Science as Argument: Implications for Teaching and Learning Scientific Thinking

DEANNA KUHN
Teachers College, Columbia University

INTRODUCTION

American science education has faltered, from just about everyone's perspective. American students rank near or at the bottom in international surveys of science achievement. Science teachers at the intermediate and high school levels report that their students are already "turned off" to science by the time they reach their classrooms, and these teachers lack conviction about their own competency in science (Easley, 1990). Teachers who even undertake science education with younger children seem to have only the most modest of goals—to keep alive the "natural" curiosity that children display.

Despite the increasing concern accorded to this state of affairs, there exists no firm sense of exactly what we would like students to acquire from beginning science education. The increasing technical complexity, specialization, and rapid evolution of knowledge in science make the mastery of any particular body of scientific knowledge an unwieldy and unsatisfying educational goal. More promising is the concept of science education as promoting a way of thinking. If the development of scientific thinking is to be a meaningful goal of science education, however, it is essential that we have a clear sense of what it means to think scientifically. Exactly what scientific modes of thought do we wish students to acquire, and how can we justify their value? The present article addresses these questions.

A major argument will be that significant benefits accrue both conceptually and practically if we treat scientific thinking not as a rarefied form of thought but instead bring it into the realm of the ordinary. Doing so does not imply that the differences between scientific and informal thought are inconsequential, nor does it imply that scientific thinking is effortless and "comes naturally," with little involvement on the part of educators. To the contrary, the conception of scientific thinking advanced here provides the educator with a challenging but clear vision of the goals of elementary science education.

Science Education 77(3): 319–337 (1993)

CCC 0036-8326/93/030319-19

CHARACTERIZING SCIENTIFIC THINKING

Concepts of science and scientific thinking have undergone dramatic change in a relatively short period of time. Not long ago, the positivist view of science as an accumulative body of factual knowledge prevailed. Over the course of just a few decades, the view has become widely accepted that it is impossible to study the evolution of knowledge apart from the cognitive processes of human knowers. On the philosophical side, we now have "naturalistic" epistemologists who allow for the role of empirical observations as an integral part of attempts to understand knowledge (Heyes, 1989). On the psychological side, we have a blossoming psychology of science, with its proponents beginning to explore all of a variety of ways in which psychological study illuminates the attempt to understand the progress of science (Gholson et al., 1989).

Within the fields of developmental and educational psychology, another set of changes has occurred. The global stage theory deriving from the ideas of Piaget has grown into disfavor and been replaced to a large extent with more process-oriented approaches (Kuhn, 1992). Attempts to base curriculum explicitly on the reasoning strategies comprising Piaget's stage of formal operations have largely disappeared, and science educators in particular have become much less optimistic about what the stage approach can offer. Teaching reasoning strategies per se has come to be seen as a narrow, even sterile, approach to science education. In its place, we now see much greater attention to the domain-specific knowledge children have about scientific topics. Research has shown that both children and adults hold a variety of naive, intuitive conceptions about the way the world works, and these naive theories provide a starting point for both studying and fostering children's scientific understanding (Posner et al., 1982; Vosniadou and Brewer, 1987, in press; West and Pines, 1985).

Both of these developments would appear to be positive ones in their implications for science education. Science need no longer be treated as an accumulation of assertions, disembodied from the human thinking that gave rise to them—assertions that a teacher labors to pass on to students, in a way unconnected to their own thinking. Hence, the scientific enterprise seems less distanced from the thinking of children than it has in the past. Scientific thinking is not a disembodied set of procedures imposed on those bold enough to seek entry into the realm of science. Rather, the thinking of professional scientists who advance scientific thought in the culture develops out of the intuitive scientific thinking of children.

To realize the potential that this connection holds, however, we must be able to characterize what doing science is, in a way that makes sense for both the child and the scientist, as well as for the adolescent or lay adult—the intermediaries to whom we can link both the child and the scientist and thereby the two latter to one another. How can scientific activity be characterized in broad enough terms to fit all of these cases?

Sciences as Exploration

One widely used descriptor is science as exploration. One of its virtues is that it appears to fit both scientist and child fairly well. But, in what way can exploration serve to characterize the activity of average adolescents and adults, the crucial intermediaries in our chain? Indeed, the problem is evident before adulthood—even before the end of childhood and certainly by adolescence. Anyone who has experience with both young children and adolescents will recognize how readily the term fits in one case and how difficult it is to apply in the other. We can literally observe young children busily engaged in exploring the world around them—finding out how things work and constructing experiments to test their limits. One is hard pressed, in contrast, to identify anything equivalent in an adolescent. What has happened to the exploration that was so easy to see during the childhood years? Do children in fact lose their scientific natures, to be reconstructed again only among the scant few who will embark on scientific careers? The dismal picture with respect to science achievement in our schools is consistent with this view, and indeed we can hear it expressed explicitly by those concerned with educating our youth in science:

> Children are born scientists. From the first ball they send flying to the ant they watch carry a crumb, children use science's tools—enthusiasm, hypothesis, tests, conclusions—to uncover the world's mysteries. But somehow students seem to lose what once came naturally. (Parvanno, 1990)

I argue that if we look at scientific thinking more deeply, this view of it is wrong in two respects: No, it does not come naturally, but, then, once you get it you do not lose it. Viewed in this way, scientific thinking is an endpoint, not a starting point, of a complex process of intellectual development.

Science as Argument

To understand scientific thinking in this way, we need an alternative, or at the least a supplement, to science as exploration. The alternative I propose here is science as argument. The objective, recall, is to link the thinking activity of scientists to that of ordinary children, adolescents, and adults, and so it is necessary to argue for the aptness of this characterization for both groups.

A briefer case can be made regarding the professional scientist because others have laid this ground well. Science is a social activity. It advances through thought processes that occur between persons, not just within them (Westrum, 1989). Those seeking to understand the evolution of scientific thought tend to have focused on the insights achieved by the lone scientist, to the exclusion of the social exchange that is the arena in which these ideas are articulated, questioned, clarified, defended, elaborated, and indeed often arise in the first place. From the positivist conception of science as absolute and accumulative, we have come to recognize that there will be no scientific method capable of detaching science from controversy, from argument. Not just the theories but even the so-called "facts" of science become argumentive constructions that must be entered into the arena of public debate.

The legal model of advocacy may be an apt one. Most often, scientific questions are posed by means of two, or sometimes three or four, competing theories. The process is one of debate, with individuals typically playing advocacy roles. To participate, an individual scientist must analyze the evidence and its bearing on the different theories as a means of argument to the scientific community in support of his or her view. Equally important, this analyzing and weighing process of argument is, in interiorized form, almost certainly an important part of what goes on in the private thought of the individual scientist. Scientists are well aware that explicitly justified arguments are needed to convince the scientific community, and they become accustomed to thinking in such terms.

Where do we find anything like these same activities in ordinary life? In the arguments people have with one another, certainly. But, as just claimed is the case for the scientist, these arguments also take an internal form. The idea that "children's thinking tends to replicate the procedural logic of the social communications in which they participate," as Damon (1990) puts it, has been put to great advantage in understanding young children's thinking, as well as their social relations. The same correspondence can be probed in the case of the more complex thinking achieved by adolescents and adults—whether we regard it in the Vygotskian framework of an interiorization from social to mental planes (Rogoff, 1990) or more in the Piagetian framework of a correspondence between the two planes (Damon, 1990).

My claim, then, will be that we can find scientific thinking in older children, adolescents, and lay adults if we conceive of it as argument. It will be useful first to examine such thought as it occurs in informal reasoning. In the next section, I describe research that explores informal or everyday reasoning in a framework of argument. I then go on to make explicit links between argumentive thinking and scientific thinking as it is traditionally conceived. Finally, the educational implications of science as argument are examined.

STUDY OF INFORMAL REASONING AS ARGUMENT

With only a few exceptions, psychologists have approached thinking as problem solving, rather than argument, which has tended to be regarded as the province of philosophers. Yet, it is in argument that we may find the most significant way in which thinking and reasoning figure in the lives of average people. Thinking as argument is implicated in all of the beliefs people hold, the judgments they make, and the conclusions they draw.

In my research on argument, like Billig (1987), I have drawn on the connection between social and individual (or rhetorical) argument. The dictionary definition of an argument in the latter sense is "a course of reasoning aimed at demonstrating the truth or falsehood of something." More commonly, an argument is regarded in its social sense as a dialogue between two (or more) people who hold opposing views. Each offers justification for his or her own view, and, at least in a skilled argument, each attempts to rebut the other's view by means of counterargument.

Although connections are rarely made between these two kinds of arguments, they in fact bear a close relationship to one another, as suggested in the preceding discussion of science as argument. In a social (dialogic) argument, at a minimum one must recognize an opposition between two assertions—that, on surface appearance at least, both are not correct. One must then connect supporting and refuting evidence to each of the assertions and, if the argument is to move toward resolution, be able to relate and weigh supporting and refuting evidence in an integrative evaluation of the relative merit of the opposing views.

Less often noted is the fact that these same skills are in fact entailed in more implicit form in a rhetorical argument, although the rhetorical argument may on the surface appear less complex cognitively. An argument in support of an assertion is an empty, indeed superfluous, argument unless one can conceive of an alternative to what is being asserted—an opposing assertion. Once two or more contrasting assertions are in place, cognitively speaking, the further challenge poses itself of relating evidence to them. Presumably, it is a weighing of positive and negative evidence that has led one to espouse the favored assertion over its alternatives. Indeed, it is just such a weighing process that is implicit when we speak of a *reasoned* argument. Thus, any reasoned argument in support of an assertion implicitly contains a full dialogic argument.

This identity between rhetorical and dialogic arguments provides a framework for exploring the nature of the less externally observable rhetorical argument. Are the elements of the dialogic argument evident when we probe the thinking underlying people's beliefs and opinions? And, is the presence or absence of these elements revealing of the quality of people's thinking? To investigate these questions (Kuhn, 1991), we asked people their views on three topics: (1)What causes prisoners to return to a life of crime after they are released? (2)What causes children to fail in school? (3)What causes unemployment? These topics were chosen as ones that people have occasion to think and talk about. The 160 participants were chosen to represent average people across the life span, beginning with adolescents (ninth graders) and including young adults in their twenties, middle adults in their forties, and older adults in their sixties. Within each age group, as well as males and females, we included subjects of two different education levels—in general, those who had high school vs at least some college education (these differences were prospective among the adolescent group). We also included a group of experts of three different types—experienced probation officers, regarded as having domain expertise in the return to crime topic; experienced teachers, regarded as having domain expertise in the school failure topic; and philosophers (specifically, PhD candidates working on their dissertations), who we regarded as having expertise in reasoning itself.

Following the framework of the dialogic argument, we asked subjects first to describe and justify their theories and then probed them regarding alternative theories, counterarguments, and rebuttals. We also presented some evidence of our own related to the topic and asked them to evaluate it. What kind of evidence could we expect lay people with no special knowledge or interest in these topics to offer? In fact, for each topic roughly 40% (averaged across topics) offered what we classified as genuine evidence. What we call genuine evidence is by no means evidence that is conclusive, nor even compelling, nor even necessarily convincing evidence. Rather, it is simply evidence that is: (1) differentiated from the theory, which we will see is an important criterion; and (2) bears on its correctness. About half the responses classified as genuine evidence refer to covariation—variation in the alleged cause corresponds to variation in the outcome, but some other kinds of reasoning familiar to cognitive psychologists appear as well, such as counterfactual reasoning, discounting, and analogy.

The following response (for the return to crime topic), in contrast, was classified in a category we termed pseudoevidence:

> *(How do you know that this is the cause?)* I think because if they commit crime they're getting attention. They'll be—the prisoner, you know, in prison—they'll be taken care of, they'll be given food and all this, and they get attention. They come out and everybody, you know. . .he was a prisoner, so that they stay away from him; they're scared and everything. So they decide that the only way they're going to be, that they'll have attention or they're going to be cared for is if they're in prison.
>
> *(If you were trying to convince someone else that your view is right, what evidence would you give to try to show this?)* The evidence I would give is that when they are in prison, they're secure. They're sure that, you know, no one's going to hurt them. Well, they're not sure no one's going to hurt them, but, you know, they know they're secure. They've got a place to eat, a place to sleep. But if they come out into the world and, you know, with unemployment and people not wanting to take anybody in that was an ex-con or something, when people reject them because of their past, they're sure to go back into the crime just to go back to their jail cell and stay in it.
>
> *(Can you be very specific, and tell me some particular facts you could mention to try to convince the person?)* Well, some facts could be that when they're in there, they'd want to make friends with other cons and stuff like that, and when they're out here, they've got to start all over again, and it's real hard for people who committed a mistake, for other people to accept that they've paid for it and everything. And then when they're out here people reject them and they look at them, you know, like they're scared of them. They don't want to stay in the world if they think everywhere they go people are going to be looking at them and feeling, you know, real insecure when they're around. So they'd rather be where people, you know, they're all the same.

How should we characterize such a response? Does the subject offer evidence for her theory? We classified it as pseudoevidence, defined simply as a scenario, or script, depicting how the phenomenon might occur. The defining characteristic that distinguishes pseudoevidence from genuine evidence is that pseudoevidence cannot be sharply differentiated from the theory itself. Hence, responses to "What causes *X*?" do not differ sharply from responses to "How do you know that this is the cause of *X*?" In this case, the subject makes an intuitively convincing case for the *plausibility* of the cause she specified leading to the outcome, yet without providing any genuine evidence that this cause is in operation in instances of the phenomenon. Her own words, in fact, establish that for her the function of evidence is to establish such plausibility: "The evidence I would give is that when they are in prison they're secure." This "evidence" does not establish that preference for life in prison *is* the cause of the phenomenon; rather, it enhances its plausibility as a possible cause.

At its most minimal, pseudoevidence simply illustrates the causal sequence. At its best, it enhances the plausibility of the causal sequence, as in the example offered here. In establishing plausibility, pseudoevidence scripts bear on causal *mechanism,* and there is much in the causal reasoning literature to suggest mechanism as a perfectly appropriate means of causal explanation (Antaki, 1988; Hilton, 1988). We thus might be tempted to regard pseudoevidence and genuine evidence as alternative explanatory styles. Yet, the difference is a more fundamental one than that of style. In the causal reasoning and attribution literature, subjects are asked to identify the causal factor in a specific past instance, making the criterion of plausibility a relevant one: Could this cause have produced the outcome? In the present context, in contrast, subjects are asked to justify the assertion that *in general* X is the cause of outcome Y. In this context, genuine evidence is superior to pseudoevidence on the grounds that it is more definitive.

What is the basis for this claim? First, plausibility is neither a necessary nor a sufficient condition for the correctness of a causal theory. Often, in the history of science causal theories that initially appear implausible have later been proven correct. Likewise, highly plausible theories have been disconfirmed. Further, a causal relation between two factors can be demonstrated in the absence of any plausible theory connecting them, for example, when a substance is found beneficial in treating a disease in the absence of an understanding of how it achieves its effect.

"Good" pseudoevidence, then, might heighten our interest in testing a causal theory (by enhancing plausibility), but it cannot tell us whether the theory is correct. In fact, because pseudoevidence can never conflict with a theory it cannot really be considered evidence at all. Instead, it should be regarded as part of the theory itself. In proposing their theories, it is reasonable to assume that all of our subjects envisioned some mechanism whereby the alleged cause produces its effect. When, in offering pseudoevidence, they elaborate their description of this mechanism, they are elaborating the theory, not providing evidence that bears on its correctness. Again, even the most plausible theories can be wrong.

A salient question thus becomes: Can subjects who offer only pseudoevidence envision an alternative to this scenario? Can they envision the possibility that it is *not* what happens? Some subjects generate an alternative theory without difficulty. Others generate what appears to be an alternative but then immediately agree with it—"That could be part of it, too"—in effect incorporating the alternative cause into their own theory. Such subjects are unable to conceive of anything that is *not* a cause. Other subjects try unsuccessfully to generate an alternative, producing something like their own theory. But, most interesting are subjects who decline:

> I don't know what they would say. I'd really have to get someone else's point of view. . .my thoughts run in this direction and that's about it.

Or

> I don't know what they might say is the reason. I don't think I'm wrong.

Or, significantly, the hypothetical other's view is simply assimilated to one's own:

> I think they'll say the same thing I'd say. I think that the majority think the way I do.

The percentage of subjects overall who are able to generate alternative theories averaged across topics is about 60%, higher than the 40% who generate genuine evidence, but, importantly, there is a significant association between the two skills, one that makes clearer the meaning of pseudoevidence. In not generating alternatives, those subjects who rely upon pseudoevidence do not call upon this pseudoevidence to perform the function it cannot—to address the correctness of a theory relative to all the others with which (if the subject conceived of the possibility) it could compete. Thus, subjects who generate neither genuine evidence nor alternative theories take their theories for granted, simply as statements about the way the world is. The theories are not reflected on as objects of cognition—as claims needing to be evaluated in the light of alternatives, as well as evidence.

To evaluate a theory against alternatives implies that it could be true or false, that is, indicates an acceptance of its falsifiability. It is the critical issue of falsifiability that we looked at in the study of counterarguments: "What could someone say to show that you were wrong?" Do subjects comprehend the evidence that would falsify their theory were they to encounter it? The success rate here is about 50%. Despite the critical role falsifying cases play in examining a theory, many of our subjects show considerable resistance to the idea. As one of them put it rather plaintively, "If I knew the evidence that I'm wrong, I wouldn't say what I'm saying." Many of those who did attempt a counterargument simply offered an alternative theory as a counterargument (e.g., "They would say it's not the parents, it's the school that causes kids to fail."), leaving the subject's own theory unexamined.

Finally, rebuttals are critical because they complete the structure of argument, integrating argument and counterargument. Only 25% of subjects (averaged across topics) achieve an integrative rebuttal. Other subjects offer only a simple rebuttal, in which they simply provide a counterargument to an alternative theory, again leaving the original theory unexamined. And some simply argue by assertion, for example, "If they said it's the school, I'd say no, it's the parents," leaving both the subject's and the alternative theory unexamined. And some simply decline, like Lois, who says, "I don't think I'd even try. (Why not?) He wants to believe it, that's fine. I'm not argumentative."

The evidence of our own that we presented to subjects took two forms. Underdetermined evidence in effect simply restated the phenomenon in the context of a specific instance, with few clues as to its cause. Overdetermined evidence, in contrast, explicitly referred to three broad families of causes without favoring any of them. Subjects typically assimilated both kinds of evidence to their own theories. "This pretty much goes along with my own view," was the prototypical response. Subjects expressed high certainty regarding their evaluations of this evidence (just as they did about their own theories). If evidence is simply assimilated to a theory, any ability to evaluate its bearing on the theory is, of course, lost. More broadly, with this loss comes loss of the ability to maintain a differentiation between what derives from one's own thought and what derives from external sources and hence control of the interaction of theories and evidence in one's own thinking.

Results across subject groups showed no significant differences by sex, or by age group, but consistent and sizable differences by education group at every age level. The other important result is that we observed by no means total, but significant generality of skills across the three topics. Although many subjects exhibit a skill on some topics and not others, the numbers exhibiting the skill for all topics or no topics are significantly greater than would be expected by chance if performance across topics were independent. This outcome is, of course, critical because it suggests that we have identified forms of thinking that transcend the particular content in terms of which they are expressed. However imperfectly, we are tapping something about the *way* people think. The expertise results support this claim. The philosophers reasoned well overall, as expected, but the domain expertise of the others did not influence reasoning ability. Parole officers reasoned no better about the crime topic than they did about the other topics, nor did teachers reason better about the school topic.

Regarding the relations of such skills to education, it is significant that these associations appear consistently at all age groups despite the fact that the skills involved are not an explicit part of the school curriculum at any age level. Most interesting is the fact that the skill differences appear among teens, when the education differences are only prospective, and we see no further development in skill when we might most expect and hope for it—between the early adolescent and young adult years. Together, these findings suggest that it is some broad, general kinds of experiences associated with education—not all of which takes place inside school—that are responsible for these differences. Within school, it is possible that academic experience encourages the attitude that assertions need to be justified and alternatives considered. But, whatever these benefits are they are conferred early, certainly by the end of junior high school, and we see no further development in this respect. These conclusions are supported by similar findings regarding the relation of argumentive reasoning skills to age and education level reported by Perkins (1985) and by Voss and Means (1991).

LINKING ARGUMENTIVE AND SCIENTIFIC THINKING

What does the research described in the previous section have to do with scientific thinking? How might the challenges that people confront in informal reasoning be connected to those that arise when they are called on to reason in a scientific context? In other research (Kuhn et al., 1988, in press; Schauble, 1990), we explicitly asked lay adults, adolescents, and children to think in scientific ways. As noted earlier, the development of scientific thinking traditionally has been conceptualized as the development of strategies that operate in a more or less domain-general manner (Inhelder and Piaget, 1958). Alternatively, and more recently, it has been conceptualized as domain-specific conceptual change (Keil, 1984; Carey, 1986; Chi, in press). In our work, we focused on the specific theories a subject holds within a content domain without foregoing the search for strategic change in the ways in which the subject brings new evidence to bear on these theories. As in the informal reasoning domain, then, the problem is one of the coordination of theories with new evidence bearing on them.

Like scientists, subjects are asked to investigate a domain and draw conclusions about the causal relations that exist there. In microgenetic studies extending over several months, we begin by assessing the subject's own theories regarding the causal relations present in the domain. Over multiple sessions, subjects are then engaged in generating evidence and using it as the basis for making inferences regarding these effects. For example, in one of the problems we used (Kuhn et al., in press) several factors such as the size of a boat, weight placed inside the boat, and depth of the water may affect the speed with which model boats are towed (by a weight-and-pulley apparatus) down a narrow tank of water. In a parallel version of the problem presented on a microcomputer, variable features of race cars affect the speed they travel along a simulated racetrack. In the weekly sessions following assessment of their own theories, subjects conduct their own investigations of how these factors operate and draw conclusions. We are thus able to observe the process of theory revision, as well as the evidence generation and evidence interpretation strategies subjects employ.

The difficulties that our preadolescent (and often adult) subjects exhibit as scientists indicate that the challenges parallel those that we have identified in the informal reasoning domain. First, the subject must have the ability to reflect on his or her own theory as an object of cognition to an extent sufficient to recognize that it could be wrong. Second, the subject needs to recognize evidence that could disconfirm the theory. If a subject theorizes a causal relation between two factors (e.g., sail size and the boat's speed), to discover that the theory is wrong and the sail size has no effect the subject must conceive of the possibility that the theory could be wrong and then generate and interpret evidence that disconfirms it.

Subjects whose scientific thinking is the most rudimentary transform the task goal into one of producing an outcome, for example, the fastest boat. By middle childhood, however, most subjects have at least a tentative grasp of the concept that one can demonstrate a set of causal relations that exist among features of the boats and outcomes. These relations, however, are more strongly theory driven than evidence driven. The subject does not conceive of the possibility that theorized relations are wrong and does not generate evidence that could disconfirm them.

The classic inferential error subjects consistently make is the one originally labeled by Inhelder and Piaget (1958) as *false inclusion*—the cooccurrence of antecedent and outcome is taken as evidence that the antecedent is causally implicated in the outcome (despite the presence of additional covariates). Hence, the subject who believes the sail size to be causal takes the cooccurrence of the large sail and a fast speed as evidence of the theory's correctness. When asked to justify their inferences, such subjects are likely to mix theory- and evidence-based justifications indifferently—both point to the theory's correctness in the subject's mind and hence one is as good as the other in justifying it. In such reasoning, then, evidence (of cooccurrence) may function as no more than an illustrative instance of what the subject knows to be correct.

In the upper part of Table 1, this form of reasoning is illustrated in parallel in the scientific and informal argumentive reasoning domains (for the car problem in the former case and the school failure problem in the latter). In both cases, evidence (of cooccurrence) serves as an illustrative instance of what the subject knows to be correct. In the science domain, "car 2" is shown in parentheses, as it typically plays only an incidental part in the subject's reasoning.

Table 1
Illustration of False Inclusion and Valid Exclusion Reasoning in Scientific & Argumentive Contexts

	Science	Argument
	Theoretical Belief: Muffler > Speed	Theoretical Belief: Family Problems > School Performance
False inclusion	Car 1— Muffler, large wheels Fast [Car 2— No muffler, small wheels Slow] (What have you found out?) The muffler makes a difference. *(How do you know?)* Because car 1 has a muffler and it goes fast. *(What about the wheels—do they have anything to do with it?)* No, they just go around to make the car move.	*(How do you know that family problems are the cause of children doing poorly in school?)* My neighbor has always had marital problems and she and her husband finally got a divorce and her child has done very poorly in school.
Valid exclusion	Car 1— Muffler, large wheels Fast Car 2— No muffler, large wheels Fast *(What have you found out?)* The muffler makes no difference because here you have it and here you don't and the car *still* goes fast.	*(What evidence would show you were wrong?)* If you looked at families who have problems and found that their children did just as well in school as children whose families don't have problems.

To ever transcend this form of reasoning (and many subjects, in both domains, never do), one must master *exclusion* reasoning, illustrated in the lower part of Table 1 for both domains. Exclusion is essential to effective scientific reasoning because it allows one to eliminate factors from consideration. Exclusion (inferring the absence of a causal effect) poses more of a challenge than inclusion (inferring the presence of a causal relationship) for several reasons. First, and most fundamentally, is the domination of affirmation over negation—the presence of something is more salient than its absence and, for this reason, both scientific and lay theories pertain more often to the presence than the absence of causal relations. Second, the belief that a factor is irrelevant often leads subjects to ignore it in their investigations. In so doing, they forego the possibility of encountering disconfirming evidence and, hence, ever revising this belief.

Third, the ways in which uncontrolled evidence (of antecedent–outcome correspondences) can be used as the basis for inclusion do not work for exclusion. As illustrated in the upper part of Table 1, inclusion is often (although unjustifiably) based upon a single instance (of cooccurrence of antecedent and outcome). Although such reasoning is sometimes observed, it is harder for a subject to draw the conclusion that a causal effect is absent based upon a single instance without at least some implicit comparison. When inferences are based upon multiple instances, a similar asymmetry between inclusion and exclusion arises. After a number of (uncontrolled) multivariable antecedent–outcome instances have been generated, a subject may make an inductive inference, for example, "the big boats go faster than the small ones." Such inferences are sometimes correct, but they are particularly prone to belief bias because subjects are remembering antecedents and outcomes and can do so selectively. In any case, a comparable strategy is not as readily available in the case of exclusion, for example, noting that cars that lack mufflers tend overall to be no faster or slower than those with mufflers. Again, the bias is toward noting the presence, not the absence, of effects. A valid exclusion inference can be made based upon just two instances, as illustrated in the lower part of Table 1, but it requires that other variables be held constant so that they do not contribute their own effects and invalidate the comparison.

Yet another and perhaps the most important reason that exclusion poses a formidable challenge is that it competes with the ever-present and more compelling temptation of false inclusion. The lure of false inclusion is particularly great because these inferences are so often theoretically motivated. The subject who believes that the muffler helps the car go fast, for example, needs only to observe a car having a muffler that goes fast and it becomes difficult to resist the inference that the muffler at least contributed to the car's speed and to attend to the pattern of evidence over multiple instances that would demonstrate the muffler's noncausal role.

Often, in our microgenetic studies the insight is achieved, only to be lost again. Randy, for example, starts with the incorrect belief that the muffler affects speed. After a long period during which he conducts experiments that are not capable of disconfirming his belief, at the sixth session he designs and correctly interprets a valid experiment showing that a car with and a car without a muffler yield the same speed. "No, the muffler doesn't matter," he concludes. "I just had a feeling it might help to push it along." But, he then goes on to comment on a third car, one that achieved the maximum speed and happened to have a muffler. Here, Randy tells us, the muffler might have helped "just a little bit."

In the logbooks that were provided for subjects to record information, in contrast to the professional scientists and college students we observed, who for the most part systematically recorded particular constellations of features and their observed outcomes, in 11-year-old Jamie's logbook we find the assertions, "With big wheels will go slower because it takes more time for the wheels to go around" and "A tailfin would make a difference because it has more weight for the wheels to turn around." Like our subjects in the argument research who generate pseudoevidence in support of their theories, these children confuse evidence with a plausible theory than can assimilate it.

In both the scientific reasoning context described here and the argumentive reasoning context described earlier, what children or adults need to be able to do is to distance themselves from their own beliefs to a sufficient degree to be able to evaluate them, as objects of cognition. In other words, they must have the capacity and the disposition to think about their own thought. This is metacognition in the most fundamental and arguably most significant of the many ways in which this term has been used in recent years. The traditional scientific hypothesis-testing strategies of investigation and inference indeed need to be acquired as important tools, but our observations of children's struggles as amateur scientists have led us to the conclusion that the major challenge they face is not one of acquiring correct experimentation strategies but of developing the ability to coordinate their existing theories with new evidence they generate, in an explicit, conscious, and controlled way—in other words, again, to think about their own thought.

Becoming a competent scientist, accordingly, is more than a matter of acquiring formal hypothesis-testing strategies applicable to any kind of content. Attention to subject's conceptions of the content of what they are reasoning about is largely absent in traditional studies of scientific reasoning and casts much of this research in a new light. For example, the failure to control variables traditionally has been interpreted as as failure to attend to these variables and therefore to recognize the possibility that left uncontrolled they may exert their own effects on outcome. Our observations, however, suggest a different interpretation. Rather than underattend to these variables, subjects more often overattend to them, in ways driven by their theoretical beliefs (Kuhn, 1989).

Conceptualizing an individual's task in scientific reasoning as one of coordinating two entities—or problem spaces, in Klahr and Dunbar's (1988) terms—rather than simply the application of a set of strategies to varying content, makes the task a much richer, more complex one. Coordination implies reciprocal adjustment. Not only must evidence serve as the basis for evaluating and possibly revising theories, but theories influence the direction and form of investigation. This holds true for professional as well as intuitive scientists. The difference between the two lies in the degree of control that is attained of the coordination process (Kuhn, 1989).

As was the case in our research on informal argumentive reasoning, significant cross-domain generality of these skills can be observed. In the study by Kuhn et al. (in press), individual case studies of change patterns revealed corresponding improvements in strategy across the car and boat problems. In current work with adult subjects, a more conventional transfer design is employed, with gains made during the first 5 weeks transferring to new problem content introduced at the sixth week. Thus, while our findings show that the content of subjects' theories cannot be ignored, they also imply that the acquisition of content within knowledge domains does not by itself explain the development of scientific thinking. The generality of change observed across domains makes it less plausible that subjects are simply gaining knowledge of the specific domains in which they are working. The strategic changes we observe have to do with *how* conclusions are drawn, not the particular content of these conclusions. Consequently, approaches that examine the development of scientific thinking only as domain-specific conceptual change (Brewer and Samarapungavan, 1991; Chi, in press) reduce their explanatory power by neglecting an important aspect of what is developing during the childhood and adolescent years in the realm of scientific thought.

What does the research that has been described in this article suggest regarding the attainment of these competencies during the normal course of development? When the evidence regarding children engaged in scientific investigation is considered in conjunction with the evidence regarding adults' informal argumentive reasoning, the picture that emerges is one of a gradual, perhaps life-long, developmental course in which both more and less advanced forms of thought coexist. Because they are inconsistent, the more and less advanced forms compete with one another in a prolonged contest that has different outcomes for different individuals.

If we accept this picture, the microgenetic evolution examined in the scientific reasoning research can be thought of as a microcosm of a much more extended developmental course. Rudimentary forms of theory–evidence coordination are apparent among kindergartners (Sodian et al., 1991) and even younger children. Imagine a 2-year-old who calls her parents into her bedroom one night with the claim that it is a ghost in her closet that is the cause of a soft "whooshing" noise that is keeping her awake. This childs understands as well as her parents that opening the closet door will provide the evidence capable of disconfirming this causal hypothesis (even though she does not understand the logic of disconfirmation in any formal or reflective sense). In the next few years, this 2-year-old will develop the understanding of false belief that has been the object of attention of the new theory-of-mind research (Feldman, 1992; Perner, 1991; Wellman, 1990). Yet, for many years, and most likely well into and even throughout her adult life, she will exhibit difficulty in bringing evidence to bear on her own beliefs in a way that reflects clear differentiation between the implications of evidence and what she believes to be true.

The development of such reasoning is an especially interesting and challenging topic precisely because, unlike, say, conservation acquisition, which is marked by the fairly discrete onset of qualitatively distinct new behavior, the development appears gradual, continuous, and multifaceted, with gains achieved not once but many times over as new content and contexts arise. Equally significant is the fact that such development has a strong metacognitive component. As our research clearly shows, it is not enough to have the competence to execute the more advanced reasoning strategies. One must know when and why they need to be used and, most important, learn how and why not to succumb to the faulty strategies that remain so seductive.

TEACHING AND LEARNING SCIENTIFIC THINKING

The major significance of the parallelism illustrated in Table 1 and explored throughout this article is that it links scientific thinking to thinking more broadly conceived. The idea that there exists a link between scientific and everyday thinking is not a new one, as indicated by the following quote from no less a thinker than Albert Einstein (1954, p. 290):

> The whole of science is nothing more than a refinement of everyday thinking. It is for this reason that the critical thinking of the physicist cannot possibly be restricted to the examination of the concepts of his own specific field. He cannot proceed without considering critically a much more difficult problem, the problem of analyzing the nature of everyday thinking.

The purpose of the present article has been to make explicit some of the connections Einstein envisioned. Doing so has implications with respect to the issues of science education raised at the outset of this article. Scientific thinking tends to be compartmentalized, viewed as relevant and accessible only to the narrow segment of the population who pursue scientific careers. If science education is to be successful, it is essential to counter this view, establishing the place that scientific thinking has in the lives of all students. A typical approach to this objective has been to try to connect the *content* of science to phenomena familiar in students' everyday lives. An ultimately more powerful approach may be to connect the *process* of science to thinking processes that figure in ordinary people's lives.

The linking of scientific and argumentive thinking undertaken in this article offers some idea of how this might be achieved. Forms of thinking have been identified that apply across many domains—domains traditionally regarded as "scientific," as well as others that appear to have nothing to do with science. As has been shown here, these forms of thought can be rigorously defined within the framework provided by the structure of argument. Although they can be defined in general form, applicable across different content domains and reasoning contexts, they are not divorced from specific content because individuals' own theoretical conceptions within a content domain provide the basis for the reasoning that occurs.

Four implications warrant mention. First, as already suggested, it is essential that students come to appreciate the *relevance* of scientific thinking. It is not just from our perspective as educators but in the minds of students that the connection must be made between scientific thinking and thinking in the broader sense. To reinforce this connection, the thinking activities in which students engage within educational settings should be situated in a broad range of content domains, extending well beyond those falling within the traditional boundaries of science. Social science topics like those employed in the argument research are particularly apt as these are topics in which average people see themselves as competent to hold opinions and make judgments. Everyone is in some sense a social scientist. In pure science domains, in contrast, thinking can be inhibited by the strong belief in their own ignorance that students often bring to the topic. Paradoxically, then, to enable students to see the significance of scientific thinking we may need to move outside of traditional science domains.

Second, if the goal is to enhance the quality of students' thinking it is essential to engage them in the practice of thinking. As discussed elsewhere (Kuhn, 1986, 1988, 1990), despite a long history of advocacy of this method in educational theory, many educational programs designed to teach thinking skills focus on teaching students about good thinking rather than engaging them in the practice of thinking. In the case of the thinking skills that have been the focus of the present article, it is not difficult to envision how students might be engaged in the practice of thinking. Their own theories on familiar social science topics could serve as a starting point, as illustrated by the argument research that has been described. Students would be asked for evidence to justify their theories and this thinking then probed, using the argumentive framework of alternative theories, counterarguments, and rebuttals described earlier. The same format could be extended to other kinds of subject matter, including topics in pure science, with new concepts, terms, and theories introduced by a teacher.

Such activities, however, would differ in an important way from the research interviews described in this article; namely, they would take place in a social context. This brings us to the third of the four implications to be noted. The argument research emphasizes the link between dialogic and rhetorical (or social and internal) argument, and this link points to social argument as a powerful vehicle for developing the kinds of thinking that have been the subject of this article. Social dialogue offers a way to externalize those internal thinking strategies that we would like to foster within the individual. Those who have examined children studying science in classrooms or other group settings note that the opportunities such settings typically afford for students to appreciate the evidence-based nature of science, in particular the coordination of models, theories, and explanations with data and the contemplation of alternative models, are scant (Munby, 1982). The potential for such settings to foster appreciation of the "norms of scientific discourse" (Eichinger et al., 1991), on the other hand, is considerable.

In exploring this potential, it needs to be kept in mind that the contexts in which students conduct discourse differ in numerous respects from those in which professional science is conducted. Reif and Larkin (1991) outline the many ways in which scientific and everyday thought and discourse differ and note the dangers of students' assimilating one to the other. The point of this article has been not to argue for their identity but rather to make a case that the *connection* (rather than the similarity) between the thinking of intuitive and professional scientists has theoretical and practical significance as important as the many differences.

The fourth and final implication is that science education of the sort contemplated here needs to have a strong epistemological component. In the argument research, we included several questions of an epistemological nature, such as whether experts could be absolutely certain regarding these issues. About 50% of our sample, whom we classified as absolutists, answered this question affirmatively. Another roughly 35%, termed multiplists or relativists, responded negatively, noting that experts disagree and therefore that nothing is certain and all opinions are of equal validity. Only 15% fell into an evaluative category, in which knowing is regarded as a process that entails thinking, evaluation, and argument. (These subjects also responded negatively regarding expert certainty but, unlike multiplists, regarded themselves as less certain than experts.)

This epistemological naivete may be a critical factor in accounting for the limited argumentive reasoning ability that people display. Without an epistemological understanding of their value, the incentive to develop and practice the skills of argument is likely to be lacking. The critical role that the development of epistemological understanding plays in science education scarcely needs to be spelled out. As a number of studies have shown (Carey et al., 1989; Songer and Linn, 1991), students' conceptions of what science is are fundamental to meaningful science education. The student who says (quoting from one of the adolescents in our argument research), "You can't prove an opinion to be wrong because an opinion is something somebody holds for themselves," lacks any basis for judging the strength of an argument beyond its power to persuade. Such students can only appreciate science in a limited way and are particularly unlikely to see its relevance to their own lives.

In conclusion, it should be noted that science as exploration and science as argument do not in fact contradict one another. The exploratory behavior that infants and children display "naturally" is indeed worth supporting and nurturing in all of the ways we know how. Gifted science educators such as Duckworth (1990) seem even to be capable of resurrecting it in adults in whom it has lay dormant for years. But, by itself it is not enough. There is more that we are trying to do, or should be trying to do, than keeping alive a "natural curiosity." Rather, the natural curiosity that infants and children show about the world around them needs to be enriched and directed by the tools of scientific thought.

Both the scientific and argumentive thinking research described here make clear that we have much to teach children and adolescents (and adults) about thinking scientifically, little of which "comes naturally." Most important, it is a way of thinking that we would like them to practice and perfect not just in their thinking about science but in all of the thinking they do. In linking scientific thinking to thinking more broadly—in bringing it into the realm of the ordinary—we stand not only to introduce students to what is significant and powerful about a scientific mode of thought but also to make them aware of the role it can play in their own lives. This may be the most significant, far-reaching, and long-lasting benefit that students take away from their learning in science.

REFERENCES

Antaki, C. (1988). Structures of belief and justification. In C. Antaki (Ed.), *Analysing everyday explanation*. London: Sage Publications.

Billig, M. (1987). *Arguing and thinking: A rhetorical approach to social psychology*. Cambridge, UK: Cambridge University Press.

Brewer, W., & Samarapungavan, A. (1991). Children's theories vs. scientific theories: Differences in reasoning or differences in knowledge? In R. Hoffman & D. Palermo (Eds.), *Cognition and the symbolic processes*. Hillsdale, NJ: Erlbaum.

Carey, S. (1986). Cognitive science and science education. *American Psychologist, 41*, 1123–1130.

Carey, S., Evans, R., Honda, M., Jay, E., & Unger, C. (1989). An experiment is when you try it and see if it works: A study of grade 7 students' understanding of the construction of scientific knowledge. *International Journal of Science Education, 11*, 514–529.

Chi, M. (in press). Conceptual change within and across ontological categories: Examples from learning and discovery in science. In R. Giere (Ed.), *Cognitive models of science:*

Minnesota studies in the philosophy of science. Minneapolis, MN: University of Minnesota Press.

Damon, W. (1990). Social relations and children's thinking skills. In D. Kuhn (Ed.), *Development perspectives on teaching and learning thinking skills. Contributions to human development.* Vol. 21. Basel: Karger.

Duckworth, E. (1990). Opening the world. In E. Duckworth, J. Easley, D. Hawkins, & A. Henriques (Eds.), *Science education: A minds-on approach for the elementary years.* Hillsdale, NJ: Erlbaum.

Easley, J. (1990). Stressing dialogic skill. In E. Duckworth, J. Easley, D. Hawkins, & A. Henriques (Eds.), *Science education: A minds-on approach for the elementary years.* Hillsdale, NJ: Erlbaum.

Eichinger, D., Anderson, C., Palincsar, A., & David, A. (1991). An illustration of the roles of content knowledge, scientific argument, and social norms in collaborative problem solving. Paper presented at the annual meeting of American Educational Research Association.

Einstein, A. (1954). *Ideas and opinions.* New York: Crown.

Feldman, C. (1992). The new theory of theory of mind. *Human Development, 35,* 107–117.

Gholson, B., Shadish, W., Neimeyer, R., & Houts, A. (Eds.). (1989). *Psychology of science: Contributions to metascience.* New York: Cambridge University Press.

Heyes, C. (1989). Uneasy chapters in the relationship between psychology and epistemology. In B. Gholson, W. Shadish, R. Neimeyer, & A. Houts (Eds.), *Psychology of science: Contributions to metascience.* New York: Cambridge University Press.

Hilton, D. J. (1988). Logic and causal attribution. In D. J. Hilton (Ed.), *Contemporary science and natural explanation.* New York: New York University Press.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence.* New York: Basic Books.

Keil, F. (1984). Mechanisms in cognitive development and the structure of knowledge. In R. Sternberg (Ed.), *Mechanisms of cognitive development.* New York: Freeman.

Klahr, D., & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive Science, 12,* 1–48.

Kuhn, D. (1986). Education for thinking. *Teachers College Record, 87,* 495–512.

Kuhn, D. (1988). Making cognitive development research relevant to education. In W. Damon (Ed.), *Child development today and tomorrow.* San Francisco: Jossey-Bass.

Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychological Review, 96,* 674–689.

Kuhn, D. (1990). *Developmental perspectives on teaching and learning thinking skills. Contributions to human development.* Vol. 21. Basel: Karger.

Kuhn, D. (1991). *The skills of argument.* New York: Cambridge University Press.

Kuhn, D. (1992). Cognitive development. In. M. Bornstein & M. Lamb (Ed.), *Developmental psychology: An advanced textbook.* 3rd ed. Hillsdale, NJ: Erlbaum.

Kuhn, D., Amsel, E., & O'Loughlin, M. (1988). *The development of scientific thinking skills.* Orlando, FL: Academic Press.

Kuhn, D., Schauble, L., & Garcia-Mila, M. (in press). Cross-domain development of scientific reasoning. *Cognition and Instruction.*

Munby, H. (1982). *What is scientific thinking?* Toronto: Guidance Centre, University of Toronto.

Parvanno, C., quoted in Elder, J. (1990, January 7). Hands on. *New York Times* (supplement).

Perkins, D. (1985). Postprimary education has little impact on informal reasoning. *Journal of Educational Psychology, 77,* 562–571.

Perner, J. (1991). *Understanding the representational mind.* Cambridge, MA: MIT Press.

Posner, G., Strike, K., Hewson, P., & Gertzog, W. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education, 66,* 211–227.

Reif, F., & Larkin, J. (1991). Cognition in scientific and everyday domains: Comparison and learning implications. *Journal of Research in Science Teaching, 28,* 733–760.

Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context.* Oxford: Oxford University Press.

Schauble, L. (1990). Belief revision in children: The role of prior knowledge and strategies for generating evidence. *Journal of Experimental Child Psychology, 49,* 31–57.

Sodian, B., Zaitchik, D., & Carey, S. (1991). Young children's differentiation of hypothetical beliefs from evidence. *Child Development, 62,*753–766.

Songer, N., & Linn, M. (1991). How do students' views of science influence knowledge integration? *Journal of Research in Science Teaching, 28,* 761–784.

Vosniadou, S., & Brewer, W. (1987). Theories of knowledge restructuring in development. *Review of Educational Research, 57,* 51–67.

Vosniadou, S., & Brewer, W. (in press). Mental models of the earth: A study of conceptual change in childhood. *Cognitive Psychology.*

Voss, J., & Means, M. (1991). Learning to reason via instruction in argumentation. *Learning and Instruction, 1,* 337–350.

Wellman, H. (1990). *The child's theory of mind.* Cambridge, MA: MIT Press.

West, L., & Pines, A. (Eds.). (1985). *Cognitive structure and conceptual change.* Orlando, FL: Academic Press.

Westrum, R. (1989). The psychology of scientific dialogues. In B. Gholson, W. Shadish, R. Neimeyer, & A. Houts (Eds.), *Psychology of science: Contributions to metascience.* New York: Cambridge University Press.

Accepted for publication 4 February 1993