The theory theory as an alternative to the innateness hypothesis.

Alison Gopnik
Dept. of Psychology
University of California at Berkeley

Address for correspondence: Alison Gopnik, Dept. of Psychology, University of California at Berkeley, Berkeley, CA, 94704 gopnik@berkeley.edu

In L. Antony & N. Hornstein (Eds). (2003) Chomsky and his Critics . New York: Basil Blackwell

One of the deepest and most ancient problems in philosophy is what we might call the problem of knowledge. There seems to be an unbridgeable gap between our abstract, complex, highly structured knowledge of the world, and the concrete, limited and confused information provided by our senses. Since the Meno, there have been two basic ways of approaching this problem, rationalism and empiricism. Each era seems to have its matched pair of advocates of each view, making their way through the centuries like couples in some eternal philosophical gavotte, Plato and Aristotle, Descartes and Locke, Leibniz and Berkeley, Kant and Mill. The rationalist approach says that although it looks as if we learn about the world from our experience, we don't really. Actually, we knew about it all along. The most important things we know were there to begin with, planted innately in our minds by God or evolution (or chance). The empiricist approach says that although it looks as if our knowledge is far removed from our experience, it isn't really. If we rearrange the elements of our experience in particular ways, by associating ideas, or putting together stimuli and responses, we'll end up with our knowledge of the world. There is both a tension and a kind of complementarity between these two ideas, and philosophy and psychology often seem to alternate between one view and the other. Rationalists can explain the abstract, complex, nature of our knowledge quite well, but they can't explain, and so deny, the fact that we learn. Empiricists can explain the fact that we learn, but they can't explain, and so deny, the fact that our knowledge is so far removed from our experience.

Chomsky made two very important contributions to this debate. The first was to cast it as an empirical, scientific, psychological question, rather than simply as a philosophical question. Chomsky has argued that we can develop a scientific account of how knowledge develops in individual human minds, and that such an account will provide an answer to the problem of knowledge. Moreover, he suggested that such an account would involve representations of the outside world and rules for manipulating those representations. These representations and rules characterize the cognitive capacities of human minds, they are the theoretical entities of cognitive psychology. That idea is at the heart of cognitive science, and underlies its success. It has been very widely influential. In fact, perhaps the greatest successes in cognitive science have been in domains that Chomsky didn't himself investigate, particularly vision, using methods, psychological experimentation, detailed computational modelling, and, most recently, neuroscience, that he has eschewed. There were other earlier sources for this idea, too, including Piaget and the Gestalt psychologists. But none of that alters the fact that, in its modern incarnation, this foundational idea of cognitive science is largely due to Chomsky.

We might call this part of Chomsky's view "cognitive naturalism", the idea that knowledge can be understood by scientific investigation of the mind. It assumes, of course, some version or other of scientific realism. If you think scientific investigation is the right course to find the truth of these questions, you must think that scientific investigation is the right course to find the truth, in general.

The second part of Chomsky's contribution was to present a particular, rationalist hypothesis, the innateness hypothesis, as the empirical answer to the problem of knowledge. Chomsky's thesis is that human minds are highly constrained innately. We can only formulate a very small set of possible representations and rules. Information from the outside world may trigger the development of those representations, and may narrow the set of possibilities even further by processes like parameter-setting, but the constraints remain unchanged throughout life. Representations and rules are not inferred or derived from the input. This rationalist answer was, of course, a radical, indeed revolutionary, departure from the empiricist assumptions of earlier psychology, particularly behaviorism. It has also been very widely influential. It has dominated Chomsky's own field, syntax, and has influenced theories of semantics (Pinker, 1989), and phonology (Eimas, 1975), and of our knowledge of everyday physics (Spelke, Breinlinger, Macomber, and Jacobson, 1992), biology (Atran, 1990), psychology (Leslie & Roth, 1993), and even ethics (Barkow, Cosmides, and Tooby, 1992).

There are a range of empirical phenomena, particularly developmental phenomena, that could be adduced to support the innateness hypothesis. But it is fair to say that these phenomena have played little role in Chomsky's own thinking and argumentation. Chomsky's arguments for innateness don't come from studying the development of language and thought in children but from considering the characteristics of adult language and thought. Chomsky's most important argument for rationalism is the same argument that Socrates originally formulated in the Meno, it has come to be called the poverty of the stimulus argument. The learning mechanisms we know about are too weak to derive the kind of knowledge we have from the kinds of information we get from the outside world.

The central argument I want to make in this paper is that these two important Chomskyan theses, cognitive naturalism, on the one hand, and the poverty of the stimulus argument, on the other, are at odds with one another. Suppose we apply the program of cognitive naturalism to scientific knowledge. If we do, and if we are scientific realists, we must believe that there are learning mechanisms that allow human minds to derive abstract, complex, highly structured, veridical, representations and rules, namely theories, from limited input, namely evidence. They are just the mechanisms we use in science. But if we believe this, then it is at least logically possible that those same learning mechanisms are involved in our development of other kinds of knowledge, such as everyday physical, , biological, psychological and even linguistic knowledge.

Moreover, recent empirical developmental research suggests that this is precisely the case. Within the last ten years the idea that there are deep similarities between scientific theory formation and cognitive development, an idea we have called “the theory theory”, has become, at the least, a serious developmental hypothesis. The cognitive abilities involved in science do seem to also be involved in everyday cognitive development. Of course, this still leaves open the question of whether this is the correct account of development in any particular domain. But it also means that the poverty of the stimulus argument no longer applies. To discover whether any particular kind of knowledge is due to innate Chomskyan devices or to theory-formation we need to do detailed empirical studies, and, in particular, detailed developmental studies. We can’t make those arguments simply by looking at the adult system.

The Theory Theory

First, I’ll present a very brief sketch of the theory theory as it has emerged in cognitive development (for more detailed recent accounts see Gopnik & Wellman, 1994, Wellman & Gelman, 1997, Gopnik & Meltzoff, 1997). The basic idea is that children develop their everyday knowledge of the world by using the same cognitive devices that adults use in science. In particular, children develop abstract, coherent, systems of entities and rules, particularly causal entities and rules. That is, they develop theories. These theories enable

. children to make predictions about new evidence, to interpret evidence, and to explain evidence. Children actively experiment with and explore the world, testing the predictions of the theory and gathering relevant evidence. Some counter-evidence to the theory is simply reinterpreted in terms of the theory. Eventually, however, when many predictions of the theory are falsified, the child begins to seek alternative theories. If the alternative does a better job of predicting and explaining the evidence it replaces the existing theory.

This account has been successfully applied to explain children's developing understanding of the physical world (Gopnik, 1988; Smith, Carey, and Wiser, 1985) the biological world (Carey, 1985; Gelman & Wellman, 1991; Keil, 1989) and the psychological world (Gopnik, 1993; Wellman, 1990; Perner, 1991). It has been applied to adults (Murphy & Medin, 1985) school-age children (Carey, 1985; Keil, 1989) preschoolers (Gelman & Wellman, 1991) and infants (Gopnik, 1988; Gopnik & Meltzoff, 1997). We have suggested, in particular, that infants are born with initial innate theories, and that they begin revising those theories even in infancy itself. There are, of course, debates about which domains and which ages the theory applies to, and the theory is in competition with other accounts, including Chomskyan innateness theories.

While the theory theory has primarily been articulated in cognitive development, it also helps to explain science itself . On this view, the success of science depends on the fact that we adults retain natural learning capacities that are most clearly evident in children (and were probably designed by evolution to be used by children). Science simply puts these universal and natural capacities to work in a socially organized and institutionalized way. Moreover, in science, we typically tackle just the problems that we have not solved in childhood, particularly problems where evidence is scarce and difficult to come by. Many of the social institutions and technological inventions of science are adapted to the special demands of these problems. Nevertheless, these institutions and inventions would be useless without the foundation of a universal set of cognitive capacities, capacities to form, test and revise theories. These capacities were designed to give us a generally veridical view of the world, just as, for example, perceptual capacities give us a generally veridical view of the world. The theory theory then is both a theory of cognitive development and a scientific, naturalistic, psychological, account of scientific knowledge (see Gopnik, 1996, Kitcher, 1988; see also some essays in Giere, 1992).

We can translate the idea of the theory theory into the Chomskyan language of rules and representations, in a way that is consistent with Chomsky's cognitive naturalism. A theory is a particular kind of system that assigns representations to inputs, in the way that the perceptual system assigns representations to visual input or the syntactic system assigns representations to phonological input. We can capture the distinctive structural features of these representations by talking about the specific abstract, coherent, causal entities and laws of the theory. The representations are operated on by rules that lead to new representations, for example, the theory generates predictions.

In a theory, however, unlike a Chomskyan grammar, the very patterns of representation that occur can alter the nature of the representational system itself. This is what makes theory-formation a kind of learning. The patterns of representation that occur can alter the nature of the relations between inputs and representations. As we get new inputs, and so new representations, the very rules that connect inputs and representations change. Eventually, we may end up with a system with a completely new set of representations and a completely different set of relations between inputs and representations than the system we started out with. This differentiates theories from other kinds of representational systems.

To borrow Neurath's philosophical metaphor, the theory-formation view sees knowledge as a boat that we perpetually rebuild as we sail in it. At each point in our journey there may be only a limited and constrained set of alterations we can make to the boat to keep it seaworthy. In the end, however, we may end up with not a single plank or rivet from the original structure, and the process may go on indefinitely.

This kind of system may sound so open-ended as to be uninteresting. But, in fact, the theory formation view proposes that the representational system will change in relatively orderly, predictable, and constrained ways. Moreover, there may still be some overall constraints on the kinds of representations that are generated, not every logically possible theory will be formulated or tested by human beings. These constraints reflect the basic presuppositions of scientific inquiry, for example, that the world has a causal structure that can be discovered. However, the range of theories that can be generated is very much wider than the limited set of representations that are possible on innateness views. It includes, for example, the theories of relativistic physics.

The theory theory vs. other empiricist alternatives

I want to argue, then, in favor of the theory theory as an empiricist alternative to the innateness hypothesis. However, my argument is different in important ways from earlier empiricist arguments. One such argument is that unless we can specify just which learning mechanisms we are talking about, and we know for sure that there aren’t any other possible learning mechanisms, we can't be certain that some capacity is innate. The poverty of the stimulus argument doesn't go through. There may be general-purpose learning mechanisms that could account for the development of a particular type of knowledge. Putnam made this sort of argument many years ago when Chomsky first formulated the innateness hypothesis. Similarly, Piagetians argue that vague, possible, general-purpose processes of "assimilation" and "accomodation" can account for cognitive development. Chomsky is right in thinking that these are weak arguments. It seems unsatisfactory to prefer an unspecified and vague general-purpose learning mechanism, a mechanism that may or may not exist, to a specific proposal for an innate structure.

Of course, there have also been more specific proposals about the kinds of learning mechanisms that could be opposed to innateness. These arguments typically have the following form. A particular mechanism explains some particular type of learning. In this type of learning, inputs are transformed into representations that are at least somewhat more abstract, complex and highly structured than the inputs themselves. The actual representations such mechanisms generate are not nearly as abstract, complex and highly structured as the representations we ultimately want to explain, either the representations of language or of everyday thought. But, the argument goes, if we just scale the system up, make the input richer, add more connections, we will get there eventually. Trust us.

This was, of course, precisely the argument of behaviorism. Operant reinforcement could get a pigeon from a series of specific experiences of lights and shocks to a representation that would allow it to perform a complex behavior (though, of course, behaviorists didn't like to talk about representations). Therefore, if the reinforcers and the stimuli and responses were complex enough, operant conditioning could get us to language and everyday cognition. It was also the argument of associationism. We could get from two specific simple experiences to a more complex representation (even on Hume's view, for example, a causal representation) by associating the simple experiences. Therefore, we could get to all representations if we made the system sufficiently complex.

A corollary of this sort of argument, of course, was that, in fact, the representations were not actually as abstract, complex or highly structured as they appeared to be on the surface. If you looked more carefully you would see that they really were far more concrete, specific, context dependent and generally similar to the input. The two strategies go hand in hand as ways of bridging the gap between input and representation.

The most recent version of this argument invokes connectionist learning systems. These systems generate often surprisingly abstract and complex representations from repeated patterns of input. The learning mechanisms these systems employ are more powerful than the mechanisms of simple association or conditioning. Nevertheless, the representations these systems generate still look very different from the kinds of representations we want to explain. In particular, they lack the componential and inferential character of the representations of ordinary language and thought. Even the strongest exponents of these views still have to resort to a scaling up argument (see e.g. Elman et al, 1996; Bates & Elman, 1992).

The theory formation alternative, in contrast, presents a very different kind of argument. If anything it's actually a kind of scaling down argument. No one would deny that scientific theories are very abstract, complex, highly structured kinds of knowledge, indeed, if anything, they seem more abstract, complex and highly structured than everyday knowledge. Nor would anyone deny that theory change is based on evidence, and so, ultimately, on sensory input. And very few people, no-one except perhaps Jerry Fodor, would accept that scientific theories are innately determined and are simply triggered by evidence. Scientific theory-formation is a kind of demonstration proof that there are learning mechanisms in the universe that are powerful enough to generate the kinds of representations we want from the kinds of input we know we have. More than that, these learning mechanisms can be and have been instantiated in human brains right here on earth.

Theory-formation isn't a vague general-purpose learning mechanism that may or may not exist. Nor is it a specific learning mechanism that may or may not be able to generate sufficiently abstract, complex, highly structured representations with enough scaling up. It's a mechanism that does exist and that does generate the right kind of knowledge from the right kinds of input.

Innate theories and starting-state nativism

There is another aspect of the theory theory that makes it different from other empiricist alternatives. The classical empiricist views, as well as contemporary connectionist views, assume that all the representational structure comes from the nature of the input and the learning mechanisms themselves. Representational structure can be reduced to input. In contrast, scientific theory change never seems to involve raw, uninterpreted data, Instead, theory change is possible because theories themselves build on, revise or replace earlier theories. The earlier theories enabled us to select and interpret the evidence that eventually led to the new theories. In fact, as early as the

1950's, in Carnap's late work, for example, this view of science as a succession of changing conceptions of the world replaced the earlier logical positivist attempt to reduce science to "sense data".

In accounts of science, the origin of the first theory in this succession of changing theories is obscure. But if we think about theory-formation as a mechanism of cognitive development the origin of the first theories is quite simple. They are the theories we are, literally, born with. We learn by modifying, revising and eventually replacing those earlier theories with later ones. The empiricism of the theory theory is an empiricism of revision, not reduction. The idea that we are born with substantive theories of particular domains is also consistent with the empirical evidence of the last thirty years of infancy research. This idea, however, is not part of most empiricist theories in psychology. Notably, Piaget himself, was only willing to grant infants reflex actions and accomodation and assimiliation mechanisms.

The reader may be reeling slightly at this point. I've been arguing that theory formation is an empiricist alternative to innateness and here I am attributing rich innate knowledge to newborn babies. There is, however, a very important difference between this kind of innateness, and the Chomskyan innateness hypothesis. On my view, which I have called "starting-state" nativism, these first theories are subject to radical revision and are indeed radically revised in the light of further evidence (Gopnik & Meltzoff, 1997). By the time we reach adulthood the theories may look almost nothing like the initial theories we had at birth, and the further revision of theories in organized science may lead to still further radical changes. These changes are not arbitrary, though, they are inferred from the evidence of the experience we accumulate in the course of development. This is very different from Chomsky's innateness hypothesis, and from classical rationalism. On that view, innateness is a claim about the constraints on the final state of the system, not about the state the system starts out with.

Again, though, while other philosophers might object to the idea of innate knowledge in principle, Chomsky obviously should not. In fact, Chomsky has made quite convincing arguments in favor of the idea that innate knowledge is possible, though his own proposals about the character of such knowledge are unlike mine.

The theory theory, then, like other empiricist theories, sees the development of knowledge as a kind of learning. The representations and rules of new theories are inferred from evidence and they change, often quite radically, as a result of our experience of the outside world. These changes lead our theories to become increasing veridical. But the theory theory doesn't face the same gap between input and knowledge as classical empiricist theories. By granting that there are innate theories, representations that are already abstract, complex and highly structured, the role of input becomes quite different. Input transforms existing abstract, complex and highly structured representations into new abstract, complex and highly structured representations, and this process can go on indefinitely. Most importantly, we have independent evidence that human beings can engage in this kind of learning, this is the kind of learning that goes on in science.

It has to be said right off that we still don't know, in detail, how these theory-formation mechanisms work, either in science or in cognitive development. We do, though, have a general outline, which is certainly more specific than the vague general-purpose mechanisms that Putnam or Piaget invoked, and at least as specific as most nativist accounts. And, in fact, quite recently, we have begun to formulate some more detailed computational accounts of how at least some aspects of theory-formation might take place.

One central problem in theory-formation, in both children and scientists, is the problem of how we infer causal structure from patterns of evidence. Over the last fifteen years or so philosophers of science and computer scientists have begun to formulate detailed, normative, computational accounts of how this can be done. (Glymour, 2001; Pearl, 2000; Spirtes et al. 2000). These accounts propose that causal structure can be represented by acyclic directed graphs, more commonly called “Bayes-Nets”. They have also demonstrated how such graphs can be accurately derived from information about the conditional probabilities of types of events. In our empirical work, we have begun to find that children as young as two-years-old use some of the same mechanisms that are proposed in these formalisms to accurately infer new causal facts (Gopnik & Glymour, in press; Gopnik et al. 2000).

Moreover, while there is a good argument for rejecting vague alternative possibilities, there is not an equally good argument for rejecting vague alternative certainties. Its not that we don’t know what the learning mechanism is like or even if it exists, we don’t know what it is like (in detail) but we do know that it exists, at least in scientific learning.

Of course, we might want to argue that our current inability to specify the nature of the learning mechanisms in science in detail is, in fact, a proof that they don’t exist. Perhaps it’s just an illusion that scientific theories provide us with veridical representations of the world based on evidence. Perhaps they are just arbitrary social constructions imposed by more powerful scientists on less powerful scientists. Just such a claim underpins the postmodern arguments against scientific realism. If the postmodernists were right, the theory theory would be doomed. But Chomsky can't accept that argument and still be a scientific realist, and realism underpins his whole enterprise (and mine and that of every other scientist who is not accepting money under false pretences). Chomsky's whole project of cognitive naturalism depends on the idea that science allows us to learn about the world by looking at evidence.

An alternative position that Chomsky might take, and sometimes indeed does seem to take, is that while science is veridical it is also intrinsically incomprehensible. Chomsky is usually both a realist about science and a naturalist about cognition. He believes that knowledge can be understood using the techniques of empirical science. But he sometimes seems to think that this is only true of some kinds of knowledge. He certainly believes that we can understand syntactic knowledge empirically, but there are times when he seems to say that we can't understand such things as scientific knowledge (or ethical knowledge etc.). This view is consistent with the innateness hypothesis. If the innateness hypothesis is generally correct, then some things are simply incomprehensible, just contingently beyond the limits of our innate capacities, and understanding how we do science might be one of them.

There seems very little good reason for this apriori pessimism, particularly since there has been hardly any empirical research on the question. In an age in which we humans have developed theories of quantum mechanics and relativistic physics, arguments about our innate inability to understand certain things are bound to be unconvincing. But even if this pessimism were accurate, the poverty of the stimulus argument would still be undermined in the same way. Given a choice between a specific proposal and an unspecified possibility we might do well to at least tentatively adopt the specific proposal. But given a choice between two mechanisms that we have good reason to think exist, we can't adopt one because there is some contingent reason that we find it difficult, even impossible, to understand the other in detail. For a realist, this would be like arguing that because we can't know about events outside the light cone, we should believe that everything happens inside the light cone. It would be an unhappy outcome for psychologists if we had to accept that we could know nothing about scientific learning, but even that would not in itself be a reason for denying that scientific learning exists or accepting the innateness hypothesis.

Phenomenological and Social Objections

So far, I've been arguing that theory-formation is at least a possibility as an answer to the poverty of the stimulus argument, and that it is not subject to the same objections as other empiricist alternatives. Indeed, Chomsky ought to accept that it is a possibility given his cognitive naturalism and scientific realism. But, of course, merely saying that it is a possible answer doesn't make it a plausible answer, let alone the correct answer. The best argument that it is a plausible and correct answer is the detailed empirical developmental research I referred to earlier. Still, to many philosophers the very idea that children could be employing the same learning mechanisms as scientists generates a reaction of shocked, even indignant, incredulity. (For some reason, I've found this initial reaction to be stronger in philosophers than in psychologists or, especially, practicing scientists, who seem to find the idea appealing and even complimentary. Its odd since you'd have thought philosophers, in particular, would, like the Red Queen, have had extensive practice in believing at least six impossible things before breakfast.) There might be specific objections to the theory theory, even if it escapes the usual objections to empiricism.

Some of these objections, though, at least shouldn't be Chomsky's objections. In fact, the objections that philosophers raise to the theory theory are often just the same objections that were raised against Chomsky's cognitive naturalism, and Chomsky's counter-arguments to those objections apply just as effectively. In fact, the arguments of theory theory advocates sound very much like Chomsky in his cognitive naturalist mode. (This is hardly surprising since most advocates of the view, including me, come from a strongly Chomskyan intellectual tradition, indeed, sociologically and historically, you might think of the theory theory as a new species resulting from the cross-fertilization of that tradition and the Piagetian tradition in cognitive development).

Two of the principal objections to the theory theory are, in fact, reminiscent of similar objections to Chomsky's account of linguistic knowledge as rules, and I would give replies that are very similar to his. Scientific theory formation, like the canonical kind of rule, (say a traffic rule) is supposed to be typically conscious and even self-conscious, scientists consciously and reflectively consider how to gather evidence and how evidence tells on their specific hypotheses. The kinds of theory formation we see in children, the kind that lead to everyday knowledge do not, on the face of it, seem to be consciously accessible in the same way. In particular, children may not consciously assess evidence and consider its impact on theories. In the same way, of course, linguistic rules are not consciously accessible in the same way as traffic rules.

But, as Chomsky himself argues, it is difficult to see why conscious phenomenology of a particular kind would play an essential role in finding things out about the world. A characteristic lesson of the cognitive revolution in general and Chomskyan cognitive naturalism, in particular, is that human beings (or, for that matter, machines) can perform extremely complex feats of representation without any phenomenology at all. It is rather characteristic of human cognition that it is largely inaccessible to conscious reflection. Why should this be different in the case of scientific knowledge?

In fact, one might be rather skeptical about how much real scientists, as opposed to the idealized scientists of philosophy, have conscious access to their theory-formation mechanisms. But even if scientists are, in fact, sometimes reflective about their theory-formation processes, it seems, at least, much too strong to say that this is a necessary condition for theory- formation in science. It seems unlikely that it is the reflective phenomenology itself that is what give scientists their theory formation capacities or that give those theories their epistemological force. Rather, it seems that some set of abstract rules and representations must underlie the flow of phenomenal experience in science, as they do in language.

A second objection, again reminiscent of the objections to Chomsky's account of language, is that science is a socially constructed enterprise rather than an individual psychological phenomenon. One might have thought, in fact, that this objection applied more strongly to language. Again, though, Chomsky's own arguments seem quite applicable. Even if science, like language, is part of a rich social context, it nevertheless, ultimately depends on individual cognitive processes in individual human minds. The social character of science at some level must boil down to individual scientists making individual decisions about which theories to formulate and to accept. Moreover, the socially oriented view of philosophy of science has always had a difficult time explaining how science gets it right. It has been difficult to reconcile with scientific realism. Just as it is hard to see how phenomenology, by itself, could lead to veridicality it is hard to see how a particular social structure, by itself, could do so. Again, pursuing Chomskyan cognitive naturalism, I would argue that some abstract, individual set of psychological rules and representations underpins the social interactions of science.

So Chomsky, at least, ought to be unpersuaded by objections based on the apparently, conscious, social character of science, in contrast to the largely unconscious, and individual character of cognitive development in children. The same arguments he applies to language can be applied to science. In both cases we can postulate that, underlying the conscious, social phenomena, there is a cognitive system, a set of rules and representations that accomplish certain ends.

Universality, uniformity and learning

Another difference we might point to between children and scientists is that children converge on roughly similar representations at roughly similar times, both in language development but also in their understanding of the everyday world. It might be objected that scientists do not always show this sort of uniform development, and that this weighs against the theory-formation view. This would be a more serious objection from a Chomskyan point of view, given that Chomsky has emphasized the universality and uniformity of language development as evidence for innateness.

The theory theory proposes that there are powerful cognitive processes that revise existing theories in response to evidence. If cognitive agents began with the same initial theory, tried to solve the same problems, and were presented with similar patterns of evidence over the same period of time they should, precisely, converge on the same theories at about the same time. These assumptions are very likely to be true for children developing ordinary everyday knowledge. Children will certainly start with the same initial theory and the same theory-formation capacities. Moreover, the evidence is ubiquitous and is likely to be very similar for all children.

Notice, however, that, for scientists, these basic assumptions are not usually going to be true. In science, the relevant evidence, far from being ubiquitous, is rare and difficult to come by, and often must be taken on trust from others. The social mechanisms of deference, authority and trust, are, like all social mechanisms, highly variable. Moreover, different scientists also often begin with different theories, and quite typically approach different problems.

In fact, when the assumption of common initial theories and common patterns of evidence, presented in the same sequence, does hold, scientists, like children, do converge on a common account of the world. Indeed, even the timing of scientific discoveries is often strikingly similar, given independent labs working on the same problem with a similar initial theory and similar access to evidence (hence all those shared Nobel prizes). This convergence to the truth itself is the best reason for thinking that some general cognitive structures are at work in scientific theory change. Scientists working independently converge on similar accounts at similar times, not because evolutionary theory or the calculus or the structure of DNA (to take some famous examples) are innate, but because similar minds approaching similar problems are presented with similar patterns of evidence. The theory theory proposes that the cognitive processes that lead to this convergence in science are also operating in children.

This example may be generalized to make another point. The uniformity and universality of development of a cognitive ability does not in itself speak to whether it is a result of an innate device or of learning. By definition, a learning mechanism with the same initial state given uniform evidence will show as much convergence, universality and uniformity as an innate device.

In fact, we have good reason to believe that this point is not lost on nature. Apparently, evolution doesn't much care whether convergence is the result of an innate device or of learning mechanisms, provided that the results are good, good either because they are functional or veridical. Closely related species of animals may use very different mixes of innate structure and learning to acquire very similar abilities. To take a very simple example, there is a great deal of variation among different songbirds in the acquisition of song. Some birds appear to have particular songs innately programmed, others rely on a tutor in a critical period, others don't have a critical period, for some the tutor must be of the same species, others will learn a "foreign" song if the tutor interacts with them in the correct way and so on.

The difference between a learning account and an innateness account is not that an innate mechanism can lead to uniform outcomes and a learning mechanism can't. It is that a learning mechanism can lead to either uniform outcomes when the evidence is uniform, or varied outcomes when the evidence is varied. The innate mechanism should always lead to uniform outcomes. In the case of the theory theory, the learning mechanism leads to generally uniform outcomes in children, where the evidence is uniform, and progressively less uniform outcomes in adults and in organized science where the evidence is much less uniform.

This is once again part of why the poverty of the stimulus argument is so central for an innateness argument. If we could point to a sufficiently powerful learning mechanism, like scientific theory formation, and also show that the input to that mechanism was relatively uniform in many cases, we could predict that in those cases the development would be as uniform and universal. The argument for innateness only works if we think there are no such mechanisms.

Theory-formation and language.

Chomsky's innateness hypothesis is intended to apply quite widely, and it has been applied quite widely. Nevertheless, Chomsky himself clearly applies it centrally to language and, particularly, to syntax. In contrast, the theory theory has primarily been applied to our everyday knowledge of the world, our everyday understanding of biology, physics, psychology and so forth. How does the theory theory relate to language acquisition?

First, the theory theory clearly can be, and has been, applied to understanding semantic and lexical capacities. In particular, there are a number of "theory theory" accounts of our understanding of many types of words. For instance some accounts suggest that many nouns encode "natural kinds" which are further defined by particular theories (Carey, 1982; Keil, 1989; Gelman & Coley, 1991, Murphy & Medin, 1985). We have shown that the acquisition of early words is, empirically, closely related to the child's developing theoretical understanding of actions, appearances and kinds (Gopnik, 1984;Gopnik & Meltzoff, 1987, 1986). Bartsch & Wellman, (1995) have shown that children's acquisition and use of mental state verbs tracks their development of successive theories of mind. Children seem to understand the meaning of the words they hear in terms of the theories they have, they treat the words of natural language the way that scientists treat theoretical terms. Moreover, rather than reflecting some fixed set of semantic primitives, children's understanding of words changes in parallel with their changing theoretical understanding of the world. Finally, language itself seems to play an important role in theory-formation. We have also shown empirically that the words children hear influence the development of their theories (Gopnik, Choi, and Baumberger, 1996).

Since semantics, by definition, relates linguistic expressions to our understanding of the world, and I have argued that our everyday understanding of the world is theory-like, this is not surprising. Moreover, in so far as semantics provides a foundation for syntax, theory formation also may play a role in syntactic development. We seem to use theory formation to develop an understanding of the meaning of words and sentences, and, as many people have argued before, that understanding might itself play an important role in developing more strictly syntactic abilities.

No one so far has applied theory formation ideas to the development of syntax itself. It is possible that syntax is simply not the sort of thing you can have a theory about. In particular, theory formation depends on a kind of realism, it depends on the idea that there is something out there that you can have a theory about, and that you can get a continuous stream of information from the world about that something out there. Interestingly syntactic development is not like this. There is nothing out there that syntactic representations are representations of, knowing a syntactic structure and having a syntactic structure are just the same thing. Unlike other kinds of knowledge, adult syntactic knowledge isn't defeasible even in principle. The phenomenon of creolization which is often invoked in support of innateness is a good example of this (Bickerton, 1981). In creoles children don't develop an incorrect account of the pidgin languagewhich is subject to correction, they simply develop a new language.

But even if theory-formation can't explain syntax for this reason, this outcome should not be an altogether happy one for Chomskyan theory. For the same arguments that say that syntactic knowledge is not a kind of theory, call into question whether it is really a kind of knowledge either. Chomsky rightly argues that philosophical attempts to define knowledge in terms of necessary and sufficient conditions are otiose, he thinks that knowledge is whatever a science of knowledge says it is. But if we take the view of cognitive naturalism we can at least say that the cognitive natural kinds that emerge from the recent research, the varieties of knowledge, look very different from the Chomskyan picture.

There is a kind of knowledge that looks very much like the classical philosophical view of knowledge, a kind of knowledge that involves defeasible representations of an independent reality. The theory theory provides a naturalized account of such knowledge. There is also a kind of "knowledge" that is very different. It may be that we use representations and rules to guide our actions and, in particular, to coordinate our actions with the action of others, but not to get to a veridical account of the outside world. In this respect syntactic abilities look much more like certain kinds of social abilities, or perhaps motor abilities or musical abilities, than like scientific knowledge or our everyday understanding of the world. This doesn't mean that we can't reach a naturalistic understanding of these abilities, or that they don't involve rules and representations, or that we can't call them knowledge if we want to. It does mean that they are very unlike other kinds of knowledge.

The innateness hypothesis may apply to these kinds of knowledge. But there also might be learning mechanisms that lead to these representations, though they might be quite different from theory formation. The answer to the poverty of the stimulus argument here would have to be a rather different one. It would depend on the richness of our everyday capacities to develop certain kinds of actions in coordination with the actions of others. In fact, those who propose alternatives to innateness for syntax depend on just such arguments (e.g. Slobin, 1985). If researchers could point to specific learning mechanisms that allowed us to coordinate our actions with those of others, given certain types of interactions, they could oppose those mechanisms to the innateness hypothesis, just as I have opposed theory formation to innateness (for an example of such an argument for a learning mechanism for phonological knowledge see Kuhl, 1994).

In any case, cognitive naturalism and the innateness hypothesis were exciting because they were supposed to explain language and knowledge in general, not just syntax. Even if Chomsky is right that syntactic structure is innate, syntax comes to look like a relatively isolated phenomenon using very different mechanisms than the rest of cognition, and even of semantic and lexical knowledge. In fact, sociologically speaking, at least, something like this seems to have come to pass. Where twenty or thirty years ago every cognitive psychologist knew the latest theories of syntax and employed them as models for cognition in general, now hardly any psychologists know even the names of the theories.

If we take the innateness hypothesis as a broad claim about the nature of language and mind then it is clearly seriously challenged by the theory theory. If we take it as a narrow hypothesis about a syntactic capacity that is quite different from other linguistic and cognitive capacities then it may survive, but it seems a rather pyrrhic victory.

On the other hand, if we think that Chomsky's real contribution was his cognitive naturalism then the influence of Chomsky's ideas is both broad and important. They have allowed us to make real progress in solving the problem of how we come to understand the world around us. Cognitive naturalism has helped us to understand the myriad devices we use to get to the truth about the world around us, from human vision to science itself. It has helped us to begin to solve the ancient problem of knowledge.

References

Atran, S. (1990). Cognitive foundations of natural history: towards an anthropology of science. New York: Cambridge University Press.

Barkow, J. H., Cosmides, L., & Tooby, J. (1992). The adapted mind: Evolutionary psychology and the generation of culture. New York, NY: Oxford University Press.

Bartsch, K., & Wellman, H. M. (1995). Children talk about the mind. New York, NY: Oxford University Press.

Bates, E., & Elman, J. L. (1992). Connectionism and the study of change. In M. H. Johnson (Ed), Brain development and cognition: A reader . Oxford: Blackwell.

Bickerton, D. (1981). Roots of language. Ann Arbor: Karoma Publishers.

Carey, S. (1982). Semantic development: The state of the art. In E. Wanner, & L. R. Gleitman (Eds.), Language acquisition: the state of the art . New York: Cambridge University Press.

Carey, S. (1985). Conceptual change in childhood. Cambridge, Mass: MIT Press.

Eimas, P. D. (1975). Speech perception in early infancy. In L. B. Cohen & P. Salapatek (Eds.), Infant perception: Vol. 2. From sensation to cognition (pp. 193-231). New York: Academic Press.

Gelman, S. A., & Coley, J. D. (1991). Language and categorization: The acquisition of natural kind terms. In S. A. Gelman, & J. P. Byrnes (Eds), Perspectives on language and thought: Interrelations in development (pp. 146-196. xii, 524). Cambridge, UK: Cambridge University Press.

Gelman, S. A., & Wellman, H. M. (1991). Insides and essence: Early understandings of the non-obvious. Cognition, 38(3), 213-244.

Giere, R. N. (1992). Cognitive models of science (Minnesota studies in the philosophy of science ; v. 15 . Minneapolis: University of Minnesota Press.

Glymour, C. (in press). The mind's arrows: Bayes nets and graphical causal models in psychology. Cambridge, MA: MIT Press.

Gopnik, A. (1984). Conceptual and semantic change in scientists and children: Why there are not semantic universals. Linguistics, 20, 163-179.

Gopnik, A. (1988). Conceptual and semantic development as theory change. Mind and Language, 3, 163-179.

Gopnik, A., & Meltzoff, A. N. (1997). Words, thoughts and theories. Cambridge, MA: Bradford, MIT Press.

Gopnik, A. (1996). The child as scientist. Philosophy of Science, 63(4), 485-514.

Gopnik, A. (1993). How we know our minds: The illusion of first-person knowledge of intentionality. Behavioral & Brain Sciences, 16(1), 29-113.

Gopnik, A., Choi, S., & Baumberger, T. (1996). Cross-linguistic differences in early semantic and cognitive development. Cognitive Development, 11(2), 197-227.

Gopnik, A., & Meltzoff, A. (1987). The development of categorization in the second year and its relation to other cognitive and linguistic developments. Child Development, 58(6), 1523-1531.

Gopnik, A., & Meltzoff, A. N. (1986). Relations between semantic and cognitive

development in the one-word stage: The specificity hypothesis. Child Development, 57(4), 1040-1053.
Gopnik, A., & Wellman, H. M. (1994). The theory theory. In L. Hirschfield, & S. Gelman (Eds), Mapping the mind: Domain specificity in cognition and culture (pp. 257-293. xiv, 516). New York: Cambridge University Press.
Gopnik, A.; D. Sobel, L. Schultz, & C. Glymour (2001). Causal learning mechanisms in very young children: Two, three, and four-year-olds infer causal relations from patterns of variation and covariation. Developmental Psychology, 37, 5, 620–629

Gopnik, A.; & C. Glymour (in press). Causal maps and Bayes nets: A cognitive and computational account of theory-formation. In P. Carruthers, S. Stich, M. Siegal,(Eds.) The cognitive basis of science. Cambridge: Cambridge University Press.

Keil, F. C. (1989). Concepts, kinds, and cognitive development (The MIT Press series in learning, development, and conceptual change . Cambridge, MA: MIT Press.
Kitcher, P. (1988). The child as parent of the scientist. Mind and Language, 3, 217-228.2.

Kuhl, PK. (1994). Learning and representation in speech and language. Current Opinion In Neurobiology, 4(6), 812-22.
Leslie, A., & Roth, D. (1993). What autism teaches us about metarepresentation. In S. Baron-Cohen, H. Tager-Flusberg, & D. Cohen (Eds), Understanding other minds: Perspectives from autism (pp. 83-111. xiii, 515). Oxford: Oxford University Press.
Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. Psychological Review, 92(3), 289-316.
Pearl, J. (2000). Causality. New York: Oxford University Press.
Perner, J. (1991). Understanding the representational mind (Learning, development, and conceptual change . Cambridge, MA: MIT Press.
S. Pinker (1989). Learnability and cognition: The acquisition of argument structure. Learning, development, and conceptual change. Cambridge, MA: MIT Press.
Slobin, D. I. (1985). The cross-linguistic study of the language-making capacity. In D. I. Slobin (Ed), The cross-linguistic study of language acquisition (pp. 1157-1256). Hillsdale, NJ: Erlbaum.
Smith, C., Carey, S., & Wiser, M. (1985). On differentiation: A case study of the development of the concepts of size, weight and density. Cognition, 21, 177-237.
Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. Psychological Review, 99(4), 605-632.
Spirtes, P., Glymour, C., & Scheines, R. (2001). Causation, prediction, and search (Springer Lecture Notes in Statistics, 2nd edition, revised). Cambridge, MA: MIT Press.
Wellman, H. (1990). The child's theory of mind. Cambridge, Mass.: MIT Press.
Wellman, H., & Gelman, S. (1997). Knowledge acquisition in foundational domains. D. Kuhn, & R. Siegler (Ed.), Handbook of child psychology (5th ed.,). New York: Wiley.