INFORMAL LOGIC
XIV.1, Winter 1992

Book Review

The Skills of Argument by Deanna Kuhn

JONATHAN BARON *University of Pennsylvania*

Kuhn, Deanna. (1991). ***The Skills of Argument.*** New York: Cambridge University Press. 324 pages.

This is an extremely important book, which should be read by anyone interested in teaching people to think better or in understanding the nature of people's attitudes or their informal reasoning. The author is a professor at Teachers College, Columbia University, editor of the journal *Human Development*, and an author of *The Development of Scientific Thinking Skills* (Kuhn, Amsel, and O'Loughlin, 1988). (Kuhn, 1989, summarizes much of that book and a little of the present book.)

The present book reports a well-done empirical study of people's informal reasoning. It goes a long way toward filling the large gap between psychologists, who study in great detail people's reasoning about mostly irrelevant problems, and teachers of informal reasoning, who (in my limited experience) develop elaborate schemes for fixing things (such as the traditional Aristotelian fallacies) that might not be broke. Before we try to teach people to reason better, we need to have not only a conception of how they ought to reason but also an account of where they go astray, and why. This account is what Kuhn helps to provide.

The Study

The study asked people to respond to three questions: "What causes prisoners to return to crime after they're released?"; "What causes children to fail in school?"; and "What causes unemployment?" The 160 main subjects ranged in age from 14 to 69 and were selected to represent equally the two sexes and two educational levels: college (or "college bound" for the adolescents) and noncollege. Five teachers and 5 parole officers were also included to examine the effect of familiarity with two of the domains, and 5 graduate students in philosophy were included to examine expertise in thinking itself.

The Basic Interview

Each subject was interviewed about these questions in two sessions. The interview began with the following questions and probes (some abbreviated or omitted)—illustrated for the crime question—which you might try to answer yourself:

1. What causes prisoners to return to crime after they're released?

1a. Anything else?

2. (*If multiple causes mentioned*) Which of these would you say is the major cause . . . .

3. How do you know this is the cause?

3a. (*Probe, if necessary*) Just to be sure I understand you, can you explain exactly how this shows that this is the cause?

4. If you were trying to convince someone else that your view [that this is the cause] is right, what *evidence* [verbal emphasis] would you give to try to show this?

4a. (*Probe, if necessary*) Can you be very specific . . . .

5. Is there anything further you could say to help show that what you've said is correct?

6. Is there anything someone could say or do to prove that this is what causes prisoners to return to crime?

7. Can you remember when you began to hold this view?

8. Suppose now that someone disagreed with your view . . . . What might *they* say to try to show that you were wrong?

9. What evidence might this person give to try to show that you were wrong?

10. (*Probe, if necessary*) Just to be sure I understand . . . .

11. (*If not already indicated*) Is there any fact or evidence which, if it were true, would prove that you were wrong?

12. (*Omit if . . . already generated*) A person like we've been talking about . . . —what might they say is the major cause?

13. (*Include an alternative cause if not given*) . . . What could you say to show that this other person was wrong?

13a. (*Probe, if necessary*) Just to be sure I understand . . . .

14. Would you be able to prove this person wrong?

15. (*If needed*) What could you say to show that your own view is the correct one?

Responses were categorized (with reliability checks) into two or three categories for each of several issues. The categories differed in success in answering the question asked. The main results were these: successful and unsuccessful answers were found in all ages and at both educational levels; overall, about half of the answers in each category were classified as successful; age and sex had essentially no effect on success; college responses were consistently more successful than noncollege (with differences on the order of 25% for each answer); familiarity with the domain had no effect; but the philosophy students gave successful responses to all questions. We can conclude that something about the early education or family background of college bound students has an effect on their reasoning ability according to these measures. Moreover, the effect is completely general and not dependent on specific knowledge of the domain. The results from the philosophy students suggest that further education *can* have an effect too—unless these students were just as good when they began high school—but the other results indicate that little improvement in reasoning occurs in high-school or college for most students in the U.S. (as found by Perkins, 1985, as well).

I shall now review the results in a little more detail. Kuhn describes her results by quoting many examples of each type of response. Then she gives statistical summaries and comparisons. I shall not attempt here to capture the richness of these data. The reader who reads only a fraction of the examples (and, really, many can be skipped) will get a good picture of the richness and variety of the subjects' approaches, which could not possibly be captured in statistical summaries. But I shall stick to some of the main results here.

Causal theories. Essentially all subjects produced some answer to the question in the form of a causal account or theory. For example, lack of employment and return to a poor environment were popular answers to the crime question. Some subjects produced several such theories, or spontaneously provided alternatives to their favored one. The college group produced more multiple theories.

Evidence to support theories. College and noncollege groups differed greatly in the provision of what Kuhn calls *genuine*

evidence, evidence that is both "distinguishable from description of the causal sequence itself" and "bears on its correctness" (p. 45). Genuine evidence consisted of covariation between the cause and the outcome (either asserted to exist or hypothetical), other relevant facts (e.g., the possibility that recidivism is familial, which can support some theories), analogy, application of general principles, or evidence against alternatives. By contrast *pseudoevidence* was "evidence by illustration." Subjects answered questions about evidence (questions 3-6) as if they were questions about the theory itself (questions 1-2). Some responses to the request for evidence were quite long, since they amounted to narratives of particular cases. Some subjects failed to provide evidence of any sort: they implied that evidence is unnecessary, made irrelevant assertions, or "cite[d] the phenomenon itself as evidence regarding its cause" (p. 82). Interestingly, most of the subjects said that the "evidence" they provide did prove the correctness of their theory, and their level of confidence was unrelated to the quality of the evidence they provided.

Alternative theories. Many subjects (fewer in the college group) were unsuccessful or only partially successful in generating alternative theories to their own (mostly in response to question 8). The most common failure was that the response overlapped or coincided with the subject's own theory. Sometimes the same words were used for the alternative (p. 106)! Some subjects denied that anyone would propose an alternative, or indicated that, while they could not think of an alternative, they would reject it if it were presented to them.

Counterarguments. Subjects gave a great variety of successful counterarguments to their own theories (mostly in response to questions 9-11), which Kuhn catalogs in some detail. Unsuccessful counterarguments, provided more often in the noncollege group, also took several forms, such as arguments that the outcome does not occur (e.g., that prisoners do not return to crime), causal theories that are consistent with the subject's theory, or remedies (consistent with their favored theory) rather than alternative theories. More interesting still are cases in which subjects simply failed to attempt a counterargument: some subjects rejected the possibility of a counterargument, e.g., because they claimed that both the antecedent (such as family problems) and consequent (school failure) are both present; other subjects implied that counterarguments would not matter because, if someone gave them one, it would have had utterly no effect on their belief. "In extreme cases, . . . the subject exhibits a kind of proprietorship over the theory that undermines its independent existence, rendering it incontestable. To challenge the theory is to challenge the subject's own self" (p. 144).

Rebuttals. Many subjects—especially the college group—were capable of saying, in a variety of ways, what kind of evidence would rebut the alternative (mostly in response to questions 14 and 15). Again, other subjects failed to generate rebuttals successfully or did not try to do so. The unsuccessful attempts failed because they did not contradict the alternative theory, because they contradicted the subject's theory, or because they simply asserted that the subject's theory was correct or the alternative was wrong. Subjects who did not try to rebut the counterargument often simply claimed ignorance. In general, subjects were more likely to be successful in rebutting an alternative theory than at providing an argument against their own theory. Some successful rebuttals—e.g., those that challenged the causal necessity of the factor in question—were based on forms of argument that could also apply to the subject's theory, even though the subject did not think of this.

Response to Evidence

At the end of the interview, subjects were asked to interpret two kinds of evidence on the crime and school issues. *Underdetermined* evidence simply presented a scenario about one individual person, without any mention of any causal factors. The *overdetermined* evidence reported a study of 25 prisoners or pupils, with no control group, but with one expert finding evidence for each of three possible causes (e.g., a school psychologist found evidence of learning problems).

Despite the fact that both sets of evidence were essentially useless, subjects tended to interpret them as supporting their theories. Noncollege subjects were more confident of these interpretations than college subjects. For the underdetermined evidence, many subjects simply asserted that their theory was correct for the case. For the overdetermined evidence, subjects tended to focus on the parts consistent with their view. Most subjects (especially the noncollege group) interpreted the evidence as agreeing with or supporting their theory. Only the philosophy students uniformly pointed out that the evidence was useless. (I wonder what they would have said about *useful* evidence, though!)

Importantly, subjects often answered questions about the *evidence* as if they were asked about the truth of their own *theory*. This was a major finding of Kuhn's earlier work (Kuhn et al., 1988; Kuhn, 1989), and it is replicated here. If this response is typical of subjects' thinking, then they are unable to evaluate evidence independently of their theories. If the evidence disagrees with the theory, they will not note the disagreement. Instead, they will either neglect the disagreement or (Kuhn found) eventually change their theory without acknowledging the role of the evidence.

This is an example of a more general point that Kuhn makes repeatedly: If people cannot reflect on their thinking, they cannot control it. I think that Kuhn is a little unclear about how this works, but I think she is basically correct. If I can try my own version of the point, it is this: If we seek to teach people to think better, we must be able to tell them what to do, and they must be able to assess whether they are doing it or not. Such learning requires a language of the sort that Kuhn uses in her interview, a language that distinguishes theories from the evidence that supports them, among other things. Students of thinking must be able to follow an instruction like, "Think of an alternative theory and ask whether the evidence supports the alternative as much as it supports your own theory."

Such instruction also requires *analysis* of thinking of the sort that subjects fail to do here. In order to learn to evaluate evidence correctly, it helps to be able to say what the evidence would imply in the absence of any prior commitment to a particular theory. Perhaps Kuhn exaggerates the *necessity* of such "metacognitive" knowledge. It might be that people can learn to think well by observation and feedback without the use of such general terms and without such analysis. Surely, though, these abilities help.

Epistemological Theories

In the middle part of the interview, subjects were interviewed about their epistemological reasoning:

1. How sure are you about what causes prisoners to return to crime?
2. Do experts know for sure what causes . . . ?
3. (*If no*) Would it be possible for experts to find out for sure . . . ?
4. How sure are you of your view, compared to an expert?
5. Is more than one point of view possible . . . ?
6. (*If yes*) Could more than one point of view be right?

Following other theorists, responses were classified (on the basis of responses

to all questions) as reflecting three kinds of implicit theories of knowledge. "Absolutist" theories held that experts could be certain of the truth and that the subject was certain of the truth too. Most subjects, paradoxically, agreed that other theories could be true too. The paradox was sometimes resolved by the assertion that the subject was correct because people are entitled to their own theories, so that the subject's theory is true for the subject. The majority of subjects were absolutist, even in the college group.

Some of the quoted responses suggest (to me) that subjects in this group interpreted the questions psychologically rather than logically. In rebutting a counterargument, for example, the subject was thinking not in terms of showing that the argument was wrong according to some standard but, rather, in terms of changing the belief of a hypothetical opponent, e.g., "If you repeat yourself parrot-like, . . . there's no way you can convince the other person . . . [but] you reduce their resistance to a point. You might be able to convince them of minor parts of your statement."

"Multiplist" theories of knowledge hold that experts are not certain and that conflicting theories can be simultaneously correct. Responses in this category often referred to personal experience or emotion as the grounds for belief. Subjects owned their beliefs, as indicated by the following responses to the question, "Would you be able to prove this person wrong?": "No, I would just be able to say I disagree with you and this is why and you can't tell me that my experience is wrong because this is what my experience was."; ". . . you can't prove an opinion to be wrong, I don't think . . . an opinion is something which somebody holds for themselves. You can't change their opinion or alter it. They have their own opinion."

"Evaluative" theorists held themselves to be less certain than experts. They held that "viewpoints can be compared with one another and *evaluated* with respect to their relative adequacy or merit" (p. 188), even if certain knowledge is impossible. Only 14% of the college group fell into this category, 5% of the noncollege group. Subjects in this category were less likely than others to be sure or very sure that their theory was correct; absolutist subjects were the most likely.

Kuhn argues that opinions held unreflectively are as good as useless. Reflection involves considering at least one alternative and finding evidence that favors one's own view more than it favors the alternative(s), or evidence that impugns the alternatives more than one's view. These are the same moves that are required when people defend their views in dialogic arguments with others. In order to engage in such reflection, "Individuals must also hold the implicit epistemological theory that treats argument as worthwhile, as a fundamental path to knowing. In other words, people must see the point of argument, if they are to engage in it" (p. 201). People who hold that everyone's opinion is equally valid have no incentive to learn the standards of argumentation and belief formation. (Kuhn's view of the role of epistemological beliefs is consistent with the findings of Baron, 1991, and Schommer, 1990.)

Concerns

Let me now discuss a few concerns I have with the study and the conclusions drawn from it. It will be apparent that none of these is very serious, but I hope that the strength of the work will be clarified by the weaknesses of the arguments against it.

Leading the witness. Kuhn might be said to lead her subjects down a garden path into overcommitment to their theories. The interview essentially forced the subjects to take a position, and it then asked them how they *knew* that "their" position was true. Certainty and ownership were presup-

posed. Of course, a sensitive subject would challenge these presuppositions, as many did. And Kuhn presents evidence that, for most subjects, the presuppositions were correct: subjects had in fact thought about the questions before they were interviewed and had formed clear views that they held with confidence. So my concern here is a methodological warning to future investigators rather than a serious worry about Kuhn's conclusions.

A related problem is that subjects who agree with the alternative theories that they generated were scored as not fully successful in generating alternative theories. Again, it is assumed that subjects must have a view. Subjects who were totally ignorant and who appreciated their ignorance would have been counted as inadequate for many answers. But, again, the data indicate that very few subjects were in this category. Subjects who generated multiple theories initially were in fact more successful than other subjects in generating an alternative that they did *not* accept.

Answering the wrong question. Many answers that Kuhn counted as inadequate were adequate answers to a different question, e.g., providing a remedy when asked for an alternative theory, or an alternative theory when asked for counterevidence. Perhaps the problem was not in subjects' thinking but, rather, in their ability to understand interview questions.

If this is true, it would not be entirely uninteresting, and it would not entirely undermine Kuhn's account. Kuhn would argue, I think, that subjects certainly knew what the words mean (especially given the fact that most questions were repeated in different ways), so, if they did not respond appropriately, it was most likely because they did not regard the distinctions conveyed by the words to be important. Misunderstanding, if it occurred, was therefore very likely a result of the very deficits that Kuhn claims to find.

Still, the *quantitative* results of this study might not be generalizable to other methods of investigation into the subjects' thinking. If, for example, subjects were asked to think aloud, they might be more likely to generate alternative theories—or less likely. I shall return to this issue.

Mechanism as evidence. Some of the responses that Kuhn classified as pseudoevidence were arguments of plausibility based on a description of the mechanism, e.g., the intervening steps in the causal chain. Kuhn regards this as an elaboration of the theory, which cannot establish the theory's correctness: "Pseudoevidence does not provide a basis for choosing one alternative theory over another. Often, of course, neither does genuine evidence, but the latter at least bears on their relative correctness" (p. 115).

Although description of mechanism is indeed an elaboration, I think it might sometimes serve as genuine evidence too, for it can bear on the correctness of the theory. (In the domains studied, essentially all evidence is probabilistic rather than decisive.) For example, when it is claimed that electromagnetic fields can cause cancer, the possibility that living tissue has small particles that respond to magnets is important: without evidence of such particles, even epidemiological evidence would most likely be artifactual, the result of unknown extraneous variables. Likewise for arguments about the effect of irradiation of foods or bovine growth hormone. (In both cases, the absence of a plausible mechanism makes most scientists unconcerned.)

Kuhn also regards the provision of an alternative theory as only partially successful when subjects are asked to provide evidence against their favored theory (p. 127). In criminal trials, however, the mere provision of an alternative account is often a powerful argument for the defense. Is this an error? I think not. If the defense cannot construct such an account, the defendant is truly more likely to be guilty. In-

nocence usually implies that an alternative account can be constructed, even if it is only somewhat plausible.

The importance of alternatives. The view that consideration of alternatives is necessary for rational belief is one that is widely held, and it is especially reasonable for the kinds of situation that Kuhn examines. However, some knowledge could result without such consideration, by methods of inference that are so secure that no alternatives need be explicitly considered. (Margolis, 1987, takes such a view.) Examples are found in mathematical proofs, but these are possibly not the only examples. Perhaps it could be argued that there is always an alternative, namely, that the proposition in question is not true, but this is trivial, and it is not what Kuhn has in mind.

Kuhn's view of the importance of alternatives is related to her presupposition that subjects have theories in which they are confident. She says, "To recognize that there are alternatives that stand in contrast to one's own theory is the first, vital step to (a) recognizing that one's theory could be wrong and, therefore, (b) seeking to evaluate to what extent it is correct. . . . Subjects who cannot envisage the possibility of alternatives to their theories, claiming that 'the majority think the way I do' or 'my thoughts run in this direction and that's about it,' cannot *know* that these theories are correct, no matter how strongly they believe in them" (pp. 114-15). Conceivably, however, a rational person might say, "I can't think of any alternatives, or I am not inclined to try, so I have no confidence in my theory, or no theory that I could call my own."

If Kuhn has missed this point, it must be an easy one to miss, for I have missed it too. In Baron (1985a; less so in Baron, 1988), I too argued for the importance of generating alternative possibilities as a way of avoiding error. But we cannot spend our lives thinking about everything. Now I think that the point is this: we should not adopt a possibility with high confidence until we have considered alternatives to it (except in those few cases in which we have high confidence in the methods of inference that led to it). We might think of two sorts of errors in belief formation, those of omission and those of commission. Errors of commission are those in which we hold a belief with confidence when we should not do so. Errors of omission are those in which we have no belief, when we ought to. (The extension of this distinction to decision making is complex, and I shall avoid it here.) The search for alternatives is important in avoiding errors of commission. Errors of omission result from the failure to search for even a single possible answer, or perhaps from the inability to think of a single possibility despite the effort. This issue leads to another concern.

Effort vs. ability. Kuhn classified responses according to whether they failed or succeeded in answering the questions posed, e.g., providing counterevidence when asked for it. Failure can result from two sources, lack of effort and lack of ability. By "effort," here, I mean simply trying to do something, not necessarily spending large amounts of scarce resources (such as time or concentration) on it. Some subjects explicitly tried to answer the question and said that they were *unable* to do so. Lack of ability, in turn, can result from ignorance or from processing difficulties such as inability to retrieve information from memory. (Ignorance itself can result from lack of effort or lack of ability.)

Of course, all sources of failure are important, but we have reason to distinguish these various sources (as argued by Baron, 1985a). Lack of effort is potentially remediable by instruction in thinking itself. Ignorance (regardless of its source) could be corrected by specific instruction in the

domain. Lack of processing capacity seems to be largely irremediable (except by biological manipulations such as stimulant drugs). It is, for example, clear that general improvements in processing ability do not result from practice itself (Baron, 1985b). If we are to take Kuhn's interview as a measure of educational outcomes, then, it is a measure that could be somewhat impure, because it is more sensitive to inability than it needs to be. It might be better not to count "I can't think of anything" as an error. There were few such responses, however, so the main results are surely unaffected by this problem.

Let me turn now to a couple of points that are not so much concerns but simply topics for discussion: the relation between thinking and argument and the role of standards.

Thinking and Argument

A major explicit assumption of Kuhn's project is that thinking is analogous to, and (at least to some extent) develops from, interpersonal, dialogic argument. Kuhn suggests that thinking can be taught through the encouragement of dialogic argument. I would suggest that this, by itself, is not enough to communicate the standards of good and poor argument. I think that these must be taught more explicitly than through practice alone. But here I want to concentrate on another issue.

Much of the interview can be seen as a scaffold for producing an argument of the sort that students might make in written assignments, that lawyers might make in briefs, and that scientists and other scholars might make in their papers. Such arguments gain strength by considering alternatives and rejecting them. Students' papers that follow this sort of outline should get higher grades than those that do not.

Kuhn is well aware that this is only one way to study thinking as it affects formation of beliefs. On the one hand, the scaffold should not be necessary. We ought to be concerned about whether people *spontaneously* try to think of alternative theories and counterarguments, as well as with their inclination and ability to do so when they are explicitly asked. Perkins, Faraday, and Bushey (1991) found widespread deficiencies in the provision of both theories and arguments on the other side, and they found that specific requests for other-side arguments dramatically improved the tendency to generate them. Kuhn, by contrast, is concerned with people's failures to generate other-side arguments when they are explicitly asked to do so. Viewed in this way, the problem with people's informal reasoning is considerably greater than even Kuhn's data suggest.

On the other hand, Kuhn did *not* provide certain kinds of support that *are* present in true dialogic arguments. When thinking involves two or more people, the other people are more likely to provide the alternatives and counterarguments that the individual might miss. (Kuhn did simulate this in question 13, in which the interviewer provides the subject with an alternative to rebut.) Viewed in this way, Kuhn (and Perkins et al.) might overestimate the deficiencies that people have.

Of course, people think by themselves *and* they argue with each other *and* they make arguments to others (in writing) when others are absent. Kuhn's results are perhaps most relevant to the last activity. The results of Perkins et al. are most relevant to the first. The relative importance of these different contexts for thinking undoubtedly depends on a person's cultural environment, although all of them are of some importance in modern Western societies.

Standards

As Margolis (1987) points out, belief formation occurs naturally and automatically much of the time. Animals form be-

liefs without much thought, just as they make decisions, and we do these things the same way much of the time. Human beings have, however, gone way beyond other animals in our ability to regulate our behavior on the basis of social norms. Human cultures have universally developed norms for all sorts of overt behavior, and norms for what beliefs people should hold. But the existence of norms for methods of belief formation (and decision making) may not be universal. Western culture has developed such norms. Kuhn's work—and the work of many others—indicates that the effort to spread such norms has been incompletely successful. Many of Kuhn's subjects seemed unfamiliar with the idea that an argument can be good or poor in some way that is more objective than simply whether or not it succeeds in changing someone's mind. That objectivity, of course, comes from these standards. What is perhaps most disturbing is that the five schoolteachers interviewed seemed to have no special awareness of such standards. Many other subjects were parents. Only the philosophy students were consistently aware of the possibility of evaluating arguments and reasoning as good or poor. If teachers and parents do not know the standards, who will insure that they are maintained?

Conclusion

As Kuhn points out, educational discourse is now awash with discussion of "change" in the direction of teaching "thinking skills" (Brown, 1991). But most of that discussion does not seem to delve into the question of just how people should think and just where they need correction. Kuhn has enriched our knowledge of the second question, and, to some extent, the first. We need more work like this, both from her and from others.

References

Baron, J. (1985a). *Rationality and Intelligence*. New York: Cambridge University Press.

Baron, J. (1985b). What Kinds of Intelligence Components Are Fundamental? In S. F. Chipman, J. W. Segal, and R. Glaser (eds.), *Thinking and Learning Skills. Vol. 2: Research and Open Questions*, pp. 365-90. Hillsdale, NJ: Erlbaum.

Baron, J. (1988). *Thinking and Deciding*. New York: Cambridge University Press.

Baron, J. (1991). Beliefs about Thinking. In J. F. Voss, D. N. Perkins, and J. W. Segal (eds.), *Informal Reasoning and Education*, pp. 169-86. Hillsdale, NJ: Erlbaum.

Brown, R. G. (1991). *Schools of Thought: How the Politics of Literacy Shape Thinking in the Classroom*. San Francisco: Jossey Bass.

Kuhn, D. (1989). Children and Adults as Intuitive Scientists. *Psychological Review*, 96, 674-89.

Kuhn, D., Amsel, E., and O'Loughlin, M. (1988). *The Development of Scientific Thinking Skills*. Orlando, FL: Academic Press.

Margolis, H. (1987). *Patterns, Thinking, and Cognition: A Theory of Judgment*. Chicago: University of Chicago Press.

Perkins, D. N. (1985). Postprimary Education Has Little Impact on Informal Reasoning. *Journal of Educational Psychology*, 77, 562-71.

Perkins, D. N., Faraday, M., and Bushey, B. (1991). Everyday Reasoning and the Roots of Intelligence. In J. F. Voss, D. N. Perkins, and J. W. Segal (eds.), *Informal Reasoning and Education*, pp. 83-105. Hillsdale, NJ: Erlbaum.

Schommer, M. (1990). Effects of Beliefs about the Nature of Knowledge on Comprehension. *Journal of Educational Psychology*, 82, 498-504.

JONATHAN BARON
DEPARTMENT OF PSYCHOLOGY
UNIVERSITY OF PENNSYLVANIA
3815 WALNUT ST.
PHILADELPHIA, PA 19104-6196
E-mail: baron@cattell.psych.upenn.edu