OFFPRINT FROM

STRUCTURAL/PROCESS MODELS OF COMPLEX HUMAN BEHAVIOR

edited by

JOSEPH M. SCANDURA
professor in structural learning
and educational design
University of Pennsylvania, USA

and

CHARLES J. BRAINERD
professor of psychology
The University of Alberta, Canada

SIJTHOFF & NOORDHOFF 1978
Alphen aan den Rijn – The Netherlands
DISCUSSION OF SELECTED ISSUES RELEVANT TO STRUCTURAL LEARNING

Joseph M. Scandura

Interdisciplinary Studies in Structural Learning and Instructional Science, University of Pennsylvania, Philadelphia, Pa. and Board of Scientific Advisors, MERGE Research Institute, Narberth, Pa., U.S.A.

A number of important comments (questions) on the structural learning theory were made (asked) at the NATO ASI, both during and after the formal lecture sessions. Some of the more important issues are discussed below.

1. **Comment.** What kind of data do you have to support the theory and more generally how are the experiments mapped into the theory?

**Response.** I am afraid that any convincing answer to this question would take far more space than is available. Perhaps the best way to answer here is to say that explicit attention has been given to making the theory operational (i.e., testable) in all of its essential aspects. In this regard, I will welcome comments and/or criticisms on the formulation in Chapter 2 and more generally in my book on problem solving.

As to the research itself, I must refer you to the rather considerable literature that has been developed over the past decade. Some of the most directly relevant and recent empirical studies are briefly summarized in the concluding section of my chapter; most are included in my book on problem solving (Scandura, 1977). Structural Learning II: Issues and Approaches (Scandura, 1977).

---

I would like to thank the following individuals who helped to regenerate comments made at the ASI and who suggested specific comments of their own for response: John Durnin, Dean McIntyre, Steve Reed, and Larry Reeker. Steve Reed and Werner Feibel, in particular, also made some very helpful and insightful comments in response to a draft of my responses.
1976) also provides a useful source since it describes some of our earlier research on rule learning and its relationship to a variety of other approaches, of which mine is just one.

2. Comment. Are you proposing the structural learning theory as simply another (Scientific) language or as a scientific theory?

Response. I am, of course, proposing the "theory" as both. As a language, what (competence/knowledge) can be represented in structural learning terms can be represented in terms of any number of other procedural (e.g., programming) languages. The structural learning "language" differs from most such languages, however, in important ways. For one thing, it is not a specific language but rather a class of languages, each having a common, prespecified form. In this sense, structural learning languages are like production systems (which are commonly used in formulating computer simulation theories). Unlike theories based on production systems, however, the structural learning theory is concerned with the explication of processes by which one identifies underlying competence (represented in terms of particular structural learning "languages"). In this sense, structural learning shares goals that overlap with those of automatic programming.

Structural learning is a theory in the sense that it purport both to explain and to predict many kinds of complex human behavior on the basis of a small number of basic assumptions. I am referring here to the (a) goal switching control mechanism, (b) fixed processing capacity and (c) fixed processing speed, assumptions that are assumed to influence the behavior of all human beings or, equivalently, to determine how all people use and acquire their available knowledge. I should point out also that, whereas these assumptions say nothing directly about specific (individual) knowledge (including higher-order knowledge), they do provide constraints that must be satisfied by any adequate representation thereof. In addition, a serious attempt has been made to make the structural learning theory fully operational (in terms of observable behavior) in all of its essential aspects.

The theory described in Chapter 2, nonetheless, is incomplete in the sense that it works from the top-down (although it can be used to account for more behavioral detail than can Piaget's epistemology, for example, or Pask's Conversation Theory). In particular, at the present time, the method of structural analysis (see Section 2.4, Chapter 2) is better thought of as a general approach to be used in identifying specific competence rather than as a theory. (At the present time, we are working on the problem of "automating" (i.e., objectifying and systematizing) the method. Definite progress is being made in this direction but much more needs to be done. Among other things we are hoping (a) to develop
criteria that would enable one to determine (a priori) whether disconfirming data are due to theoretical flaws or faulty structural analysis and (b) to apply structural analysis to the rather ill-defined domain of tasks associated with the Piagetian stage of concrete operations.

Nonetheless, structural analysis in its preliminary forms has been applied in a fairly large number of cases with some success. The results of such analyses, together with specific assumptions associated with other relevant parts of the theory, have yielded very specific predictions concerning human behavior. In fact, the empirical support generally has been far stronger than in that offered in support of traditional theories; it has been possible to predict individual behavior in specific situations with a high degree of reliability.

I wish that it had been possible to present more of our empirical results at the ASI (and in my chapter) but time and the focus on theory seemed to preclude this. Moreover, this seemed to be unnecessary as *Problem Solving* (Scandura, 1977) contains original reports of a large number of studies based directly on the theory. (Most of these studies are cited in the concluding section of Chapter 2.) In fact, we have conducted a rather large number of studies on rule learning, going back to 1959-1960 that in retrospect can be seen to anticipate many of the theoretical ideas discussed in Chapter 2. Comprehensive summaries of many of these studies are included in *Structural Learning II: Issues and Approaches* (Scandura, 1975). Although delayed in press for four years, the book still seems current and provides a useful bridge between earlier theory on complex human behavior and a variety of contemporary approaches, including my own.

It is true, nonetheless, that many of our studies have been designed primarily to test the viability of various basic assumptions in the theory and/or to demonstrate the viability of the theory as a basis for educational applications. Indirect tests of the theory in comparison with other theories have played a secondary role. In some cases, this was because many of the basic questions posed in our research simply have not been addressed by other theories. (However, see my comments concerning the type of control assumed to operate in theories based on production systems. Other scientists also are conducting research designed to contrast predictions based on the Structural Learning Theory with those based on other existing theories. Feibel, for example, is comparing Piagetian and structural learning approaches to training in formal operations. Another extremely important feature of the Structural Learning Theory is that it provides a unique view of the relation between theory and practice, including a highly efficient means of bridging the gap (Scandura, 1971; 1977, Chapters 1 and 11).
In Chapter 15 of Problem Solving, I have shown how issues addressed by other theories can be explained in structural learning terms. This exercise was useful in other equally if not more important ways. First, it indicated that formulation of experimental studies in structural learning terms could make it possible to ask sharper questions. Rather than demonstrating that various heuristic "hins" tend to facilitate problem solving, for example, one can specify more precisely not only the higher-order processes (rules) themselves but also where they will and will not be useful. Second, more precise structural analyses often suggest that inferences, concerning problem solving processes and drawn from empirical findings, often have limited generality. Although purportedly characteristic of problem solving generally, the explanations offered often have more to do with the specific tasks and competencies involved. The contention that algorithmic learning is rote whereas propositional knowledge is meaningful is misleading in this sense. (For details and discussion, see Scandura, 1977, Chapter 15, Section 2.)

3. Comment. When you asked them, two participants (Klahr and McDermott) who have used production systems agreed that production systems are better viewed as a type of scientific language, than as a theory. Nonetheless, much of the literature implies that production systems constitute a theory as well. What is your position on this?

Response. In my lecture I pointed out that the stack-type control implicit in production systems seems to assume too much on the part of the human processor. Viewed as a theory, production systems assume (essentially) that human beings are able to compose arbitrary productions (processes). We have obtained a significant amount of data (e.g., see Chapters 5 and 6 in my book on Problem Solving) which strongly suggest that this type of control mechanism cannot be assumed to be universally available to young children. Whether or not it may be assumed to be universally available to adults is problematical since I know of no data that directly addresses the issue. This and similar points are developed in my book, especially in Chapter 15 where relationships to other theories and approaches are dealt with more fully.

My position basically is that whatever control mechanism is assumed should have some chance of being universally available to humans without training. Our data show that goal switching control meets this requirement with children from the age of about seven whereas stack-type production systems do not. Put

--It becomes increasingly difficult to test this assumption (under appropriately "idealized" conditions—see 10. below) with younger and younger children. Nonetheless, we have completed an experiment
differently, not all people can or do use their available knowledge in accordance with stack-type principles.

As a consequence of adopting a control structure that is too constrained, one has less flexibility regarding higher-order rules. For example, in order to deal with the situation where two or more productions may apply at a given stage, McDermott (at the ASI) proposed a presumably universal higher-order oracle (rule) that selects the most specific one.

Interestingly enough, I too found this particular solution appealing when I first faced the problem of selection in formulating the structural learning theory. Moreover, we reported data (although not deterministic data) supporting the idea (e.g., see Scandura, 1971, pp. 41-44).

This assumption, however, did not share the rather nice deterministic (fully predictive) properties of the remainder of the theory. Consequently, I gave up that assumption in favor of one that allowed more flexibility. In particular, although people on the average may tend to select the more specific rule, this is not necessarily the case. Indeed, any one of us can here and now program himself so as to pick whichever kind of rule might be desired in any particular case. Goal switching control, as in the structural learning theory, does not impinge on that freedom. In this view, selecting the more specific of two or more rules corresponds to just one type of higher-order (selection) rule that individuals might learn and/or use in deciding which of two or more alternative rules to use. (In order to make more specific predictions about the alternatives (rules) individuals will select, it is necessary to specify the higher-order selection rules governing each subject's behavior. Presumably, this might be accomplished via testing as with respect to any other rule, e.g., see Scandura, 1977, Chapter 5.)

Interestingly enough, adopting this more flexible point of view regarding selection not only made it possible to account for

with four, five, and six year olds; that data too support the hypothesis of universality (of goal switching control) (Scandura, 1977, Chapter 5).

3 Determinism per se is not an essential feature of the theory (e.g., see discussion of discrete versus continuous representation, Scandura, 1977, pp. 326-328). Equivalence classes of observables that vary continuously, for example, cannot be well defined but require probabilistic formulation (i.e., as "fuzzy" sets). Nonetheless, the problem of such selection is independent of the above and did lend itself to deterministic formulation (with regard to discrete observables).
a wider variety of behavior but it actually brought about greater
theoretical parsimony with regard to the assumed mode of control.
(For more detailed discussion see Chapter 2 and especially
Scandura, 1973, Chapter 8.)

In many ways the most important difference, perhaps, is the
fact that the structural learning formulation is designed to sat-
sify the requirements for individual differences measurement as
well as for representing competence and cognitive universals.
In addition, considerably more attention has been focused on the
systematic construction of competence.

In spite of these differences, production systems are perhaps
closer to the structural learning formulation than any other common
type of representation used in cognitive science. More particu-
larly, one might convert production systems into structural learn-
ing systems by introducing the following changes: (1) At a global
level replace stack-type control with goal switching and explicitly
introduce higher-order rules (productions), not on a limited or ad
hoc basis but in the more general sense that I have proposed. (2)
Generalize the operations in productions so that they include
branching as in structural learning rules. (This is important for
diagnostic purposes.) (3) Attach a range to each production.
Ranges (of rules) play a central role in goal switching. (4) De-
fine all data structures dynamically in accordance with the rules
which operate on them. (Specifically, the test conditions in
rules define the effective data structures.)

These are, of course, rather fundamental changes. It is
questionable under these conditions whether one would want to use
the term "production system" (which has a well-defined mathematical
meaning). The differences become even larger when one takes into
account the role of memory (e.g., see 10 below).

4. Comment. What is the difference between a procedure and
a rule? You sometimes seem to use the two terms interchangeably.

Response. Technically speaking, a rule is a triple consist-
ing of a domain, a range, and a restricted type of procedure. For
present purposes, an unrestricted procedure may be thought of
simply as a set of well-defined operations and decisions, each of
which can be carried out mechanically, in a specified order. Most
of the essentials are easily represented as flow diagrams. How-
ever, whereas unrestricted procedures may generate and subsequently
call (use) new procedures, or set new goals for itself, this is not
allowed of the (restricted) procedures in rules. What is done, in
effect, is to factor the universal control mechanism out of the
procedures, thereby allowing a fully modular representation of
competence. This restriction also is crucial for purposes of
individual differences measurement.
The characterization of a rule as a triple goes back a long time in my own thinking (to a large amount of research on rule learning in the early/mid 1960s, see Scandura, 1968, 1970, 1976 for summaries). Explicitly restricting the procedures in rules is more recent (although it had been done informally from the beginning). Unfortunately, my earlier usage could be confusing in this sense. I hope that these comments may help.

5. Comment. Many of us are ready to accept the goal switching control mechanism but it seems to deal only with a small part of what is needed (in any particular case) to explain complex human behavior.

Response. In one sense this observation is correct but in another sense it is not. On the one hand, the control mechanism itself plays a relatively small part in accounting for specific behavior. The specific content (competence) involved is predominant. On the other hand, the number of possible content (problem) domains is indeterminately large. This is why I believe that the really important problems lie at a level above the actual representation of knowledge/competence. In particular, we must decide on which of the many possible ways of representing competence to use and also must come to grips with the difficult problem of how to construct such representations for given content.

Because it is universal, goal switching control plays an important role in solving both of these important problems. (1) It places critical constraints on the form in which competence should be represented. (2) Even more important, it places constraints on how we might go about finding such representations in general. Structural analysis has been designed to satisfy such constraints and in this sense plays a fundamental role in explaining complex human behavior. Nonetheless, structural analysis at present still requires a considerable amount of experience and intuition. Making the method more systematic and objective has high priority in our present work.

6. Comment. What are higher-order rules and what are not higher-order rules? Are there higher-order rules other than composition?

Response. First of all, let me reemphasize: A rule is a rule is a rule. All rules are characterized as triples, consisting of a domain, a restricted type of procedure, and a range (e.g., Scandura, 1966, 1968, 1970; also see Chapter 2 and Glossary). Domains and ranges consist of conditions to be satisfied, respectively, by inputs and outputs. (In the latter case, the conditions correspond to what the knower expects the rule to generate. In the former case, the conditions specify the entities to which the knower would apply the rule. As I mentioned in my lectures, the procedures in rules
are more general than the operations in productions. However, they are more constrained than unrestricted procedures of the sort sometimes used in artificial intelligence. In particular, these restricted procedures allow branching but do not allow for the possibility of generating new procedures during the course of a computation and then turning around and using the new procedures. The latter capability corresponds to the goal switching control mechanism which, because of its hypothesized universality over human beings, is partialled out of the representation of individual rules.

The descriptor "higher-order", then, is not a characteristic of rules per se but rather is a relative term. A rule is higher-order relative to another if it operates on (a structure containing) the other rule. In this sense, higher-order rules are like functions defined on functions (in the mathematical sense of the term "function"). Higher-order rules in my sense definitely do not correspond to the term "higher-order rule" as used more popularly to denote rules, say, that are higher in a learning hierarchy (than other rules). Higher-order rules in the latter sense correspond to outputs generated by my higher-order rules.

Let me give a simple example. Consider rules \( A \rightarrow B \) and \( B \rightarrow C \), together with a higher-order composition rule that operates on pairs of rules such that the output of one serves as the input of the other. (Notice that a pair is a structure.) Clearly, this composition rule is higher-order relative to the \( A \rightarrow B \) and \( B \rightarrow C \) rules in the sense that I have described. When applied to these rules, the composition higher-order rule generates the new composite rule, \( A \rightarrow B \rightarrow C \). This composite rule is of a "higher-order" relative to \( A \rightarrow B \) and \( B \rightarrow C \) in the popular sense but it clearly is not the same as the composition rule itself (i.e., a rule output is not the same as the rule which generates that output).

In this regard, some have failed to properly distinguish this notion of higher-orderness and generality. In particular, higher-order rules need not be more general than some other rules (including the rules on which they operate). The two concepts are quite distinct. Thus, a higher-order rule, just as any other rule, may be highly general or overly restrictive. The latter has nothing to do with higher-orderness per se.

With regard to the second question, composition rules are only one kind of higher-order rule. There are higher-order rules which operate on rules and generate new rules that are more general. Others generate analogous rules (e.g., by substitution of conditions and/or operations); still others operate on structures containing more than two rules and select one of them. As I mentioned at the ASI, there are also higher-order rules that operate on complex
rules and generate more automated versions of these rules. Higher-order rules of the latter sort, presumably, play an important role in improving performance generally and in bringing about automation, or what Landa calls "post-algorithmic" knowledge. (Some contemporary investigators apparently would equate higher-order rules of this type with "basic learning mechanisms." This is too restrictive a view, I think, and therefore wrong.) Then, of course, there are higher-order rules which combine these and other types. A listing of many of the more basic type of higher-order rules is included in Scandura (1973, Chapter 5). Higher-order rules that bring about automatization were referred to as "elimination" rules (p. 104).

(One specific example of a higher-order generalization rule is given in my response to Comment 7.)

7. Comment. The structural learning theory is based heavily on rules. Those of us in education, however, know that simply teaching students rules is not sufficient. Some students may use the rules perfectly well when told to use them but flounder when left on their own. They do not seem to know what to do.

Response. This comment raises an important question concerning the difference between knowing a rule and knowing when to use one. In fact, this distinction has been very much a part of my own thinking from the beginning in characterizing what a rule is (e.g., see Scandura, 1966, 1968, 1970). As noted in Response 4, a rule includes a domain, a (restrictive type of) procedure (operation), and a range. All are necessary. Domains, in particular, are conditions that must be satisfied in order for rules to be applied. (In the structural learning theory, ranges also play an important role with respect to the goal switching control mechanism.)

In particular applications, it is often easier to identify the operation of a rule than the domain of applicability. This is particularly true of higher-order rules, more informally called "heuristics." Consider the higher-order rule shown in Figure 1 (from Scandura, 1973, 1977). (Note: In applying structural analysis, it is important to specify domains of the solution rules associated with the sampled problems; see Chapter 2, Sections 2.4.) Although the operation (flow diagram) is specified in this case, the domain strictly speaking is not. In the study in which this higher-order rule was used, the experimental inputs were triples of instances such as

1 \rightarrow 3
4 \rightarrow 12
5 \rightarrow 15
Figure 1. Flow diagram for the higher-order division rule, which acts on instances and generates a general rule of the form \( n \rightarrow an \).

The higher-order rule operates on such triples and generates as outputs rules of the form \( "n \rightarrow an + d" \). But, this higher-order rule will not work with all such triples. For example, consider

\[
\begin{align*}
1 & \rightarrow 1 \\
4 & \rightarrow 16 \\
5 & \rightarrow 25
\end{align*}
\]

In order for the rule to succeed universally the domain would have to be limited to triples of the form

\[
\begin{align*}
1 & \rightarrow a + d \\
m & \rightarrow am + d \\
m + 1 & \rightarrow am + a + d
\end{align*}
\]

Whereas it would be possible to devise such a rule in this case, and in others even more complex (e.g., see Scandura, 1977, Chapters 3 and 4), I do not believe for the most part that human knowledge is that algorithmic. In general, I believe that human knowledge, especially of the higher-order variety, consists of rules with only more or less accurate domains. (The theory of "fuzzy sets," therefore, could be useful in formalizing portions of the theory (also see footnote 2 to issue 3). As noted briefly, in 10, however, I prefer to maintain a sharp distinction in the theory
between structural and incidental factors. This approach provides a useful bridge between theory and laboratory findings, on the one hand, and development and applications on the other. For discussion, see Scandura, 1977, especially, Chapters 1, 10, and 11.)

This does not mean, however, that one cannot have a theory of how people use and acquire "imperfect" as well as perfect knowledge. Indeed, psychologically speaking, I do not think that there is any difference between the two. In the structural learning theory, at least, it makes no difference whether the rules of competence, or of knowledge, are perfect or imperfect. Their use and acquisition are governed by the same principles.

Now, mathematicians and computer scientists frequently refer to quite a different kind of imperfection. This imperfection is mathematical in nature and follows directly from a well known theorem by the logician Church. Namely, there exist classes of tasks (problem domains) for which no algorithmic solution exists.

In short, there is an important difference between mathematical constraints and psychological relevance. No serious psychologist ever thought that anyone could know all that might be known. It is, in fact, the introduction of imperfect as well as perfect higher- and lower-order rules which provides the needed flexibility and makes it possible to "account" for problem domains of often unclear scope. The non-algorithmic and often unpredictable nature of such knowledge is what gives structural accounts a "life-like" and creative character. On the other hand, it is important to know that some task domains are sufficiently complex that no strictly algorithmic account is possible. In such cases (as with most realistic task domains) one must rely in evaluating alternative rule accounts on empirical comparisons of relative generating power.

8. Comment. From reading your chapter and from your lecture, it appears that you use the term "competence" to refer to rules that characterize the capabilities of given populations of subjects. When you talk about some of the applications, however, I sometimes get the feeling that you are talking about what someone else, for example, a teacher, has in mind as to what should be taught.

Response. Fair enough! In fact, both are possible. Moreover, as long as certain requirements as to atomicity are met (see Section 5) one cannot differentiate between the two insofar as the subject's performance capabilities are concerned—unless, as I pointed out in the chapter, one modifies (i.e., enlarges or refines) the problem domain in question.

When I say that one cannot differentiate between competence in these two senses, what I am referring to is the fact that in
either case one can identify individual behavior potential equally well as long as the atomicity requirements are met (Section 5). What these requirements do is to ensure, for example, that what the teacher wants the student to learn is represented in terms of unitary elements (i.e., atomic capabilities relative to the subjects). Effectively, then, learning proceeds in atomic (all or none) sized units from where the student is.

The distinction between how a student might be inclined to perform a given task and how the teacher would have him do it is important, however, where the elements are not atomic. In particular, where the teacher adopts an expository mode of instruction the intentions involved (i.e., the particular rules/procedures that the teacher would have the student use) might not be the same as those preferred (or understood) by the child. Under these conditions, before learning can commence, the teacher and the learner must agree that the two procedures do the same things. (As I see it, this sense of agreement is identical to that which plays such an important role in Pask's theory in bringing about what he calls "understanding." In that case, of course, Pask is concerned not just with procedures for performing tasks but with logical derivations of procedures for solving such tasks. His derivations, in effect, correspond to special kinds of higher-order rules.)

9. Comment. I think I understand how you operationally define individual knowledge by using idealized prototypes (as a basis for comparing individual behavior on selected test items). In effect, you measure individual knowledge relative to the given prototypes. Then, as I understand it, you talk about new learning as being based on this individual knowledge. But, in solving problems, subjects often make use of "extra-domain" knowledge. How do you deal with this problem?

Response. This is an important question but fortunately one for which I think there is an answer, at least there is an answer if one is committed to an operational theory (of complex human behavior). By way of commenting, let me acknowledge first that it is popular today in many circles to talk about how people integrate new knowledge with existing cognitive structures. Indeed, I make the same sort of assumption, for example, in talking about memory storage. Moreover, an adequate theory should shed light on the cognitive mechanisms people use (e.g., control mechanisms, etc.).

I strongly disagree, however, with the assumption that any proper understanding of human cognition must take into account "all that the learner knows" (in some usually ill-defined sense) that might be relevant. In general, I believe that this hope is forlorn. It is impossible a priori to identify all that any given
individual knows that might conceivably be relevant to some given domain of desired learning or performance.

It is, I would propose, impossible to know exactly what anyone knows. Operationally speaking, it is only possible to determine individual knowledge relative to some given content domain, or more exactly, relative to the competencies associated with a content domain. Well, then, what happens to that extra-domain knowledge that we have assumed may influence behavior on the domain? The answer briefly is that that knowledge would be implicit in the individual rules of knowledge ascribed to an individual via the diagnostic assessment procedures proposed in the theory. These rules, whether or not they satisfy the "clinicians" sense of all that is relevant, presumably interact in accordance with the hypothesized goal-switching control mechanism and other (relevant?) constraints imposed by the theory.

(Note: The constraints relevant in any particular case will depend on the desired level of predictive detail, see Chapter 2, Section 5.9.)

As a very simple example, suppose that we are only interested in whether or not subjects in a population are able to convert a given number of yards into the appropriate number of inches. Furthermore, suppose that two different subjects are able to perform the tasks equally well, but that one knows how to perform the conversion directly by multiplying by 36. The other presumably has to use "extra-domain knowledge" to derive an appropriate rule. The point is that, if we are interested only in performance on yards-to-inches problems, then the former rule of competence would serve equally well as a basis for testing both subjects. Insofar as problems in that domain are concerned (where we are only concerned with success or failure, say), it would not make any difference whether the subject knew a rule for performing on the task directly or had to first derive one.

Of course, one might argue that this domain is arbitrarily restricted and that being able to derive such a rule certainly has implications for solving other kinds of (extra-domain) problems. This is certainly true. Moreover, a similar statement can be made of any other problem domain, no matter how complex. But, to check such implications, one would have to test subjects on extra-domain problems, thereby effectively redefining the content domain.

This is a perfectly valid thing to do. Notice, however, that this simply reemphasizes the importance of my main point. Namely, one can only determine (operationally define) individual knowledge relative to a given problem domain. Extra-domain knowledge, in effect, derives from unintentional (and inappropriate) attempts to use competence associated with one problem domain to explain/predict behavior with respect to other domains. When we talk about
extra-domain knowledge, then, what we really mean is that we have not adequately specified the scope of the problem domain in question, that we have not specified all of the competence associated with that domain (and subject population), or both.

10. Comment. Your brief discussion of stochastic and deterministic theorizing confuses me. Are you really saying anything new?

Response. An adequate answer to this question would involve more than simply repeating the brief remarks I made at the conference. Since I first argued in favor of deterministic theories in structural learning (Scandura, 1971), a number of investigators have refined and extended the notion in a number of important ways (e.g., Hilke, Kempf, & Scandura, 1977; Reulecke, 1977; Scandura, 1977a; 1977b, Chapters 1, 7, and 10).

First of all let me emphasize that I did not invent deterministic theories in behavioral science nor for that matter do I think that all human behavior can be explained in strictly deterministic terms (e.g., see footnote 3 to issue 3). (Indeed, most structural/process theories are basically deterministic.) Although space prohibits and other publications make it unnecessary to go into all of my arguments here, perhaps one example will suffice: Among other things, I am proposing that we adopt a different methodology both in formulating and in testing deterministic structural/process theories.

Present practice generally is to simply convert deterministic theories into corresponding probabilistic ones and to test those. Thus, if the theory says that a student should be capable of performing task A after learning B, a typical experiment might involve training on B and testing subjects who receive such training to see if they perform better on A than students who do not receive such training. Instead of eliminating factors which may modify the effects (on A) of learning B, we simply randomize over these factors.

What I am proposing, in this regard, is that we maintain a sharp distinction in testing our theories between deterministic (e.g., B) and stochastic (other factors) influences. Deterministic influences are those that should apply in essentially all situations with all individuals under appropriate "idealized" conditions. Stochastic influences, in turn, refer to deviations from such conditions. An important implication of maintaining such a distinction, and I shall only mention it here, is that it provides a direct and highly efficient basis for applying laboratory results in the real world (e.g., in education). (See Scandura, 1977, Chapter 1 and 11.)

To give you a better feeling for what are meant by "idealized
conditions," let me give a simple example. In one of our experiments, we taught children rules for converting, say, from measure A to B and from B to C. We also taught some of the children a higher-order composition rule (restricted to a suitable domain). Then, we tested them on A-C problems. Theoretically, according to the control mechanism outlined in Chapter 2, one would predict that every subject who learned all three rules would succeed on the A-C problem under "idealized conditions." The question is: What are the "idealized conditions?"

In a word, idealized conditions serve as boundary conditions which must be met if the theory is to apply. Thus, in the structural learning theory, the control mechanism is the sole determinant of what happens in a given situation only under certain very specific conditions. (1) One must know not only what relevant rules have been learned, but which are immediately available to the human processor. What does this mean operationally? It means that training must proceed to a certainty or as close to a certainty as possible. Just as a physicist would try to create a perfect vacuum in comparing the rates of fall of an iron ball and a feather (see Chapter 1, Scandura, 1977), one would try hard to insure that assumed rules are indeed learned and available to the subject at the time of testing. One would not want to set a learning criterion (e.g., three consecutive successes) independently of what the circumstances require. The critical thing is "to get as much of the air out" as possible. (2) Notice also that having learned a rule at some previous time, even to a very high criterion, and having it immediately available in the processor (in working memory) is quite another thing. Empirical conditions should be set up to come as close as possible to guaranteeing availability. That is, the experiments should be run under "memory-free" conditions. This might be accomplished, for example, by having cues for the assumed rules immediately and at all times in front of the subjects. No secrets!

Another proposed universal constraint has to do with processing capacity. Here, I refer to the widely accepted belief that people can only process a small number of cognitive elements at the same time. In the experiments we have run under "memory-free"

---

4 Werner Feibel has commented to the effect that "idealized conditions" have much in common with the Gaiven notion that the most important factor to be controlled in research is the level of reasoning involved. I would certainly agree with this observation (see Chapter 2, Section 2.5, Molar Domains). What I am proposing, however, is more general (and more precise): Namely, there are any number of appropriate idealized conditions—they are best thought of as boundary conditions which must be satisfied if a (deterministic) theory is to provide a complete explanation of the phenomena.
conditions, all relevant and/or trained rules were assumed to be uniformly available. There were no limitations on processing capacity.

Therefore, to obtain information about the control mechanism, unencumbered by processing capacity, one also must be sure that the available rules do not take up too much space in the processor. Otherwise, the processor may become overloaded in the process of shifting control, testing rules and/or using them (i.e., rules).

In order to reduce processing load one may try to insure some degree of overlearning (with respect to the assumed rules). Overlearning has the effect of chunking rules and other cognitive elements into a smaller number of units. One might also minimize the role of processing capacity, for example, by providing subjects with paper and pencil and all of the time that they need.

In summary, insuring idealized, memory-free conditions amounts to insuring the availability of assumed rules (even "error rules, see above) and eliminating the role of processing capacity (and other constraints on cognition). In experiments run under such conditions, we have been able to get close to 100 percent correct prediction. If the conditions are changed, then the results will deviate from perfect prediction just to the extent that the conditions deviate from the required idealized conditions (see Scandura, 1977, Chapters 7 and 11). The principle is exactly the same as would be the case in comparing the rates of fall of an iron ball and a feather. Their rates of fall may be expected to deviate from being equal just to the extent that one has failed to create a perfect vacuum.

There are, of course, many empirical situations where processing capacity (and other cognitive constraints) may play an important role. In such situations, one has two choices. One can admit failure—an inability to account perfectly for the behavior, and turn unashamedly to stochastic theorizing (although in our view, preferably maintaining a distinction between deterministic and stochastic factors, e.g., Scandura, 1977, Chapter 7). Or, one can attempt to add more structure to the deterministic partial theory so that it accounts perfectly for behavior in a greater variety of (but never all) empirical situations (see Scandura, 1977, Chapters 1 and 7). This is exactly what introducing a fixed capacity processor does to the structural learning theory.

Finally, let me caution against inappropriate empirical comparisons of stochastic and deterministic theories: "One might be tempted to propose, that a stochastic theory provides a better account of certain data than a deterministic theory, if the stochastic theory does a better job of predicting average behavior,
which it is designed to do, than a deterministic theory predicts the non-idealized behavior of individuals in specific situations which it is not designed to do" (e.g., see Scandura and the discussion symposium in Spada & Kempf, 1977).

11. Comment. I have read your chapter and heard your short lecture on the topic but I still do not really understand how you would eliminate long-term memory. Can you clarify this?

Response. First of all, let me say what I do not mean by long-term and short-term memory. (1) I am definitely not referring to the retention interval in question. (For a time, these terms were used in the literature to refer both to structural characteristics of the human information processor and to memory over short and long periods of time. As far as I can tell, the latter usage seems to have dropped out of the seminal literature.) (2) In referring to short- and long-term memory, neither am I referring to physiological structures which may underlie human memory. For example, although the "sensory register" plays a central role in the thinking of most cognitive psychologists about memory, the Structural Learning Theory does not attempt to explain its functioning, except to acknowledge that sensory information is somehow made available by the human organism (i.e., the body) to be operated on via the hypothesized cognitive constructs.

Consequently, the Structural Learning Theory, even the enriched form involving memory, applies by definition only to situations where the experimental conditions (exposure times, etc.) are sufficient to insure that sensory information is available to the "sensory register" long enough for the hypothesized cognitive/perceptual processes (constructs) to take over. More particularly, my main concern in this area has been with cognitive constructs, that can be operationally defined in terms of overt behavior and that are useful in accounting for a wide variety of human memory behavior.

With regard to your specific question regarding long-term memory, I do not mean to imply that one can eliminate long-term memory with nothing to take its place. Rather, I am arguing that traditional formulations concerning the relationships between permanent and working memory (as short- and long-term memory have come to be called) put the emphasis in the wrong place. The figure below depicts the standard (earlier) view in which working and permanent memory are viewed as separate entities with some means (arrows) of moving between them.
The distinction between permanent and working memory also can be conceived in a manner that is quite analogous to the model proposed by Tom Landauer. In this case, permanent memory is conceived as a vast data bank, only part of which is active at any given point in time. In different theories, of course, various kinds of data structures are postulated as are various ways of activating new elements (and deactivating others). Here again, at this global level of description, Landauer's model provides an excellent prototype and this perhaps is why it appeals to me so much even though I believe it has serious limitations.

The basic problem, of course, is to determine the nature of the interaction between permanent memory and working memory. In the standard view (i.e., the first class of models proposed), the arrows between working memory and permanent memory serve this purpose and ordinarily would be assumed to be part of neither permanent nor working memory. The very nature of this conceptualization tends to imply that there are special and distinctive means (singular or plural) of getting from working memory to permanent memory and vice versa. One also could argue, of course, that the arrows are simply place holders for processes that are themselves contained in permanent memory and/or working memory. In this case, however, the form of representation would contain extraneous information and thereby tend to mislead.

Alternatively, the problem may be posed as one of postulating mechanisms by which new data structures are activated and by which others are deactivated. Is this accomplished by processes located in permanent memory? Processes in working memory? Processes in neither? Or, processes in both?

In the Structural Learning Theory, the basic memory elements are called rules, structures and goals to be achieved. (In view of the discussion below, recall that structures are the entities on which rules operate and may contain or even consist of rules; see Glossary.) These, then, are the data structures. The theory further postulates that all processing takes place in working memory or, equivalently, with respect to those rules, structures, and goals that are active at any given point in time. In particular, recall that in the theory rules may act on other rules (more properly, on structures containing rules).

A major problem for the theory, then, as with all theories,

5 See Chapter 3.
is to detail mechanisms by which new elements may become activated and old ones deactivated. Just from the way the structural learning theory has been formulated, it is easier to specify precisely how new elements become activated than it is to specify how old ones become deactivated. Specifically, any time a new element is generated from already available (active) elements, that new element is assumed to be activated automatically. I would like for the theory to be equally definitive and precise with regard to deactivation but, so far, I have been only partially successful in this regard. My views on this are sketched in my chapter, Chapter 7 of Scandura (1977), and in earlier publications (Scandura, 1971; 1973a, Chapter 10; 1973b). Suffice it to say here that goals, rules, and structures are discarded (deactivated) when, during attempts to solve problems, they no longer play a role in the process. Unfortunately, there has been little research dealing with this issue. The most complete although slightly dated conceptual discussion probably is still in Scandura (1973a, Chapter 10) although important additional points are made in Scandura (1976; 1977, Chapter 7).

I also should reemphasize that memory storage in this view consists of tying new information in with old (both must be active simultaneously) in the sense of deriving a new rule (qua structure, including a domain, a range, and restricted procedure) which when applied to the old information generates the new. (For more details, see Scandura, 1973a, Chapter 10.) Retrieval, then, simply amounts to regeneration to-be-retrieved information (structures) from what is active at the time of recall. Notice, in particular, that there is no need to talk about going through some large fixed data base (permanent memory). Rather, active rules are applied to active structures. (The latter may be large but nonetheless they are not fixed once and for all. They are defined operationally by the rules which operate on them. Further, the rules themselves may be very simple—corresponding to mere facts—with the emphasis placed on the actions involved (as in all procedural accounts) rather than on the states.)

In general, retrieval will be automatic (according to the assumed control mechanism) to the extent that information active during storage is active at recall. This ties in with the well-known observation that success during retrieval is a direct

---

5 In the standard view, this problem corresponds to specifying the mechanisms by which information moves between working and permanent memory. Such mechanisms frequently go under descriptions like, "If the information is rehearsed long enough, or if enough attention is paid to it, it will be recorded in permanent memory." The implication usually is that such information can be reactivated with higher probability than otherwise.
function of the extent to which we reinstate the conditions under which learning (storage) took place. (The extreme case would be to represent both the effective cue and the to-be-recalled item. In general, of course, the goal would be presented in more general terms.) The complexity/comprehensiveness of the structure(s) to which a memory element is connected during storage also should play an important role in memory since this would be related to the variety of retrieval situations in which regeneration would be possible.

Landauer deals with the activation/deactivation problem by introducing a random walk over time. In fact, time seems to be the only structural factor that enters into his formulation. Landauer postulates a random walk (and I think rightly so) precisely because the theory omits mention of and consequently cannot deal precisely with the influence of specific structural factors (e.g., rules and higher-order rules). Accordingly, one might expect his theory to succeed in situations where structural relationships vary widely, in an essentially random fashion, and where times of presentation and retrieval are the primary consideration. Notice, in this regard, that the traditional memory experiments to which Landauer has applied his theory are exactly of this sort. I would postulate, then, that his theory would fail just to the extent that specific structural factors have been identified and enter into memory situations in predetermined (i.e., manipulatable) ways.

To make the discussion somewhat more concrete, suppose I ask you to name the states bordering on West Virginia. In this case, as in all verbal discourse (e.g., where a memory task is presented to the subject by verbal means), the first task confronting you is to assign some meaning to the sounds streaming from my mouth (or, in this case, to assign meaning to the graphemes on this page).

Presumably, since you are actively listening (reading) to what I am saying (have written), there is active in your processor an interpretation rule for performing this task. This rule, presumably after many years of training and learning, will be both general in its applicability and in many respects automatic. In any case, the end result of applying such an interpretation rule will be a set of structures containing the intended retrieval goal, structures (information and relationships associated, via the interpretation rule, with such words as "states" and "bordering"), and rules (presumably including the interpretation rule itself). In this case, the goal situation would be represented by a relation satisfied by states bordering on West Virginia. Other active information would include a set of rules (a structure) denoting my directive to name states—together, of course, with the interpretation rule itself.

This active information, presumably, provides the basis from
which you will try to generate the needed information. The activated structures, for example, might include rules for naming states given a map (e.g., of the Mid-Atlantic Appalachian Region of the United States). Other naming rules that might be denoted by my directive include driving experiences in that region (e.g., sequences of states traveled through on various journeys through/into West Virginia), and/or even historical facts (e.g., West Virginia was once part of Virginia). If the activated set of rules includes one that is adequate, you would be able to name the states: Pennsylvania, Maryland, Virginia, Kentucky, and Ohio.

When some such rule is not immediately available, of course, one would have to be derived as postulated in the theory. In the latter case, given the variety of possible higher-order rules that might be involved, it would be presumptuous to attempt a definitive analysis. Moreover, all of the subjects with whom I have tried this task seemed to have available an appropriate map reading rule.

The major problem seemed to be that of having an adequate visual image of the region in question (i.e., a map to read). The given information, recall, activated a "map reading" rule but at best it only supplied (cued) a map of West Virginia perse. According to the postulated control mechanism, control presumably went to an appropriate domain goal. If you used the same (map generating) rule that I did, you would have started with West Virginia and tried to reconstruct a map of the region. The main problem, then, as before, is to have activated, or to activate, a rule for generating such a map. In this regard, notice that such rules are implicit in the rela-

In this regard, I would make two parenthetical comments: (1) Higher-order rules tend to be applicable in broad varieties of situations and, hence, are more likely to be available in any given situation than other rules on which they act. Higher-order rules also tend to be simpler. (2) The number of levels of derivation that might be required even in the most extreme cases, as I pointed out at the ANSI, may on the average tend to be much smaller than one might think. For example, think of any well-known figure in the United States (say the President) and ask yourself how many first-hand acquaintances you would have to go through in order to identify a sequence of individuals, each of whom is acquainted with the "next" figure, forming a chain between yourself and the specified person. Except in extreme cases, this can be accomplished usually with no more than three or four links. For further discussion of these issues, and a number of other examples, see Scandura, 1973a, pp. 305-310, especially pp. 314-317.
tion "bordering on," which is an integral part of one effective goal situation. They correspond to what are otherwise called search rules. (Note: Relations are formally equivalent to sets of functions/rules.)

Notice also, that (map generating) rules of this sort might incorporate many facts (e.g., the names of various states). If you find it bothersome to put so much structure into the rules, let me point out that this is formally the same as putting the structure in a data base. More particularly, there is little difference between thinking of simple (rule) operations (e.g., associations) which generate names, given images of the states, for example, and storing the names in given locations corresponding to the images. At one level, it is just a question of terminology. My reasons for preferring the former representation are that it is theoretically more parsimonious (e.g., it has closer ties with the memory-free theory), more practical (e.g., operational), and more general (it allows for individual differences measurement).

The task used in the memory experiment of Chapter 6 of Problem Solving (Scandura, 1977) was much simpler to analyze and, there- by, was more suitable for studying basic questions regarding the control mechanism. Nonetheless, this task had a similar structure. In that case, subjects were presented with an A→C task, together with an A→B rule and a higher-order composition rule. In order to solve the task, they first had to regenerate the needed B→C rule (the domain goal of the higher-order composition rule requires an A→B and a B→C rule). This was accomplished via a retrieval rule operating on a two-dimensional grid, the grid playing a role analogous to that of the image of Appalachia. (It is important to emphasize that this experiment was conducted under "memory-free" conditions. The subjects were provided with memory prompts to minimize the possibility of exceeding their capacity for processing information. In effect, while new elements were assumed to be activated during subjects' processing, deactivation was not considered.)

Perhaps surprisingly, this same type of analysis can even be applied to such routine memory tasks as, "Tell me the name of Joe Young's wife." Here, the goal situation obviously is the name of Joe Young's wife. The directive, "Tell me the name" denotes a set of (known) naming rules, only some of whose domains include Joe Young's wife (and hence might be applied).

To summarize, all cognition may be assumed to involve the activity of a small number of elements in memory. Everything else is reconstructed from this small set of elements. Consequently, if one puts enough structure into the active elements, as is the case when one introduces rules and higher-order rules (and/or the structures on which they operate), there is no need to consider a per-
manent long-term memory per se. Thus, in most other theories in which activation plays a central role (cf. Landauer’s theory), one needs to conceive of a permanent memory because activation is controlled by factors (e.g., a random walk) that have nothing to do with the (presently) active elements themselves. In the Structural Learning Theory, the notion of a permanent memory is superfluous except perhaps as a cumulative record of what has been active at some point in time.

Let me emphasize that I did not come to this still emerging view arbitrarily. Rather, my theorizing has been forced in the above direction by my goals. I have tried to come up with a theory that has a number of important properties. First and foremost, I have been concerned with both universality and inherent simplicity. That is, I want a theory that is broadly applicable and that is based on a very small number of intuitively feasible assumptions. Equally important, of course, I want a (memory) theory that is empirically valid. In this regard, whereas there are only a limited number of directly related experiments (e.g., see Scandura, 1977, Chapter 7; 1973a; 1973b; 1971), the memory theory appears to be consistent with a variety of common observations and basic results in the area (e.g., see Scandura, 1973a, Chapter 10).

Perhaps more important, this view of memory follows naturally from the prior considerations that led to the memory-free theory itself. I have in mind here particularly those parts of the theory pertaining to the goal switching control mechanism. The strong, rather direct deterministic results we have obtained, suggesting the universality of this control mechanism, seem to provide a good foundation for an enriched theory. This and related issues (e.g., how to identify adequate bases upon which to build behavior theories) are discussed in the Addendum to Chapter 11 and also in Chapter 1 of Problem Solving (Scandura, 1977). A reasonably good and accurate historical record of how this view of memory evolved can be obtained by reading the corresponding sections, first, of Scandura, 1971; then, of Scandura, 1973a, Chapter 10; next, of Scandura, 1973b; and Scandura, 1976, pp. 315-316; and finally, of Scandura, 1977.

I might add in this regard that it is only in the final book (on problem solving) that I have had the "courage" to suppress entirely the role of permanent memory. Nonetheless, I want to emphasize that I consider this aspect of the theory still to be very much in an evolutionary state. While I am generally satisfied with the overall form, many fine details remain; more important, empiri-

---

Some would say a "meta-theory," since it largely involves constraints that must be satisfied by memory theories rather than a specific predictive theory attached to particular content.
cal research has hardly begun (however, see Scandura, 1977, Chapter 7; also 1971, 1973a).

Given current circumstances, one might reasonably ask whether my "theory" of memory is simply another language for saying the same thing as other contemporary memory theories. At a schematic level, as we have seen (cf. the two classes of theories above), the answer is "yes." Given any theory of one type it is possible to reformulate it as a theory of the other type (the aesthetic appeal and heuristic value of the two formulations, of course, might differ considerably).

When one considers more of the detail, however, the structural theory of memory differs from others in important ways. Within the second (present) class of models, for example, Landauer's theory is complementary to mine in that it deals with the kinds of unstructured tasks characteristic of traditional learning and memory research. The value of the structural learning formulation only becomes apparent where the to-be-remembered material is highly structured in prescribed and/or understood ways.

For another thing, as developed in my theoretical chapter, the level of detailed representation (process atomicity) required in order to achieve a truly exacting and operational account of human cognition goes far beyond anything available in the published (experimental) literature on memory. In principle, therefore, if not in practice, a theory satisfying these requirements could provide relatively exacting predictions. Testing such predictions, however, might well require the development of new empirical paradigms.

12. Comment. Can you pinpoint any specific predictions that might be tested at the present state of the art?

Response. Yes, I think so. Short of the above (level of detail), there are more general questions that might be asked. Although I have not seriously attempted to identify a comprehensive list of such questions, the following may suffice to give some of the flavor.

(1) One basic difference between my theory and others with which I am familiar is that rules are explicitly included in the processor. Hence, if the human processor has a fixed capacity, as most of us believe, the more rules in the processor at any given time, in general, the fewer (nonrule) elements might be retained simultaneously. In the production system models of Newell and Simon (1972), on the other hand, the productions are held in permanent memory. Working memory consists solely of assertions (on which productions act). Since the latter type of theory is prototypic of most contemporary memory theories, the indicated
distinction could provide the basis for fairly clear tests.

Intuitively, all other things being equal, it would appear that the larger the number of different rules involved in solving a problem, the greater the memory load. Moreover, although citing previous results would clearly be suspect, there are some data which suggest that active rules might in fact take up space. As Posner and Ross (1964) showed, for example, the number of transformations and their size affect the amount of untransformed material that is lost from store. In Scandura (1973a, Chapter 7), where this study is discussed along with others, this result was interpreted to mean that memory load depends on the complexity of the rule needed to solve a task. The conditions of this study were not that unlike those that would be obtained where subjects are required to use varying numbers of different transformations (rules) of the same complexity in order to solve given problems.

(2) Unlike other memory theories, there are no fixed mechanisms in the Structural Learning Theory by which new elements are activated/deactivated. This is all under the control of the individual rules involved. A suitable test of these assumptions would help to clarify the issue.

I have already pointed out in this regard that Landauer’s theory should have difficulty explaining experimental results just to the extent that factors other than time play a distinct and manipulatable role in storage/retrieval. I would add here that the Structural Learning Theory provides one way of explaining/predicting the effects of such structural factors. In particular, the theory provides a potential basis for detailed studies of how memory is influenced by the respective conditions operating during storage and retrieval (e.g., by the specific rules involved). There can be mismatching during storage and retrieval, for example, where meanings are stored but verbatim recall is required, and vice versa (see Scandura, 1973a, Chapter 10).

(3) More generally, the structural learning theory might provide a useful (operational) and more precise basis for determining relationships between the types and complexity of the structures to which to-be-remembered elements are tied during storage and the variety of situations in which retrieval may be expected to succeed.

Overall, the Structural Learning Theory not only differs from most other contemporary memory theories at a global level, it also differs from most such theories in the level of detail which it allows (even where this level of detail is difficult to realize in practice). From the latter perspective, for example, the well known result of Bransford and Franks would be considered interesting. It would not, however, be considered to be particu-
larly basic. What these authors found, essentially, was that college students tend to store by meaning and to combine where possible the meanings of individual statements.

These normative data would not be adequate from the present perspective. Rather than embarking on exacting experiments, with elaborate controls, while randomizing over content and averaging over subjects, the preferred structural methodology would be to first gain more detailed specification of the prototypic competence associated with particular content (and subject populations). Only then would hypotheses be subjected to behavioral tests (perhaps based on individual data).

It is important to emphasize in this regard that I am not simply proposing informal clinical observations (as part of structural analysis) in place of rigorous behavioral experiments. Far from it. The constraints imposed by the Structural Learning Theory (i.e., what counts as a rule, structure, higher-order rule, etc.) direct the observer's attention in very particular directions. The point basically is that by following such a methodology we can formulate our theories with far greater precision than seems currently to be the case in much of cognitive psychology. Perhaps more to the point, when one searches for prototypic rules, as one must in following the structural approach, the results are often at variance with what one gets following the usual normative paradigm (i.e., where the viability of theories is tested vis a vis group data). For further discussion, see Scandura (1977, Chapters 1 and 15, respectively).

In the structural view, serious experimental study should begin in earnest only after more informal methods break down and/or fundamental theoretical issues arise that need empirical clarification. In the latter regard, many of the most fundamental questions that I consider most basic with regard to the completeness and internal consistency of the Structural Learning Theory itself do not play a role in most other memory theories and, hence, are questions that seem to have been avoided.

A good deal of further clarification and empirical work, for example, will be required to clarify the mechanisms by which active elements in the processor become deactivated. Similar questions exist with regard to exactly how memory load is to be calculated, especially where a number of rules are presumed to be interacting. Also, how is memory load determined when goals are being used to test available rules, etc.? Empirical/theoretical questions such as these can only be asked, let alone answered, on the basis of fairly precise theoretical formulations. My "meta-theory" of memory, together with the structural approach that I have described and illustrated in my book for getting at specific competence, provide one way of attaining this level of detail.
In concluding this rather long answer to a very important and still only partially resolved problem, let me make one further point. The level of precision needed in representing rules, if one is to be able to operationalize memory load (see Sections 4.2 and 5 in my theoretical chapter; also see Scandura, 1977, Chapter 7), might appear to pose problems in operationalizing the rule construct (i.e., in assessing behavior potential, see Section 5; also Part 4 in Scandura, 1977). In the case of process atomic rules, for example, what a subject knows depends on the test context (e.g., on what is active). Moreover, each subject's behavior itself effects what is active.

Indeed, these problems could be insurmountable were it not for the fact that (rule) constructs in the theory are represented in a distinctly modular fashion. In particular, each rule of competence, no matter what level of precision is involved, can be "measured" separately with respect to individual subjects. The fact that such measurement (testing) will hold only under predetermined conditions does not appear to me to be a fundamental problem (i.e., an in principle limitation). The essential thing is that regularities in behavior exist and hold up in circumstances that can be predetermined. In view of the "West Virginia" illustration above, for example, one of the most basic kinds of regularities might involve the way(s) in which discourse is interpreted by various people.

To be sure, following the above prescription for behavioral research would require a large amount of exacting analyses together with rigorous experimentation. Moreover, attempting to capitalize on all aspects of the theory in practical application could pose serious problems as to feasibility. Granted, we may need to "cut corners" in particular applications and, for some purposes, even in some kinds of basic research. Only by knowing what a complete understanding of relevant phenomena would take, however, will we be in a position to know in any given application which corners may be safely cut.

13. Comment. I think I understand how you propose to do away with permanent memory. You put a lot of the structure in the individual chunks (rules/structures).

Response. Roughly speaking, yes.

14. Comment. But just what does this buy you?

Response. What it buys is flexibility, something that appears to be absolutely necessary if we are to accurately capture complex human behavior. In most theories of memory, for example, it is either explicitly or implicitly assumed that there is a single mode or way of accessing permanent memory from working memory and vice versa. I believe that people routinely make this transition in any
number of ways depending on the context and, specifically, on what is in working memory at the time. In the theory, the only way new entities enter working memory is via the entities already there (e.g., by "retrieval" rules acting on other rules).

Equally important, it seems intuitively clear, and the available data tend to support this, the entities in working memory may vary from extremely global to highly specific. For example, we often tend to think about problems globally first, especially when we get into new areas, before we gradually "home-in" on the needed specifics. In thinking about mathematics or physics, for instance, or for that matter in mulling over almost any deep scientific problem, one typically senses going through a period of "warming-up." Put differently, the initial units available in working memory exist only at a high level of abstraction. (As illustrated above, their presence might be cued by the then operating internal and external contexts.) These units presumably interact as specified in the theory and in the process may generate (i.e., recreate) specifics—i.e., needed aspects of these more global units. These specifics, in turn, may generate others.

Although the theory is described more fully in my book on problem solving, I should reemphasize that many theoretical details still need to be worked out in this area. Empirical testing has barely begun.