the formal apparatus of the theory that it would prove impossible to even characterize them using some other formalism.³⁸

Referring back to Gopnik & Glymour, it seems overwhelmingly likely that the CBNs refer to System-1 processes. As discussed in section II.iii above, CBNs certainly meet System-1 criteria of being tacit, fast, and computationally cheap.³⁹ Moreover, the fact that conceptual change in early childhood seems likewise to occur in this tacit and relatively automatic manner (across a variety of cultures and cognitive abilities) bespeaks of the dominance of System-1. The empirical results of Gopnik, et. al. (2002, 2003, 2004) significantly contribute to a clearer understanding of the subtle and intricate complexities of System-1 processes, but this only tells half of the cognitive story, when it comes to describing scientific practice.

Granted, Gopnik & Glymour (2002) seek to account for the fundamental *form* of dynamical processes characteristic of inferential patterns used in science, and not the content. Though the contents and even the *particular* form of System-2 processes are undoubtedly highly culture-specific, their underlying *general* form seems to be innate. The latter innate form may appear far more difficult to detect in the face of the

³⁶ Giere (1988)

³⁷ I am borrowing an analogy from dynamical systems theory.

³⁸ Just to name one example: trying to solve for the degree of length contraction of a rigid rod in Minkowski geometry is a problem obviously constituted by this particular metrical geometry. Change the geometry (for instance by projecting down to a 3D manifold endowed with the Euclidean metric) and you’ve lost the problem.

³⁹ Gopnik and Glymour discuss how computationally cheap learning algorithms can be implemented across CBNs. “[W]e can construct algorithms that will arrive at the correct causal structure if they are given information about the contingencies among events. These systems can learn about causal structure from observation and intervention.” (2002, 129) Moreover, concerning their tacit characteristics, “[a]dults [as well as children] seem to unconsciously consider the relationships among possible causes...[by] consider[ing] alternative causal graphs when they decide how much causal influence event has on another event.” (ibid.)

overwhelming evidence of innate System-1 processes, since one may assume System-2 processes evolved relatively recently. Consider, for example, the phenomenology of **Q1**. One need not know algebra to figure it out, it just takes some common sense or practical reasoning. (Carruthers (2005), ch. 7) In other words, the cautionary and deliberative response emblematic of System-2 appears to be an innate *tendency*, and hence not entirely constituted by sociolinguistic tools and norms (like algebra in this example). The fact that this tendency is fragile, however, is underscored by “the assumption that System-2 is vulnerable to interference by competing cognitive activities...[since] manifestations of intuitive [System-1] thought that are normally inhibited may be expressed when people are placed under cognitive load.” (Kahneman (2002) 482)

Kahneman (2002, 473) cites some of the characteristic features inhibiting the efficacy of System-2, which include:

1. Time constraints
2. Concurrent involvement in different tasks
3. Good mood
4. Switching evening tasks for ‘morning people,’ and vice versa.

Such strongly negative correlations between the above-mentioned cognitive and physiological factors versus System-2 efficacy suggest that dynamical properties of System-2 are fundamentally innate. For what is significant is the fact that the correlation between System-2 and factors 1.)-4.) is *strong*, not so much that it is negative.

Overall, the theorizing-theory may be equivocating a *mentality* with a *theory of mind*.⁴⁰ Gopnik & Meltzoff (1997) speak at great lengths about the mentality of children. But to associate the mentality of early childhood learning with theories of conceptual change in the sciences, whether in the specific form of CBNs or otherwise, bespeaks of such a possible equivocation. In effect what this does is to homogenize the characteristic dynamics of learning from childhood to adulthood. But conceptual change in the sciences involves System-2/System-1 interaction effects, whereas early childhood learning is accounted for by System-1 processes. CBNs may provide a valuable key in revealing the tacit and inaccessible aspects of System-1 learning, which produce “highly accessible impressions...control[ling] judgments and preferences.” (481) Nevertheless,

⁴⁰ Another reason why Carruthers (2005) prefers the notion ‘mind-reading’ instead of ‘theory-theory’ (ch. 4, 5)

the slow and deliberate character of conceptual change in the sciences indicates that the aforementioned System-1 impressions can be “modified and overridden by the deliberate operations of System-2.” (ibid) It is doubtful that anything like this occurs among toddlers.

Gopnik et. al focus primarily on *radical* conceptual change in the sciences, i.e., ‘revolutionary science.’ (Kuhn (1962)) My illustration above, on the other hand, pertains to normal science. During periods of crisis brought about by radical conceptual change, System-2 processes admittedly can take a back-seat, since there exists no coherent framework of norms by and through which such a system can have its say. Nersessian (2002) for example, argues along similar lines: System-1 processes involving the formation of quasi-visual mental models play a crucial role in periods of radical conceptual change in the sciences. Nevertheless, as I shall argue below, in times of crisis System-2 still serves in some essential capacity.

IV.) A more nuanced account of System-2/ System-1 dynamical interaction

In this section I sketch a notion that cognitive processes underlying science may in principle be quantified in terms of a degree of influence System-2 has with respect to System-1. This degree of influence, in turn, is stable enough to distinguish the kind of theorizing activity taking place in a given scientific discipline.⁴¹ First, I’ll list a few clarifying points:

- I am *not* saying that science narrowly construed (Carruthers (2002a), 74) is somehow unique simply because of its relatively rare occurrence in known history.⁴² Like Carruthers (2002a) I focus on the question concerning the

⁴¹ Adult practical reasoning also involves two-systems reasoning. (Carruthers (2005), ch 7) Practical reasoning, however, in all its flexibility and particular idiosyncrasies, may prove recalcitrant in the face of some such systematic attempt of categorization. The degree of influence System-2 has over System-1 may vary stochastically: Practical reasoning operates on an essentially case-by-case basis, as opposed to the relatively controlled and constrained cognitive activities of science.

⁴² I.e., arising sometime in late sixteenth-century Europe, with its programmatically very specific agenda involving an interesting theoretical admixture of neoPlatonism and proto-positivism. Such an admixture is apparent, for instance, in Francis Bacon’s fascination application of the cryptographic methods to the Book of Nature metaphor. Moreover, this cultural phenomenon introduced an equally unique empirical methodology based not Aristotelean non-interventionism (except in the case of astronomy) but rather on an Archimedean engineering strategy. (Pescic (2000))

“[n]ature and extent of cognitive changes [and processes] *necessary for science to begin.*” (ibid., italics added).

- I am advocating here a plural-aspects metaphysical position. That is to say, I view the phenomenon of science as exhibiting a potentially infinite set of aspects characterized in terms of ‘weight parameters’ ascribed to the System-2/System-1 dynamic. The degree of influence System-2 has over System-1 represents a unique *aspect* of the phenomenon of scientific activity. Abstracting away these aspects leaves one with an oversimplified view of science.

To motivate my intuitions, I fix a particular research tradition, and examine it diachronically. For purposes of simplicity, I will use the hackneyed case of celestial mechanics. Consider the following table:

Celestial Mechanics

Paradigm/ theoretical framework	System-2/System-1 interaction dynamic
Ptolemaic era (approx. 15 th cent.)	System-2 \Leftrightarrow System-1
Late Newtonian era (mid 19 th cent.)	System-2 \Rightarrow System-1
Early Relativistic era (early 20 th cent.)	System-2 \Leftarrow System-1

Table IV.1

What Table IV.1 aims to show is the relative dynamic preponderance, i.e., the ‘weighted equilibria’ shared by the two systems. Consider the phenomenological and empirically adequate scheme of Ptolemaic astronomy.⁴³ From this cognitive-historical standpoint, one could claim that when conducting ‘normal science’ in this adolescent research tradition, System-1 and System-2 are in a roughly equally-weighted dynamical equilibrium: The mathematical framework (essentially Fourier methods, in contemporary terms) imposed upon by System-2 does *not* strictly constrain the loose set of heuristics delivered up by System-1. The constraint is just one of empirical adequacy.

⁴³ Just as many physicists today adopt a similar instrumentalist stance when working in the research tradition of the ‘Renormalization Group Program’ (RGP), so the fifteenth century astronomers had a variety of various boot-strapping tools and heuristics (epicycles, equants, etc) no one really took metaphysically too seriously. Nevertheless, in its heyday Ptolemaic astronomy was empirically adequate, because of the relatively imprecise extant measurement and observation techniques.

As the above example suggests, I submit the claim that all quantitatively-based research traditions in their ‘adolescent’ stage, (i.e., when their mathematics is thought to be primarily just a collection of algorithms for prediction) share this “50/50” cognitive dynamical feature. This is chiefly because the empirically adequate mathematical frame of the adolescent research tradition acts as a mere ‘scaffolding,’ and not as a robust architectural framework. That is to say, the scaffolding consists of algorithms predicting observed data, but offers no robust explanations of such observations. The research tradition must await a more mature future stage, in which a coherent metaphysical infrastructure would evolve, offering consistent interpretations and explanations.⁴⁴

As mentioned in the previous section, when doing normal science, System-2 imposes constraints which guide the trajectories (generated by System-1) in a given solution space. In the case of an adolescent research tradition, the sole constraint System-2 imposes is one of empirical adequacy. This gives System-1 heuristics a relatively free hand in problem-solving activity. These heuristics, for instance, can take the form of mental models, i.e. structural and functional analogues embodying representations of causal structures connecting relevant spatial and temporal events.⁴⁵

In the case of adolescent research traditions, like Ptolemaic astronomy or the present-day Renormalization Group Program (RGP), the mental models guide the quantitative solution procedures. For example, Ptolemaic astronomers used ‘equants’ and ‘epicycles’ in the boot-strapping quantitative schemes to fit the observed data, just as contemporary RGP physicists employ ‘dressed parameters’ in a similar manner.⁴⁶ Epicycles and equants are based on mental models of circular motion, as dressed

⁴⁴ As Kuhn (1962) had already pointed out, this may require radical revisions of the theoretical framework.

⁴⁵ Nersessian ((2002), 141). Echoing Gopnik & Glymour’s (2002) guided speculations on the relevance of CBNs in modeling tacit causal inferences, Nersessian writes: “Advocates of mental modeling argued that the original capacity developed as a means of simulating possible ways of maneuvering within the physical environment...little [however] is known about the nature of the generative processes underlying the construction and use of mental models.” (140) Nersessian goes on to distinguish modal from amodal systems generating mental models. The former are analogues of perceptual states, whereas the latter are arbitrary transductions from perceptual states, arising in language-based systems. Visual-based modal systems seem best described by System-1, whereas language-based amodal systems pertain most likely to System-2.

⁴⁶ Mathematical complexities aside, the similarity is based on approximation series methods. Ptolemaic astronomy is based on Fourier series techniques, whereas RGP on perturbation series.

parameters are based on models of subtracting finite quantities.⁴⁷ The key point is that such mental models generated by System-1 are not reified by System-2, since the research tradition offers no metaphysical story in which these models can secure their place in some explanatory framework of the phenomena under study. Hence, in such adolescent traditions, the researchers' use of System-1 based reasoning in the form of model heuristics is flexible and pragmatic. The Ptolemaic astronomer, for instance, is always free to add more equants and epicycles to fit the observed data.

Consider, on the other hand, the second case of the mature research tradition of Newtonian celestial mechanics in the latter half of the nineteenth-century.⁴⁸ Here, when conducting normal science, the dynamical reciprocity is heavily weighted in the direction of System-2. This is primarily because the tradition possesses a functioning metaphysical and explanatory infrastructure. There remains not much 'wobble-room' left for System-1 heuristics, except in the trivial cases of students forming their own 'scaffolding,' or transient mental models. These models mainly comprise the students' effort to assimilate the theory's abstract mathematical structure, held in place by its coherent metaphysical picture.

Learning and doing normal science in such a mature research tradition entails that one should constantly monitor System-1's intuitive contents, which can lead one astray in the face of the tradition's mathematical and explanatory architecture. The latter are the contents of System-2. For example, when an astronomer tries to solve a problem involving the trajectory of a planet in this mature research tradition, he or she immediately defers to a very abstract and powerful mathematical framework, with a formalism that can simultaneously keep track of multiple constraints and evolutionary equations.⁴⁹ Moreover, this mathematical framework is lent a consistent explanatory character, through precisely quantifiable concepts such as kinetic and potential energy,

⁴⁷ Dressing a parameter involves the mathematically unwarranted maneuver of 'subtracting infinities.' In other words, the act is based on the illegal maneuver: $0 \text{ "=" } \infty - \infty$, based on the (false) analogy with $0 = x - x$, for any finite quantity x . (L'Hopital's Rule shows that $\infty - \infty$ produces an indeterminate result.) The unwarranted move is also based on the intuition that 0 lies 'halfway' between $-\infty$ and $+\infty$ on an infinite number line.

⁴⁸ Including all elegant and powerful reformulations and 'Least-Action Principle' generalizations of Newton's Laws, introduced by Euler, Lagrange, Hamilton, and Jacobi.

⁴⁹ Stated more technically, the astronomer solves a collection of differential equations (Euler-Lagrange, or Hamilton-Jacobi) elegantly generated from a formalization of the system's action (expressed as an integral over time of the energy of the system.)

action-angles, generalized coordinates, holonomic and scleronomic constraints, etc. These latter concepts, due to their highly abstract but metaphysically and quantitatively precise character, often prove themselves elusive and counterintuitive to System-1 heuristics: at a certain point, the student must drop the scaffolding.⁵⁰

Consider the third case, i.e., during the crisis of conceptual change of the transition from Newtonian to relativistic celestial mechanics. During this transition period, the balance tilts heavily in favor of System-1. After the Newtonian framework was dethroned or fundamentally called into question, but before relativistic mechanics became the dominant theoretical framework, scientists had no choice but to heavily rely on thought-experiments and System-1 based heuristic models.

This is generally true, because visual models and thought experiments occur commonly in periods of radical conceptual change: Such reasoning makes evident and abstracts the constraints of existing representational systems, i.e., the source domain of knowledge. But such reasoning also gives, “in the light of constraints provided by the target problem, effective means of integrating constraints from multiple representations such that novel representational structures result.” (Nersessian (2002), 145) The key point is that the research tradition in crisis lacks an explanatory framework, i.e. System-2 contents, by which target problems can be represented, and hopefully solved and explained. So System-1 must do all the footwork, through its use of model-based reasoning, analogies, and thought experiments, in an attempt to bridge the target domain with the source.⁵¹ “The cognitive literature agrees with the position that analogies employed in conceptual change are not ‘merely’ guides to reasoning but are *generative* in the reasoning process in which they are employed.”⁵²

⁵⁰ John von Neumann mentioned that one never “understands” concepts in mathematics (and one may assume, highly abstract mathematical physics), “[s/]he just gets used to them.” In my argument above, this translates as System-2 doing the bulk of the work, since System-1 has no intuitive basis to initially grasp such highly abstract and mathematically subtle notions.

⁵¹ Perhaps this is what Werner Heisenberg meant, when he was reported to remark that “before a theory can axiomatically reformulated, its significant problems must be solved first by dirty means alone.” Presumably the latter refers to System-1 heuristics.

⁵² Nersessian (2002) 146, italics added. Nancy Nersessian basically argues that the provisional modeling or ‘scaffolding’ phase is instrumental in conceptual change. For instance, James Clerk Maxwell was only able to formulate his electromagnetic theory when he first developed transitory mechanical models of ‘gears’ and ‘vortices’ in an aether, the latter being an analogue of an ideal fluid. Once Maxwell’s equations constituting his theoretical framework were set in place, he dispensed with the scaffolding.

Examples abound of System-1 dominance in the case of the transition to relativistic mechanics. For example, in the case of celestial mechanics, Einstein briefly flirted with a ballistic model of photon propagation, in an effort to account for the anomalous gravity-lens effect. The ballistic model allowed for photons to have a distribution of velocities. Einstein later abandoned the crude heuristics of the ballistic model, as he found, among other things, its equations too messy. (Torretti (1983), ch 2) But the resulting mathematical ‘ugliness’ (or lack of systematic unity) of the ballistic model is emblematic of System-1 heuristics: An anomalous effect involving the bending of light around the planet Mercury is provisionally modeled by a source domain of knowledge, namely, what is known about the distributed velocities of nonzero mass particles executing ballistic trajectories. Later, of course, such heuristics were swept aside and replaced with abstract and precise mathematical and metaphysical notions of general covariance and invariance in the theory of General Relativity.

Another example involves System-1 heuristics immediately superseding Special Relativity (SR): Prior to SR, Lorenz-Fitzgerald electromagnetism (LFE) relied on a variety of models of charged particles propagating in a hypothesized aether. The contraction equations derived in the LFE survive as an honorific in the Special Theory of Relativity, though derived entirely from abstract principles of the First and Second Postulates of Special Relativity.⁵³ And as is well known, the very conceptual foundations of relativity theory had their origins in System-1-type thought experiments.⁵⁴

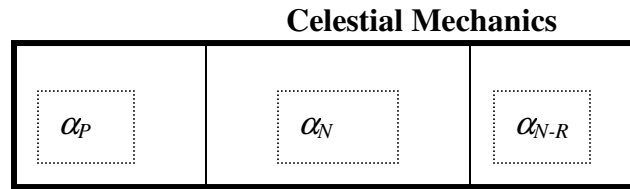
As the above example of celestial mechanics shows, certain domains of science can be classified in terms of weighted dynamical equilibria shared between the two systems. I am couching my claims in the overall framework of dynamical systems theory. The difficult, but in principle not impossible, problem however remains one of developing a coherent and consistent method for quantifying such weights. For purposes of this discussion I define the relation:

⁵³ Which state: 1) The laws of physics should remain invariant across all inertial reference frames. 2) The speed of light is invariant in all inertial reference frames.

⁵⁴ In the case of Special Relativity, Einstein’s thought experiments involved light-signals emitted from passing trains. In the case of the derivation of the General Equivalence Principle in General Relativity, Einstein’s thought-experiment involved an elevator in free-fall in a locally uniform gravitational field.

Defn. IV.1: $(\text{System-2}) \leftrightarrow_{\alpha} (\text{System-1})$ iff System-2 and System-1 are in dynamic equilibrium such that there exists some α , such that: $W[\text{System-2}] = \alpha$, and $0 < \alpha < 1$, for some generalized weight mapping W .⁵⁵

Defn IV.1 obviously entails that \leftrightarrow_{α} is an equivalence relation.⁵⁶ Hence the *degree* of continuous variation of parameter α partitions the theoretical domain into equivalence classes, indexed according to α . These equivalence classes form natural kinds. For example, the domain of celestial mechanics as depicted in Table IV.1 is partitioned into three equivalence classes indexed according to the respective weight parameters $\alpha_P, \alpha_N, \alpha_{N-R}$, in which the indices refer to the Ptolemaic, Newtonian, and Newtonian-Relativistic transition, respectively:



Aside from perhaps generating a more quantitatively-based version of cognitive-historical studies of science, this plural aspects conception based on relative weights may capture certain essential cognitive features of scientific theorizing with greater descriptive and explanatory power, than other attempts to describe scientific theorizing ‘as a whole’.

⁵⁵ W is probably best described in terms of a Category-Theoretic functor, because of the non-trivial structural properties shared by any attempt to systematically quantify over such dynamical properties of the two systems. Without loss of generality, DefnIV.) is equivalent saying $W[\text{System-1}] = \beta$, where: $\alpha + \beta = 1$.

⁵⁶ **Reflexivity:** For any System 2 (SYS_2) it’s obviously in dynamical equilibrium with itself. I.e., : $SYS_2 \leftrightarrow_{\alpha} SYS_2$ iff SYS_2 and SYS_2 are in dynamic equilibrium such that, $W[SYS_2] = \alpha$ where: $0 < \alpha < 1$.

Symmetry: If System 2 is in dynamical equilibrium with System 1, with weigh α , then the converse holds with weight β . I.e., : $SYS_2 \leftrightarrow_{\alpha} SYS_1$ iff SYS_2 and SYS_1 are in dynamic equilibrium such that, $W[SYS_2] = \alpha$ where: $0 < \alpha < 1$. $\Rightarrow SYS_1 \leftrightarrow_{\beta} SYS_2$ iff SYS_1 and SYS_2 are in dynamic equilibrium such that, $W[SYS_1] = \beta$ where: $0 < \beta < 1$ and $\beta = 1 - \alpha$.

Transitivity: If System 2 is in dynamical equilibrium with System 1, with weigh α , and if System 1 is in dynamical equilibrium with System 1’ with weight α' then System 2 is in dynamic equilibrium with System 1’ according to the composition rule of mappings. I.e., : $SYS_2 \leftrightarrow_{\alpha} SYS_1$ iff SYS_2 and SYS_1 are in dynamic equilibrium such that, $W[SYS_2] = \alpha$ where: $0 < \alpha < 1$. & $SYS_1 \leftrightarrow_{\alpha'} SYS_1'$ iff SYS_1 and SYS_1' are in dynamic equilibrium such that, $W[SYS_1] = \alpha'$ $\Rightarrow SYS_2 \leftrightarrow_{\gamma} SYS_1'$ iff SYS_2 and SYS_1' are in dynamic equilibrium such that, $W[SYS_1] = W'[W[SYS_2]] = \gamma$

As a final remark, in the case of synchronic studies of science, my scheme would distinguish salient cognitive differences which other accounts may gloss over. For example, in the case of contemporary microbiology, certain natural kinds of sub-theoretical traditions (whether molecular or functional) would be classified according to different weights α . This would provide an indication of the degree of quantitative maturity that each subspecialty has arrived at. Last of all, this partitioning schema is open-ended in breadth (since α is a continuous parameter) and in depth (one could partition sub-specialties into sub-sub-specialties, etc.)

V. Conclusion

Alison Gopnik et. al's research (2002-2005) offers valuable information and insight into the underlying complexities of System-1 processes that are active during early childhood learning. However, the authors' continuity claims and analogies of childhood learning with respect to conceptual change in science omits what I consider is the essential role played by System-2 in the latter. What distinguishes scientific activity is the unique degree of interaction between System-2 and System-1. Such interaction, in my opinion, is stable enough that with respect to a particular research tradition, the dynamical equilibrium System-2 shares with System-1 is characterized in terms of degrees. I have endeavored to show a possible strategy for quantifying the degrees in terms of weight parameters in equilibrium dynamics. These hypothetical weight parameters can serve to characterize unique aspects of scientific activity in such a manner that any researcher in a cognitive-historical or cognitive-philosophical program would deem as salient.

VI. References

Arbib, Michael A. & Hesse, Mary B. (1986) *The Construction of Reality*. Cambridge: Cambridge University Press.

Carrier, M, Massey, G., Ruetsche, L, eds. (2000) *Science at Century's End: Philosophical Questions on the Progress and Limits of Science*. Proceedings of the Pittsburgh-Konstanz Colloquium in the Philosophy of Science. University of Pittsburgh Press/ Universitaetsverlag Konstanz.

Carrier, Richard (2003). "Fundamental Flaws in Mark Steiner's Challenge to Naturalism in *The Applicability of Mathematics as Philosophical Problem*." Internet Infidels. URL = < http://www.infidels.org/library/modern/richard_carrier/steiner.html >

Carruthers, et. al., eds. (2002) *The Cognitive Basis of Science*. Cambridge: Cambridge University Press.

Carruthers, Peter (2002a) "The Roots of Scientific Reasoning: Infancy, Modularity, and the Art of Tracking," in Carruthers, et. al., eds. (2002) 73-98.

Carruthers, Peter (2005) *The Architecture of the Mind: Massive Modularity and the Flexibility of Thought*. (in press)

Cherniak, Christopher (1986). *Minimal Rationality*. Cambridge, Mass: MIT Press.

Dehaene, Stanislas. (1997) *The Number Sense: How the Mind Creates Mathematics*. Oxford: Oxford University Press.

De Mey, Marc. (1992) *The Cognitive Paradigm: An Integrated Understanding of Scientific Development*. Chicago: The University of Chicago Press.

Faucher, et. al. (2002) "The baby in the lab-coat: why child development is not an adequate model for understanding the development of science," in Carruthers, et. al., eds. (2002) 335-361. On-line pdf: URL = < <http://rucss.rutgers.edu/ArchiveFolder/Research%20Group/Publications/BABYfinalt.htm> >

Fodor, J (1983). *The Modularity of Mind*. Boston: MIT Press.

Fodor, J. (2000). *The Mind Doesn't Work That Way*. Boston: MIT Press.

Galison, Peter (1997). *Image and Logic: A Material Culture of Microphysics*. Chicago: University of Chicago Press

Geary, David C. (2005) *The Origin of Mind: evolution of Brain, Cognition, and General Intelligence*. Washington, DC: America Psychological Association.

Giere, Ronald. (1988) *Explaining Science: A Cognitive Approach*. Chicago: University of Chicago Press.

Gopnik, Alison & Meltzoff, Andrew, N. (1997) *Words, Thoughts, and Theories*. Cambridge, Mass: Bradford Books, MIT Press.

Gopnik, Alison & Glymour, Clark (2002). "Causal Maps and Bayes Nets: A cognitive and computational account of theory-formation," in Carruthers, et. al., eds. (2002) 117-132.

Gopnik, Alison (2003). "The Theory-Theory as an Alternative to the Innateness Hypothesis," *Chomsky and His Critics*, Antony & Hornstein (eds.) New York: Basil Blackwell (in press). On-line pdf:

URL = < <http://ihd.berkeley.edu/chomsky.pdf> >

Gopnik, Alison & Schulz, Laura. (2004a) "Mechanisms of Theory Formation in Young Children," *Trends in Cognitive Science*, vol. 8, n. 8 (371-377).

Gopnik, et. al. (2004b). "A Theory of Causal Learning in Children: Causal Maps and Bayes Nets," *Psychological Review*, 111, 1, 1-30. On-line pdf:

URL = < <http://ihd.berkeley.edu/psychreview.pdf> >

Gopnik, et. al. (2005) "Inferring Hidden Causes" *Proceedings of the Twenty-Fourth Annual Meeting of the Cognitive Science Society* (in press). On-line pdf:

URL = < http://ihd.berkeley.edu/unobserved_cogsci_4.0.pdf >

Giere, Ronald N. & Feigl, Herbert, eds. (1992) *Cognitive Models of Science. Minnesota Studies in the Philosophy of Science* (vol. XV). Minneapolis: University of Minnesota Press.

Kahneman, Daniel (2002). "Maps of Bounded Rationality: A Perspective on Intuitive Judgment and Choice." Nobel Prize Lecture. On-line pdf:

URL =< <http://nobelprize.org/economics/laureates/2002/Kahnemann-lecture.pdf> >

Kant, Immanuel (1783). *Prolegomena to Any Future Metaphysics*. Translated by Paul Carus, and revised by James Ellington (1977). Indianapolis: Hackett Publishing Company.

Kitcher, Philip and Salmon, Wesley, eds. (1989), *Scientific Explanation*, vol. 13, *Minnesota Studies in the Philosophy of Science*, 3-219. Minneapolis: University of Minnesota Press.

Kitcher, Philip (1989). "Explanatory Unification and the Causal Structure of the World," in Kitcher & Salmon, 410-505.

Kuhn, Thomas (1962) *The Structure of Scientific Revolutions*. University of Chicago Press.

Marr, David (1982). *Vision : A Computational Investigation into the Human Representation and Processing of Visual Information*. San Francisco: W. H. Freeman

Nersessian, Nancy (2002). "The cognitive basis of model-based reasoning in science," in Carruthers, et. al., eds. (2002) 133-153.

Pesic, Peter. (2000) *Labyrinth: Toward a Quest for the Hidden Meaning of Science*. MIT Press

Petitot, et. al. (1999) *Naturalizing Phenomenology: Issues in Contemporary Phenomenology and Cognitive Science*. Stanford: Stanford University Press.

Salmon, Wesley (1989) "Four Decades of Scientific Explanation," in Kitcher and Salmon, 3-219.

Schneider, Hans Julius (2000). "Metaphors and Theoretical Terms: Problems in Referring to the Mental," in Carrier, Massey, Ruetsche, eds. (2000) 193-216.

Shimony, Abner (1993) *The Search for a Naturalistic World View: (vol. 1) Scientific Method and Epistemology*. Cambridge: Cambridge University Press

Spohn, Wolfgang (2000). "Bayesian Nets Are All There Is To Causal Dependence," (in M.C. Galavotti et al. (eds.), *Stochastic Dependence and Causality*, Stanford: CSLI Publications) 157-172. On-line pdf:

URL = < http://www.uni-konstanz.de/FuF/Philo/Philosophie/Mitarbeiter/spohn_files/wspohn41.pdf >

Steiner, Mark (1998) *The Applicability of Mathematics as a Philosophical Problem*. Cambridge, Mass: Harvard University Press.

Thagard, Paul (1992). *Conceptual Revolutions*. Princeton, NJ: Princeton University Press.

Torretti, Roberto (1983) *Relativity and Geometry*. NY: Pergamon Press