5 Everyday Reasoning and the Roots of Intelligence

D. N. Perkins
Michael Faraday
Barbara Bushey
Harvard Graduate School of Education

Be reasonable! That familiar imperative serves as just one reminder of the value we vest in people's rational conduct of their own lives and their dealings with others. We want ourselves and those with whom we interact to make sound decisions that reflect thoughtful exploration of the alternatives and the factors that recommend one over another. We want to eschew silly, hasty, and misguided beliefs in favor of a sound picture of the world. Although we may, on occasion, want to take a little fling and do something unreasonable, on the whole we all want to be reasonable.

But how well do we succeed? And, if we do not always succeed so well, on what causal factors does success, failure, or a shade in between depend? Such questions provide an opportunity to engage the issue of the nature of intelligence in its broadest sense. Apart from the technical senses that psychologists have attached to the notion (cf. Baron, 1978; Cattell, 1963; Gardner, 1983; Guilford & Hoepfner, 1971; Jensen, 1984; Sternberg, 1985), intelligence preserves a natural language meaning of great scope. Roughly speaking, intelligence is whatever a person has that empowers the person in some crosscutting way to perform intellectual tasks well. (*Crosscutting* is somewhat negotiable and, in fact, variable from theorist to theorist, but great skill in only a very narrow domain does not normally count as a sign of intelligence.) Accordingly, being reasonable, as an intellectual endeavor, plainly calls on quite different resources that might count as parts of intelligence broadly construed.

WHAT IS IT TO BE INTELLIGENT?

Consider what rival roots of intelligence might contribute to being reasonable. Perhaps good reasoning in everyday life reflects principally general intelligence

other countries. Would a nuclear freeze between the United States and the Soviet Union make much of a difference in the probability of world war, considering that several other countries, such as China, Iran, and India, also have nuclear weapons? Further, is it possible that a world war could be carried on without nuclear weapons? In this fashion, an agile reasoner would proceed to develop a thorough map of all facets of the issue.

Of course, a situation model can be faulty. For instance, a simplistic response to the nuclear freeze issue just mentioned might be: "A nuclear freeze wouldn't reduce the possibility of world war, because everyone knows that the Russians would cheat. The freeze wouldn't last, because we couldn't trust them." Terse though this may seem, subjects in the line of experimentation discussed here often say little more. Typically, faulty situation modeling quickly affirms one's snap judgment in response to a vexed issue. As this example shows, two basic criteria for evaluating situation models are bias and completeness. As to bias, the example argument is entirely one-sided. As to completeness, the argument does not even elaborate thoroughly the position it adopts; there are many other ways, besides devious Russians, by which a nuclear freeze could fail to reduce the risk of war.

How does this notion of situation modeling contrast with a formal view of reasoning and its pitfalls? Simply in that the problems of bias and completeness do not arise in a formal context. If you have a one-sided, well-formed argument in a formal context, it does not reveal bias; it is a *proof*. If you have a well-formed argument in a formal context, you do not need five other strands of argument to strengthen the case; a proof is a proof. To put it another way, bias and incompleteness in the situation modeling sense are not *formal weaknesses*. Formal weaknesses—and indeed most *informal fallacies* (cf. Engel, 1976: Fearnside & Holther, 1959)—represent inferential gaps: For instance, A does not imply B if there is some shortfall of entailment in A. But problems of bias and completeness are "sins of omission." A may well constitute evidence for B, but what about all the other strands of argument? And what about additional information, rather than a logical weakness, that may impugn the A–B relation?

As these points show, thinking of informal reasoning as a process of situation modeling reveals weaknesses of reasoning with a different character from those that are prominent in formal logic. Whether this distinction plays out in an illuminating way depends on what data say about the reasoning difficulties that people do, in fact, encounter.

A TURN TO METHODOLOGY

If you want to examine people's informal reasoning, what do you do? What do the data consist in? How are they collected? How are they evaluated? Although these questions inevitably allow many answers, more than a decade of work in our laboratory has led to a fairly refined methodology for probing informal reasoning. The answers to the broad questions raised at the outset derive principally from the data this methodology has yielded. Although individual studies differ in their details, most share the broad characteristics outlined in this section.

Issues Employed

The general methodology for assessing informal reasoning involves administering a procedure that asks subjects to think about everyday social and political issues. The issues employed in the research are current and genuinely vexed, allowing several substantial arguments on both sides of the case. Pilot testing has confirmed the vexed nature of the issues for the subject populations, demonstrating that a variety of arguments on both sides appears with no strong consensus. In addition, the issues have been chosen not to depend on subjects' background knowledge, which varies greatly across the subject population; they prove accessible to junior high school students as well as adults. To give the flavor of the issues, we offer the following two questions:

1. Would providing more money for public schools significantly improve the quality of teaching and learning?
2. Would a nuclear freeze agreement signed between the United States and the Soviet Union significantly reduce the possibility of world war?

Outline of the General Procedure

The three-part experimental procedure was designed to probe subjects' reasoning performance and capacity. We first presented an issue and asked for the subject's snap judgment on the question, confidence in the snap judgment, degree of interest in the issue, and amount of time previously spent thinking about the issue.

Second, subjects were asked to think about the issue, reach a position if they felt comfortable doing so, and write down their reasons. Guidance was minimal: Subjects were urged to take their time and to try to be complete in thinking about the issue and recording their thoughts, regardless of whether they seemed important. The wording of instructions was carefully chosen to avoid triggering an argumentative stance and to favor complete, evenhanded reflection. For example, the use of the word "argument" was avoided altogether. This unstructured phase of the procedure provided a naturalistic measure of informal reasoning performance with a nudge toward completeness, which served as a baseline for assessing gains from scaffolding.

The third phase of the procedure provided subjects with generic scaffolding to help them fully and systematically deploy their capacities to reason. *Scaffolding* is a term commonly used to describe how a skilled individual can help a less

skilled one to manage a performance by supporting the learner's efforts when the learner is at a loss and hanging back whenever the learner proves able (Greenfield, 1984b; Rogoff & Gardner, 1984). In the research, the questions and instructions used for scaffolding were metacognitive principles that anyone could employ to guide his or her own thinking. The scaffolding emphasized countering the difficulties of bias and completeness salient in the situation modeling perspective on informal reasoning. The scaffolding consisted of prompts to be complete, seek reasons on both sides of the case, and impartially assess their importance.

The point was to see how much further subjects could develop their initial arguments with the help of scaffolding. On the one hand, perhaps their initial arguments would turn out to rise close to some sort of capacity ceiling, with scaffolding having little impact. On the other hand, subjects might prove able to develop their arguments much more extensively in response to the scaffolding. This result would, in turn, suggest that, with a better metacognitive repertoire, subjects could construct much richer arguments on their own.

Over the course of several studies, this procedure was administered in oral as well as in written form, depending on the particular experimental context or question. When reasoning was assessed in a classroom, protocols were written. Written protocols allowed for group administration and self-pacing. When the procedure was administered to individuals in the laboratory, formats were either written or oral. Administration of an oral procedure involved an experimenter's reading instructions and prompts to a single subject and either taping the subject thinking aloud or having the subject write out his or her thoughts. A direct comparison of oral and written reporting of arguments indicated no decisive quantitative or qualitative advantage of either mode.

In addition to the reasoning procedure, each subject completed a short-form IQ measure, either the Quick Word Test (Borgatta & Corsini, 1964) or the Slossen Intelligence Test (Slossen, 1981). The measure of IQ made it possible to investigate the relation of verbal IQ to everyday reasoning.

Scoring Methodology

Two kinds of scoring were employed. Numerical quality measures, consisting of both counts and scales, gauged the extent to which arguments were well developed or subject to bias and incompleteness. Second, error analyses examined the subjects' arguments for qualitative weaknesses in reasoning and classified these lapses. The quality measure scoring was performed by two judges working independently; they coscored a random subsample of the data on which interjudge agreement was based. Judges were blind as to which were solo and which were scaffolded performances. In general, interjudge agreement was high and statistically highly significant.

Analysis of the quality measures focused on comparing the original unstructured reasoning performance to the scaffolded performance. One kind of analysis centered on certain counts. Each unstructured and scaffolded argument was scored for completeness based on the number of "my-side" and "other-side" arguments. My-side arguments are those that support subjects' initial snap judgments, whereas other-side arguments are those that oppose *them*. *Argument* is used here to denote a line of argument, which might be more or less elaborated. For instance, an argument that more money would abet student learning because schools could buy better books would count as one line of argument, regardless of whether the subject elaborated about what kind of books they might be. My-side and other-side arguments were summed to give a "both-sides" count. Separate counts were taken of the initial arguments and the scaffolded performance.

In addition to the various counts, each performance was given two ratings for overall quality, one reflecting the treatment of the subject's side and a second reflecting the treatment of the other side of the case. This holistic rating used a five-point scale ranging from 0 to 4. On this scale, 0 stood for no response at all, and 4 was assigned when most major arguments on the topic were given with good elaboration and connection to other issues. By using this scale, the scoring could incorporate considerations of soundness of the arguments not captured by mere counts. In fact, however, in the statistical treatment of the results the holistic ratings and the counts proved highly intercorrelated and disclosed essentially the same patterns.

The error analysis was conducted on one large sample of arguments gathered from subjects ranging from the first year of high school to the fourth year of graduate school and beyond. Two judges critiqued in natural language one argument from each of more than 300 subjects. The judges were familiar with argumentation in general, including formal and informal argument. They were encouraged to describe the weaknesses they saw in any terms that seemed appropriate to them. Concurrently, a classification system for categorizing weaknesses in everyday argument was developed; the judges were not familiar with this system when they critiqued the subjects' arguments. Subsequently, their criticisms were categorized according to the classification system to yield a picture of what sort of lapses, as naturalistically identified, occurred with most frequency in the sample. This categorizing was accomplished by multiple judges, again with a satisfactory degree of accord.

IS EVERYDAY REASONING GENERALLY SOUND?

Common sense may be as common as its name avers. Many have urged that although people may falter in studies of formal logic, they manage the reasoning demands of everyday life with élan. To be sure, one-sided and underdeveloped arguments seem commonplace in, for example, political speeches, newspaper editorials, and everyday conversations. But possibly they stand out because of

their weaknesses; the larger part of everyday reasoning is, perhaps, entirely satisfactory. If everyday reasoning is generally sound, the factors that account for variations in it become a much less interesting issue. What, then, does our decade of research disclose?

Our findings suggest strongly that situation modeling marred by incompleteness and bias is the norm rather than the exception. When asked to reason about vexed political issues of some currency, subjects consistently produce sparse and one-sided situation models. In one study (Perkins, 1985a), first-year high schoolers generated a scant 1.8 lines of argument on their preferred side of the case and 0.6 lines of argument on the other side of the case in response to controversial questions. One might attribute this lackluster performance to the youth of the subjects. However, in the same study, first-year college students and first-year graduate students performed similarly. The college students produced 2.9 arguments and 1.1 objections, and the graduate students produced 3.3 arguments and 1.3 objections.

At least, the graduate students' performance might appear reasonably rich in my-side arguments. But how many lines of arguments are needed to explore an issue? From piloting, we know that at least six lines of argument for each side of the case were presumably within easy access for these subjects. Later scaffolding research bears this out. Therefore, one can conclude that the performances, in general, evinced faulty situation modeling in terms of both bias and depth.

These points emphasize problems of biased and incomplete reasoning. What about the error analysis, which focused on naturalistic critiques of the subjects' arguments? The categorization of critiques revealed that at least three fourths of the identified weaknesses had to do with bias and incompleteness. They were, in the phrase employed earlier, sins of omission, in which the reasoner's line of argument was all right so far as it went but in which the reasoner should have and could have considered other causal chains that would have challenged the line of argument or provided arguments on the opposite side. Thus, the results offer broad support for the notion that everyday reasoning is troubled principally by the bias and incompleteness problems of situation modeling (Perkins, in press).

If informal reasoning is not well done in general, at least the results seem to suggest that it is done better with maturation and education. Older, more educated subjects showed better performance on most of the numerical quality measures. However, this contrast does not mean as much as one might, at first, hope. The effects of education and maturation were examined by comparing the scores of first-year students with those of fourth-year students at the high school, college, and graduate school levels (Perkins, 1985b). At only the high school level did a gain in lines of argument reach statistical significance, and here the gain was still small: 0.4 lines of my-side argument in 3 years. Figures for the whole sample disclose a mean rate of gain in lines of argument of 0.1 per year of education. Other measures demonstrated similarly slow growth rates.

The minimal difference between first- and fourth-year scores *within* high school, college, and graduate school argues that the contrasts *across* high school, college, and graduate school are principally selection effects. In other words, admission processes successively filtered out those with lesser ability, so that, by graduate school, the students as a whole performed much better than students in high school. However, the same graduate students would have performed nearly as well when in high school as they did in graduate school.

Perhaps, of course, formal education is the wrong context in which to seek growth in informal reasoning. People might learn everyday reasoning in the "school of life." To test this hypothesis, the performance of people who had been out of school for a number of years was examined (Perkins, 1985b). Correlations between reasoning measures and years of life experience were negligible. As a broad generalization, the adult, out-of-school subjects performed at about the same level as the school subjects of similar educational achievement.

In summary, the overall pattern of results from several studies argues that everyday reasoning is neither very good nor likely to improve much with maturation, education, or experience of life, at least not beyond the first year of high school, the earliest period sampled. Of course, this conclusion calls for clarification on several points. If maturation, education, and experience of life do not make much difference, what does? Could motivation be the key factor? If so, perhaps people reason well enough when it matters, although in desultory fashion in our experiments. What about the role of the knowledge base? What about the impact of general intelligence? We now turn to these questions.

DOES UNDERMOTIVATION ACCOUNT FOR FAULTY SITUATION MODELING?

Motivation deriving from interest and personal investment in the issues at hand might contribute substantially to good reasoning performance. Indeed, if the results outlined earlier simply reflect low motivation on the part of the subjects, then they present nothing to worry about: People might reason well enough when given good reason to reason. But is motivation the key variable, as one might suppose? The issue of motivation has been part of the series of studies outlined here since the beginning, and at various points along the way the issue of motivation has been specifically addressed.

Subjects in the studies found the issues presented to them moderately interesting. On a rating scale in which 4 was "very interested" and 1 was "not at all interested," across studies most subjects rated their interest at 3, "somewhat interested." Of course, subjects may not have wanted to disappoint the experimenters with their disinterest. But even presuming inflated ratings by subjects, if interest were an influence on reasoning, interest should still correlate with performance. On the contrary, interest consistently failed to correlate significantly with measures of performance. How thoroughly and with what balance a person explored an issue bore little relation to the interest level the person reported.

Subjects' ratings of interest aside, in the one-on-one oral interviews it was

possible to gauge directly how hard the subjects were trying, a difficult matter to judge in the strictly paper-and-pencil procedure utilized for many of the studies. We found little evidence of shyness or inhibition on the part of the subjects; they seemed interested and motivated. During the oral interviews general prompts, such as, "Can you say more?" were given. Subjects appeared willing to try to say more but often found that they had simply "run dry" on an issue. (However, with generic metacognitive prompts subjects produced balanced, quite complete situation models, as we discuss later.)

Of course, these indices of interest do not treat true personal involvement. One might expect better situation modeling in cases in which people are deeply personally invested in an issue. To investigate this question, a sample of 39 adults was enlisted. Subjects were recruited by posters soliciting people in the midst of a major decision concerning such issues as employment, health, education, family, or marital status. The subjects were interviewed about their reasoning on their personal issues as well as on one of the usual social issues. As one might expect, these subjects did, in fact, produce significantly more my-side (total of 4.7) and other-side (total of 4.4) arguments about their own decisions than when asked to reason about a social question ($p < .001$, two-tail t-test) (Perkins, in press).

These results show both a more thorough and a more balanced situation modeling of personal issues. So, when it really counts, do people reason perfectly well, making the measures reported earlier not as discouraging as they might seem? Unfortunately, the picture is more complex than that, for several reasons.

First of all, the good pro–con balance of the subjects' personal arguments can be discounted as artifactual, a limitation of the sampling design we were aware of from the outset: These subjects were located for having vexed personal issues to reason about. *Vexed* means that one finds compelling reasons for both sides. Thus, the present study gives no information on whether people tend to develop more balanced situation models for very important personal decisions. It is perfectly possible that, in everyday life, people approach many important decisions with bias, never see their genuine vexedness, and take hasty action on the basis of simplistic situation models.

The data do suggest that when people perceive a personal issue as both important and vexed, they develop a better elaborated situation model than they evolve for the social issues used in most of our studies. A key factor in this seems to be time commitment. Subjects reported spending an enormous amount of time thinking about their personal issues and talking about them with others—an average of 125 hours. They reported that the average amount of time they had spent thinking about the social issue before the investigators raised it with them was 10.5 hours. What do we make of this? On the one hand, it is encouraging to find that people do develop better situation models in such circumstances. On the other hand, one cannot be too impressed that a greater time investment by a factor of 10 yielded situation models that were richer only by about a factor of 2.

It should be added that, although the personal situation models were better elaborated, the situation models for the social issue were predictive of variations in elaboration. Performance on the personal issue correlated significantly with performance on the social issue ($r = .36$, $p < .01$). This is not a very high correlation, but it is about the same as the test–retest correlation between different social issues. Accordingly, the figures suggest that reasoning about the personal and the social issues tapped the same underlying competency, with enormous differences in preparation time accounting for more elaborated personal situation models.

A final point to bear in mind is that many circumstances in life that call for careful, responsible reasoning are not major life decisions. What political candidate to vote for, what position to take on a referendum, what attitude to take toward a troublesome employee or a demanding boss, what car to buy, and what course of action to recommend to a troubled friend are all situations that invite careful thought but that are not likely to secure 125 hours of painstaking attention or even the 10.5 hours reported for the current social issues. So, the motivation of deep personal involvement, insofar as it promotes better situation modeling, cannot be counted on in many situations in which such modeling is needed.

In summary, what can be said about undermotivation as an explanation for sparse and biased situation modeling? First of all, given an occasion that maintains thinking about an issue for a moderate period (a school board meeting, a family discussion about whom to vote for, our experimental setting), variations in topic interest may not account for variations in modeling. At least, across a variety of studies, no such influence emerged. Second, personal involvement in an issue perceived as vexed does yield a gain of about a factor of two in completeness of situation modeling albeit at a cost of enormous time investment. There is no evidence one way or another as to whether people are more likely to take a balanced view of issues that are very important to them, personally. Finally, many occasions that call for careful reasoning are not crucial life decisions. With all these considerations in mind, it seems that motivation, although plainly important, probably does not fully account for sparse and biased situation modeling or allow us to be sanguine about the quality of reasoning in everyday life.

DOES LACK OF KNOWLEDGE ACCOUNT FOR FAULTY SITUATION MODELING?

Recent research on expertise has emphasized the role of context-specific schemata in performances that might be taken to involve mostly general reasoning abilities (cf. Chase & Simon, 1973; Glaser, 1984; Rabinowitz & Glaser, 1985). With this in mind, it is natural to wonder whether the variations in subjects' reasoning as gauged in the experiments described heretofore might simply reflect their knowledge of the issue. To be sure, an effort was made to choose issues that

students would not actually study as part of formal education. Nonetheless, older and more educated persons would be more likely to know more about the matter by chance, and individuals would, of course, vary in their degree of chance knowledge.

In investigating the role of knowledge, we presumed that the knowledge most directly relevant would be prior experience in thinking about the issue. One might object to this on the grounds that more prior thinking might mean a more settled position on the issue. However, more prior thinking on a controversial issue, much of which would likely be in conversational contexts with more than one side represented, ought to produce more familiarity with the structure of arguments on both sides and thus a better situation model, however settled one's position. Consequently, subjects were asked whether they had thought about the issue at all before. If they responded positively, they were asked to indicate how much total time they had spent thinking about it in terms of number of minutes up to an hour or number of hours.

In general, the studies did not show significant correlations between this "prior-thought" variable and everyday reasoning performance, as measured by the counts of lines of argument, degree of one-sidedness, and quality ratings. In one study, a significant correlation ($r = -.34, p < .001$) appeared between prior thought and the number of prompts given by the experimenter to keep the subject on the topic under consideration. But this relationship did not correspond to the quality of argument per se. However, it is understandable that a person who has thought extensively about an issue is likely to be able to zero in on it better than a person who is exploring it for the first time.

In sum, no evidence emerged for a relation between the quality of situation modeling and prior familiarity with the issues. But how could it be that prior knowledge does not make a difference? The results suggest that knowledge of the issues did not vary extremely widely in the subject population; had it done so, surely situation modeling would have been affected.

DO VARIATIONS IN GENERAL INTELLIGENCE AFFECT SITUATION MODELING?

As mentioned earlier, the studies reviewed here have found little improvement in everyday reasoning ability in relation to either age or education. Such findings suggest that skill in everyday reasoning may be primarily a function of innate intellectual competence. The relation of IQ to everyday reasoning skills speaks to this issue.

Regression analyses of pooled student and pooled nonstudent data from the major study reported by Perkins (1985a) provided estimates of the role of IQ. For the student sample, sizable correlations emerged between IQ and total number of arguments ($r = .47, p < .001$) and between IQ and other-side arguments ($r = .32, p < .001$), the latter considerably smaller than the former. The nonstudent sample yielded a correlation between IQ and total number of arguments of a somewhat lesser magnitude ($r = .29, p < .06$) and a nonsignificant relationship between IQ and other-side arguments ($r = .10$, n.s.). Subsequent studies revealed similar correlations (r = approximately .3) between my-side arguments and IQ.

The performance-versus-competence study, in which subjects were provided with generic cues to stimulate performance, revealed a curious relationship between IQ and argument production. Although IQ correlated significantly with my-side arguments produced without prompting ($r = .53, p < .02, n = 20$), its correlation with other-side arguments was nonsignificant and negative ($r = -.18, n = 20$). This study was relatively small, but it points up an intriguing pattern of correlations with IQ that is corroborated by data from other studies. To pursue the issue, we pooled the performances of 99 students ranging from first-year liberal arts undergraduates to first-year law school students and graduate students. Analysis of the pooled data disclosed significant correlations between IQ and my-side arguments ($r = .37, p < .001$) but virtually no correlation between IQ and other-side arguments ($r = .08$, n.s.).

In sum, certainly IQ makes a contribution to well-elaborated situation models, but there is an important caveat. It appears that people selectively tap their IQ when reasoning. The pattern that emerged shows weak to moderate positive correlation between intelligence as measured by IQ and my-side arguments, and virtually no correlation between IQ and other-side arguments. Only in one sample of the large Perkins (1985a) study do we find significant correlations between IQ and other-side arguments, and these are considerably lower than the correlations between IQ and my-side arguments. In effect, people invest their IQ in buttressing their own case rather than in exploring the entire issue more fully and evenhandedly.

DOES KNOW-HOW ABOUT REASONING IMPROVE SITUATION MODELING?

Both the problem of one-sidedness and the fact that IQ measures account for only a modest part of the variance in situation modeling demonstrate room for a major role for know-how about reasoning. On the one hand, if people knew better how to handle situation modeling, perhaps they would do a better job of it. On the other hand, it is also possible that people typically reason close to some kind of performance ceiling. They are doing as well as they can, at least when not afforded the luxury of many weeks to think out a problem, as with the subjects who were making major life decisions.

This question was examined by a number of studies in which the authors sought to modify the momentary or long-term know-how of subjects by giving

directions or teaching strategies (Perkins, Bushey, & Farady, 1986). For example, a study was designed in which an experimenter presented high school subjects with an issue and instructed them outright to give as many reasons as possible, even if they seemed insignificant, and to give reasons on both sides of the case. The subjects' performance following these instructions was then compared to their previous, unaided performance. In light of the initial performance, subjects' simply being told to generate many reasons and to attend to the other side of the case evoked little change in number of my-side arguments but a substantial increase (from 0.8 to 2.0, or a gain of 150%) in other-side arguments. This study demonstrates that some know-how about reasoning can be transmitted and improve situation modeling simply by changing the broad demand characteristics of the task. It is interesting to note, however, that just asking for more my-side reasons did not yield gains.

If explicit directions can improve momentary performance so easily, at least on the other side of the case, what about longer term instructional programs that support and encourage this sort of reasoning more than conventional instruction? With this and related questions in mind, we examined the impact on students' reasoning of a term in a high school debate class; the first year of a liberal arts program noteworthy for its efforts to develop general reasoning skills; a semester at a graduate school of education, focusing on students taking a course that encouraged exploratory reasoning; and the first year of law school.

The preinstruction levels of performance were lackluster and resembled those discussed earlier (Perkins, 1985a) of conventional education at the high school, college, and graduate school levels. On the posttests, only the debate class and the liberal arts program showed significant gains. These gains were modest: one line of my-side argument for the debate class and 0.6 lines of my-side argument for the liberal arts program. Nonetheless, such figures are substantially greater than the average rate of 0.1 lines of argument per year of education from the earlier studies. Interestingly, the gains appear only on the my-side part of the performance, even though, at least in the debate class, emphasis constantly fell on preparation for arguing either side of the case.

Another study (Perkins et al., 1986) specifically examined gains in situation modeling in response to metacognitive prompts. Subjects reasoned about an issue until they had no more to say. They were pressed on this point, to be sure that they were not holding back reasons that they did not think worth mentioning. Then the investigator offered generic advice. For instance, when subjects mentioned no reasons on the side of the case opposite their own, the investigator said, "You've mentioned some reasons why . . . [the conclusion]. Can you think of any reasons, even though you might disagree with them, why someone might say . . . [opposite conclusion]?". This and other prompts not specific to the issue served as scaffolding to support subjects' continued exploration of the issue. The scaffolds consisted of questions or instructions that, in principle, a person could ask himself or herself.

When the investigator provided the generic know-how after subjects thought they had run dry, they were able to improve on their initial performances dramatically in terms of both bias and completeness. In response to scaffolding, my-side arguments increased by 109% (3.8 new lines of argument), and other-side arguments increased by 700% (4.9 new lines of argument). Because people can balance and elaborate their situation models with generic prompts from an investigator, it follows that, in principle, they can likewise scaffold themselves to perform better. Their premature running dry, despite the experimenter's general urging to say more if they could and their apparent willingness to comply, suggests that they did not know how to scaffold themselves to be more generative.

But can people learn to scaffold themselves to reason more fully? To address this question, a short, high-school-level course in situation modeling was devised. The course consisted of 16 class sessions given at the rate of four 1-hour lessons per week. This was viewed as quite a short intervention, considering that we hoped to affect well-entrenched habits of mental organization. Class content consisted of a variety of exercises, designed to facilitate generativity and attention to both sides of the case. For instance, subjects analyzed the arguments contained in brief essays, wrote short arguments, and learned several strategies to facilitate skilled situation modeling. In essence, this intervention sought to provide basic know-how about generating arguments that are both evenhanded and thorough.

The intervention had its principal impact on bias in situation modeling. Along with a small, nonsignificant increase in my-side arguments, subjects showed a significant increase in other-side arguments, producing one other-side argument before instruction but two afterwards ($p < .001$, one-tailed t test). This is not all that much, but it is more than an order of magnitude greater than the growth rate of reasoning performances without any special intervention. We also believe, in retrospect, that much stronger interventions are possible through even more direct attention to strategies and task demands than these teaching experiments offered. It should be added that there is no presumption of lasting effects or transfer beyond the instructional context. The principal aim of the experiment was to demonstrate simply that change in solo performance is possible. Given evidence that it is, we argue that lasting, transferable change is possible with sufficient effort.

These results join with a larger corpus of findings that suggests that cognitive abilities of various sorts can be enhanced by instruction that emphasizes metacognitive awareness and the use of strategies. For example, in the area of reading, Palincsar and Brown (1984) have demonstrated striking gains by students of low reading ability in their retention and transfer of ability to a variety of settings of reading. Schoenfeld (1982; Schoenfeld & Herrmann, 1982) has reported substantial improvement in college students' mathematical problem solving through an intervention that emphasizes heuristics and metacognitive self-

monitoring. Further results of this sort may be found in, for instance, Nickerson, Perkins, and Smith (1985).

In summary, a variety of experiments argue that situation modeling can be improved by direct or even semidirect instruction. Educational intervention can improve the completeness of situation modeling and biased situation modeling as well. The scaffolding study, in particular, forecasts that quite dramatic improvements are possible if subjects learn to scaffold themselves fairly fully in reasoning about an issue. Good situation modeling is, in substantial part, a matter of know-how.

WHY PEOPLE REASON AS THEY DO: A MAKES-SENSE EPISTEMOLOGY

This survey of several factors that impinge on situation modeling has turned up a couple of puzzles that call for resolution. First of all, it appears that people could easily develop situation models much better than they do. Why the shortfall? Second, intelligence in the psychometric sense—IQ—appears not to contribute as straightforwardly to good situation modeling as one might suppose. Why not?

The previous section suggests a broad answer to both of these questions. Good situation modeling involves a certain amount of know-how. You need to know about the traps of one-sided thinking, ways to provoke yourself to think more thoroughly when you seem to be running dry, and so on. The shortfall in situation modeling reflects a lack of this metacognitive know-how. The vexed relation to IQ also is explained: More raw brain power, of the sort that IQ supposedly measures, does not necessarily yield better situation modeling unless it is wisely deployed. The know-how to do so is needed, too.

Okay as far as it goes, this explanation in terms of know-how nonetheless seems fundamentally dissatisfying. The know-how in question is not that esoteric. Why do people not figure it out for themselves? Why are they not taught it in schools? Why do they not pick it up from parents and friends? When nearly every person almost effortlessly learns to speak a native language and navigate in the complex environment of his or her town or city, why does some relatively simple know-how about good situation modeling come so hard? It is almost as though there were something in the way.

Perhaps what is in the way is a competing standard for good situation modeling. When people treat vexed issues like nonvexed issues and generate the barest of situation models, presumably they have satisfied their criterion for "true" or adequate situation modeling. The criterion appears to be that the model "makes sense": It hangs together well and displays high congruence with one's most prominent prior beliefs. Notice how little of a situation model this standard requires. A bald assertion of a claim without reasons can hang together well and display high congruence with prominent prior beliefs. In such a case, the claim seems self-evident. Even if a person has to work a little harder to build a model,

the person only has to get to the point of telling one story about the situation that weaves together the facts in one way, from one point of view, congruent with the person's prior beliefs. Then the model "makes sense." When sense is achieved, there is no need to continue. Indeed, because further examination of an issue might produce contrary evidence and diminish or cloud the sense of one's first pass, there is probably reinforcement for early closure to reduce the possibility of cognitive dissonance. Such a makes-sense approach is quick, easy, and, for many purposes, perfectly adequate.

To put a name to this syndrome, we have written before of people having a *makes-sense epistemology* (Perkins, Allen, & Hafner, 1983). A person's epistemology refers to the person's tacit or explicit grounds for belief. A makes-sense epistemology is one in which the primary criterion of truth is making sense in the sense just described. This is, of course, not an epistemology that a person harbors explicitly as an overt philosophy. It is a default epistemology, a pattern that the mind falls into as the simplest, more-or-less functional thing to do.

What makes this makes-sense epistemology so robust? A number of factors can be suggested. First of all, it is quick, easy, and, for many purposes, perfectly adequate. The criteria of hanging together and matching prominent prior beliefs provide a fairly strong filter against bad models, particularly if one has had considerable experience in the domain in question. Second, the makes-sense epistemology suits well the character of the human organism as an information processor of significantly limited capacity (cf. Newell & Simon, 1972): People can stop thinking about something after achieving superficial sense rather than pressing on to more complications. Third, the makes-sense epistemology provides a line of ego defense. If one thinks beyond what makes superficial sense, one may find oneself pressed to question cherished beliefs about who one is or what the world is like (cf. Paul, 1986, 1987).

Two puzzles were introduced earlier: why people do not build better situation models when they easily can do so and why intelligence in the sense of IQ does not contribute more straightforwardly to good situation modeling. A first answer points out that metacognitive know-how is crucial to good situation modeling but is apparently lacking. But, why? The notion of a makes-sense epistemology helps us to explain why. The need for this metacognitive know-how is not very salient, because a makes-sense epistemology keeps people pretty happy with their beliefs and adequately functional most of the time.

HOW PEOPLE COULD REASON BETTER: A CRITICAL EPISTEMOLOGY

It is useful to have a contrasting notion to the makes-sense epistemology just outlined. One might speak of a *critical epistemology* (cf. Perkins et al., 1983). A critical epistemology incorporates higher standards for good situation models. It is not enough for a particular story about a situation to hang together: One must

consider what other, rather different stories might also hang together. It is not enough for a particular story to match one's prominent prior beliefs: One should check one's data base of information and experience more thoroughly for inconsistencies. Often, one should seek further information. As we construe it, a critical epistemology includes not only this broad know-how about which criteria to use for a good situation model but also heuristic know-how about how to build a better model, for instance, how to think of more reasons, construct counterexamples, and so on. In addition, a critical epistemology incorporates epistemic feelings and values about objectivity, fair play, the importance of taking multiple perspectives, and so on. In similar spirit, Scheffler (1982) wrote of the importance of "cognitive emotions" in guiding cognition, and Paul (1986, 1987) wrote of the importance of "strong sense" critical thinking, in which the thinker willingly and objectively engages controversies between competing value and belief systems.

Of course, by describing this critical epistemology we do not mean to create an either-you-have-it-or-you-don't dichotomy between a makes-sense and a critical epistemology but rather to define a direction of development. An epistemological continuum exists, reflecting the soundness of the justifications offered for the adequacy of situation models. These justifications range from bald, intuitive assertions on the extreme makes-sense side to more complex, critical models in which an issue may be looked at from a variety of perspectives and may be adjusted for inclinations in the direction of overgeneralization, bias, and other common pitfalls of reasoning. Obviously, one can have a critical epistemology in various respects and to various degrees. Also, everyday experience teaches that we, as individuals, vary from occasion to occasion in the care we take as practicing epistemologists. Sometimes, harried and hurried, the wisest person falls into the pattern of a makes-sense epistemologist. Likewise, not very agile thinkers, in circumstances that neither arouse their prejudices nor provoke indifference, may display much more of a critical epistemology than they usually do.

As noted earlier, a makes-sense epistemology works pretty well most of the time. Perhaps it is worth a moment to stand back from our enterprise of trying to explain why people reason as they do and contemplate how people should reason and why. Critical epistemologists clearly work harder than makes-sense epistemologists. Why should they bother? A makes-sense model is often adequate. Or is it? Makes-sense epistemology is obviously adequate for making the plethora of little decisions that arise during the course of a life. But people do not simply adjust or fit in to a life. The more vigorous and compelling strands of modern psychology and philosophy hold that reasoning is, in fact, constitutive of what we take our lives to be. We do not just stub our toes on reality; we also constitute or generate it. So, a makes-sense epistemology is constitutive of a makes-sense sort of reality, in which a makes-sense sort of life is perfectly at home.

A makes-sense epistemology becomes inadequate, though, when an unusual situation arises in which there is little obvious congruence with one's current

beliefs. To make sense of such dilemmas, a hard-core makes-sense epistemologist must work to avoid seeing incongruities even when they virtually abound. Makes-sense epistemologists must conceive a life course that is relatively simple and straightforward, and they struggle inefficiently when messy problems arise. Unfortunately, into each life a few messy problems do fall, be they puberty, divorce, obstreperous children, car repairs, or death. Moreover, modern times are fraught with complex decisions about environmental, economic, and other issues that affect everyone. Because it tends to gloss over incongruities, a makes-sense epistemology leaves people ill-equipped for decisions they have to make in such serious contexts. Because it is inadequate for dealing with problems that afford no solutions that feel right, a makes-sense epistemology constrains people from effectively helping either themselves or others when the needs are greatest.

And what of critical epistemologists? Critical epistemologists must also make sense. To fail to do so is to become incoherent to oneself and to one's fellow travelers, which is no help to anyone. But sense that is critically made is generally richer in its implications, in the constraints that it recognizes, and in the options that it suggests than the sense to which people usually aspire. The endeavor of critical thought approaches a situation not with the goal of avoiding options and possibilities but with an inclination to create them through a deep understanding of the situation.

The habits of the critical epistemologist equip him or her for even the most difficult of decisions. These decisions cannot, of course, be made easily, but by being able to generate several alternatives and consider more than one point of view, a person can feel satisfied that he or she has made a reasonable decision. A well-reasoned decision is less likely to produce surprise or impotence in the face of its consequences. Further, if we grant that people do really make the meaning that is experienced as one's life, then it is clear that critical epistemologists, individually and collectively, generate qualitatively different kinds of lives than makes-sense epistemologists. Critical epistemology leads to the construction of experience that is richer in possibilities and more manageable. Whether critical epistemologists are *happier* we do not venture to say. But it seems plain that their better models of the world afford more perspective, variety, and control over fate and fortune.

THE ROOTS OF INTELLIGENCE

What are the roots of intelligence, broadly construed? This question has been the leitmotif for the present review. Everyday reasoning is plainly an important manifestation of intelligence. To examine what factors in the psychological makeup of the individual make everyday reasoning more or less effective is to examine the roots of intelligence.

This review of research has sought to disentangle several likely contributing

factors: interest in issues, knowledge of specific issues, IQ, and metacognitive know-how. In embarking on such an enterprise, one can imagine a worst-case scenario in which the findings indicate that everything contributes moderately in a thoroughly entangled way, and no clear discriminations can be made. Indeed, as a point of logic, such an outcome might seem almost inevitable. How could interest, knowledge, IQ, and metacognitive know-how *not* each count enormously? Even to ask the question is to evoke analogies like asking whether motors or wheels or gasoline are more important to a car: All seem transparently part of the system that gets the result. However, despite these discouraging signs, fairly clear answers have emerged about the relative contributions of these components to everyday reasoning. What are these answers, and why do they come out the way they do?

Prior Knowledge

The findings indicate a relatively minor role of prior knowledge about particular issues to account for differences in good situation modeling. But how can this be, particularly considering the contemporary research on expertise that has emphasized the importance of domain-specific schematic repertoires? In part, the result is an artifact of experimental design: Issues were sought in which no subject group was likely to have high expertise. However, there is a broader reason as well. By its very nature, everyday reasoning constantly engages people in reasoning about problems in which they lack expertise. We have to vote for candidates, consider referenda, advise friends, deal with squabbles in the work place, make major purchase decisions, and so on. We cannot possibly accumulate real expertise in all these things. In other words, everyday reasoning, because of its eclectic reach, intrinsically involves reasoning as best one can with one's general knowledge and experience about issues in which one probably does not have a great deal of specific expertise. Moreover, the teaching and scaffolding experiments show that the same general knowledge base can be brought to bear much more or less fully on an issue, depending on how vigorously and evenhandedly a reasoner probes.

Interest

The results suggest that interest plays a moderate role in building situation models but not nearly as important a role as one might suppose. How can this be? First of all, interest does not necessarily imply good situation modeling: A person may have strong interest yet bring a strong initial bias to a situation, or a person may have strong interest and no great initial bias yet fail to perceive how a situation is vexed and hence not engage in a serious effort to build a careful situation model. Moreover, even if interest straightforwardly promoted good situation modeling, life presents an abundance of situations in which we ought to

reason well yet do not have strong interest—or, at least, not self-interest. The motivation, thus, has to come from a sense of responsibility.

IQ

The results show a definite influence of IQ on situation modeling, but a somewhat oblique influence. Higher IQ correlates with more generativity on one's preferred side of a case. However, situation models produced by individuals with higher IQs tend to be more biased, not less, and still fall well short of thorough elaboration even on the preferred side of the case. How can this be? As discussed earlier, the sort of potency measured by IQ can be deployed more or less wisely. Having a high IQ gives no guarantee of one's using it well any more than having alot of horsepower under the hood of your car guarantees your driving the car well.

Metacognitive Know-How

The results argue that metacognitive know-how about the demands, opportunities, and pitfalls of situation modeling has a very substantial influence. Attitudes of objectivity, fair play, and so on, together with this know-how, comprise a critical epistemology that fosters good situation modeling. Moreover, such know-how can be taught, and such attitudes can be fostered.

How is it, then, that we do not see as much of this critical epistemology as we would like in everyday reasoning? Our suggestion is that a makes-sense epistemology, with its tacit criteria that models hang together and match prominent prior beliefs, dominates much of everyday reasoning and, indeed, serves tolerably well in many situations. This makes-sense epistemology stands in the way of people feeling a strong press for and trying to develop a more sophisticated critical epistemology.

Here, of course, is where education can help. In normal educational practice, there is very little direct teaching of a critical epistemology—in fact, there probably is some undermining of it. Yet the opportunity to do better is clear. As noted earlier, not only the present line of investigation but a number of others have produced results suggesting that cognitive skills in general can be improved. Moreover, in many ways, the know-how and attitudes of a critical epistemology are fairly accessible—much simpler in outline than trigonometry. Finally, an effort to teach something for which many people feel no great need is hardly out of place in educational practice: A good deal of education seeks to teach people things they feel no great need to know, such as trigonometry. Indeed, the people may often be right. In the case of a critical epistemology, however, we urge that they are wrong. Because a makes-sense epistemology means that people are not likely to seek out a critical epistemology for themselves, it becomes the responsibility of education to add the critical edge to everyday reasoning.

REFERENCES

Baron, J. (1978). Intelligence and general strategies. In G. Underwood (Ed.), *Strategies in information processing* (pp. 403–450). London: Academic Press.

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

Borgatta, E. F., & Corsini, R. J. (1964). *Quick Word Test*. New York: Harcourt, Brace & World.

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology, 54*, 1–22.

Chase, W. C., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4*, 55–81.

Engel, S. M. (1976). *With good reason: An introduction to informal fallacies*. New York: St. Martin's Press.

Ennis, R. H. (1986). A taxonomy of critical thinking dispositions and abilities. In J. B. Baron & R. S. Sternberg (Eds.), *Teaching thinking skills: Theory and practice* (pp. 9–26). New York: W. H. Freeman.

Fearnside, W. W., & Holther, W. B. (1959). *Fallacy: The counterfeit of argument*. Englewood Cliffs, NJ: Prentice-Hall.

Gardner, H. (1983). *Frames of mind*. New York: Basic Books.

Glaser, R. (1984). Education and thinking: The role of knowledge. *American Psychologist, 39*, 93–104.

Greenfield, P. M. (1984). A theory of the teacher in the learning activities of everyday life. In B. Rogoff & J. Lave (Eds.), *Everyday cognition: Its development in social context* (pp. 117–138). Cambridge, MA: Harvard University Press.

Guilford, J. P., & Hoepfner, R. (1971). *The analysis of intelligence*. New York: McGraw-Hill.

Jensen, A. R. (1983). The nonmanipulable and effectively manipulable variables of education. *Education and Society, 1*(1), 51–62.

Jensen, A. R. (1984). Test validity: *g* versus the specificity doctrine. *Journal of Social and Biological Structures, 7*, 93–118.

Newell, A., & Simon, H. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Nickerson, R., Perkins, D. N., & Smith, E. (1985). *The teaching of thinking*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction, 1*, 117–175.

Paul, R. (1986). Dialogical thinking: Critical thought essential to the acquisition of rational knowledge and passions. In J. Baron & R. Sternberg (Eds.), *Teaching thinking skills: Theory and practice* (pp. 127–148). New York: W. H. Freeman.

Paul, R. (1987). Critical thinking and the critical person. In D. Perkins, J. Lochhead, & J. Bishop (Eds.), *Thinking: The second international conference* (pp. 373–403). Hillsdale, NJ: Lawrence Erlbaum Associates.

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

Perkins, D. N. (1985b). Reasoning as imagination. *Interchange, 16*(1), 14–26.

Perkins, D. N. (1986a). Thinking frames. *Educational Leadership, 43*(8), 4–10.

Perkins, D. N. (1986b). Thinking frames: An integrative perspective on teaching cognitive skills. In J. B. Baron & R. S. Sternberg (Eds.), *Teaching thinking skills: Theory and practice* (pp. 41–61). New York: W. H. Freeman.

Perkins, D. N. (in press). Reasoning as it is and could be. In D. Topping, D. Crowell, & V. Kobayashi (Eds.), *Thinking: The third international conference*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Perkins, D. N., Allen, R., & Hafner, J. (1983). Difficulties in everyday reasoning. In W. Maxwell (Ed.), *Thinking: The frontier expands* (pp. 177–189). Hillsdale, NJ: Lawrence Erlbaum Associates.

Perkins, D. N., Bushey, B., & Farady, M. (1986). *Learning to reason* (Final report for grant no. NIE-G-83-0028). Cambridge, MA: Harvard Graduate School of Education.

Rabinowitz, M., & Glaser, R. (1985). Cognitive structure and process in highly competent performance. In F. D. Horowitz & M. O'Brien (Eds.), *The gifted and talented: Developmental perspectives* (pp. 75–98). Washington, DC: American Psychological Association.

Rogoff, B., & Gardner, W. (1984). Adult guidance of cognitive development. In B. Rogoff & J. Lave (Eds.), *Everyday cognition: Its development in social context* (pp. 95–116). Cambridge, MA: Harvard University Press.

Scheffler, I. (1982). In praise of the cognitive emotions. In I. Scheffler (Ed.), *Science and subjectivity* (2nd ed.) (pp. 139–157). Indianapolis, IN: Hackett.

Schoenfeld, A. H. (1982). Measures of problem-solving performance and of problem-solving instruction. *Journal for Research in Mathematics Education, 13*(1), 31–49.

Schoenfeld, A. H., & Herrmann, D. J. (1982). Problem perception and knowledge structure in expert and novice mathematical problem solvers. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8*, 484–494.

Slossen, R. L. (1981). *Slossen Intelligence Test*. New York: Slossen Educational Publications.

Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. New York: Cambridge University Press.

Toulmin, S. E. (1958). *The uses of argument*. Cambridge, England: Cambridge University Press.