R. PETER POLSON
partment of Psychology
iiversity of Colorado
)ulder, Colorado 80309

R. PATRICK SUPPES
epartment of Philosophy
anford University
anford, California 94305

R. DAVID SWINNEY
epartment of Psychology
ufts University
ledford, Massachusetts 02155

)R. KURT VAN LEHN
Cerox Palo Alto Research Center
33 Coyote Hill Road
'alo Alto, California 94303

1 Problems of Methodology in Cognitive Science

James R. Miller
University of Colorado

Peter G. Polson
University of Colorado

Walter Kintsch
University of Colorado

The conglomerate that makes up cognitive science includes portions of psychology, artificial intelligence, and linguistics, as well as anthropology, philosophy, and neuroscience. There can be no doubt that, in some of these disciplines, considerable progress has occurred during the last decade and that, therefore, cognitive science is a field of genuine promise today. But in what sense does cognitive science transcend these individual disciplines? At the moment, attitudinal differences in methodology sometimes preclude productive interactions. Of course, each discipline should have its own research style, and there can be no question of imposing some common methodology on cognitive science as a whole. But we must aim for a better understanding of the goals and methods employed by other disciplines and an appreciation of the strengths and limitations of these methods. After all, cognitive science is built on the premise that these methods complement each other, thereby permitting us to advance research in cognition beyond our own discipline's boundaries.

But when, in fact, is cognitive science research really interdisciplinary? Suppose a linguist makes an excellent case for the proposition that anaphora rules should apply to indexed noun phrases at a quasi-syntactic level intermediate between the autonomous syntactic part of the grammar and the logical formalism; he then shows that this is not the case in several current AI language comprehension programs, and dismisses them. Is this cognitive science? Alter-

natively, consider a program for text comprehension that makes bridging inferences when there are gaps in the text. The author points out that people can make such inferences, too, and therefore asserts the "psychological reality" of the program. Is this cognitive science? Or, if a psychologist shows that reaction times to passive questions are longer than reaction times to active questions, does this establish the "linguistic reality" of the passive transformation? These examples are not real ones, but neither are they fanciful. Is this limited extension of research from one domain to another what we mean by interdisciplinary work?

Before addressing what we perceive as potentially serious methodological problems, we need to remark on the rather broad view of methodology taken here. We are concerned as much with methods of observation, data collection, and analysis, as with problems of theory construction and evaluation. Indeed, it seems that the latter are primary; once you have decided what sort of theory you want, the empirical methods simply follow.

We first make a few brief observations about the history of cognitive science, because there are some problems we have inherited and have not yet come to terms with. We then enumerate a number of issues that we think deserve attention and that seem to spring from long-standing, unresolved historical conflicts. These sections comprise our diagnosis of what is wrong with cognitive science. We then suggest some treatments to set things right again. We explore how cognitive science could make more effective use of experimental methods and we discuss the opportunities and dangers that stem from the heavy reliance on introspective reports within cognitive science. Most importantly, we take a look at the nature of theory construction in cognitive science, and consider some of its puzzles and contradictions. It is certainly not the case that we have all the answers, but we like to raise questions, even at times in an intentionally pejorative manner, in the hope that the discussions will help us to find some of the answers.

The Dual Origins of Cognitive Science

Cognitive science, as it is practiced today, has two distinct historical roots. It derives from one scientific tradition that emphasizes objectivity and the study of behavior from the outside, and from another that is subjectively oriented and that has proposed to study mental life from the inside (Aebli, 1980). The information processing tradition of cognitive psychology on the one hand, and action theory and purposive or intentional descriptions of behavior on the other, are representative examples of these incompatible trends with cognitive science. Essentially, it is the same conflict that was acted out in a more extreme form earlier in this century by behaviorism and phenomenology.

Some of cognitive science is based on self-observation. We observe a process in ourselves, analyze it, then recognize it again when we see it in others. We understand what happens in others because we know what happens with ourselves. Cognitive theory, thus, is an externalization of processes gained through self-knowledge. The data to be accounted for are direct, simple observations; complicated operationalizations and statistical manipulations would only obscure this straightforward characterization. It is cognitive science from the inside.

The antithesis is the black-box approach that characterizes much of current information-processing psychology. The organism is a black-box, whose input and output can be studied, and we must infer often complex information-processing mechanisms in order to relate the two. These mechanisms are conceived as fixed, interacting components, typically described in rather mechanistic terms. For example, our information-processor might have a short-term memory with a capacity of 7 ± 2 units, various pattern matching devices, a high sensory threshold, and so on.

This is very different from a person who has goals, purposes, and a conscious mind. The contrast is perhaps best expressed by their opposing views of verbal communication. Central to the information-processing approach is the information transmission metaphor. On one end, we have a sender, on the other the receiver, each of which may be decomposed into complex information-processing components; between the two, there is a channel with certain properties, through which information flows from the sender to the receiver. The alternative conceptualization holds that there are two conscious minds, each in a certain state; one is trying to influence the other, that is to change its state, by providing it with certain cues (spoken words, gestures), from which the receiver will infer what the sender meant. Here there is no flow of information anywhere (Hörmann, 1981).

We argue that the problem for cognitive science is to find the right synthesis of these approaches. Neither is satisfactory in itself, but their synthesis must be something other than thoughtless confusion.

METHODOLOGICAL QUESTIONS

The Rejection of Intuition

Despite the growth of cognitive science over the last 10 years, many experimental psychologists are still puzzled by some of the goals of cognitive science. Particularly bothersome is the fact that a theory in cognitive science often amounts to an explication and externalization of intuitions. An anecdote may help to focus this issue. Suppose, like some of us, you are working on a model of text comprehension. We are all able to comprehend text and we have fairly good intuitions about some aspects of the comprehension process. What is the function of a model under these circumstances? In part, it is to make these intuitions explicit, in part to fill out these components of the process where our intuitions fail. For instance, one component where intuitions have proven to be reliable is

the following: If you read a text that you can comprehend well, you can predict quite accurately which portions of that text will be recalled later and which will be forgotten. One of the feats of the model we have been developing (Miller & Kintsch, 1980), is that it is also able to make such predictions, just about as successfully as people do. Recently, an experimental psychologist reviewing a paper describing such recall predictions rejected them as being trivial: All they show is that people recall the important passages and forget the irrelevant detail, as everyone knows and as alternative theories also predict. (Indeed, any theory that wouldn't predict this trend should be rejected.)

A fundamental problem revealed by this criticism is the collision between an approach that views behavior from the *outside* with one whose goal is to specify and explicate behavior on the *inside*, including peoples' intuitions about comprehension. However, such general intuitive knowledge about comprehension is implicit, vague, and cannot be used reliably to troubleshoot comprehension problems nor can it improve the readability of a textbook chapter. The model discussed above is not trivial simply because it makes some predictions anybody can make, just as the linguistic analysis of a sentence is not useless simply because it parses the sentence into constituents any naive speaker could identify. Explaining intuitions is not a bad goal for a theory.

On Mistaking Intuitions for Theory

Self-observation and intuition, and task analyses based on them, are important tools in cognitive science. But it is important to realize that they provide a starting point for research, not an ultimate goal. Intuitions are not in themselves a cognitive theory; rather, they need to be explained. Cognitive scientists do not always seem to appreciate this point. Let us invent an anecdote to illustrate this point. Imagine a psychologist delivering a paper entitled, "Everyday planning techniques: How to bake a cake." He relates how he personally planned this procedure, and then, after reporting protocols of 15 other subjects, identifies three major cake-baking strategies, through which he claims 85% of the statements in the protocols can be explained. Among traditional psychologists, such a paper would be greeted with derision. Some cognitive scientists might applaud. Both reactions miss the point: The paper is inadequate because it offers no theory; it fails to explain, because it merely restates intuitions, and offers them as explanations. On the other hand, such a paper might provide a valuable task analysis, which is a necessary step in formulating a theory of planning strategies, that could account for the behaviors in question.

Describing intuitions is the first step toward explaining them, but no more. What does constitute an explanation? What is the difference between the model that explains prose recall and the planning strategies for baking a cake that merely describe what people do without explaining that process? The comprehension model is an information-processing model: It identifies certain processes and mechanisms (a short-term memory buffer, cyclical processing, memory retrieval) that interact to produce comprehension. As a result of this interaction, certain parts of the text will be recalled later while others will not be recalled. In a sense we are not directly explicating intuitions in this case at all. Instead, what we have are two ways of understanding a phenomenon like text recall: One is by intuitions that we as scientists share with every literate adult; the other is by explaining behavior as a result of particular information-processing operations that are specified in our model.

The point of this example is not that all explanations must involve information-processing notions. Any kind of general theoretical framework, not just that of information-processing psychology, could, in principle, provide an adequate scientific explanation. That is, our cake-baking strategies could have been related to any concrete, specific theory of planning—whatever the basic terms of that theory. In the cake-baking example it is the atheoretical treatment of the task that we object to, not the teleological terminology.

Are theories at different levels—a strategy explanation versus an information-processing account—real alternatives that can be compared and tested against each other, or are they merely different ways of talking about the same thing? A somewhat idealized description of a recent dissertation at Colorado illustrates the dilemma (Caccamise, 1981). When subjects are asked to generate ideas relevant to a given topic, they occasionally repeat ideas they have mentioned before—more frequently under some experimental conditions than under others. The question of interest concerns the nature of these repetitions. Consider first an information-processing model of idea generation from long-term memory, that gives a reasonably accurate account of these repetitions. In such a model, repetitions indicate an interference process: once an idea is generated, its accessibility in memory increases, and the retrieval process is more likely to follow this same route because of its increased accessibility. Thus, idea repetitions are explained as consequences of the retrieval dynamics.

An interesting alternative views the same phenomenon as a planning task. Ideas are repeated because the subject wants to do so, and not because of interference effects over which he has only partial control. The subject plans the repetition because it might serve as a starting point for a process of elaboration that did not occur when the idea was mentioned the first time. This is certainly a very reasonable intuition, namely, that a repeated idea may lead to a chain of associations that were not produced the first time.

Thus, the information-processing model accounts for precisely the same observations as the planning model. Which model is right? The appealing intuition or the mechanistic model? The mechanistic model clearly "explains" the data; does the strategy model? If so, how do we tell which explanation is better? Are they really alternatives, or is the one a description of processes, while the other describes our intuitions? It can't be that easy, because there is a clear incompatibility between an interference process and a planned elaboration! Can you

have both—one to describe the dynamics of an automatic retrieval process and the other to describe a consciously controlled planning strategy that makes use of that retrieval process? We have returned to the conflict between cognition as mechanistic information-processing and cognition as intentional behavior.

Conceptual Confusions, New Insights, and a Latter Day Tower of Babel

One thing that can enormously upset the more meticulous observers of the cognitive science scene is the terminological inflation and confusion. If only words were explanations, especially fuzzy words that can mean anything and everything! We have only to invent three more terms as all-encompassing as "schema" and cognitive science will be ripe for Nobel prizes. But young sciences are notoriously hard to pin down to terminological purity, and terms like "schema" obviously have their uses.

More interesting, and perhaps more serious, is the confusion between purposive and mechanistic language that characterizes much of the writing in cognitive science. As if it were the most natural thing in the world, purposive terminology has been imported into an information-processing framework: subgoals are stored in short-term memory; unconscious expectations are processed in parallel; opinions are represented propositionally; the mind contains schemata. Is it a sign of conceptual weakness, or merely an excusable sloppiness in the use of language? Or is it not confusion at all, but a new synthesis that cognitive science has achieved? Can the methodology of artificial intelligence overcome the distinctions between information-processing terms and purposive concepts? It is not the same to say that we now can talk about intentionality in information-processing terms (a claim we believe to be valid) and to simply disregard the distinction! We need to explore whether or not the distinction has any merit today. If the answer is no, we need to be able to tell people why; if the answer is yes, the current practice in cognitive science is badly in need of a change.

But even if we were much more careful about our language than we are, cognitive science has a serious communication problem. At one time theories were eminently describable, because they were simply described verbally. The spectre of fuzzy-headed thinking was always present, of course, and substantial amounts of handwaving could be (and often was) concealed by a carefully designed paragraph. The precision of mathematical models helped to remove some of this fuzziness. One needed more than a passing familiarity with mathematics to understand these models, but this was nothing that some time spent with a couple of books on probability theory and Markov chains could not cure. As the topics of interest to psychologists grew increasingly complex and knowledge-laden, these abstract mathematical techniques became less useful, and simulation models, based on the technology (if not the methodology) of artificial intelligence, have come to dominate this aspect of cognitive psychology. As

powerful as these models are, and as hopeful as the outlook on understanding the true nature of cognitive processes has become as a result of these models, significant communication problems have appeared.

The communication problem that characterizes computer models cannot be solved as easily as that of mathematical models. It is very difficult to understand how a simulation model works without gaining experience in constructing models of this type, and this requires access to extremely expensive computer systems, as well as training in a particularly technical area. This problem is compounded by the multiple and incompatible versions of LISP or other languages that are convenient for designing AI systems: It is very difficult for programs to be exchanged between researchers at different institutions that use different computers, or even different operating systems running on the same computer. Even with appropriate tools and experience, the difficulty of understanding another person's program is well known among programmers of any computer language. It may be possible for a program to be a reasonable instantiation of a theory, but it would be completely unreasonable to publish a copy of the program's code and rely upon the reader's ability to infer the properties of the theory from this listing. Practically speaking, this is rarely done anyway; instead we have returned to describing these models verbally, and the opportunities for hand waving and fuzzy thinking (or at least fuzzy descriptions) have returned as well. In fact, there is an even greater opportunity now for hand waving, because potentially significant aspects of a model can be implemented by programming tricks that bear no relation to what is known about cognition. Has a new Tower of Babel arisen through our own doing?

So far, we have only complained about what is wrong with cognitive science, trying to trace the reasons that lead some of our colleagues (in all disciplines) to complain that cognitive science is in fact something less than a true science. Now, we shall try to offer some constructive suggestions—about the role of experimentation in cognitive science, about the development of a methodology for taking introspective reports, and most important of all, about the nature of theorizing in cognitive science.

METHODOLOGICAL SUGGESTIONS

Sometimes An Experiment Might Just Be The Right Thing to Do!

Insofar as cognitive psychology is a part of experimental psychology, it shares with the latter its preference for the use of the experimental method. In linguistics and AI, the psychologist's preference for experimental data seems to be regarded as perverse, sometimes with benevolent tolerance, sometimes with ridicule. One of the questions we want to raise here concerns the use of the experimental

method—when is it proper, when is it necessary, what does it have to offer to cognitive science? What are its limitations?

The first point to be established is that introspection alone is not enough. For many psychologists, the traumatic collapse of introspective psychology at the beginning of this century is unforgettable, and they hardly need to be convinced. Linguistics, AI, and philosophy could learn many arguments against introspection from the history of psychology, but learning from history is a surprisingly rare occurrence. If people can easily be mistaken about their own states of pain (Dennett, 1978), it is hardly advisable to rely completely on introspection. Furthermore, many of the processes cognitive scientists are interested in are unconscious ones, such as comprehension and motor behavior. Similarly, the elementary laws of learning and forgetting simply do not lend themselves to introspective study, and we are forced to use experimental methods for their exploration.

The question then becomes when and how to use experimental methods. The attitude of some experimentalists, that nothing is proven unless there are some experimental results about it, is surely unacceptable, and probably one of the reasons why so many non-psychologists reject experimentation altogether. There are, indeed, too many trivial experiments, but that is a poor reason for neglecting a powerful methodological tool. Experimentation is the potential contribution that psychology can make to cognitive science.

Experiments can be used for two somewhat different reasons: for exploration and for hypothesis testing. Experimental psychologists, traditionally have been taught to use experiments almost exclusively for the latter purpose. One formulates a hypothesis, sets up an experiment to test it, and then decides whether or not the results of the experiment permit the hypothesis to be rejected. Very sophisticated procedures are available for designing experiments properly, and there exists an extensive inventory of statistical techniques to analyze their results efficiently and effectively. There are numerous experiments of this type that have made substantial contributions to our knowledge. But the hypothesis testing method has been greatly oversold to psychologists. Many hypotheses tested with great skill and effort are not worth testing: they are uninteresting, trivial, isolated from any theory, or tenuously related to the interesting aspects of the theory, or they are simply superfluous because we already know the outcome.

The use of experiments as systematic exploration procedures, where introspection or field observation fail, is less standard. Sometimes, it is useful to set up "interesting" experimental conditions, just "to see what happens." Sometimes, an investigator may want to explore systematically the boundary conditions of a phenomenon. Sometimes, he is merely concerned with demonstrating control over a phenomenon. The machinery of hypothesis testing is irrelevant in such cases (though not infrequently pseudo-hypotheses are invented by the successful investigator after the fact). But experimental exploration is becoming more and more important as cognitive science enters fields where the problems are unsuitable for introspection, the first crop of problems that lent themselves to introspection having been exhausted.

What needs to be developed more systematically is an experimental methodology for exploration. The hypothesis testing paradigm is well established, but its reflexive use by some psychologists has contributed to divorcing that discipline from important scientific developments. On the other hand, experimental methodology is underdeveloped in precisely those areas where it is needed most: the systematic exploration of new research issues, and the global evaluation of theories. More will be said about this later.

Protocol Analysis: Disciplined Introspection

In parallel with the further development of experimental methodology, a great deal of attention should be devoted to introspective methods. How introspection is used makes a difference. It is possible to be misled, as the history of psychology shows. At this point, some very promising beginnings for a post-behavioristic introspective methodology have been made, as described by Ericsson and Simon (1980).

Verbal reports are data in the same sense as key presses, response latencies, patterns of eye fixations, and sequences of hand motions. Thus, verbal reports are to be explained by correspondences between observation performance and predictions of a model just like any other class of data. In addition, some investigators argue that they are a preferred source of information about cognitive processes because of the richness and the density of information that can be obtained from verbal reports concurrent with the performance of a complex task.

Ericsson & Simon's classification scheme identifies two dimensions of verbal reports. The first is the time of verbalization. This dimension distinguishes information that is reported while in the focus of attention (i.e., while being held in short-term memory) versus information reported after completion of the experimental procedure. The other dimension concerns the relationship between the attended to or retrieved information and the content of the requested verbal report. This relationship can range from a direct report of the information to reports requiring complex or ill-specified transformations of this information.

Ericsson and Simon provide a theoretical analysis of the processes involved in generating various kinds of verbal reports, making use of widely accepted assumptions concerning the structure and dynamics of a human information-processing system, in particular, short- and long-term memory. Data take on meaning in the context of a particular theory of the process under study. Relationships between measurements taken in the laboratory and entities specified by the theory are defined by that same body of theoretical ideas. What Ericsson and Simon have shown is that our current body of theoretical knowledge in cognitive psychology is powerful enough to put our use of verbal reports as data on a sound theoretical and logical foundation.

When we talk about introspective methodology, the main concern is when and how far we can trust protocols. But, as we have just seen, some progress is being made in that respect. A more serious problem is the matter we discussed earlier,