Structural Approach to Instructional Problems

JOSEPH M. SCANDURA

University of Pennsylvania

ABSTRACT: This paper introduces a structural/process/systems approach to instructional science and shows either how what have been classified as "open problems" have already been solved in fact or in principle or how such problems might reasonably be attacked by building on existing structural theory. The following questions are considered: (1) How are behavioral objectives and knowledge related? (2) What form should a performance test theory take? (3) Why is it that some people can skip prerequisites whereas others cannot? (4) How do people discover new information? (5) Are content and performance analysis distinct or is reconciliation possible? (6) How might horizontal decalage in Piagetian psychology be analyzed in structural terms? (7) How might one apply the structural approach in adaptive instruction?

For many years, educational psychology and the application of psychology to education were largely synonymous. Courses in educational psychology at colleges and universities often consisted of watered-down versions of principles adapted from the various subspecialties of academic psychology. Traditional learning theory, for example, played one predominant role. Measurement played another.¹

But times do change—in this case in at least three major ways. First, increasing attention has been given to task structure. This change came about gradually, prodded initially by military training research. Beginning during World War II, the urgent need to train personnel to perform complex tasks, and to do so quickly, resulted in the recruitment and involvement of psychologists in the design of efficient training systems. As a result of such work, the central importance of task variables, and particularly the need to develop effective ways of analyzing tasks in behavioral terms, became increasingly apparent. Second, in academic psychology over the past 10 years, we have seen an inexorable shift from peripheral stimulus–response (S–R) approaches to research on human learning, toward increasingly cognitive orientations. This shift has been reflected in educational psychology, especially during the past 5 years. Third, individual differences measurement has undergone a major shift from normative to criterion-referenced testing. This shift is particularly apparent in those areas of educational measurement that bear most directly on instruction.²

Concurrent with these changes in emphasis, there has been rapid growth in instructional design (i.e., in the development and application of systematic techniques for constructing educational materials and methods). Indeed, there are signs that instructional design of the 1970s and 1980s may play as significant a scientific and practical role in education as did normative testing in the decades following Binet and the old Army Alpha.

Each of the aforementioned changes and related technologies can play a useful role in educational science. Nonetheless, each of them alone fails in some important way to capture the full essence of teaching and learning. To properly understand the teaching–learning process, the phenomenon must be treated as a whole. It will not be adequate to develop theories that just deal with one or another aspect of teaching or learning (e.g., content analysis, task analysis, cognitive processes, criterion-referenced testing) without also taking into account overall relationships.

¹ Child development and personality have also occupied important positions in educational psychology, but with the exception of some Piagetian concerns, these do not enter into our discussion.

² This does not mean to imply that advances have not been made within the normative test tradition (e.g., Guilford, 1967). Thus, for example, Speedie, Treflanger, and Feldhusen (Note 1) have developed an instructional model which deals with (normative) relationships between human abilities and efficient problem solving.
I am definitely not, however, recommending that we revert to the use of vague block diagrams, and certainly not to rhetoric, as a substitute for serious theory. Few would deny that content, cognition, and individual differences all play a central role in understanding the teaching-learning process. What must be done is to spell out the relationships among them in precise theoretical terms that lend themselves to scientific explanation, strong prediction, and control.

Although some important beginnings in this direction have been made by structural/process/system (structural) theorists (e.g., Bormuth, 1970; Landa, 1974; Pask, 1975; Scandura, 1973a), these developments have been largely overlooked by American psychologists. This is unfortunate, not only because structural theories provide unifying conceptual foundations for a wide range of behavioral concerns but also because they provide potentially practical bases for solving important instructional problems.

Moreover, although working independently, a number of structural theorists have come to similar conclusions. The works of the English cyberneticist-engineer Pask (1975) and myself (Scandura, 1973a), for example, share a number of important features, although the scientific language used and the specific mechanisms differ considerably (cf. Pask, 1975, p. 304; Scandura, 1976a). The independent work of Landa (1974) in Moscow is also compatible, although the emphasis here is not so much with theory as with the use of algorithms as a scientific language for describing various aspects of teaching and learning. In this sense, it is a close scientific cousin to the "set-function language" (Scandura, 1967, 1968).

Given the relatively rapid progress that has been made, it may be desirable at some point to explore more fully the relative advantages and relationships among the most advanced work in this area. However, since much of this work at the present time is familiar to only a few specialists, I shall not go into that here.

This article is concerned primarily with my own theory and approach (e.g., Scandura, 1971a, 1973a, 1973b, 1974d, 1976b, in press-b). This theory is potentially compatible with more pragmatic developments in American instructional science (e.g., Gagné, 1970), but it appears to be more rigorous and general. More important for present purposes, it also appears to provide a basis for resolving and/or for attacking a number of important open problems in instructional theory. In the following sections, I identify some basic and generally familiar questions concerning instructional science and attempt to show either how they have already been solved in fact or in principle or how they might reasonably be attacked by building on this theory.

The following questions are considered: (1) What is the relationship between behavioral objectives and knowledge? (2) What form should a performance test theory take? (3) Why is it that some people can skip prerequisites whereas others cannot? (4) How do people discover new information? (5) What is the relationship between content and performance analysis (as in learning hierarchies)? Are these forms of analysis distinct or is reconciliation possible? (6) Is horizontal decalage (in Piagetian psychology) destined to remain nebulous, or can it be pinned down in structural terms? (7) How might one apply the structural approach in adaptive instruction?

To provide some perspective for what follows, let me first sketch the general nature of the structural learning theory. Special aspects of the theory will be developed as needed in subsequent sections.

The Theory of Structural Learning: A Brief Overview

It is impossible within the space of a few pages to summarize adequately even the main features of the structural learning theory. The earliest presentation (Scandura, 1971a), although somewhat outdated, is still perhaps the best introduction, and my book (Scandura, 1973a) is the most thorough available treatment. My forthcoming book on problem solving (Scandura, in press-b) provides perhaps the clearest version along with important extensions and applications to education.

Nonetheless, let me try to present a not too misleading overview. The structural learning theory (cf. Scandura, 1971a, 1973a, 1976b, in press-b) provides a unifying theoretical framework within which to view the concerns of the teacher or competence researcher (e.g., the artificial intelligence, linguistics, subject matter specialist), the cognitive scientist, and the individual differences specialist.

It would be presumptuous, of course, to suggest that the to-be-outlined theory provides a satisfactory account of all such phenomena. Nonetheless, in some respects, and in certain subareas, the theory has achieved a degree of rigor and empirical support that warrants serious attention by psychologists, including potential critics. In other respects, however, the theory represents hardly more
than a sketchy map and a hunting license with the promise of plentiful game.

The theory is basically a relativistic one. What individual subjects know and what they are able to do are always judged relative to some predetermined content domain and subject population. In the theory, the term *content domain* is used in a broad sense and, in principle, may encompass anything from simple arithmetic to language or moral behavior. As a first approximation, content domains may be thought of as sets of potentially observable stimulus (input) situations and corresponding responses (outputs) that happen to be of interest to some observer (e.g., teacher). One such set might consist of geometry construction problems and their solutions (Scandura, Durnin, & Wulfeck, 1974). Similarly, the subject population might be either multicultural or highly homogeneous, and its appropriate characterization may equally well involve ethnicity and/or age (e.g., developmental level). In most applications to date, both the content domains and the subject populations have been relatively well delineated but, in principle, this is not an essential limitation.

The main point is that the subject population places definite constraints on the processes (rules) that may be introduced to account for such content. The processes that collectively make it possible "to solve" problems in a content domain are referred to as rules of competence. Collectively, the set of such rules is called a competence account of the content domain.

In principle, any given task (problem) can be solved in any number of ways. In practice, however, only a small number of alternatives will normally be compatible with how a knowledgeable member of the target population might solve it; for example, German children use the equal-additions method of subtraction whereas American children use borrowing. Similarly, children at Piaget's preoperational stage judge "more or less" in different terms than do 7–10-year-olds. Such constraints severely limit the (theoretically) infinite number of rule sets with which one might account for a given content domain. (With most nontrivial domains, it may be difficult to find even one really adequate account and, in general, it can be proven that there are content domains where no such account exists. This important mathematical fact, however, although interesting in its own right, is of questionable behavioral significance. First, the corresponding mathematical theorem cannot apply to human behavior because the totality of such behavior to date is finite. The theorem holds only with respect to certain pathologically infinite domains. Second, the theorem implies that there are content domains about which perfect knowledge is not possible. But, then, we always knew that anyhow. The problem for the behavioral scientist is to find the best possible competence account.)

It is assumed in the theory that what an individual subject does and can learn depends intricably on what is already known. More particularly, it is assumed that the human information processor may be adequately characterized in terms of (a) universal characteristics of the processor itself and (b) individual knowledge that is judged relative to the competence in question (i.e., relative to the competence associated with given content domains and subject populations to which the individuals belong). Justification for the latter contention is relatively complicated and cannot be considered here. The interested reader is referred to Scandura (1973a; and, especially, in press-b). Suffice it to say here that, as in all information-processing theories, individual knowledge (behavior potential) is represented in the theory in terms of processes or rules. This matter is discussed more fully in answers to Questions 1 and 2.

Control mechanisms are among the most important universal characteristics. Control mechanisms serve to tell the organism which processes (rules) to use and when to use them; they are essential in all information-processing systems, whether man or machine. The omission of such mechanisms in Tolman's theory is what prompted Miller, Galanter, and Pribram (1960) to criticize Tolman's rats for being deeply engrossed in thought with no capacity to act.

Whereas all complete information-processing theories make a distinction between process (rule) and control, control in most cases either plays a distinctly subordinate role (e.g., Newell & Simon, 1972) or is distributed among a variety of different control mechanisms whose coordination in turn is often left unspecified (e.g., Pascual-Leone, 1970). In contrast, the structural learning theory postulates a goal-switching control mechanism that makes minimal assumptions about the processor but that, nonetheless, has been shown adequate to account for a wide variety of behavior. This mechanism is hypothesized to be common to all humans and to apply universally to all specific knowledge. It is described incidentally in the answers to Questions 3 and 4.

A second general characteristic, which has been considered in testing the theory, is processing capacity. Again, almost all contemporary informa-
tion-processing theories assume in one form or another that "working memory" has a limited capacity. In the structural learning theory, working memory is assumed to hold not only data (the stuff on which rules operate) but rules themselves. While capacity per se is assumed to be fixed (although it may vary over individuals), the memory load associated with any given task depends directly on the process used in attacking it. Thus, whereas it may be impossible to multiply large numbers in one's head using the standard algorithm, many people know short-cut processes that enable them to perform successfully. The theory also allows for the inclusion of other general constraints, such as processing speed, and so on, but this part of the theory has not yet been developed.

Each universal characteristic of the human information processor says something about behavior but not all there is to say. Accordingly, one can conceive of a succession of deterministic partial theories, each of which in turn says progressively more about human behavior. Each partial theory is deterministic in the sense that it deals with the behavior of given subjects in particular situations. Deterministic predictions may be expected to hold, however, only in situations that satisfy appropriate boundary conditions; for example, the partial theory involving the control mechanism fully accounts for behavior only in situations in which all relevant knowledge may be assumed to be readily available. To the extent that memory, or processing capacity, for example, is involved, theoretical predictions can be expected to deviate from obtained results. The idea is directly analogous to the situation with the inclined plane law of elementary classical physics; for example, this law allows one to calculate the force needed to move a given cart up an inclined plane but only where the inclined plane is perfectly smooth and the wheels on the cart are frictionless. Deviations from prediction may be expected just to the extent that the inclined plane is bumpy and/or that friction otherwise plays a role.

In effect, the structural theory of cognition is a top-down theory. Progressively, more structure may be added to the theory by adding more and more universal characteristics. Thus, adding processing capacity to the "memory-free" theory, which involves only the control structure, makes it possible to account for behavior under a wider variety of conditions.

The above implies a particular (structural) approach to theory construction that is an essential ingredient of the structural learning theory. I shall not, however, discuss this issue further or the important implications it has for testing behavioral theories. Suffice it to say that each partial theory must be tested under appropriate idealized conditions (in the same sense that the inclined plane law must be tested using smooth inclined planes and frictionless wheels). This issue was discussed originally in Scandura (1971a), with later refinements, elaborations, and extensions by Hilke, Kempl, and Scandura (1976) and Scandura, (1976b, in press-b).

In contrast to general cognitive constraints, specific knowledge is assumed to vary over individuals. The theory shows how competence, corresponding to the knowledge had by an idealized member of some population, may be used to define opera-

**Figure 1.** Schematic representation of major components of the structural learning theory. \( (S_i - R_i) \) denotes a set of stimulus situations and responses, and \( \{r_1, r_2, \ldots, r_n\} \) and \( \{r'_1, r'_2, \ldots, r'_m\} \) denote sets of rules. The \( r_1, r_2, \ldots, r_m \) denote either rules or structures [sets of rules].
tionally the knowledge had by actual individual members of that population. The rules of competence serve effectively as "rulers," or standards against which individual knowledge may be measured. An introductory description of this process is given in answer to Question 2.

One final point: The theory sketched above is not fully operational in one important sense. Because predictions in the theory depend directly on the competence associated with a given subject population and content domain, and because the number of different content domains and subject populations is indeterminately large, it is essential that a fully operational structural theory include a theory (systematic method) for identifying arbitrary competence. (Contrast this requirement with linguistics, where the search essentially is for a single competence theory.) Although a fully adequate solution to this important problem is well beyond current reach, the constraints imposed on competence by universal characteristics of the human information processor make it possible to proceed in an at least quasi-systematic manner. (This is depicted by the vertical arrows in Figure 1.) A preliminary method for constructing competence theories underlying given content domains is considered in Sections 5 and 7. Further work in this direction is reported in Scandura (in press-b).

To summarize, the structural learning theory is concerned not only with content (competence), cognition, and individual differences but also with the interrelationships among them. Some of the more salient features of the theory will become clearer in the following discussion.

1. Behavioral Objectives and Knowledge

WHAT IS THEIR RELATIONSHIP?

As is well known, specifying educational objectives in behavioral terms tells what it is that the learner must be able to do after learning. On the other hand, the term knowledge as used in the structural learning theory refers to an underlying rule (procedure/algorithm/relational net) construct which reflects a potential for behavior.

Perhaps the most basic inadequacy of traditional behavioral approaches to education is that specifying behavioral objectives does not tell what the learner must learn or what the teacher must teach. Specifying only behavioral objectives themselves leaves the "guts" out of learning. Consider an observation made in an individually prescribed learning environment by Bob Davis (Note 2), an innovator in mathematics education. Davis became interested in a child who had learned to place the decimal point in adding numbers according to his own system. Given "4 + .3", the child would respond "7". Similarly, he could correctly add "2 + .7". His system worked fine if the decimal point was to the left. But if asked to add "3 + .2", for instance, the child would respond "5.4".

Why? Not surprisingly, answering this question requires more than knowing the behavior to be expected in adding decimals. It is necessary also to specify what (competence) a person must know in order to add them. In particular, in the above example, one must specify what rule would lead a child to respond correctly to the first two instances and incorrectly to the third.

To argue that the behavioral objective observed by Davis may have been poorly formulated would miss the main point. Specifying underlying competence is important for several reasons: (a) Given any class of tasks (e.g., a behavioral objective), if there is one rule that will generate a solution for each task, then there are any (countably infinite) number of other rules that will do the same thing. (This is a mathematical fact that can be easily proven.) (b) In practice, there is often more than one feasible rule for generating behavior associated with a behavioral objective. Which rules are feasible depends on the "culture" in question; for example, borrowing is the preferred method of teaching subtraction in American schools, whereas in German schools the method of equal additions is used. (c) The selection of one or another (or all) feasible rule(s) has important and direct implications for instruction. (As we shall see in the next section, it also has direct implications for performance testing.)

In teaching logical reasoning in reading (e.g., Lowerre & Scandura, 1973), for example, it is possible to emphasize syntax (e.g., All A are B, x is an A; therefore x is a B) and/or semantics (e.g., Venn diagrams). Whereas syntactic rules of inference are perfectly adequate within their assigned domains, there is little basis for positive transfer to other inference rules as is the case with inference rules based on meaning. In the latter case only, the processes (e.g., using Venn diagrams) by which one combines meanings of individual premises and checks to see if conclusions follow from them are the same irrespective of the particular inference rules involved (see Scandura, in press-b, chap. 12; for a more general but analogous treatment, see Scandura, 1971b, chap. 3).
In short, unlike behavioral objectives qua behavior objectives, which are devoid of underlying competence, competence and behavior are intimately connected in the structural learning theory—indeed, in most of the newer structural-algorithmic approaches to education (e.g., Gagné, 1970; Landa, 1974; Pask, 1975). Specifying behavior alone is not sufficient; the competence (rules) that makes that behavior possible must also be specified.

To minimize misinterpretation of the intended scope, or applicability, of the rule construct, three cautions seem in order.

(a) Although the precise specification of competence in terms of rules is usually accomplished via some formal, relatively low level (i.e., detailed) computer language, this does not mean that rules necessarily must be represented with this precision or in this degree of detail. Furthermore, competence need not always be presented in terms of discrete entities. Continuous quantities and operations may be required as well (e.g., coordinating pressure on a gas pedal in driving an automobile with a stick shift; making use of kinetic feedback in wrestling; and mentally rotating objects in space).

In effect, whereas rules (procedures/algorithms) in computer programming are typically represented in terms of fixed computer languages, the linguistic elements used to represent rules of human competence/knowledge are more varied. Specifically, the units (operations and decisions) of which knowledge rules are constructed, and correspondingly the linguistic elements used to represent these units, vary in nature and scope according to the intended population of human information processors. Generally, the more sophisticated the population, the larger and more varied are the linguistic elements that can properly be understood (as units). In turn, the larger and more varied the units, the easier it is to represent competence. It is this characteristic, flexibility, that leads to the broad applicability of structural representations, including applicability to what initially might appear to be intractable tasks.

Consider, for example, the ability to read critically, that is, the ability to detect logical (or other) relationships among statements in a paragraph. It would be difficult indeed to detail all of the operations and decisions involved in encoding and interpreting individual morphemes as well as in determining grammatical and logical interrelationships. However, given that subjects can properly understand individual statements, the task becomes much easier; for example, the individual meanings can be represented as regions in Venn diagrams and the interrelationships, as set membership, intersections, unions, and complements of such regions (e.g., see Scandura, 1971b, chap. 3; Scandura, in press-b, chap. 12).

(b) The representation of competence in terms of rules is quite independent of how competence is to be imparted to children. The same competence frequently can be acquired by telling or by self-discovery, by symbol juggling or by concrete manipulation. In general, the way in which information is presented to the child depends on factors other than the particular competence in question; for example, it may depend on whether the teacher during the course of learning wants the student also to gain experience in discovery (and thereby to learn how to make discoveries in related situations, Scandura, 1971b, 1973a). (Choices regarding method of instruction are usually made on other grounds; see Section 7 where the assignment of values to objectives/knowledge and relationships to instructional methods is discussed.) The main point is that if one knows precisely what it is that one wants a child to learn (and not just the behavior the child is to evidence as a result of learning), then one can facilitate learning far better than if one does not.

(c) "Knowing" the subject matter content involved is not equivalent to specifying the relevant competence. The former refers to an intuitive understanding and ability to use the content, whereas the latter refers to the ability to suitably describe or illustrate such understanding and ability. This distinction between knowing something and describing that competence is quite analogous to that, for example, between being able to solve mathematical problems and being able to tell someone else how to solve them.

In summary, it would appear that specifying behavioral objectives is not equivalent to specifying underlying competence. The former tells only what the learner is supposed to be able to do; the latter tells in addition what the learner must know or learn in order to do it. Moreover, there is an important difference between specifying competence so that humans can understand it and specifying competence in a form that can be interpreted by computers. In the former case, competence is specified in terms of units at whatever level is appropriate to the target population. In the latter case, competence is specified at a fixed level of analysis. In the next section, we shall see that the
level of representation also has important implications for performance testing.

2. Performance Test Theory

WHAT FORM SHOULD IT TAKE?

Let me now turn to the widely recognized need for better measures of specific behavioral competencies (e.g., see Resnick, 1972), and especially for better measures based on a sound theory of performance testing (e.g., see Glaser, 1973).

Mastery testing (e.g., Bloom, 1973) represents an important advance over normative testing insofar as instruction is concerned. It provides information not only concerning the relative capabilities of two or more individuals but also concerning the specific capabilities of individuals. Criterion-referenced testing represents a refinement of mastery testing in which the conditions of testing and of mastery are defined more precisely. The introduction of item forms into criterion-referenced testing by Hively and his collaborators (e.g., Hively, Patterson, & Page, 1968) goes further in this direction, as does their later development at the University of Pittsburgh (e.g., Ferguson, 1969). Item forms make it possible with paper-and-pencil tests to pinpoint, more precisely than in simple criterion-referenced testing, just what kinds of items within a given task domain a person can deal with effectively and what kinds he cannot.

None of the above forms of criterion-referenced testing, however, deal with the relationship between behavior and competence. The structural learning theory (Scandura, 1971a, 1973a), on the other hand, provides an explicit way of dealing with this relationship. Specifically, rules of competence introduced to account for performance on given behavioral objectives provide an instrument of sorts with which to measure human knowledge. The theory tells how, through a finite testing procedure, one can identify which parts of to-be-taught rules individual subjects know. The rules in a very real sense serve as rulers of measurement and provide a basis for the operational definition of human knowledge.

Let us briefly consider how this may be accomplished (for details, see Durnin & Scandura, 1973; Scandura, 1973a). The flow diagram in Figure 2 depicts a rule (procedure/algorithm) for subtracting numbers. This rule is broken down into steps that are so simple that each individual subject may be assumed able to perform either perfectly on each step or not at all. We say that each com-

---

Figure 2. Subtraction algorithm.
Component step acts in atomic fashion (i.e., acts as a unit). In line with the above discussion, it is worth emphasizing that what are atomic units relative to one population may not be atomic units with respect to another (e.g., less sophisticated) population.

Because success on any path of a procedure depends on success on all atomic components, each path through the procedure also acts in atomic fashion. Furthermore, there are only a finite number of behaviorally distinct paths. We do not distinguish paths according to the number of repetitions of loops, because the same cognitive operations and decisions are required regardless of how many times a given loop is traversed in carrying out a given "cognitive computation."

Collectively, the paths of the subtraction rule impose a partition on the domain of column subtraction problems; that is, they define a set of distinct, exhaustive, and homogeneous equivalence classes of subtraction problems such that each problem in any given class can be solved via exactly one of the paths.

One path through the subtraction algorithm is represented schematically in Figure 3, along with two-column subtraction problems to which that path applies. Notice that Operation (arrow) 1 says to go to the right-most column. The second node asks whether the top number is greater than the bottom number. (The first node = START.) Since the answer is yes, Operation 2 is applied (i.e., the bottom number is subtracted from the top number). Next (Node 3), we ask if there are any more columns. If there are, we proceed to the next column (Operation 3) and repeat. Otherwise we stop.

The fact that each path is associated with a unique subclass of column subtraction problems makes it possible to pinpoint through a finite testing procedure exactly what it is that each subject knows relative to the initial rule. It is sufficient for this purpose to test each subject on one item selected randomly from each subclass. Success on that item, according to our assumptions, implies success on any other item drawn from the same equivalence class, and similarly for failure.

Individual knowledge (or behavior potential), then, may also be represented in terms of rules—specifically, in terms of subrules of the given rules of competence. The knowledge attributed to different individuals, however, may vary even if only one rule of competence is used to access behavior potential. For example, suppose a subject succeeds on only one path and fails on the others. Then his knowledge would be represented by that path (which is a subrule of the initial one). When two or more paths are involved, a combination of the paths would be used to represent the knowledge.

Fortunately the above discussion is not just a theoretical exercise. Although the variety of task domains considered is limited, a significant amount of supporting data has been collected over the past several years, with subjects ranging from preschool children to PhD candidates. Given a class of tasks, the general form of each study went as follows: (a) One or more rules were identified which were both adequate for generating solutions to each of the tasks and compatible with the way a knowledgeable or idealized member of the target population might be expected to solve them. (b) These algorithms singly and/or collectively were used to partition the class of tasks into equivalence classes. (c) Subjects in the target population were tested on two items (tasks) from each equivalence class (and item form). (d) Performance on one item from each equivalence class (item form) was used as a basis for predicting success or failure on the other (second) item.

With highly structured tasks run under carefully prescribed laboratory conditions, and given performance on initial items, it has been possible to predict performance on new (second) items with over 96% accuracy (Scandura, 1973a; Scandura & Durnin, in press-a). When testing took place under ordinary classroom conditions, with the subjects run as a group, the predictions were accurate in about 84% of the cases (Durnin & Scandura, 1973).

In the latter study, the equivalence classes determined via the structural/algorithmic approach were compared with item forms identified by Hively et al. (1968) and Ferguson (1969). Whereas the levels of prediction on success items were approximately the same, the algorithmic/structural
approach was significantly better in predicting failures. Equally important, this level of prediction was obtained with half as many test items—with an even greater increase in efficiency possible through the use of conditional testing (Durnin & Scandura, 1973).

Moreover, the use of item forms has been limited to paper-and-pencil tests. And as noted by Gagné (Note 3), item forms have intrinsic limitations with regard to non-paper-and-pencil applications such as job analysis. The structural approach is not limited in this way. The direct relationship between molarity of atomic rules and sophistication of population allows for broader applicability. In job analysis, for example, it would make little sense to attempt a molecular analysis of arithmetic skills in order to judge accounting skills, or of writing syntax in evaluating professorial capabilities. Although the impatient reader may have some doubts, minimal capabilities can reasonably be assumed of all bona fide professionals. Thus, all trained accountants presumably can add a column of figures, and all experienced PhDs have at least minimal writing capabilities. Hence, it is sufficient to consider only those molar competencies (atomic rules) that distinguish among individuals in the target population—for example, the ability to set up and administer efficient accounting systems for companies of various types.

There is one further major advantage of the structural approach to assessing behavior potential: It fills an important need in making individualized instructional decisions. Quoting Glaser (1973), "techniques need to be developed for analyzing properties of individual performance frequently enough and in enough detail for individualized instruction decisions" (p. 563). Or as suggested by DiVesta (1973), we need a deeper understanding of the relationships between objectives and what individuals know relative to these objectives.

The algorithmic/structural approach makes it possible to identify precisely not only what individuals can and cannot do, as is the case with item forms, but also what the learner does and does not know relative to the particular rules involved. A simple basis for instructional decision making follows directly: Assume the paths the learner already knows and concentrate on those that he or she does not.

In summary, it would appear that any viable theory of performance testing must take into account underlying competence. Rules of competence associated with populations of subjects provide not only a basis for measuring individual knowledge and for providing remedial instruction but also a basis for selecting appropriate test items. Furthermore, because the appropriate level of rule representation varies directly with population sophistication, it is often practicable to analyze even complex task domains (at a level of analysis that is sufficient for assessing the behavior potential of individuals in the population).

The interested reader is referred to the literature for information regarding the consolidation of knowledge (Scandura, 1974a, in press-a), hierarchical relationships among paths and the conditional testing this makes possible (Durnin & Scandura, 1973; Scandura, 1973a), a possible basis for assessing sentence production abilities (Carroll, 1975), the use of sets of rules for assessment purposes (Scandura, 1976b, in press-b), and the assessment of skilled performance in which response measures (e.g., latencies) more refined than success/failure are required (see Scandura, 1973a, chap. 8; in press-b, chaps. 2 & 15).

3. Skipping Prerequisites (Hierarchical Transfer)

WHAT IS ITS SOURCE?

The importance of prerequisites in learning is widely recognized (e.g., Gagné, 1970). Teaching, however, is not just a question of presenting prerequisites according to a predetermined hierarchy. As Resnick (1972) has noted, "individuals appear to vary widely in their ability to 'skip' over prerequisites" (p. 9). Why is this so? Can we develop ways of determining ahead of time whether individuals can skip prerequisites? Can we determine which ones can be skipped? Further, is it possible to teach people how to transfer? If the ability to transfer not only can be taught but also can be tied in with problem solving and "generative skills" (Bruner, 1964), so much the better.

Stated simply, one fundamental question underlying these concerns is the following: When subjects are presented with a task where they know rules corresponding to all (lower order) components of a solution, why is it that some of them succeed whereas others fail? Consider, for example, two children who know or who have been taught one rule for converting yards into feet (e.g., multiply by 3) and another rule for converting feet into inches (e.g., multiply by 12). Suppose these children are presented with the task of converting yards into inches, and one child succeeds but the other fails. Why? Clearly the yards-to-inches
problem can be solved by combining the two known rules, the rule for converting yards into feet and the rule for converting feet into inches. But how does the child know how to combine the given rules? Knowing two component rules is surely not logically equivalent to knowing when and how to use them.

A basic assumption in the structural learning theory is that new rules are derived by application of certain rules to other rules. (In previous writings, I have referred to the former as being of higher order relative to the lower order rules on which they operate; cf. Scandura, 1971a, 1973a.) Presumably, then, the child who succeeds on the yards-to-inches task does so because he knows something that the other child does not. He knows a higher order composition rule, which includes in its domain (of rule pairs) the yards-to-feet and feet-to-inches rules.

Higher order composition rules may operate on arbitrarily large classes of rule pairs of the form A → B, B → C. Depending on the sophistication of the subject population, these classes may include only simple linear conversion rules, as in our example, or in addition, rules involving electronic trouble shooting, flying airplanes, etc.

With young children, it seems most reasonable to limit the domain to similar conversion problems. Given a deutsch-mark-to-U.S.-dollar's rule and a U.S.-dollar-to-Luxembourg-franc's rule, for example, such a higher order composition rule would generate a composite rule for converting deutsch marks into Luxembourg francs. This composite rule would consist of two components, one for converting deutsch marks into U.S. dollars and the following one for converting U.S. dollars into Luxembourg francs. As we see in the next section, higher order rules are not limited to just combining lower order component rules.

To avoid possible confusion, it should be noted that higher order rules correspond roughly to plans (e.g., Miller, Galanter, & Pribram, 1960), heuristics (e.g., Polya, 1962), and higher level strategies and not to a popular use of the term higher order to indicate tasks and rules that are higher in learning hierarchies (e.g., Gagné, 1970), or to Miller et al.'s (1960) TOTE hierarchies (cf. Sandura, 1973a). All rules, higher and lower order alike, are represented in the same way (e.g., as flow diagrams, labeled directed graphs, or relational nets). The descriptor “higher order” is not a property of rules per se but rather of the use to which they are being put at a given point in time. If rules are operating on other rules, they are referred to as being of relatively “higher order”; if they are being acted upon, they are said to be of relatively “lower order” (Scandura, 1974b).

While knowing both requisite higher order rules and requisite lower order rules is a necessary condition for solving composite tasks, this is not sufficient. In order to effectively use available rules to derive solution rules and to solve problems, some type of control mechanism also is needed to determine how each rule is to be used and when. Control mechanisms in most contemporary information-processing theories operate within the context of goal-directed human information processors (e.g., Newell & Simon, 1972; Pask, 1975). According to perhaps the most commonly made assumption, for example, the rules (productions) available to a subject are assumed to be listed in a cyclical stack. Given a problem (given plus goal), the rule at the top is tested to see if the given is in its domain. If it is, the rule is applied and put at the bottom of the stack. Then, the next rule is tested in the same way, only this time against the output of the previously applied rule. If a test element is not in the domain of a rule, the rule is simply put at the bottom of the stack and the next one is tested.

This control mechanism has much in its favor. It is simple. It is directly compatible with linguistic grammars and formal mathematical systems (Scandura, 1973a). And it has been used successfully in simulation studies of human performance (e.g., Newell & Simon, 1972). Until recently, how-

3 Heuristics, as well as plans, strategies, and other higher order rules, can in principle be formulated as rigorously and as algorithmically as any other rules. In my own work, this has been done in studies involving simple trading games and number series (Scandura, 1974d), geometry construction problems (Scandura et al., 1974), and algebraic proofs (Scandura & Durnin, in press-b). Moreover, the growing field of artificial intelligence is devoted largely to just this problem.

The commonly used definition of “heuristics” as rules of thumb refers to the difficulty of specifying a domain of applicability, precise or otherwise, which guarantees solution of all problems where a heuristic (and other rules) may be involved. Put differently, people can and do learn imperfect rules as well as perfect ones. Any viable theory of human cognition must allow for both kinds.

4 In many contemporary information processing theories, no unique control mechanism is assumed (e.g., Pascual-Leone, 1970; the hierarchical artificial intelligence models of the Massachusetts Institute of Technology). Such theories, however, suffer from the disadvantage of being either idiosyncratic and/or of moving effective control to a higher level. In the former case, there may be as many mechanisms as programmers, and theory becomes a very individualistic enterprise. In the latter case, if there is more than one control mechanism, some control mechanism is needed to tell the others what to do and when.
ever, the question remained open of whether or not the mechanism accurately reflects the way in which people use their available knowledge.

In a study reported by Scandura (1973a, 1974c), 7-9-year-old children were trained on simple $A \rightarrow B$ (e.g., marks to dollars) and $B \rightarrow C$ conversion rules and tested on an $A \rightarrow ? C$ problem. If subjects automatically use available rules as specified in stack control, they would be expected to succeed. That 24 of 30 subjects failed on the $A \rightarrow ? C$ problem strongly suggests that the hypothesized stack-type mechanism cannot be assumed to be uniformly available (without training), at least not with young children.

In the structural learning theory (Scandura, 1971a, 1973a), the hypothesized control mechanism rests on the assumption that control shifts among various higher and lower level goals automatically in a predetermined manner. For present purposes, we may think of the mechanism informally, operating as follows: Given a task (e.g., the given in a problem and its goal) for which a subject does not have a solution rule immediately available, control is assumed to switch automatically to the higher level goal satisfied by (solution) rules that do apply. With a higher level goal in force, subjects are assumed to select from among available and relevant higher order rules in the same way as they would with any other goal. In effect, they try to derive a solution rule that does apply in the given situation. Where no higher order rules are available, the theory assumes that control moves to still higher level goals. Conversely, once a higher level goal has been satisfied, control is assumed to revert to the next lower level where newly derived (learned) rules may be applied.

With this control mechanism in mind, the 24 subjects in the study (Scandura, 1974a) who failed on the $A \rightarrow ? C$ problem were randomly divided into two groups of 12. One group was trained on the higher order composition rule identified above. The other group served as a control. After the higher order rule training, all subjects were trained on a new pair of $A \rightarrow B$, $B \rightarrow C$ rules, which the subjects had never seen before. Then they were tested on the corresponding $A \rightarrow ? C$ problem (which was also novel). This time essentially all of the experimental subjects succeeded whereas all of the control subjects again failed.

These results are perfectly in accord with the hypothesized goal-switching control mechanism. Given the $A \rightarrow ? C$ pretest problems, for example, notice that the 24 failure subjects did not have a solution rule immediately available, nor apparently did they know an appropriate higher order rule. They only knew $A \rightarrow B$, $B \rightarrow C$ rules (e.g., rules for converting yards into feet and feet into inches). Under these conditions, they failed uniformly.

After training on the higher order composition rule, the experimental subjects fared uniformly well. Presumably, according to the goal-switching mechanism, control automatically switched to a higher level goal satisfied by $A \rightarrow B \rightarrow C$ rules. With the higher level goal in force, the experimental subjects presumably selected from among their available rules in the same way as they would with any other goal. Because they were assumed after training to have a higher order (composition) rule available, this rule presumably was applied to the available component $A \rightarrow B$, $B \rightarrow C$ rules, thereby generating an $A \rightarrow B \rightarrow C$ rule. In turn, the higher level goal was satisfied, and according to the theory, control reverted to the original goal. At this level, the child again tested his available rules. This time, however, the newly derived $A \rightarrow B \rightarrow C$ rule was available, so it was selected and applied, and the task was solved.

I must caution that this simple control mechanism is an idealisation and applies only in situations where memory is not a factor, and specifically where all of the requisite higher and lower order rules are learned perfectly and are active in "working" memory (see Scandura, 1971a, 1973a, in press-b). Perhaps surprisingly, however, this limitation has not proved to be as critical as one might expect. Empirical support has been strong, although not deterministic, even under "real-world" conditions. Ehrenpreis and Scandura (1974), for example, found that higher order rules underlying a college course for teachers could be identified in a systematic manner and that instruction on such rules had a highly positive effect on prespecified kinds of "far transfer." Furthermore, the degree of transfer was directly related to the degree to which the test conditions deviated from the ideal (Scandura, in press-a, in press-b).

The interested reader also is referred to Scandura (in press-b), where the above "first approximation" control mechanism is generalized to account for motivation (rule selection), problem definition (subgoal formation), and rule retrieval and where related empirical support is reported. Other studies are planned as this article goes to press.

In addition to the goal-switching mechanism, the extended theory takes into account other characteristics of the human information processor, for example, the limited capacity of human subjects to
process information (Voorhies & Scandura, 1974). This fixed capacity keeps the number of levels in goal switching, and the number of rules that may be tested at any given point in time, within very strict limits.

To summarize, learners who are able to skip prerequisites do so because they have already learned a higher order rule by which they can solve higher level tasks without instruction. In effect, higher order rules, heuristics, or strategies, whichever term might be preferred, correspond to Gagné’s (1962) instructions, by which learners may progress from one level in a learning hierarchy to a higher level. Single higher order rules, however, generally operate in large classes of different hierarchies (cf. Ehrenpreis & Scandura, 1974).

The significance of higher order rules can be seen in another way. Whereas failure on a superordinate task (as used by Gagné, 1970) invariably implies failure on one or more prerequisites, success on all prerequisites in traditional learning hierarchies does not necessarily imply success on superordinates. Whereas the former is a deterministic effect, the latter is not. In the latter case, higher order rules appear to provide the missing link toward determinism (Scandura, in press-a, in press-b).

Given obtained support for the hypothesized goal-switching control mechanism, the distinction between lower order rules and higher order rules becomes one of temporal use rather than of substance. (The equivalence of data and program in computer science is a direct analog.) The only formal distinction that needs to be made regarding level is that between specific competence (rules) and control. In addition to rules and higher order rules, every information processor, whether human or otherwise, requires some control mechanism to determine how available rules are to be put to use.

4. Learning Processes

HOW DO PEOPLE MAKE DISCOVERIES?

Specifying how a child is to learn a bit of content necessarily has the effect of specifying what processes the child is apt to learn along with the content. In particular, what is learned in discovering simple rules? Or, put differently, what might a teacher do to make it more likely for a learner to discover simple rules? A general rule of thumb might be that the more experience a person has in making such discoveries the better he will get. While probably correct, such an answer is relatively vague and would be of limited help in the explicit design of instruction.

According to the structural learning theory, all learning takes place by applying existing rules to generate new ones. But in the case of discovery, do higher order rules actually exist? To answer this question, consider the task of generating an output of 8 – 7, which is analogous to the given instances 1 → 3, 2 → 6, and 3 → 9. In order to solve this task, the child must first “discover” a common rule by which the output of each instance can be generated from the corresponding input.

The higher order generalization rule given in Figure 4 is adequate for “discovering” such rules. Applying this higher order rule to the three instances above, we (a) “look at” instance 2 → 6 (or 3 → 9); (b) perform 6 ÷ 2 = 3 (or 9 ÷ 3 = 3); (c) ask if the quotient (3) equals the output of 1 → 3; since the answer is yes, we (d) write n → 3n (i.e., “multiply by three”); (e) if the answer had been no, the problem would have been identified as outside the domain of the higher order rule.  

Figure 4. Flow diagram for the higher order division rule, which acts on instances and generates a general rule of the form n → an.

---

5 More complete schematic analyses of both “discovery” and “expository” learning are given in Scandura (1973a, chap. 7).
Strong behavioral support has been found for this analysis. In a study involving this and other higher order generalization rules, Scandura (1974c) found essentially perfect coordination between theoretical predictions and data; that is, the data were totally consistent with the assumption that the above control mechanism determines how people use rules, including higher order generalization rules. The experimental subjects succeeded on transfer problems, which required discovering simple rules, when and only when they knew both the requisite instances (degenerate rules) and a generalization rule that included the instances in its domain.

5. Content Versus Performance Analyses

Are they distinct or is reconciliation possible?

When subject matter has been involved, the discussion so far has centered on performance analyses. There have been parallel developments in subject matter analysis, however, often called content analysis. These developments have attracted considerable attention, in part because they appear to reflect many of the early concerns of subject matter specialists with structure (Bruner, 1960).

On the surface, at least, these two approaches appear quite different. The aim of performance analysis is to specify in terms of behavior the instructional objectives associated with particular content. As we have seen, one may also identify the competencies necessary to perform the behavior in question (see also Gerlack, 1975; Landa, 1974). In addition, given a particular behavioral objective, and the desired competency underlying it (at least implicitly), task analysis may be used to derive a learning hierarchy of prerequisite competencies. This can be accomplished by asking Gagné's (1962) now familiar question, "What must a learner know/be able to do in order to . . . ."

In general, however, subject matter concepts cannot be fully represented in terms of single behavioral objectives or learning hierarchies associated with same. Even such a relatively simple concept as "number" has a variety of interrelated aspects (e.g., assigning numbers to sets, counting, relations between adding elements and counting). Moreover, the prerequisite relation is not the only important way in which concepts may be related (e.g., consider the relation between the cardinal and ordinal aspects of number). It is not surprising, therefore, that whereas the behavioral emphasis of performance analyses has appealed to behavioral scientists, the approach has frequently been criticized by subject matter specialists as leading to fragmented curricula (e.g., Scandura, 1971b, chap. 1).

Content analysis, on the other hand, is concerned primarily with identification of the major "concepts," and the relationships among these concepts, that are characteristic of particular subject matter units. In content analysis, prerequisite relationships among concepts are typically represented in a heterarchical manner (Merrill & Gibbons, 1974). Thus, instead of having the most basic concepts at the bottom, as in learning hierarchies (Gagné, 1970), with progressively more complex concepts as one moves toward the apex, there are in hiearchies a variety of entry points and exits (e.g., see Figure 5). Whereas content analysis reflects conceptual relationships in a relatively simple manner, it also has important limitations. It does not, for example, lend itself to individual differences measurement in the same direct way that performance analysis does.

The apparently complementary advantages and disadvantages of performance and content analysis have not gone unnoticed. Indeed, the two forms of analysis have been combined in a variety of ways, both on theoretical (e.g., Pask, 1975) and on pragmatic (e.g., Merrill & Boudewill, 1973) grounds. Thus, content analysis may be used to map out general relationships, with performance analysis used to operationalize critical concepts in behavioral terms.

Simply combining content and performance analysis, however, has some major limitations:

Such an approach would appear to have much in its favor. In addition to minimizing obvious educational limitations, combining content and performance analysis has the further advantage of paralleling current thinking in cognitive psychology in which a distinction is frequently made between propositional and algorithmic knowledge (cf. Greeno, 1973; Scandura, 1974d). Propositional knowledge is represented in terms of relational nets of the sort obtained as a result of content analysis. Algorithmic knowledge is represented in terms of rules (procedures or flow diagrams) and corresponds to performance analysis.

The distinction between propositional knowledge and algorithmic (rule) knowledge has little to do with representation per se; both can be represented formally in terms of rules. It depends largely on the intended use. The term propositional is used if the knowledge is viewed as static—as something to be operated on. The term algorithmic refers to the use of knowledge as an operator. In general, nontrivial, propositional (relational) nets correspond to psychologically meaningful sets of rules or structures.
1. Relationships among (sets of) paths of relational nets, which correspond to higher order rules, are not usually or easily represented; for example, relationships among the parallel operations $a^{\pm 1}$, $b^{\pm 1}$, $c^{\pm 1}$ in Figure 5 are left unspecified. The same can be said of the more complex relation expressing the fact that the paths $\sin^{\pm 2}$, $\cos^{\pm 2}$, and $\tan^{\pm 2}$ in Figure 5 can be derived in the same way, respectively, from paths "substitute 1 $\circ c^{\pm 1} \circ a^{\pm 1} \circ$ substitute 2," "substitute 1 $\circ c^{\pm 1} \circ b^{\pm 1} \circ$ substitute 3," and "substitute 2 $\circ a^{\pm 1} \circ b^{\pm 1} \circ$ substitute 3." (Note: "$\circ$" means "followed by.""") In effect, parallel operations and relations are often scattered throughout relational nets with nothing specifically to indicate their relationships. (Pask's, 1975, theory of conversations also attends to the above concerns, although the language, emphases, and mechanisms of the theory differ substantially from the structural learning theory.)

2. Representing knowledge in terms of comprehensive relational nets essentially eliminates the need for any nontrivial, superordinate control mechanism. Aside from questions pertaining to working memory and attention, there is no question as to how units of knowledge interact, because there is only one unit of knowledge. In view of the discussion in Sections 3 and 4, however, it would appear that control mechanisms may play an important role in learning.

3. The fact that one can represent aspects of concepts, relations, and even higher order relations in behavioral terms, after they have been identified, says little about how to do it. Representation in terms of relational nets does not direct the analyst's attention toward higher order relations, so that representation in terms of behavioral objectives normally has an ad hoc character. For similar reasons, potential relationships to other content (such as can be represented by higher order rules) are never represented until (and rarely after) the new content has been made explicit.

Limitations 1 and 2 stem, I believe, from representing knowledge/content in terms of unified but largely static relational nets, and minimizing (if not ignoring) the operational aspects of knowledge. As an alternative to this type of representation, knowledge/content in the structural learning theory is represented in terms of discrete units, called rules, and of sets thereof, called structures.

It is worth noting parenthetically in this regard that different operations, even different parallel operations, constitute different rules, even where they have the same domains (e.g., consider $a^{\pm 1}$, $b^{\pm 1}$, and $c^{\pm 1}$, all of which operate on $a^2 + b^2 = c^2$). Conversely, a single operation may be attached to more than one domain, each combination of which constitutes a different rule. Such differences allow considerable flexibility in representing knowledge. For instance, incorrect knowledge can be represented just as well as correct knowledge (e.g., in school mathematics, one can represent the fact that

---

Figure 5. Relational net representing the Pythagorean theorem, some basic trigonometric identities, and their relationships.
many algebra students allow division by all numbers, even though division by zero is meaningless in ordinary arithmetic). Whereas the above distinctions are sometimes incorporated in relational net representations (as in Figure 5), this is rarely the case with higher order relationships. Higher order rules are especially convenient for this purpose.

Nonetheless, the simplicity provided by relational nets is an important advantage, which should not be overlooked. The relational net of Figure 5, for example, corresponds to a (sub)set of rules, or *structure*, which may serve as a unitary, psychological entity. Like relational nets, structures (rule sets) may be operated on as units by other rules. In this sense, they come to play a role directly analogous to that of Miller's (1956) "chunks," or Minsky's (Note 4) "frames." They are directly compatible, respectively, with the "substimuli" of Scandura (1973, chap. 10) and the "semantic knowledge" of Scandura (1973, chap. 6). In effect, relational networks may be viewed, psychologically speaking, as unitary complexes of rules, including higher order rules, which serve as idealized "chunks" upon which other rules may operate.

With regard to Limitation 3, and especially in view of the (partly) content-specific nature of task domains, it is fortunate that there are general, quasi-systematic methods (Scandura, 1974a) for identifying the higher order and other rules underlying given task domains (Scandura et al., 1974). Recently these methods have been extended to include higher order generalization/restriction and problem definition rules (Scandura, in press-b; Scandura & Durnin, in press-b). For illustrative purposes, let me briefly summarize a simple form of this method as applied to the above unit on trigonometric identities. The first step is to select a representative sample of tasks from the indicated domain. I shall consider only relationships between trigonometric identities and the Pythagorean theorem \((a^2 + b^2 = c^2)\). Suppose we choose the following tasks:

\[
\begin{align*}
1) & \quad a^2 + b^2 = c^2 \\
2) & \quad \sin^2 x + \cos^2 x = 1 \\
3) & \quad a^2 + b^2 = c^2
\end{align*}
\]

The second step is to identify rules for solving each sampled task, keeping in mind the constraints mentioned in Section 2. This gives

\[
\begin{align*}
1) & \quad \text{Divide by } a^2 \circ \sqrt{x^2 + y^2} \circ (x/y)^2 \circ \text{div substitute } 1; \\
2) & \quad \text{Substitute } 1 \circ (x/y)^2 \circ \sqrt{x^2 + y^2} \circ \text{mult by } c^2; \\
3) & \quad \text{Divide by } a^2 \circ \sqrt{x^2 + y^2} \circ (x/y)^2 \circ \text{div substitute } 2.
\end{align*}
\]

Although the above rule set accounts for the sampled tasks (i.e., there is a rule in the set that can be used to generate the desired output to each input), there are other tasks associated with the domain that cannot be accounted for; for example, consider

\[
\begin{align*}
4) & \quad 1 + \cot^2 = \csc^2 \quad \Rightarrow \quad a^2 + b^2 = c^2.
\end{align*}
\]

The third step is to identify parallels or patterns among the rules in the initial rule set. In particular, notice that the steps in Solution Rules 1 and 2 are reversed (including the inverse operations of multiplication and division).

Such parallels almost invariably reflect the presence of higher order rules. In turn, the identification of higher order rules, together with more basic lower order rules, often makes it possible not only to derive the initial solution rules but also to derive additional rules as well; for example, consider the second order rule set consisting of a higher order rule, \(h\), which reverses steps (in rules to which it is applied), together with Solution Rules 1 and 3.

Collectively, these rules can be used to derive solution rules for both Tasks 2 and 4. Thus, application of \(h\) to Rules 1 and 3, respectively, yields Solution Rules 2 and 4.

\[
\begin{align*}
(4) & \quad \text{Substitute } 2 \circ (x/y)^2 \circ \sqrt{x^2 + y^2} \circ \text{mult by } a^2.
\end{align*}
\]

The rule set, thereby, is said to *account* for Tasks 2 and 4. Hence, Solution Rule 2 may be eliminated as redundant. The derived rule set also accounts, of course, for Tasks 1 and 3.

Moreover, this rule set may be extended to account for a much broader range of tasks including such tasks as \(\sin^2 x + \cos^2 x = 1\) and \(a^2 + b^2 = c^2\). This may be accomplished by decomposing the solution rules into their components and introducing the higher order composition rule of Section 3. In the process, even Solution Rules 1 and 3 may be eliminated as redundant. It is this almost incidental (typically large) gain in generative power from which the structural method of analysis derives its power.

To summarize, the structural/algorithmic method of analysis provides a potential basis for bridging the gap between current content and performance analyses. (a) The representation of knowledge/
content in terms of rules, and sets thereof, appears to retain the advantages of performance and content analyses and to overcome the aforementioned limitations. (b) Rule sets are compatible with the control mechanism described above. They thereby reflect the close interdependence between competence (content) and cognitive theories. (Relationships between content and individual differences measurement have been described in Section 2.) (c) There are quasi-systematic methods of structural/algorhmitic analysis that make it possible to represent concepts and relations, including higher order ones, in behavioral terms. These methods may in principle be applied to any content.

It should be emphasized that the possibility of structural analysis does not solve educational problems in itself. Although rule-based analyses may be possible in principle, and will probably become increasingly available through future research, rigorous and comprehensive structural analyses require considerable time, effort, and resources. For many purposes, it may be sufficient to use less sophisticated methods, which nonetheless are compatible with the general approach (cf. Ehrenpreis & Scandura, 1974; Lowerre & Scandura, 1973; Scandura, Durnin, Ehrenpreis, & Luger, 1971). For example, in introducing teachers to the need for higher order processes, it may be sufficient initially to simply identify and illustrate various kinds of basic learning processes without any serious attempt to detail underlying rules (cf. Scandura, 1971b).

6. Human Development

How might horizontal decalage be analyzed in structural terms?

Current stage theories of human development are often suggestive regarding education, but they have not been sufficiently developed to provide anything close to an adequate basis for making specific instructional decisions concerning individual children. Nonetheless, initial steps have recently been taken in this direction. In 1972, Scandura showed how failure to identify competencies underlying number conservation has led to incorrect interpretations of data addressed to the task of “proving Piaget wrong.” Klahr and Wallace (1973) have shown how the competencies underlying conservation of quantity can be represented in terms of (computerized) production systems. Siegler’s (1975) analysis of balance beam tasks is directly analogous to the performance test theory described in Section 2.

To date, however, competence analyses of developmental phenomena have been limited to performance and have been restricted to particular (usually conservation) tasks. In contrast, many important developmental phenomena are more global in nature and involve a variety of tasks. Consider the Piagetian notion of horizontal decalage. Horizontal decalage refers to the irregular performance that is frequently observed across different task domains during transitions to new developmental stages (e.g., number versus length conservation tasks). As Beilen (Note 5) states, the issue of so-called horizontal decalage, or generalization across concept domains, has been extensively discussed in the Piagetian literature, and a number of studies by non-Genevans are addressed to this issue. As the data show, there is both generalization and variability; sometimes variability within a stage takes a consistent form, sometimes not. The issue is a difficult one for stage theory, and investigations sympathetic and unsympathetic to stage theory will be dealing with it for a long time.

The question here is whether horizontal decalage can be usefully analyzed in structural terms? Unfortunately, no hard answers to this question can be given in the absence of relevant empirical information (i.e., serious attempts at and behavioral evaluation of such analyses). On the other hand, there are no a priori reasons why analysis along the lines described would not work.

As outlined in Section 5, structural analysis would involve drawing a representative sample of conservation tasks and devising competence accounts consistent with the target populations. In particular, such accounts ought to be consistent with how a true and uniform conserver might be expected to deal with the tasks. (If there is more than one possibility, of course, each could be included and evaluated behaviorally.) Thus, a child who conserves length and area would be presumed to have acquired a general rule, applicable to both conservation tasks, and/or to have available a higher order rule by which relationships among areas can be generalized from linear relationships.

The competence account, in turn, might be used to identify the relevant knowledge available to individual children as described in Section 2. A child who conserves length but not area might be found not to have available either a generalized conservation rule or a higher order rule relating linear and two-dimensional quantities. Conversely, once having pinned down the availability of specific
higher and lower order source rules, it would be possible to make predictions concerning performance on new conservation tasks (derivable via the source rules).

In short, the theoretical machinery is available for guiding the analysis of developmental phenomena in structural terms and for evaluating such analyses behaviorally. Whether or not such guidance may prove sufficient, in its present form, to derive behaviorally valid accounts of horizontal decalage remains to be seen (see Scandura, in press-b, for latest developments).

7. Optimization and Adaptive Instruction

ARE ANALYTIC APPROACHES POSSIBLE?

Although I have, in Section 6, barely defined the problem, and although there are many important instructional problems with which I have not dealt, I hope enough examples have been presented to indicate the broad potential of the structural learning theory. Nonetheless, let me just sketch one more potential and fundamentally important area of application, that pertaining to optimization in adaptive instruction.

In schematic form, my approach contains the same ingredients as most existing theories of instruction (see Atkinson, 1972): (a) an underlying theory of learning, (b) a way of measuring student progress, (c) the assignment of values to educational objectives and costs (e.g., time/money) to instruction, and (d) some means of maximizing value per unit cost. There are, however, major differences, both overall and within each area; for example, whereas most other formal models are stochastic in nature and most assume that all subjects learn “from scratch” (Atkinson, 1972; cf. Kemp, 1974; Pask, 1975), the structural theory is deterministic and rests on the assumption that learning depends inextricably on what the learner knows.

The first step in extending the structural theory to instruction is to analyze the task domain. Given a class of tasks, the method of analysis (Scandura, 1974a) is a simple extension of that described in Section 5 and involves (a) sampling a wide variety of tasks; (b) identifying a set of solution rules (R) for solving the tasks (as idealized subjects in the target population might); (c) identifying parallels among the solution rules and devising higher order rules that reflect these parallels; (d) constructing more basic rule sets, which incorporate the higher order and other rules, and correspondingly eliminating solution rules made unnecessary by the higher order rules; (e) testing and refining the resulting rule set on new problems; and (f) extending it when necessary so that it accounts for both familiar and novel tasks in the domain. This method may be reapplied to the obtained rule set and repeated again as many times as desired. Each time the method is applied, the resulting rule set tends to become more basic in two senses: (a) the individual rules become simpler, and (b) the new rule set as a whole has greater generating power (i.e., it provides a potential basis for solving a greater variety of tasks; see Scandura, 1973a, pp. 114–117).

Once a basic rule set (B) has been identified, the entering knowledge of individual students relative to each rule might be determined using the assessment procedures described in Section 2. In addition, learning is assumed to take place as the learner interacts with the teaching environment, according to proposed control mechanisms. What is learned at each stage depends both on what is presented to the learner and on what he knows. The changes from stage to stage are cumulative.

More specifically, given a basic rule set B (which without loss of generality can be thought of as some student’s entering knowledge), it is possible to determine by algorithmic means (e.g., by computer) the problems that can be solved and correspondingly the rules that might possibly be learned by applying available higher order rules to other available rules (Scandura, 1973a). At any given stage of learning (due to information-processing capacity [Voorhies & Scandura, in press], time, and other prespecified cognitive limitations), only certain of the potentially solvable problems can be solved. Let B² denote the rule set that might be learned (at a given stage) by a person who knows exactly the rules in B. Recursively, then, Bⁿ denotes the rule set immediately learnable given the rules in Bⁿ⁻¹. Each rule in Bⁿ represents a unit of knowledge that might potentially be acquired by a learner who, on entering a curriculum, knows only the rules in B. In general, Bⁿ will be a far more encompassing and powerful rule set than the initial rule set R from which B, and ultimately Bⁿ, are derived. (It is this feature which accounts for “creative” potential.) It is important to notice that the ability to solve problems associated with Bⁿ comes about gradually as a result of solving
sequences of simpler problems associated with $B$, $B', \ldots, B^{n-1}$.

In order to talk about the optimization of instruction, two additional things must be done. First, educational objectives must be identified and values assigned to them. For present purposes, we can think of the objectives as behavioral objectives or rules (or classes of same) with the proviso that objectives corresponding to higher order rules are explicitly included. Clearly, objectives will vary in importance depending on what the teacher or curriculum constructor values most. Weights may be assigned to the various kinds of objectives to reflect these values. The second addition concerns the various costs of instruction. The time required to teach various types of rules, via various types of instruction, would provide one natural measure.

Given (a) a rule set $B$, which represents the knowledge available to a learner on entering a course, (b) the values assigned to the rules in $B^n$, and (c) the costs of presenting tasks and/or of expository rule instruction (which in principle can be calculated relative to what each learner knows at each point in the instruction), it is possible to determine the total cost and educational value associated with any given sequence of instruction. Sequence totals, then, may be compared to determine optimums (e.g., value per unit cost) for various initial conditions on the task domain $B$, values, and costs. As a very simple example, suppose $B = \{r_{ab}, r_{bc}, r_{cd}, 0\}$, where $r_{ab}$ is a rule for converting from measure $a$ (e.g., yards) to measure $b$ (e.g., feet); $r_{bc}$, from $b$ to $c$; $r_{cd}$, from $c$ to $d$; and $o$ is a higher order composition rule which operates on certain pairs of rules (e.g., $r_{ab}$ and $r_{bc}$) to generate their composites (e.g., $r_{ac}$, a rule that converts $a$ to $c$).

If the learner is initially presented with the task of converting a given measure $a$ to the appropriate number of $d$ units, and then the task of converting from $a$ to $c$, the learner will fail according to the theory on the first task; its solution (rule) requires composing all three component rules, whereas the available higher order rule can compose only two at a time. The subject would succeed, of course, on the second task and, in the process, acquire a rule for solving any $a$ to $c$ task. However, if the learner is first presented with the $a$ to $c$ task, and then the $a$ to $d$ task, he will succeed on both and in the process learn one rule for converting from $a$ to $c$, and another for converting from $a$ to $d$. (Once $r_{ac}$ is learned, $r_{ad}$ can be generated by applying $o$ to $r_{ac}$ and $r_{cd}$.) If each learned rule is given a value of 1 and the time cost for each task is assumed to be 5 minutes, then while the time required for each sequence would be 10 minutes, the educational value of the second instructional sequence would be twice the first.

In effect, the optimization of instruction, in this view, consists of finding optimal trade-offs between the sum of the values of the objectives achieved and the total time required for instruction. Optimization will involve balancing gains against costs (e.g., high instructional costs against highly valued objectives)—a type of cost-benefit analysis.

As this article goes to press, a thesis student, Wallace Wulfkeck, has extended along the above lines our previous analysis of geometry construction tasks (Scandura et al., 1974) and has programmed the system on the computer. He hopes to compare empirically the analytically generated sequences with sequences selected by students during the course of learning. 8

Although little more can be said at the present time, some sense of the potential of this line of work can be seen from the following. Suppose we have general information concerning the relative values to be assigned to various kinds of objectives associated with a curriculum and concerning the relative costs of various instructional methods that might be used. Assume that we want a rule of thumb for specifying the type of curriculum best suited to our needs.

In general, the assignment of values to objectives may vary as to homogeneity. The possibilities range from giving high value to relatively few objectives, with low value to the others, to moderate value over a broad range of objectives. At one extreme, for example, high value might be placed on computational skill in arithmetic, with little concern for meaning. At the other extreme, a broad range of arithmetic abilities might be given equal weight. Similarly, one might place high value on historical facts in a social studies course with little attention to other factors or, at the opposite extreme, give essentially equal weight to a broad range of social objectives. In Figure 6, at the risk of oversimplification, the possible assignments of values between these extremes is represented by a single dimension.

8 Wulfkeck's (1975) dissertation has since been completed. A summary report of the theoretical rationale and the empirical work has been prepared by Scandura and Wulfkeck and will appear in Scandura (in press-b).
Instructional time costs are represented similarly. In this case, the dimension ranges from high costs for all but a few kinds of instruction to more evenly moderate costs for a broader range of instructional forms. Thus, at one extreme, circumstances might require specific, teacher-directed instruction because of the relatively high costs of other forms. At the other extreme, a broad range of possibilities, as in learner-directed instruction, might be equally viable.

For illustrative purposes, a number of typical curriculum forms are plotted in Figure 6. Although one should not take too seriously the exact placement of these types, the figure does show how arbitrary curriculum forms might be viewed in these terms. The Summerhill school, which is well known for its broad-based, permissive atmosphere, is placed at the lower right. In this case, almost any educational objective may be deemed equally worthy, and differential costs attributable to different forms of instruction are largely ignored (i.e., they are all considered to be equal). In contrast, military training, with its high premium on the efficient acquisition of particular skills, is placed in the upper left. In short, the effectiveness of various types of curricula depends on the values assigned to possible educational objectives and the costs (or availability) of various kinds of instruction. The Summerhill version of an open classroom is a suitable form of instruction, but only when there is a broad range of desired objectives and largely equal costs for various forms of instruction.

In summary, it would appear that analytic approaches to optimization in adaptive instruction are possible. Although the problems remaining to be solved are formidable, they do not appear intractable.

**Conclusion**

Let me conclude with a statement of faith, arising from my own experiences. In common with most educational psychologists, I do not believe rhetoric to be the answer to instructional problems, and in
common with some (e.g., Rothkopf, 1973), I do not believe either that sheer volumes of empirical research will do. Rather, I concur with the belief of a growing core of individuals (e.g., DiVesta, 1973) that our greatest need is broad, fundamental theory tailored to the needs of education. There is nothing so practical as a good theory.

To my mind, and to the minds of most structural/system theorists (e.g., Landa, 1974; Pask, 1975), this means a theory that integrates content, cognition, and individual differences into a unified system and in which both the teacher and the learner play an integral role. Such theory should be operational and sufficiently precise so as to generate empirically falsifiable hypotheses. The structural learning theory was designed with these characteristics in mind. And while the theory is still far from a finished product, it offers one way of resolving a number of fundamental problems in instructional science.

REFERENCE NOTES


REFERENCES


Scandura, J. M. On higher order rules. *Educational Psychologist, 1974, 10*, 159-160. (b)

Scandura, J. M. The role of higher order rules in problem solving. *Journal of Experimental Psychology, 1974, 6*, 984-991. (c)


Scandura, J. M. A deterministic approach to research in instructional science. *Educational Psychologist*, in press. (a)

Scandura, J. M. *Problem solving: A structural/process approach (with instructional implications)*. New York: Academic Press, in press. (b)


