Beyond the Purely Cognitive: 
Belief Systems, Social Cognitions, and 
Metacognitions As Driving Forces in 
Intellectual Performance*

ALAN H. SCHOENFELD
The University of Rochester

This study explores the way that belief systems, interactions with social or experimental environments, and skills at the "control" level in decision-making shape people's behavior as they solve problems. It is argued that problem-solvers' beliefs (not necessarily consciously held) about what is useful in mathematics may determine the set of "cognitive resources" at their disposal as they do mathematics. Such beliefs may, for example, render inaccessible to them large bodies of information that are stored in long-term memory and that are easily retrieved in other circumstances. In other cases, individuals' reactions to an experimental setting (fear of failure, or the desire to "look mathematical" while being videotaped) may induce behavior that is almost pathological—and at the same time, so consistent that it can be modeled. In general, such "environmental" factors establish the context within which individuals access and utilize the information potentially at their disposal.

Protocols illustrating these points are presented and discussed. A model based on an axiomatization of students' beliefs about plane geometry is outlined, and is shown to correspond closely to their problem-solving performance. A framework is offered for analyzing problem-solving performance at three qualitatively different levels: access to cognitive resources stored in LTM, executive or control decision-making, and belief systems.

This discussion is one of two whose purpose is to delineate a series of psychological and methodological issues related to the use of verbal methods (clinical interviews and protocol analyses) for research into human problem-solving processes. Both works are based on the same foundation, the

*This is the second revised version (January, 1983) of a paper presented at the annual meetings of the American Educational Research Association, New York, April, 1982.

The research described in this paper was supported by grant number SED 79-19049 from the National Science Foundation.
premise that "purely cognitive" behavior is extremely rare, and that what is often taken for pure cognition is actually shaped—if not distorted—by a variety of factors. The companion work, "Making Sense of 'Out Loud' Problem-solving Protocols" (in press-b), discusses a number of variables that affect the generation and interpretation of verbal data. These included the number of persons solving a problem, the nature of the instructions to verbalize, and how comfortable the subject feels in the experimental environment. This paper tries to place such methodologies in a much broader context, in an attempt to explicate some of the "driving forces" that generate the behaviors that we see. The thesis advanced here is that the cognitive behaviors we customarily study in experimental fashion take place within, and are shaped by, a broad social-cognitive and metacognitive matrix. That is, the tangible cognitive actions produced by our experimental subjects are often the result of consciously or unconsciously held beliefs about (a) the task at hand, (b) the social environment within which the task takes place, and (c) the individual problem-solver's perception of self and his or her relation to the task and the environment. I shall argue that the behaviors we see must be interpreted in that light.

This is an exploratory discussion, an attempt to characterize some of the dimensions of the matrix within which pure cognitions reside. The discussion takes place in two parts. The first part outlines the three qualitatively different levels of analysis that I think may be needed to fully make sense of verbal data, even when one's intentions are "purely cognitive." These levels are described in the next section, and a brief analysis of some protocols from that perspective is then given. In the second part, the discussion is broadened, and I try to suggest some of the dimensions of the matrix. Much of what follows is highly speculative, and a good deal of the "evidence" is anecdotal. The idea is to point out some of the pitfalls in current lines of inquiry and to map out some useful directions for future inquiry.

THE ANALYSIS OF VERBAL BEHAVIOR

Background and Framework

I wish to suggest here that three separate categories of analysis may be necessary in order to obtain an accurate characterization of subjects' problem solving performance from the analysis of "verbal data" that they produce while solving problems. The outline of those categories is given in Fig. 1.

There are, of course, many levels of analysis beyond those discussed here. At the microscopic level, see Monsell's (1981) review of what he calls the "nuts and bolts of cognition": representations, processes, and memory
Category I: **Resources** (Knowledge possessed by the individual, that can be brought to bear on the problem at hand)
- Facts and algorithms
- Relevant competencies, including the use of routine procedures, "local" decision making, and implementing "local" heuristics

Category II: **Control** (Selection and implementation of tactical resources)
- Monitoring
- Assessment
- Decision-making
- Conscious metacognitive acts

Category III: **Belief Systems** (Not necessarily conscious determinants of an individual's behavior)
- About self
- About the environment
- About the topic
- About mathematics

Figure 1. Three qualitatively different categories of knowledge and behavior required for a characterization of human problem-solving.

mechanisms. At the very macroscopic level, there is a broad set of social cooperative behaviors within which "real" problem-solving actions often take place. "Real world" problem-solving, too, is beyond the scope of this study. Here we shall focus on analyzing the protocols obtained from students under relatively ideal laboratory situations.

The three categories listed in Fig. 1 will be described later. As background, however, it is important to characterize some of the defining properties of decision-making in the first two categories, "resources" and "control." Roughly, the distinction is as follows. A control decision is a global choice, one that in a substantive way affects the direction of a problem solution and the allocation of resources to be used in a solution. Such "strategic" decisions include selecting goals and deciding to pursue or abandon particular (large-scale) courses of action. In short, they are decisions about what to do in a solution. Suppose, for example, that a student working on a problem decides to calculate the area of a particular region, or to "look at an easier related problem." If doing so will occupy, say, five or more of the allotted twenty minutes for solving the problem, that is a control decision: spending that much time in that endeavor may "make or break" the solution. Once such a decision has been made, then any decisions about how to implement that choice—for example, whether to calculate the dimensions of the region by trigonometry or analytical geometry or make a selection of which easier related problem to explore—are matters of resource selection. These "how to" decisions will be called "tactical" choices. Note that in the case of "choosing an easier related problem," the
implementation of a problem-solving heuristic is considered a tactical matter. This is nonstandard. A brief elaboration of the three categories follows.

**Resources.** This category is quite broad, including as subcategories the range of facts and procedures that are available to the individual for implementation in a problem solution. Many of the relevant issues are characterized in Simon's (1979) review article. Simon, primarily concerned with psychological and AI simulations of expert problem-solving performance in semantically rich domains, describes the key issues as follows: "The central research questions are two: (a) how much knowledge does an expert or professional in the domain have stored in LTM [long-term memory], and (b) how is that knowledge organized and accessed so that it can be brought to bear on specific problems?"

To begin with, one needs to know what domain-specific knowledge is accessible to the problem solver. If (see Protocols 1 and 2) a student is solving a straightedge-and-compass construction problem from plane geometry, does he or she know that the radius of a circle is perpendicular to the tangent line at the point of tangency? Whether the student chooses to use that fact is another matter and will be discussed later. But (clearly) a solution that depends on that particular piece of knowledge may evolve in radically different ways if the student does or does not have it, and an evaluation of the solution depends on an adequate characterization of the knowledge base. Similar comments apply to procedures relevant for the solution of the problem. In the example just cited, does the student know how to construct a perpendicular to a given line through a given point? If the student does not recall the construction, does he or she know that it can be done, so that deriving the construction is a possibility? Or must that too be discovered? These factors determine the potential evolution and characterization of a problem-solving session.

After the question of the possession of factual and procedural knowledge comes the question of access to it. The student may know that similar triangles have certain properties, for example, but will the student "see" or even look for similar triangles in a particular circumstance? Much "expert" performance in given domains is attributed to the possession of certain problem-solving schemata; this is, indeed, the foundation of much AI research. Questions of how to represent such "compiled" knowledge are open. Among the approaches to representation "particularly worth describing [are] the predicate calculus, production systems, semantic networks, and frames" (Walker, 1981). All of these approaches take as given that there are certain regularities in experts' perceptions of problem situations and of appropriate behavior in them. This perspective is substantiated in various ways in the literature, for example, with experimental results that experts in physics (Chi, Feltovich, & Glaser, 1981) and mathematics (Schoenfeld & Herrmann, 1982) see through the "surface structure" of
problems to perceive "deep structure" similarities and approach the problems accordingly. Moreover, students develop problem schemata that may or may not be consistent with those of experts (Hinsley, Hayes, & Simon, 1977; Silver, 1979), and these schemata change with experience (Schoenfeld & Herrmann, 1982). (For a characterization of the role of schemata in students' mathematical problem-solving performance, see Silver, 1982.)

There is yet one more level of tactical behavior, that of implementing certain problem-solving heuristics. Examples of these will be seen in Protocols 1 and 2. In a sense, these are nearly on a par with domain-specific schemata. For example, "it is useful to assume that one has the desired object and then to determine the properties it must have" is a heuristic typically valuable in straightedge-and-compass constructions. Its domain-specific implementation (draw the figure and see what properties it has) is quite similar to the implementation of domain-specific schemata, such as "look for congruent triangles when faced with a problem of this nature." These heuristics, like the other categories of knowledge already described, fall into the category of tools potentially accessible to the problem-solver.

An inventory of these tools provides a characterization of what the problem-solver might be able to use in solving a problem. Since we are dealing with the "real" behavior of students rather than the idealized behavior of experts, there are no guarantees that these resources will be called upon, even if it is appropriate for the problem-solver to do so: this is where our analysis diverges from Simon's, as quoted earlier. Observing which of the tools potentially accessible to the problem-solver are selected or discarded, how such decisions are made, and how such choices affect the problem-solving process as a whole, constitutes the next level of analysis.

"Control" knowledge and behavior. In my "Episodes" paper (in press-a), two students are asked to determine the characteristics of the largest triangle that can be inscribed in a given circle. They guess that the equilateral triangle is the solution, and set out to calculate its area. The calculations get rather messy, and they are still calculating when the 20-minute videotape runs out. When they are then asked what good having the area of the equilateral triangle will do them, they cannot say. Yet their entire solution was determined by their decision to undertake the computation.

This is an extreme (although not atypical) example of what might be called an executive or control malfunction: one bad decision, unmonitored and unchecked, dooms an entire solution to failure. What the students actually knew, and what they might have done if given the opportunity to use that knowledge, becomes a moot question. So long as they pursued that computation, whatever else they knew was useless to them. In contrast, the same paper offers a protocol taken from a mathematician working on an unfamiliar problem in geometry. He generates at least a dozen potential "wild goose chases," but rejects all of them after brief consideration. With
some clumsiness, he solves a problem the students did not—although he began working on the problem with much less domain-specific knowledge than the students “objectively” had at their disposal. It can be argued that the expert’s success and the students’ failure were due, respectively, to the presence and absence of productive “metacognitive” behaviors.

One of the early researchers to stress the importance of metacognition as a major factor in cognitive performance, Flavell (1976) characterized it as follows:

I am engaging in metacognition . . . if I notice that I am having more trouble learning A than B; if it strikes me that I should double-check C before accepting it as a fact . . . metacognition refers, among other things, to the active monitoring and consequent regulation and organization of these processes to the cognitive objects on which they bear. (p. 232)

For the most part, research in artificial intelligence has not dealt directly with issues of metacognition as they are characterized here. This is a subtle point, since many of the terms used in metacognition overlap with those used in AI (see Brown’s definition in a later section). But the usages differ. Consider, for example, skilled problem-solving in physics as modeled by production systems (Larkin, McDermott, Simon, & Simon, 1980). The idea is to model competent behavior in sufficient detail to be able to select the “appropriate” behavior, a certainly enormous task. But issues of the type that humans encounter when working on such problems—“I’ve been doing this for five minutes, and it doesn’t seem to be getting me anywhere; should I, perhaps, take an entirely different perspective?”—are not the focus of such programs. They model behavior where such problem-atic performance is not a “problem.”

Likewise, there are difficult issues of strategy selection in any reasonably sophisticated program. But the use, for example, of “conflict resolution strategies” to determine precisely which production will “fire” when the conditions for more than one production have been met, still operates at a very different level than the one under consideration here. Few programs deal with planning and monitoring at that level, although there are many “planning” programs. Typical planning procedures call for leaving sequences of actions unspecified until one is constrained to specify their order, and checking for conflicts when one does so. A standard example is Sacerdoti’s (1977) task, “paint the ladder and the ceiling.” If one tries to proceed in that order, painting the ladder precludes painting the ceiling. “Planning” means specifying actions in efficient temporal order. Sacerdoti’s “nets of action hierarchies” are designed to allow for fleshing out plans in such a way that such impasses are avoided. This whole perspective, however, assumes that one works in domains where plans are there to be “fleshed out”—certainly not a universal condition in problem-solving.
One model that does not make such assumptions is the Hayes-Roth's (1979) model. That model is many-leveled and, if it is appropriate, shifts rapidly from considerations at one level (do B before A, instead of the other way round) to another (revising the entire plan structure because of an unforeseen major difficulty). This "opportunistic" model is highly structured, but also highly data-driven. It is open to the idea that one piece of new information may cause one to see everything that came before in a new light, and call for major revisions; that each piece of information, and the current state(s) of affairs must be continually evaluated and acted upon. To my knowledge, few other programs model this kind of behavior, which appears frequently when people work on tasks that are not routine for them.

There are, however, some programs that specifically separate what might be called "tactical knowledge" and "tactical strategies." For example, Bundy and Welham (1981) describe a technique called metalevel inference, in which

- The separation of factual and control information enhances the clarity of the program and makes it more modular.
- All the power and flexibility of inference is available for controlling search. (p. 189)

This perspective is at least sympathetic to the separation of "tactics" and "strategies" described above. It is an alternative to the production model system, where the decision-making resides in the nature of the productions. By and large, AI discussions of "control" have been much narrower than the mechanisms I wish to invoke here.

There has recently been much discussion of metacognitive issues in the psychological literature. The bulk of such research has focused on metamemory (one's awareness of how he or she stores and retrieves information), and much of that work has been developmental; see Brown (1978) for an overview of the relevant literature. Research suggests that the use of self-regulation is a large component of older children's successful memory performance (Brown & DeLoache, 1978). Speaking in general, Brown (1978) describes metacognitive behaviors as "those attributed to the executive in many theories of human and machine intelligence: predicting, checking, monitoring, reality testing, and coordination and control of deliberate attempts to solve problems. I believe that these are the basic characteristics of thinking efficiently in a wide range of learning situations." One can hardly disagree. Moreover, the converse must be stressed. Just as the presence of such behaviors may promote efficient problem-solving, the absence of them may doom problem-solvers to failure. Discussions of metacognition in
broad problem solving domains are rare. See Schoenfeld (in press-a) for one attempt in that direction, and Silver (1982b) for an overview of the issue. Note that metacognitive acts are generally taken to be conscious. The sequel will argue that unconscious determinants of cognitive performance must be taken into account as well.

Belief systems. Ulric Neisser (1976) begins his article with the following dialogue. It was taken from Cole, Gay, Glick, and Sharp's (1971) study of cognition in a Liberian people called the Kpelle.

Experimenter: Flumo and Yakpalo always drink cane juice (rum) together. Flumo is drinking cane juice. Is Yakpalo drinking cane juice?

Subject: Flumo and Yakpalo drink cane juice together, but the time Flumo was drinking the first one Yakpalo was not there on that day.

Experimenter: But I told you that Flumo and Yakpalo always drink cane juice together. One day Flumo was drinking cane juice. Was Yakpalo drinking cane juice that day?

Subject: The day Flumo was drinking the cane juice Yakpalo was not there on that day.

Experimenter: What is the reason?

Subject: The reason is that Yakpalo went to his farm on that day and Flumo remained in town that day. (Cole et al., 1971, pp. 187-188)

The point Neisser (1976) wishes to stress is that the subject's answers are intelligent, although they are not directly responsive: "The respondents do not accept a ground rule that is virtually automatic with us: ‘base your answer on the terms defined by the questioner.’ People who go to school (in Kpelleland or elsewhere) learn to work within the fixed limitations of this ground rule, because of the particular nature of school experience" (p. 136). There are, Neisser argues, many dimensions to "intelligence" beyond the types of (academic) intelligence measured by IQ tests, the (artificial) intelligence studied in psychological laboratories. Of course anthropologists take that as given (see, e.g., Cole et al., 1971, or Lave, 1980) and some cognitive scientists have urged that the range of cognitive investigations be broadened substantially (e.g., Norman, 1980).

The quoted dialogue serves to make another point as well, one that bears directly on current methodological issues. In the dialogue we see a clash of belief systems, where the participants see the "ground rules" for their exchange in rather different ways. Were the experimenter to declare the subject " unintelligent" because he did not answer the questions as they were posed, we would argue that he missed the point: the responses must be interpreted in the context of the social environment that generated them. The Kpellan's "rules for discourse" are that he should provide accurate information when asked a question. Thus, when pressed by the experimenter.
he suggests that the experimenter consult Yakpalo for an accurate description of what happened. Clearly, it would be inappropriate to evaluate these responses as "pure cognitions." I shall argue here that the same point holds in many of our methodologically "clean" laboratory studies, and that much of what we take to be "pure cognition" is often shaped by a variety of subtle but powerful factors. These factors may include the subject's response to the pressure of being recorded (resulting in a need to produce *something* for the microphone), his or her beliefs about the nature of the experimental setting (certain methods are considered "legitimate" for solving problems in a formal setting, others not), and the subject's beliefs about the nature of the discipline itself (is mathematical proof useful, for example, or a waste of time?). This network of beliefs provides the context within which verbal data are produced, and an understanding of that context is essential for the accurate interpretation of those data.

It should be clear that these comments are not meant as a blanket *a posteriori* challenge to the accuracy of studies that have relied upon the interpretation of verbal data. It may well be that the issue of belief systems is moot in a number of contexts—for example, in the analysis of experts' verbal protocols for purposes of constructing artificial intelligence programs. Experimenters tend to find their subjects among their colleagues, who are generally familiar with and sympathetic to the methodologies being used for protocol collection. It is unlikely, therefore, that an unsuspected difference in belief systems between experimenter and subject will result in the misinterpretation of the verbal data. The situation may be quite different, however, when students are the source of that data and the task at hand is to interpret (in the large) what they have produced. A miscellany of examples that document this point are offered in the sequel. Some less "impressive" but more typical protocols are discussed, from the perspectives at all three levels, in the next section.

**A Discussion of Three Problem-Solving Protocols**

Appendix 1 gives a protocol obtained from two students working on a straightedge-and-compass construction problem in plane geometry, recorded the second day of a problem-solving course. The students were friends, and felt comfortable working with each other. They were both college freshmen, and had both just completed a course in first-semester calculus. They had taken the "standard" geometry courses in high school. Appendix 2 gives a protocol recorded by the same pair of students a month later, after the intensive problem-solving course (see Schoenfeld, 1982, for a brief description). Geometric constructions were one of the topics discussed in the course. The students had read Chapter 1 of Polya's *Mathematical Discovery* (1962),
and worked perhaps a dozen construction problems. Appendix 3 gives a protocol obtained from a professional mathematician who had not “done” any plane geometry for a number of years. The protocols are themselves quite eloquent. The discussion is brief, serving to illustrate some of the points made in section 2. Each of the comments made here needs to be elaborated in far greater detail.

I would like to begin with the outline of a model of students’ behavior on problems like the one worked in Appendix 1. The model characterizes an almost pre-Socratic, purely empiricist perspective toward straightedge-and-compass constructions in plane geometry. The behavior of my students, mostly college freshmen who had taken a full year of geometry in high school and completed at least one semester of calculus, has been remarkably consistent with the behavior predicted by the model, which is based on the five following empiricist axioms. We shall first see how the model works, and then discuss the degree to which the students really believe what it suggests.

I. Insight comes from very accurate drawings. The more accurate the drawing, the more likely one is to derive useful information from it.

II. Hypothetical solutions come from the dominant perceptual features of the drawings. Plausible hypotheses are ranked by their simplicity or “intuitive apprehensibility”: If you can “see your way” more clearly to the end of one plausible construction than another, the first will be ranked higher and tested first.

III. Plausible hypotheses are tested *seriatum*: hypothesis 1 is tested until it is accepted or rejected, then hypothesis 2, and so on.

IV. Verification is purely empirical. Hypotheses about constructions are tested by performing the indicated constructions. If the construction appears to provide the desired result, then it is correct.

V. Mathematical proof is irrelevant to both the discovery and (personal) verification process. If absolutely necessary (i.e., the teacher demands it), one can probably verify a result using proof techniques. But this is simply “playing by the rules of the game,” verifying, under duress, things that one already *knows* to be correct.

Let us now consider the problem given in Protocol 1. Construct the circle that is tangent to the two lines given in Fig. 2, and that has the point P as its point of tangency to the top line. Among the features of this problem that may catch the student’s attention are:
F1: The radius of the desired circle is perpendicular to the top line at the point P (a recalled fact).
F2: The radius of the desired circle is perpendicular to the bottom line at the point of tangency.
F3: By some sort of perceived symmetry, the point of tangency P' on the bottom line is directly opposite P.
F4: Any "reasonable looking" line segment originating at P and terminating on the bottom line is likely to be the diameter of the desired circle.
F5: Again by perceived symmetry, the center of the desired circle seems to be half-way between the two lines, and thus on the angle bisector.
F6: The center of the circle lies on the arc swung from the vertex that passes through P.

Of these six features, F4 and F5 are perceptually dominant (and F6 is generally invoked only after F5, when one tries to identify which point on the angle bisector is the center). See Fig. 3.

Figure 2. Initial drawing for Protocol 1.

Figure 3A. The feature F4 dominates: Which point on the bottom line is the (hypothetical) endpoint of the diameter?
Various combinations of the features listed above yield hypothetical solutions to the problem. For example, F4 combines with F1, F2, and F3, respectively, to generate the following hypotheses:

The diameter of the desired circle is 

- H1: the line segment between the two lines that is perpendicular to P 
- H2: the segment from P perpendicular to the bottom line 
- H3: the segment from P to P’

Likewise, F5 combines with F1, F2, F3, and F6 to yield the following:

The center of the desired circle is at the intersection of the vertex angle bisector and...

- H4: the perpendicular to P 
- H5: the perpendicular from P 
- H6: the segment from P to P’ 
- H7: the arc from the vertex that passes through P

Finally, the nondominant features F1, F2, and F3 combine to yield

- H8: the center of the circle lies on the intersection of the perpendiculars to P and P’

Axioms I through V, along with some implementation rules, serve to predict behavior. Four such rules follow.

1. As suggested above, the features noted by the students during a solution determine the set of candidate "solutions." Thus H4 through H7 become candidates only when F5 is noted, etc.
2. Students' initial sketches determine the plausibility (and thus the ranking) of candidate solutions. If a rough sketch resembles Fig. 3a (as in Protocol 1), H3 ranks high and is likely to be the first hypothesis tested by the constructor. An initial sketch like Fig. 4 serves to rule H3 out of contention, however.

3. F4 is the default condition when F5 is not noted.

4. Unless initial sketches suggest the contrary, H8 (being the combination of three nondominant features) receives the lowest initial ranking of candidate hypotheses.

![Figure 4. This rules out H3.](image)

To see how the model (which is still in its formative stages) works, let us suppose that a student working on the problem observes F1, F2, and F3 in order, and fails to notice F5. The predictions regarding the students' behavior are as follows.

In the absence of F5, we have F4 as a default condition. Since F1 was the first feature heeded, F1 and F4 combine to yield H1. The student will conjecture H1, and try to "prove" it by picking up straightedge and compass and performing the construction. When that construction fails, the student will hypothesize H2, unless (a) it is empirically rejected as implausible because the center of the circle would be "too far to the right," or (b) the failure of H1 causes the student to reject the use of perpendiculars. When F is observed, both H3 and H8 are potential candidates. Whether or not the student has observed H8, H3 will be tested first, with straightedge and compass. If H8 has not been generated, the student will report being "stuck." If it has been, the construction will be performed and the student will report having succeeded. Although he or she may feel uncomfortable about not knowing why the construction works, the student will not doubt that it does. Whether or not the problem-solving attempt was successful, half of the allotted time will have been spent with straightedge and compass in hand. No active mathematical derivation (proof) will have been undertaken.
At this coarse level of detail, the model is remarkably robust. I could offer any number of protocols in which the students I recorded behave almost exactly as the model predicts, and could make a fairly strong argument that the students—perhaps unconsciously—adhere to the spirit as well as to the letter of the model. With a few pairs of students and the same methodology (see Schoenfeld, in press-b, for an extensive discussion), the reader can generate his or her own. Instead I have chosen to examine a more complex protocol, one that is richer than most and raises more subtle questions of analysis. This protocol is better than average (!) in a number of ways. The students work well together and concentrate on the problem at hand for a full twenty minutes. They exhibit few of the “situational difficulties” (duress at being recorded and its effects of performance) next discussed, and their “control” behavior is better than most (see in contrast Protocols 1 and 2 in Schoenfeld, in press-a). The following is a brief running commentary.

T begins by sketching in the desired circle, and the two students make a clear attempt to guarantee that they understand the problem statement. This is respectable control behavior, in contrast to the impulsive actions taken by most students in similar circumstances. By Item 4, the sketched-in circle has been erased. This may be because the rough sketch is not considered “legitimate” for working the problem as suggested in Axiom 1: see Item 10.

In Item 5, the feature F4 is noted and the associated conjecture is made in accordance with the model. The dialogue is atypical in two ways. L and T do not see F2 or F3, and are thus deprived of the opportunity for an empirical verification. But (Items 6 and 8) there is also some evaluation of their conjecture, which is certainly respectable control behavior. The way out of their impasse is empirical (Item 10), and the students spend two and a half minutes with straightedge and compass in hand.

The construction “looks right” (Item 11) but, again atypically, they have some qualms about generalizing from one example. They try to exploit a related problem in Items 14–24. Then five minutes (Items 25–41) are spent in empirical work, which suggests that their hypothesis be rejected. The accurate diagram apparently cues the recall of some relevant information (Item 38), and their rejection is substantiated with a mathematical argument. Here, incidentally, empiricism and proof work together as they should: It is certainly good practice to obtain ideas from empirical explorations, and then subject those ideas to more rigorous scrutiny. My objection is only to the “pure empiricism” suggested in the axioms earlier described.

In Item 43 comes the belated recognition of F1, which again is combined with F4 to generate H1. The enthusiastic jump into implementation (Items 45–50) may be in part a result of desperation, which may also explain their declaration that using a ruler to draw a right angle is “legal” (Items 62–63). Yet items 56–67 and 61–63 say a great deal about students’ percep-
tions of the nature of "being mathematical." Conjecture H1 is again evaluated empirically, and the control functions are again relegated to performing post-mortems (e.g., Items 80–83). There is again a reference to the related problem (Item 84) and (Item 92) an indication that their approach to that problem was also trial-and-error. The solution degenerates from there.

Let me contrast two competing explanations of Protocol 1. From one point of view the students are "empiricists of last resort." They have no idea whatsoever of how to approach the problem and turn to accurate drawings in the hope of finding some inspiration. One need not invoke their "beliefs" at all to explain their behavior. The opposite point of view is the one represented (or perhaps more accurately, caricatured) in Axioms I through V of the model.

I think the most accurate explanation takes from both of the above. It seems clear (e.g. Item 10) that the students turn to straightedge and compass, because they see nothing else to do. As the solution evolves, they become more and more dependent on the (only) tools at their disposal. Thus one could argue that, due to the nature of the environment, they become empiricists during (and perhaps only for the duration of) the solution. On the surface, some other protocols support this interpretation. In one, two students begin a discussion of a construction with "But we cannot prove that that works. Of course, we could construct that." By the time they have finished their construction, they announce "So, we've proved it," apparently convincing themselves that "construction is proof" while they solved the problem.

Were it indeed true that the students had no other ways to solve this problem, the "last resort" explanation would stand on its own. But the fact is that students L and T, and the two just quoted above, knew more than enough to solve the problem "mathematically." After the videotaping sessions were over, I asked the students to solve "a few more geometry problems" for me. The problems are given in Figs. 5 and 6.

Students L and T solved the two problems, albeit somewhat clumsily, in less than five minutes. The two students quoted solved them in less than two minutes each. Of course, the solutions to these two problems comprise the solution to the problem discussed in Protocol 1. Thus these students could have solved the given problem easily, as could most of my students. Yet they did not take any steps in this direction—largely, in my opinion, because they did believe Axiom V, here reproduced below.

Mathematical proof is irrelevant to both the discovery and (personal) verification process. If absolutely necessary (i.e., the teacher demands it), one can probably verify a result using proof techniques. But this is simply "playing by the rules of the game," verifying under duress things that one already knows to be correct.

That believe may well be implicit (they may never have enunciated it as such), perhaps unconscious. It does, however, serve to explain their be-
The circle in the figure below is tangent to the two given lines at the points P and Q. Prove that the length of the line segment PV is equal to the length of the line segment QV.

![Figure 5. The first proof problem.](image)

The circle in the figure below is tangent to the two given lines at the points P and Q. If C is the center of that circle, prove that the line segment CV bisects angle PVQ.

![Figure 6. The second proof problem.](image)

behavior. A student who does not think of proof as a useful intellectual tool is about as likely to call upon it when solving a problem as a person who does not believe in "psychic phenomena" is likely to think of telekinesis as the reason that a paper on his desk has just moved. Belief in Axiom V, then, explains why the students—in spite of their actual knowledge—"have no idea whatsoever of how to approach the problem, and turn to accurate drawings in the hope of finding some inspiration." Their implicit rejection of proof establishes the context that allows for the evolution of their overt empiricism during the solution process.

At the other end of the spectrum, consider Protocol 3. In that protocol a mathematician (who had not done any geometry for nearly a decade) works on the problem the students allude to in Item 14. It is essentially the same problem. A number of factors may contribute to the mathematician's success: better control behavior, more reliable recall of relevant facts, and (not to be underestimated) more confidence. But most important is the basic approach that he takes: he derives the information he needs through
the use of prooflike procedures. Note that the mathematician looks for congruence—"There've got to be congruent triangles in here"—long before there is a conjecture to verify. Rather than being an afterthought or a means of post facto verification, mathematical argumentation ("proof") is a means of discovery for him. (It was Pólya, I believe, who defined geometry as the art of right reasoning on wrong figures. This is clearly the mathematician's perspective, and is antithetical to the empiricist stance.) The non-empirical nature of the mathematician's approach is made emphatically clear in the last line of the protocol, where performing the construction is relegated to the status of an afterthought. Once he has derived it, the mathematician is certain that the construction will work.

In Protocol 2 we see an indication of the "Intermediate" status of the students after a month of problem-solving instruction. The course focused on heuristic and executive problem-solving strategies. Some of these are evident in the protocol, and some were present before the course. Proof was often discussed in the course, but in the usual way: "Yes it seems that way, but how do you know it will always be true?"

Objectively the students' behavior in this protocol compares favorably with their behavior in Protocol 1, along all three of the dimensions outlined in the framework given. Their recall of relevant facts (e.g., that the radius of a circle is perpendicular to any tangent at the point of tangency, Item 69) is more assured, and is called into play at appropriate times. Domain-specific procedural knowledge is also more accurate, and they are confident about their abilities to perform the appropriate constructions. However, these were not disabling factors in Protocol 1 and only tell a small part of the story.

There is a telling difference in their performance at the heuristic level. A few years ago, that difference would have tempted me to attribute their success to the heuristics that they had learned. They draw a picture of the goal state to determine what properties it has (Items 14ff.), look at extreme cases (Items 34-46), consider only obtaining partial fulfillment of the conditions (Item 52), and so on. The first of these heuristics alone might have guaranteed success in Problem 1. However, there is a good deal more.

Their control behavior is quite good, as it was in Protocol 1. They monitor and assess both the state of their knowledge and the state of the solution with some regularity (e.g., Item 71), and avoid the kinds of "wild goose chases" that often guarantee failure for less sophisticated students. Here, in fact, control behaviors become a positive force in the evolution of the solution. At the very beginning (Item 20), empiricism is put in its place. Time constraints are taken into account: in Item 63 the expedient of using the markings on a ruler is acknowledged as "illegal" but used anyway—They could bisect the line if they had to. They know that they are supposed to prove that their constructions "work," and predict early on that they can "do it with similar triangles and things" (Item 72). In this context proof is
still regarded as a means of *a posteriori* verification, employed after one "knows" the answer. The convincing comes by means of good sketches and "gut feeling," however, not by perfect constructions. "Proof by construction" is clearly put to rest in Item 78.

It is tempting, then, to argue that the control strategies serve as enabling factors, allowing the students to employ their tactical knowledge with some success. Certainly the absence of efficient control behaviors would have sabotaged their attempts. However, I think the only reason that the control strategies could work effectively is that they were now operating within the context of a new set of beliefs regarding proof and empiricism. One can conjecture that without this change in belief systems their control behavior would still resemble their behavior in Protocol 1—even if, say, they had been given a review of basic facts and procedures, and taken a course that stressed "executive" problem-solving skills in the abstract. One's beliefs establish the context within which one (a) selects from among his "resources," and (b) employs them.

This brief discussion serves merely to raise a host of questions. It is not meant to minimize the importance of tactical or strategic knowledge, but to indicate that a third and often hidden level of analysis must also be taken into account when one analyzes problem solving behavior. In various problem settings, one or another level of behavior may predominate: access to resources in AI "expert" simulations, control behavior in "wild goose chase" solutions, and belief systems in cases like Protocol 1. In most cases, a comprehensive explanation of behavior will require the analysis of all three. But this is only the beginning.

**TOWARD A BROAD PERSPECTIVE**

While the previous section raises some questions about the interpretation of verbal data, it does not at all challenge their legitimacy. The discussion is predicated on two assumptions: protocols like those in Appendices 1 through 3 provide an accurate reflection of the cognitions and behaviors of the people who produced them, and models of behavior based on data from such protocols will, in turn, reflect the subjects' behavior with some accuracy. In the case of the model and protocols discussed here, I am reasonably confident that this is the case. In general, I am much less sanguine about the "legitimacy" of verbal data, even of some data obtained in methodologically "clean" settings.

Of course this issue is not new. Methodological battles were waged, for example, over the legitimacy of introspection as a means of characterizing cognitive processes. "We have also long known, both from experiments and everyday experience, how subjects' behaviors are affected by expecta-
tion, context, and measurement procedures. The notion that there can be ‘neutral' methods for gathering data has been refuted decisively” (Ericsson & Simon, 1981, p. 17). That point granted, the question then becomes one of the intrusiveness of various experimental methods. For example, it is generally acknowledged that asking subjects to analyze their problem-solving processes while they work on problems does have measurable effects on performance. However, the current literature indicates that sufficiently “bland” instructions may not have a measurable effect on data gathered in the laboratory: Subjects who are instructed simply to “talk out loud” as they solve problems, and not to interpret or explain, will yield essentially the same performance that they would have if they were not speaking out loud (Ericsson & Simon, 1980). These results have generally been interpreted as signifying that protocols obtained under suitably “neutral” speak aloud methodologies result in accurate reflections of “natural” cognition. That interpretation is unjustified, as indicated by the following. In 1978 I made a series of recordings of students solving the following problem out loud.

Estimate, as accurately as you can, how many cells might be in an average-sized adult human body. What is a reasonable upper estimate? A reasonable lower estimate? How much faith do you have in your figures?

The problem is a particular favorite of mine, an excellent task to use for examining the way that people access and use the information that they have stored in long-term memory. It can be solved without any special technical information. One wants good estimates for “average human body volume” and “average cell volume,” under the assumption that there are such things. Since there will be a huge amount of guesswork on cell volume, body volume can be approximated roughly: a box with dimensions 6’ × 6” × 18” will be close enough (probably within a factor of two) to the actual average. (A more accurate figure can be obtained by taking an estimate of average body weight and converting it to volume. But there is no need to be so precise: that degree of specificity is an indulgence.) Approximating an average cell’s volume calls for more guesswork. We can see the markings of a ruler down to 1/32”, so perhaps 1/50” is a lower limit to what we can see clearly without “help.” Cells were discovered with early microscopes, which (considering that magnifying glasses probably have about 5 power) must have been greater than 10 power and less than 100 power. So a “canonical cell,” which we can take to be a cube, must be between 1/500” and 1/5000” on a side. The rest is arithmetic.

My first set of subjects were junior and senior college mathematics majors. The students knew me reasonably well and were familiar with my
work. Some had done protocol recording themselves as parts of senior projects. I took all of the appropriate precautions to set them at ease for the recording sessions, and recorded them working on the problem one at a time (See Schoenfeld, in press-b, for a representative protocol).

Typically, the students quickly determined that volume (rather than weight) was the quantity to compute. After brief consideration they decided to compute body volume before cell volume. They then began the most extraordinary detailed and time-consuming computations. Generally, an "average" body (most often their own) would be approximated by a series of geometric solids whose volume was rigorously calculated. For example:

- and now a leg... a cone might be more appropriate. And the base of my leg is approximately 6 or 7 inches in diameter so you would have \((3\frac{1}{2})^2 \times \pi\) and the height would be... what is my inseam size, about 32 or 34. So you've got to have a 34 and it's a cone, so you've got to multiply it by one third.

In sharp contrast to their meticulous calculations of body volumes, the students' estimates of cell size were remarkably crude and brief. They were not accompanied by estimates of how accurate they might be. For example: "All right, I know I can see 1/16 of an inch on a ruler, so say a cell is 1/100 of an inch on a side." The time spent on cell volume was a small fraction of the time spent on body volume. These results, though puzzling, were remarkably consistent.

Later in the year I began making recordings with pairs of students solving problems together. I recorded perhaps two dozen pairs of students, who solved the same problem after receiving nearly identical instructions. Not once did a pair of students demonstrate the kind of behavior I have just described. With hindsight, it became apparent that the behavior in the single-student protocols was not a reflection of their "typical" cognitions. Rather, their behavior was pathological—and the pathology was induced by the experimental setting itself. This problem used the students, because they had no idea how to approach it. Feeling on trial to produce something for a mathematics professor, they responded to the pressure by doing the only mathematics they could think of under the circumstances: computing volumes of solids. This, at least, was demonstrating mathematical behavior! (The students in two-person protocols manage to dissipate the environmental pressure between themselves, and thus avoid extreme manifestations of pathology.)

I have dwelled on this example at length, because it indicates the subtle difficulties inherent in protocol analysis. When I discovered the social causes that I now believe explain the students' behavior, I was on the verge of writing a paper describing (a) their surprising inability to make "order of magnitude" calculations, and (b) their poor allocation of strategic resources in problem-solving. In hindsight this "purely cognitive" explanation of
their verbal data makes no more sense that it would make sense to assign a
close IQ score to the Kpellan native quoted earlier, as a result of an "objective"
basis of his responses to the experimenter's questions. We need not
travel to Liberia; clashes in the "rules of discourse" between experimenter
and subject occur here in our own laboratories.

Since the length of this discussion has already grown out of hand, the
rest will be very brief. My intention is to sketch out some of the dimensions
of the matrix within which "pure cognition" resides. A broad outline of it,
given in the form of a mathematical cross product, is given in Fig. 7.

The column on the left of Fig. 7 represents an "objective" description
of the problem setting, the product of the two columns on the right the set
of "driving forces" that operate in and on the setting. We take one column
at a time.

<table>
<thead>
<tr>
<th>SETTING</th>
<th>KNOWLEDGE, BELIEF, AND VALUE (KBV) SYSTEMS</th>
<th>DEGREE OF AWARENESS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual (Self)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cognitive structures: access to facts, procedures, and strategies</td>
<td>KBV about self</td>
<td>Unaware</td>
</tr>
<tr>
<td>Task</td>
<td>KBV about facts</td>
<td>Aware but nonreflective</td>
</tr>
<tr>
<td>Environment</td>
<td>KBV about procedures</td>
<td>Locally aware and reflective (monitoring and assessment)</td>
</tr>
<tr>
<td></td>
<td>KBV about strategies</td>
<td>Reflexive abstraction</td>
</tr>
<tr>
<td></td>
<td>KBV about task</td>
<td></td>
</tr>
<tr>
<td></td>
<td>KBV about environment</td>
<td></td>
</tr>
</tbody>
</table>

Figure 7. The dimension of the matrix within which "pure cognition" resides.

The first column is familiar. In the best of circumstances, this is all
that one need be concerned with. "Task variables" can be described ob-
jectively, and the environment as well. "Cognitive structures" are the focus
of customary laboratory investigations: facts, procedures, and strategies.
Under the assumption that laboratory investigations provide an accurate
reflection of problem-solving behavior, the investigator's focus can be on
the overt manifestations of these cognitive structures. In this context the
issue is more delicate: One must (somehow) ascertain the set of facts, pro-
cedures, and strategies that are potentially accessible to the problem-solver.
The same holds for control behavior.

The second column deals with belief systems. Some ideas about belief
systems have reached the level of folk wisdom: for example, the notion that,
through perseverance, a person will turn the belief in his (or her) ultimate
success into self-fulfilling prophecy. A student's belief in his (or her) ulti-
mate failure will affect the verbal data one obtains as well: I have videotapes
of students who never seriously engaged themselves with a problem, in
order to later rationalize what they saw as their inevitable failure. (This has been admitted to me, long after taping, by more than one student.) Beliefs about the very nature of facts and procedures will determine students' performance. The student who believes that mathematical knowledge must be remembered will be stymied when a particular object (say a procedure for constructing a line parallel to a given line) is forgotten, while another who believes that the procedure can be derived will act rather differently. The effects of beliefs on selection and implementation of resources were considered in the discussion of Protocol 1. And the effect of beliefs about the environment (one must produce mathematics when one is solving problems for a mathematics professor!) were the causes of the pathological examples that began this section. These examples barely scratch the surface, of course. But this point is that if we wish to describe behavior as it occurs, we must worry about such things. (It should be noted that belief systems can be modeled. Kahneman and Tversky's (1979) computational models of subjective decision-making are, in essence, models of individuals' beliefs about what they can afford to win or lose.)

The third column reflects the degree to which the individual is aware of his or her KBV systems. Now again in artificial intelligence, this column is superfluous: the construction of most idealized systems depends on the clear elucidation of resources and control structures. The literature of metacognition, as I understand it, has focused on conscious knowledge about knowledge and its relation to intellectual performance. But as we saw in the discussion of Protocol 1, a person's unconscious beliefs can shape his (or her) behavior. Moreover, the fact that the individual is unaware of having those beliefs makes them all the more intractable: he or she does not see the difficulties that they cause. The more that one is aware of the factors that "drive" his behavior, the more potential there is for changing that behavior for the better.

DISCUSSION

This discussion has covered much territory, often at breakneck speed, and many topics that deserved extensive development and discussion could only be alluded to in passing. Despite this, I hope the main theme has been clear. The question of how to interpret "verbal data" is particularly complex and subtle. The examples discussed earlier were chosen to illustrate the various ways that factors "beyond the purely cognitive"—for example, the (perhaps not consciously held) beliefs with which one approaches a particular task or the interaction of the individual with the task environment—can "drive" people's behavior as they generate "out loud" problem solutions for us. Great care must be taken in interpreting those solutions.
It may well be the case, as reported in Ericsson and Simon (1980), that with sufficiently bland instructions, individuals’ performance in the laboratory may not be measurably changed by their speaking out loud as they solve problems. But as we saw with the “cells problem,” the behavior that they produce may be completely abnormal—even if it is so consistent that it can be modeled with great accuracy. Thus certain kinds of verbal data cannot be taken at face value. That is, it is not safe to assume that the behavior we see in the laboratory (even under apparently well-designed experimental circumstances) provides an accurate reflection of behavior outside it. The discussion of Protocol 1 raises even more delicate issues. For the sake of simplicity, let us assume that their “out loud” discussion provides an accurate reflection of their conscious processes. The most straightforward and most likely interpretation of Protocol 1, on its own merits, has the two students as “empiricists of last resort”: The development of their empiricism seems a natural consequence of the fact that they (quite clearly, from watching the videotape) have no idea of how to go about solving the problem. Yet we later discover that they could indeed solve the problem—in less than five minutes. The deep issue in trying to make sense of Protocol 1 lies in trying to understand why the students did not call upon (or, as it appears in the videotape, even think of calling upon) this obviously relevant knowledge. Even if we accept the protocol as an accurate “trace” of their efforts, these verbal data—in and of themselves—are misleading.

As I noted earlier, there are contexts in which these “other than purely cognitive” concerns appear to be moot. In early AI efforts, for example, subjects were often deliberately chosen to have no semantic knowledge about the objects they dealt with (e.g., that “horseshoe” means “implies” in symbolic logic). The idea was to model efficient symbol manipulation in circumstances where one did not rely on “understanding the meaning of the symbols” in order to obtain results. In the more recent modeling of expert performance in domains such as physics, the subjects from whom protocols were gathered generally shared the perspective of those gathering the protocols; clashes in the “rules of discourse” and the resultant misunderstandings were most unlikely to occur. Thus it may have become second nature for researchers in AI to assume that verbal data provide not only accurate, but relatively complete, reflections of cognitive processes. As the modeling of cognitive processes extends to larger subject pools (including students, for example) and to other domains (applications in education, for example) these assumptions can no longer be taken for granted. The degree to which verbal data reflect actual cognitive processes has been discussed. Extrapolations from the laboratory to the “real world” have not, but it should be noted that these are even more tenuous. For example, Jean Lave (1982) reports that people’s use of arithmetic in everyday situations does not correlate well with their scores on paper-and-pencil tests of it. Dick Lesh (1982)
reports that students' problem-solving behavior when dealing with "real" problems bears little or no relation to their "academic" problem-solving behavior. Neisser (1976) argues the point in general.

As the scope of cognitive research and its applications broadens, it will be necessary to develop more comprehensive frameworks for the analysis and interpretation of verbal data. I have tried here to indicate three qualitatively different levels for the analysis of such data, each of which illuminates a somewhat different aspect of intellectual performance. For the sake of simplicity, I have pointed to examples where one level provides the primary "key" to understanding what takes place in a problem-solving session. Examples of protocols where "access to resources" forms the primary level of analysis are those gathered from experts working on routine problems in familiar domains, e.g., those in Larkin et al. (1980). Examples of protocols where "control behavior" provides the key to success or failure are discussed extensively (in Schoenfeld, in press-a). While the model of belief systems given above is somewhat of a caricature, it is easy to find protocols that behave in accordance with this "purely empiricist" viewpoint; I think it is clear that one's belief systems play an important part in determining one's cognitive performance.

Most protocols, of course, will not be so straightforward to analyze. In discussing Protocol 1, for example, we saw that belief systems established the context within which the students worked. Within that context, however, we saw the evolution of the students' perspective: they became increasingly dependent on constructions, and may have finished that problem session with a different viewpoint regarding the way to work on geometry problems than the one they had when they began the session. This is an example of what Lesh (1983) would call the students' "unstable conceptual models." Protocol 2 is far more complex, offering what I suspect is a more typical example of the dynamic interplay among the three levels. Providing a relatively complete explanation of such behavior will be no easy task. I hope that the framework discussed here provides some of the tools for doing so.
APPENDIX 1: PROTOCOL 1

Problem worked the first week of instruction, by students L and T (college freshmen who had completed one semester of calculus).

You are given two intersecting straight lines, and a point P marked on one of them, as in the figure below. Show how to construct, using a straightedge and compass, a circle which is tangent to both lines and has the point P as its point of tangency to one of the lines.

1. T: reads the problem. Oh, OK. What you want to do is that (sketches in a circle by hand), basically. OK, how?
2. L: Now, OK, we have to find the center.
3. T: Of what?
4. L: Of the circle. We are trying to find the circle, right? If we did that then we could... oh, and the radius of course.
5. T: All right, well we know the point of tangency on this line is going to be right here (points to P). What we need to find is where the point of tangency is going to be on this other line, I think. So we can find the diameter in which case we can find the center.
6. L: Is that—that's not necessarily true, is it? Is it true that if you have a circle like that (see right), and then that (points with finger) would be the diameter. You know what I mean? Or maybe you couldn't have it that way...
7. T: The circle has like—no, you don't have a diameter running up through there. No, we have to find the diameter from the point of tangency on this line to the point of tangency on this line, wherever it lies.
8. L: No, wait: the point of tangency, the point of tangency here, would the line connecting those two points be the diameter? It seems that you could maybe construct one where it wouldn't always work.
9. T: Wait, but see, I don't know, we're not drawing it (i.e., sketching it) the right way.
10. L: Wait, do you want to try drawing it (with the compass) and see... (2½ minutes elapse in empirical work. A reasonably accurate drawing results.)
11. L: So, maybe it looks like it might be opposite, see?
12. T: But would that be true for any triangle? Oh, but see...
13. L: I'm confused. I don’t think it would be. Let’s say you had your radius over here and you went like that. I don’t think that could be... OK, I think there could be, there is a possibility.

14. T: Remember on the first problem sheet we had to inscribe a circle on a triangle? Could you do that? I couldn’t.

15. L: I couldn’t either.

16. T: We're in pretty sad shape. But just say we draw a triangle even though we don’t know how to do it. We will draw a triangle anyway.

17. L: So how’s that going to help?

18. T: Because we don’t have to inscribe it actually. We just have to have something to help us (visualize it). (Draws an apparently arbitrary third line.)

19. L: Although...

20. T: Does that do anything?

21. L: Not at this point, I don’t think. Maybe further along if we need a radius we could—but I don’t think it does anything now.

22. T: We’ve gotta do something. With what we have, you just can’t do it, right? We don’t have enough lines or whatever there.

23. L: OK, we need a center and a radius. So how do we locate the center? It has to do with... I think it has something to do with, could we do this?

24. T: No, maybe you have an equilateral triangle.

25. L: Wait, let me just try this. (Begins to expand compass.) What are you doing?

26. T: Don’t you want to see if it’s true? If you have a center way out there, because it may not connect. Don’t you see? (sketch at right)

27. L: I’m pretty sure it won’t. I don’t think it will.

28. T: But if it won’t make a circle, then that means this circle is ours (points back to earlier sketch). The one we have to deal with. You know what I mean?

29. L: I see what you mean. Like try to draw a circle out here like going through this point. See, it won’t. It won’t work because in order for it to work... (another few minutes with the compass. The dialogue has to do with their attempts to draw a very accurate figure, so that they can draw conclusions from it.)

30. T: OK, so that’s what we’re doing, right? We don’t need it that big.

31. L: Yeah, wait, you couldn’t because it is going to go through (the point P). I think it does have to be, right...

32. T: If we have these two points that’s definitely our diameter going through it. Now we can draw...

33. L: But neither is it a tangent.

34. T: That’s just what I was going to say. Can we draw these two lines so that—see you can’t for in order for this to cut through this, it’s too shallow, it’s shallow...
36. T: OK as soon as this—OK, make this a tangent.
37. L: In order for this to be, do you think it's going to be tangent to...
38. T: No, because, because we know this one is not going to... I want to see if like we make this a tangent. You see what I mean? But that doesn't look like a diameter either. Well, I don't think that's it. Of course it couldn't be because a diameter is going to be when it's parallel, isn't it?
39. L: That's the diameter.
40. T: OK. That's not going to help us (laughs).
41. L: You figured that out.
42. T: Right.
43. L: Can we construct one parallel to it? (Looks at original diagram.) But then we still don't know the center. (pause)
Could we just draw a perpendicular?
44. T: Yeah, that's what I was just going to say. If we draw a perpendicular line to this and just call that the diameter it will work from there. And then it should touch if it's perpendicular. It should be tangent at one point, shouldn't it?
45. L: Right!
46. T: Shouldn't it?
47. L: Yes!
48. T: Won't it?
49. L: Yes!
50. T: OK, draw a perpendicular, oh good.
51. L: Does one know how to do that with a compass? Do you?
52. T: This is a right angle, so... (uses the corner of the ruler).
53. L: OK. That's perpendicular. OK. Doesn't look it but it is.
54. T: That's our diameter.
55. L: So if we say this is the point of tangency...
56. T: So we can bisect this to find the center, right? So call it center C. Maybe we should have done our steps.
57. L: That's all being unmathematical, completely disorganized.
58. T: OK, back to the drawing board.
59. L: I don't know how.
60. T: Me either.
61. L: OK, if we just use the ruler with the little numbers on it here.
62. T: Or isn't that legal?
63. L: Sure it's legal (does by hand). Now we have the radius, now we just draw it.
64. T: Uh, oh, do we know, we have to see if this is going to work. I know! Ugg.
65. L: My guess is, I think it's not. But we'll try.
66. T: I would think, though, it would have to, though, wouldn't it?
The radius is shorter as... I don't know. Well, let's see what happens when it goes through there. Somehow it doesn't look perpendicular, though, doesn't it? See this line isn't straight relative to the page which is why it doesn't look perpendicular.

Oh right, but...

It looks good. Now we can tell something. Maybe, I think this tells us the point of tangency has to be way more (points to right). I think.

(Three minutes of constructions)

What circle was this one? Yup, that was a right angle. Oh, darn it.

OK, so the radius has got to be smaller because it's going outside of this line. So it's got to be a little smaller and the center has got to be up and over, like here...

But how do we...

But I don't know how to do that, without doing it until it comes out right.

Yeah.

That was dumb. By doing that we were saying that no matter what this line looked like, then it looked like this, if we dropped a perpendicular we could do it and we could get the diameter for that angle and still expect to do it. You know what I mean?

Yeah, I don't think it will work for any angle though.

Yeah, that's what I mean.

Yeah, well, we goofed again.

Well the only thing I can think of to do is what we did in class the other—well, what we were supposed to do, you know. The triangle thing, trying to inscribe it.

Wait, we know...

I know, that's the problem. We don't know how to do it.

I don't know what to do.

Alright, we are going to have to try something else.

Alright, what are we, what were those sort of things we tried with triangle one? Cause maybe we could... do the same thing with, on a smaller scale.

I got absolutely nowhere.

Yeah.

But I was trying to do things like, bisect this side.

Yeah, I did that.

It didn't work.

Yeah, let's see what we have here. We want to inscribe a circle in this right triangle.
96. T: Why do you want to do a right triangle?
97. L: I don’t know. It just is one. Oh, I blew it now, no. The ends don’t matter because we’re, you see, we want to inscribe it. We’re putting in the extra conditions, because it doesn’t have to touch this line. It doesn’t have to—oh, I don’t know.
98. T: I don’t think that will get us anywhere.
99. A: OK, guys...
100. Both: We give up.

APPENDIX 2: PROTOCOL 2

Problem worked after problem-solving course.

The common internal tangent to two circles is the line which is tangent to both, but has one circle on each “side” of it, as in the picture to the right.

You are given three points A, B, and C as below. Using straightedge and compass, you wish to construct two circles which have the same radius, with centers A and B, respectively, such that the common internal tangent to both circles passes through the point C. How do you do it? Justify.

2. L: Wait, I have to read this. Ummm.
3. T: What we want basically is this, circles and a line something like this that is going to pass through here (makes sketch).
5. T: Like that.
6. L: Except they have—where is it—have the same radius—
7. T: Uh huh.
8. L: —so it isn’t going to look like that.
10. L: But, OK. Wait, I’ve got to think for a second.
(erasing to draw again.)
11. L: OK, wouldn’t it—no, maybe not.
12. T: What?
13. L: No, that was dumb. Let me think.
   (pause)
14. L: Umm, should we try and draw it maybe, how it would
   be to see what the relationship of C is to the two circles,
   since that's not drawn.
15. T: Right.
16  I: You know how I am with compasses—go ahead.
17. T: Well, how am I supposed to draw it?
   (draws with a compass)
18. L: I've made this too big because they're going to overlap
   one another with that radius.
20. L: Just draw (i.e., sketch) it—you don't have to use the
   compass. Just draw it—just draw—no, no, no.
21. T: OK, and I'll make my circles better. (unclear). OK.
   (unclear)
22. T: What are you going to do?
23. L: I just want to see what it would look like more accurately
   (draws with compass).
24. T: Why?
25. L: Just so I could see (unclear) but you can think out loud
   if you have an idea. OK. Can you think of anything?
   (finishes sketch)
26. T: Umm. These two radii are the same, right?
27. L: Yep. Except it doesn't look the same, does it?
28. T: That's the way you put your centers in the center.
29. L: (unclear)
30. T: (unclear) OK. These two centers have to like—do you
   know what I mean?
31. L: No. Wait, what am I looking for now?
32. T: (rereads problem) Why don't we first just try to... 
33. L: If we can find (unclear) (pencil placed at center point)
34. T: All right—if you just have the two centers and you go
   over—say the radius—the radius will have to be half
   way in between the centers. All right, and then...
35. L: Say—wait—what-what-what?
36. T: If we just try to draw the two circles and the tangent
   line without worrying about point C for right now.
37. L: Right.
38. T: OK. Since they have to be of equal radius—the radius
   will be half way between the two centers? It's like the
   tangent line would be like this.
39. L: I don't get this about the radius being half way between
   two centers.
40. T: Me neither.
41. L: I don't get what you mean. How's the radius half way—
   I don't get what you mean.
42. T: If it was like this and the tangent line would just be
   (unclear)
43. L: OK, yeah.

*She meant to say that the length of the radius in this extreme
  case was half the distance between the centers of the two circles.*
44. T: OK? These two have to be the same length.
45. L: Right.
46. T: And the thing that is going to determine how long they are is the angle on this line. What I mean like if they are exactly—half way in between the two centers then the line is vertical.
47. L: Right.
48. T: If we make it somehow shorter right here and here, the circles would be like this, and the tangent would be on a slant like this.
49. L: OK. Ummm.
50. T: We have to figure out how they go through point C. So...
51. L: I don't know either.
52. T: Can we just start with C and draw a line through it somewhere and then make the circles tangent to it?
53. L: No.
54. T: Or—
55. L: No, we're given the centers.
56. T: We're also given C.
57. L: Uh huh. But just drawing the line can't guarantee you could end it with something like this if you just drew the line here. Ummm. Isn't there another way we can characterize the line? Find the locus.
58. T: Ummm.
59. L: This might not work for all of them, but, look here, doesn't this look like—that's just like the center?
60. T: That's just what I was going to measure.
61. L: Ummm. Because if we did that, we were given points A, B, and C.
62. T: Yes (looks at her sketch) that crosses it, too. That's exactly what we're going to do.
63. L: All right—wait, we're not allowed to use a ruler, but—yeah, divide it in half.
64. T: Yaaah, bisect.
65. L: Why don't you actually do it...
66. T: Let's try it on here since we're not sure.

(Begins new sketch)

67. L: Wait, I think it was the other line. (unclear) Just connect point B. We're going to have to drop a perpendicular from B to the line.
68. T: What are you doing that for?
69. L: Because this is perpendicular and that's what the radius would be, a perpendicular and from A coming to the line also.
70. T: Right.
71. L: OK. I don't know why this works, I mean, I just seem to see it, you know.
72. T: I think we can do it with similar triangles and things so let's just make sure it works (unclear).
73. L: We can do it here, too—this isn't a very nice compass.
We're running out of time (whispering). Draw faster, draw faster.

I can't—this is hard.

Draw faster anyway.

I didn't construct it right.

Well just draw it—it'll work.

Oh, wait, maybe I did actually. OK, that's the radius then.

Right.

Perpendicular. Then we just have to draw—I think that's just the right thing.

That'll do it, that'll do it—wait, we've got to draw—OK, we did it. We've got to show why. We have to show that these—the reason that these are half way in between these two points is because—angle side—we have to show that—what this side...

Like we have an angle.

But what are we trying to show—we want to show why this is in between A and B.

Right.

So we want to show that this is equal to this—that they...

Both: ...are congruent.

OK, we have that. We have...

...an angle and a side. How do we know...

And we need to show that this side is compared to that side. And...

(to A): Must we prove why something works or just show you the construction?

If you can justify it, I would be happy.

OK, let's try to justify it.

Now the angle...

We know, I mean, r is equal to r so it is just like...

We have these angles, so this angle equals this one.

After a few minutes, and with some slight confusion, they prove that their construction has the desired properties.
Problem worked by a professional mathematician, by himself.

Using a straightedge and compass, inscribe a circle in the triangle given below. (The inscribed circle is a circle that lies inside the triangle and is tangent to all three sides of it.)

1. Reads problem.
2. All right, so the picture's got to look like this, and the problem is obviously to find the center of the circle...
3. Now, what do I know about the center? We need some lines in here. Well, the radii are perpendicular at the points of tangency, so the picture's like this...
4. That doesn't look right, there's something missing. What if I draw in the lines from the vertices to the center?
5. That's better. There've got to be congruent triangles in here...
6. Let's see, all the radii are equal, and these are all right angles—and with this, of course, this line is equal to itself (marks "X" on the figure), so these two triangles (the two at the lower left vertex) are congruent. Great.
7. Oops, it's angle-side-side, oh no, it's a right triangle and I can use Pythagoras or hypotenuse-leg, or whatever it's called. I'm OK. So the center is on the angle bisectors.
8. (Turns to investigator) I've solved it. Do you want me to do the construction?
REFERENCES


Lave, J. What's special about experiments as contexts for thinking? The Quarterly Newsletter of the Laboratory of Comparative Human Cognition, 1980, 2, 86-91.


Schoenfeld, A. Making sense of "out loud" problem solving protocols. *Journal of Mathematical Behavior*, in press. (b)


