DOCUMENT RESUME

ED 099 419 TM 004 085

AUTHOR Passmore, David Lynn
TITLE Validation of Learning Hierarchies for Objective-Based Instructional Systems.
SPONS AGENCY Massachusetts Univ., Amherst. Center for Occupational Education.
PUB DATE Aug 74
NOTE 30p.; Paper presented at the Annual Meeting of the American Psychological Association (New Orleans, Louisiana, August 1974)

EDRS PRICE MF-$0.75 HC-$1.85 PLUS POSTAGE
DESCRIPTORS Criterion Referenced Tests; *Instructional Design; *Instructional Systems; *Learning Processes; Learning Theories; Methods; *Objectives; Research Methodology; Research Needs; Task Analysis; *Validity
IDENTIFIERS *Learning Hierarchies

ABSTRACT

The need for an empirically defensible means of sequencing instruction appears to have been the primary motivator for research into learning hierarchies. Also, valid learning hierarchies could act as congealing forces in individualizing instructional systems by providing psychological roadmaps for diagnosing students' preinstructional skills and monitoring students' progress through instructional systems. Four methods for generating candidates for learning hierarchies are available: introspection, formal analysis, observation, and statistical "fishing." Experimental transfer of training studies and statistical studies have been used to evaluate posited hierarchies. White recently made significant modifications of these standard evaluation methods to improve the internal validity of research into learning hierarchies. Several external validity issues remain to be resolved before scarce resources should continue to be allocated for additional learning hierarchy research. (Author)

ED 099419

BEST COPY AVAILABLE

VALIDATION OF LEARNING HIERARCHIES FOR OBJECTIVE-BASED INSTRUCTIONAL SYSTEMS

by

David Lynn Passmore

University of Massachusetts/Amherst

Paper presented at a Symposium on
Recent Developments in Criterion-Referenced Assessment
(R.K. Hambleton, Chairman)
Sponsered by the Psychometric Society at the
Annual Meeting of the
American Psychological Association
New Orleans, Louisana

August 30, 1974

TM 004 085

Acknowledgements

The preparation of this paper was supported by the Center for Occupational Education, University of Massachusetts/Amherst. The careful bibliographic work and advice of Elizabeth Marchant is gratefully acknowledged. This paper will be published as a tehcnical report of the Laboratory of Psychometric and Evaluative Research at the University of Massachusetts.

VALIDATION OF LEARNING HIERARCHIES

FOR OBJECTIVE-BASED INSTRUCTIONAL SYSTEMS

David Lynn Passmore
University of Massachusetts/Amherst

Focus of the Paper

Gagne (1968) considered the selection of a sequence for presenting learning events to be the most important decision in the design of an instructional system. This paper describes the nature of instructional sequencing research completed in the last 10 to 12 years that has been based on Gagne's studies of learning hierarchies. The purposes of this paper are to:

(1) synthesize the major reasons for interest in learning hierarchy research;

and (2) briefly describe methodologies available for this type of research.

This paper is organized under three major topics. First, factors in American intellectual heritage and in the development of instructional systems theory are related to the emergence of interest in learning hierarchy research. Second, methods for generating and evaluating learning hierarchies are reviewed. Finally, the previous two headings are summarized and some prescriptive statements are made concerning the future of learning hierarchy research.

Reasons for Interest in Learning Hierarchies

Intellectual Heritage

Forty years before he became president, John Adams was a young

schoolmaster in Worcester, Massachusetts. He kept a now-famous diary in which he tried to synthesize his ideas and experiences. Everything went into his record--the weather, teas, and dinners, conversations, his thoughts on sermons he heard and books he read--but the welter of detail organized itself into an unplanned, but discernible, rhythm.

During these apprentice years in Worcester, Adams enacted a quiet and private drama. He was recently graduated from Harvard, not yet embarked on a career, and vulnerable to alternating moments of despondency and hope. But, even in the privacy of his diary, he translated his problems into very public and traditional categories. The diary became a sort of morality play in which Adams' elations and fears paraded behind masks of abstract ideas about human nature and the world at large. On one side was the ancient notion of man's erratic frailty. On the other was a confidence, almost as ancient, in the fundamental rationality of God's creation. For Adams, and other Americans confidence in the orderlines of nature provided assurance of the continual and vigilant providence of God.

Belief in the orderliness of nature has been a kernel idea in our intellectual history from Greek times to Adam's to the present. It has been an organizing concept in the sciences. This belief is evident in Darwin's *On the Origin of the Species*, in Malthus' *Essay on the Principle of Population*, and in other important essays. It has been deeply rooted in the core of American tradition in the physical and natural sciences as well as economics. Education was also influenced by this thinking. If nature operates in an orderly manner, then the most effective methods of conducting the teaching/

learning enterprise could be uncovered using the scientist's tools. The study of educaticn as a science is founded on these principles.

Horace Mann and other early nineteenth century school reformers evangelically strove to channel generalized faith in education into support for the common school. The next generation of schoolmen, more bureaucrats than evangelists, believed that regimentation and standardization were necessary to unify support for public education. This belief led to the search for the "one best way", a *rational* system, for organizing struction. The zeal which characterized this search is evident in Philbrick's (1885) hyperbole: "If America devised the best school desk, it must go to the ends of the civilized world [p.58]."

At the beginning of the twentieth century, reform-conscious America discovered "scientific management' with the help of Frederick W. Taylor (1911). "Scientific management" principles injected the concept of efficiency into American business practices by developing methods for maximizing industrial outputs for a minimum of inputs and, thereby, minimizing the cost of the enterprise and maximizing the return on capital invested. Callahan (1962, esp., p.23 but also *passim*) documented the spread of applications of "scientific management" dogma to household tasks, family duties, church functions, and, ultimately, the process of education. Not only were the most effective instructional methods pursued, but the dimension of cost-consciousness was also added. The study of the economics of education is founded on these principles.

Armed with tools such as systems theory, educational tests, and cost/effectiveness analysis, instructional systems experts entered

the educational research arena in the liturgical spirit of the educational efficiency movement. One research problem attacked by these research workers was the determination of the most effective and efficient ways to sequence learning events.

Development of Instructional Systems Theory

One macro-level issue in curriculum design is the determination of the proper sequence among elements of a set of desirable educational experiences. For example, should students receive instruction in American History before they study civics? Should instruction in English grammar precede the study of English literature? Should students learn to weld in school-based occupational programs before they engage in on-the-job welding training programs? These broad sequencing questions, as well as antecedent questions concerning curricular scope and content, may find answers through professional analysis and public debate of the aims and means of schooling. Unique, consensus solutions to the problems posed by these and other general curriculum questions have not emerged although gallons of academic blood have been shed by curriculum theorists in arguments over substitutable problem solutions.

Rather than deal with molar curricular issues, instructional systems technicians have pursued answers to more micro-level problems. One popular touchstone of quality in current instructional systems theory is the requirement that desirable learning outcomes for students exiting an instructional system be stated in precise, intersubjectively observable terms (Lawson, 1974). Instructional systems using this approach are often called objective-based instructional systems.

Designers of objective-based instructional systems are confronted

with a large number of technical problems, one of which is selecting the temporal order of presentation of learning events. For example, must instruction in addition of single digit numbers precede lessons in the addition of numbers with multiple digits? This sequencing problem is analogous to the curriculum theorists' sequencing dilemma. Naturally, the two problems differ in [illegible]e. Also, instructional systems designers' sequencing problems should be less elusive, less value-laden, and more answerable by direct empirical evidence than are similar sequencing problems faced by curriculum workers.

Standard solutions to the problem of sequencing learning events in objective-based instructional systems have been provided for various school subject matter areas by: (1) identifying important behaviors that must be mastered for competence in a particular subject matter area; (2) stating intended learning outcomes, or educational objectives, which precisely state the behaviors which must be mastered as well as the degree to which and conditions under which the behaviors are to be exhibited by students; and (3) determining whether students' mastery of one objective depends upon their mastery of other objectives. Gagne (1970, Ch. 4) assigned the term "learning hierarchy" to designate the set of dependencies among component skills within a learning task. In objective-based instructional systems, precisely stated educational objectives define the component skills of learning tasks. An important assumption is that it is thought to be more appropriate to organize instruction by using functional dependencies among objectives in a learning hierarchy than it is to use logical, intuitive sequencing strategies (Cf. however, Briggs, 1967 for other sequencing viewpoints).

Learning hierarchies discovered through this three-part analysis serve at least three functions in objective-based instructional systems. First, from a design angle, they might suggest the most efficient ways to sequence learning events. Second, a diagnosis of a student's mastery of the component skills in a learning hierarchy may guide decisions concerning the student's entry into the instrucational sequence. This acknowledges that students differ with respect to their mastery of the component skills in the learning hierarchy as they enter the instructional system and provides a mechanism for individually-tailoring instruction to the student's pre-instructional skills. Third, learning hierarchies are a tool used in criterion-referenced assessment plans for monitoring student progress through objective-based instructional systems. So, as Gagne (1968) asserted, a valid learning hierarchy could be a most useful device for individualizing an objective-based instructional system.

It follows that specialists in instructional systems as well as in educational measurement have shown particular interest in learning hierarchy research. Instructinal systems specialists' interest in instructional design issues motivates their study of learning hierarchies. Robert Gagne and his associates performed pioneering research into the development and validation of learning hierarchies (Gagne & Bassler, 1963; Gagne, Mayor, Garstens, & Paradise, 1962; Gagne & Paradise, 1961; Gagne & Staff, 1965). Gagne's studies seem to have crystallized the methodological directions for subsequent efforts in learning hierarchy research by instructional systems specialists (White, 1973, p. 367).

Some educational measurement specialists are interested in the unique contributions that criterion-referenced assessment theory could make in the diagnosis of, and subsequent instructional placement on the basis of, relevant pre-instructional skills. Some of these specialists also believe that any strong criterion-referenced assessment theory should be applicable to making decisions concerning learners' progress through instructional sequences (Lindvall & Cox, 1969).

Most of the work completed on criterion-referenced assessment theory has had, at best, a pre-scientific character (e.g., Cox, 1970; Cox & Vargas, 1966; Popham & Husek, 1969). However, more coherent specifications of measurement theory necessary for criterion-referenced assessment were presented by Hambleton and Novick (1973) and Swaminathan, Hambleton, and Algina (in press). The decision-theoretic base of the Hambleton, *et al.* approach articulates well with the typical needs of existing instructional system monitoring operations reviewed by Hambleton (in press).

Among the novel approaches to monitoring student progress through the use of criterion-referenced tests in objective-based instructional programs is the technique of branched, sequential, or tailored testing (Lord, 1968; Ferguson, 1969; Wood, 1973). Tailored testing has been defined as a strategy for testing in which the sequence of test items (Spineti & Gambleton, 1973, p.2).

In summary, it appears that the need for an empirically defensible means of sequencing learning events in instructional systems has been the primary motivator for research into learning hierarchies. Also, valid learning hierarchies could act as congealing forces in individualizing instructional systems by providing psychological

roadmaps for diagnosing students' pre-instructional skills and for monitoring students' progress through instructional systems.

Methodologies Available for Generating and Evaluating Learning Hierarchies

A three-part process for conducting sequencing research in the design of objective-based instructional systems was outlined in the previous section. In the present section, attention is focused on the third part of this process which is the determination of whether students' mastery of one objective depends upon their mastery of other objectives. Specifically, methods for generating candidates for learning hierarchies and for evaluating these positied hierarchies are reviewed. Similar methodological reviews may be found in Briggs (1967), Resnick (1971), Resnick and Wang (1969), Smith (1972), Walbesser and Eisenberg (1972), and White (1973, 1974b).

The first two parts of the instructional sequencing research process are treated lightly in this section. Methods for identifying important behaviors that must be mastered for competence in a particular subject matter area and for explicity stating intended learning outcomes are not directed reviewed even though they are important steps in instructional sequencing research.

Techniques for selecting important behaviors to be learned in an instructional system are usually classified under the topic of task analysis. Some methods for task analysis also yield plausible candidates for learning hierarchies and will, therefore, be reviewed in this secion. Conventional wisdom in explicitly stating intended learning outcomes is recorded in Ammerman and Melching (1966),

Gronlund (1970), and Mager (1962). It will be suggested that fixation on this conventional wisdom may subvert the discovery of valid learning hierarchies.

Generating Candidates for Learning Hierarchies

A provisional ordering of instructional objectives must be posited before learning hierarchy research can proceed for a particular instructional system. How may these candidates for learning hierarchies be generated? Each researcher uses some method to complete this task even though the range of available methods have not been well documented. And it could be disastrous to operate without a hunch about the starting point for this task. Consider that there are $k!$ possible *linear* orderings for k objectives and that the number of conceivable hierarchies capable of being generated from k objectives is much larger. The sources of hunches about plausible hierarchies may be grouped into four categories: (1) introspection; (2) formal analysis; (3) observation; and (4) statistical "fishing."

Introspection. One popular method for generating candidates for learning hierarchies is to ask the question, "What would an individual need to know or do to display competence in this subject matter?" This same question, paraphrased from Gagne and Paradise (1961, p.4), is applied again to the original answer and, then, successively to subsequent answers. This method of questioning produces a learning hierarchy characterized by general, molar behaviors at the top of the hierarchy and more specific, subordinate, and molecular behaviors near the bottom of the hierarchy (e.g., Gagne, Mayor, Garstens, & Paradise, 1962, p.4).

This introspective method of hierarchy generation most frequently

appears in the literature with the instructional system developer the most prominent actor in this process. There are variations on this theme. For example, subject matter experts (Payton, 1971) and students (Kaplan, 1964) have been asked to play introspective roles to generate learning hierarchies.

Formal analysis. Formal analyses of subject matter domains provide another source of learning hierarchies. Good examples are the industrial manufacturing and construction domains developed by the Industrial Arts Curriculum Project (Towers, Lux, & Ray, 1966). Each of these subject matter domains was organized as a hierarchical taxonomy. Levels of each hierarchy proceed from general to more specific elements of manufacturing and construction practice. Within each level, elements were devised to be individually inclusive and mutually exclusive.

Hierarchies supplied by the introspective method previously discussed are developed by imagining the learner's interaction with the subject matter. In contrast, the method of formal analysis removes consideration of the learner and, instead, applies the logical rigor of taxonomic inquiry that has been developed for the natural sciences (see Gregg, 1954).

Observation. Learning hierarchies may also be posited by observing the natural order in which learners acquire behaviors. For example, a problem could be assigned to learners and, then, the milestones in their progress toward the problem solution could be recorded. Perhaps observational methods similar to those described by Piaget and Inhelder (1969) for researching hierarchies among developmental tasks might be fruitful in learning hierarchy research.

While the method of observation introduces a realistic glimpse at learners as a source of learning hierarchies, it must be remembered that the sequence which learners choose to solve a problem may prove to be neither effective nor efficient. And maximum effectiveness and efficiency in facilitating learning are primary quality criteria for instructional systems.

Statistical "fishing." Also, candidates for learning hierarchies are often captured through statistical studies which may be best characterized as fishing expeditions. The prior absence of a hypothesized network of educational objectives is a distinguishing feature of these types of studies. The main aim of these statistical studies is to suggest empirically-based, *post hoc*, structural hypotheses about the relationships among a set of behaviors.

Many of the numerical techniques, such as hierarchical cluster analysis (Tryon, 1958; McQuitty, 1960; Ward, 1961) and hierarchical factor analysis (Coombs & Satter, 1949; Palmer & McCormick, 1961), applied to military task analysis could be used to fish for provisional candidates for learning hierarchies. Baker (1972), Chenzoff (1964), Cotterman (1959), Folley (1964), Morsh, Madden, and Christal (1961) reviewed some of the more standard numerical approaches to the development of hierarchical taxonomies of behavior. A conference report by Moss and Smith (1970, esp. papers by Ammerman, Christal, and Silverman) described the application of military task analysis techniques to the construction of vocational education curricula. In addition, a rich and suggestive source of numerical taxonomy methods that could be applied in learning hierarchy research may be found in the natural science literature (For starters see, Sokal &

Sneath, 1963; Sneath & Sokal, 1962). Driver's (1963) survey of numerical classification methods in anthropology might also be useful.

One prominent line of thought in learning hierarchy research is that functional dependencies among instructional objectives can be discovered by examining the Guttman (1944) scalability of the test items measuring the behaviors specified by the objectives. Since the success of Guttman's procedure demands a linear ordering of items which measure a unidimensional entity, some researchers have preferred to use Lingoes (1963) multidimensional extension of Guttman's procedure to uncover non-linear orderings among educational objectives. Airasian and Bart (1971) and Bart and Kruz (1973) developed a more parsimonious technique for discovering hierarchies among test items. Applications of the Bart *et al.* ordering-theoretic method were presented in Airasian and Bart (1974), Bart (1972a, 1972b) and Bart and Airasian (1972).[1] Because it is relatively new and could have strong applications learning hierarchy research, the Bart *et al.* method is briefly reviewed.

---

[1]A computer program which implements to ordering-theoretic method and is compatible with the CYBER 76 model computer is available from W.M. Bart, Psychological Foundations of Education, 330 Burton Hall, University of Minnesota, Minneapolis, Minnesota, 55455.

A universe of response patterns to two binary-scorable (0= answer incorrect; 1= answer correct) test items is enumerated below:

$$v_1 = (0,0);$$

$$v_2 = (1,1);$$

$$v_3 = (1,0);$$

$$\text{and} \quad v_4 = (0,1).$$

Bart, *et al.*, described vectors v_1, v_2, and v_3 as *confirmatory* response patterns; performance on the first test item is considered prerequisite to performance on the second item. On the other hand, response vector v_4 is said to *disconfirm* the prerequisite relationship between items one and two.

Suppose that two items were administered to N examinees. If the percent of the N examinees exhibiting a disconfirmatory [illegible] pattern is less than some tolerance percent, then item one is said to be prerequisite to item two. A low tolerance percent, often one to five percent, is recommended.

The Bart, *et al.*, ordering theory may be generalized to I items. The observed percent of disconfirmatory response vectors for all two item pairs in a set of I items may be displayed in an I x I matrix indexed along the rows and columns by the item numbers. The prerequisite relationships among any of the I items may be determined by examining the disconfirmatory response matrix and can be portrayed by means of a line graph joining the prerequisite item pairs. The line graph is the candidate for further validation as a learning hierarchy.

Evaluating Posited Hierarchies[2]

An ordering of objectives posited for an instructional system must be evaluated to determine whether it should be rejected, modified, or adopted. Experimental transfer of training studies as well as statistical studies have been used in this evaluation task. Both types of studies were extensively reviewed and heavily criticized by White (1973, 1974a, 1974b) who, as a result, presented (White, 1974c; White & Clark, 1973) significant modifications of earlier paradigms for research into learning hierarchies.

Standard evaluation methods. A transfer of training effect is the action that learning one task has upon subsequent learning or performance of another task (Andreas, 1972, p. 439). Networks of positive -- that is, beneficial and facilitating -- transfer among a set of instructional objectives are frequently sought in the evaluation of learning hierarchies for instructional system design. The definition of a hierarchical relationship as positive transfer among learning events was used in Gagne's pioneering studies which set the tone for the learning hierarchy research efforts that followed.

How might mastery of one instructional objective facilitate the mastery of another objective? Several hypotheses are reasonable. First, some *identical* behaviors may be required to master both objectives. Work habits or specific subject matter skills are examples of identical behaviors that might precipitate transfer effects. Second,

---

[2] A review of papers by Airasian and Bart (1974), Durrell (1974), Huntley (1974), Phillips (1974), Roudabush (1974), Smith (1974), and Taylor, Bart, and Rosehel (1974) could have contributed to the comprehensiveness of this section but copies of these papers were unavilable during the preparation of this paper.

positive transfer may occur when *similar* stimuli and responses are involved in the mastery of both objectives. The psychological mechanisms operating to induce transfer in this case would be stimulus and response generalization.

Transfer of training studies are usually conducted in controlled settings. In the application of such studies to learning hierarchy research, factors such as warm-up effects and transfer of general work habits must be partitioned from the transfer of specific subject matter skills to unequivocally answer the question "What is learned?". Murdock's (1957) evaluative review of numerous designs used in transfer experiments might be helpful to learning hierarchy researchers as may the review and critique of methods for measuring transfer effects provided by Gagne, Foster, and Crowley (1948).

In addition to experimental studies, the same methods described as methods for statistical "fishing" for learning hierarchy candidates are also applied in the evaluation of posited hierarchies. Rather than using these statistical techniques to hypothesize hierarchies, the focus, instead, is on testing hypothesized hierarchies.

An interest in validating learning hierarchies not only implies that information will be obtained to discern the appropriateness of the posited hierarchies, but also that the hierarchies will be improved if necessary. However, the range of decisions and decision-making strategies for improving learning hierarchies have not been carefully delineated. For example, one decision that could be made on the basis of experimental or statistical evaluative evidence is that the objectives need to be reordered. Another decision that could be made based

on the same data might be that the objectives need to be subdivided, refined, and, then, reordered. But an explicit method for rationally choosing between these competing decisions is lacking but could be a fruitful line of future research. Methodologies independently developed for the formative evaluation of educational products might suggest solutions to this problem (see Passmore, Asche, & O'Kelley, 1974).

Incidentally, Hively (1972), asserted from a radical behaviorist's viewpoint that traditional methods of stating instructional objectives (so-called "behavioral" objectives) usually lead to the specification of too many classes of behavior within an objective with one result being muddled and untestable hierarchies among objectives. Hively, Patterson, and Page's (1968) use of item forms was an attempt to clarify more homogeneous intended learning outcomes for instructional systems. So, following conventional wisdom in the statement of instructional objectives may lead to problems in the evaluation of any posited hierarchy among these objectives.

White's modifications. White (1973, p. 371; 1974, p.1) felt that most evaluations of laarning hierarchies were hindered by the following six problems:

(1) small sample size;

(2) imprecise specification of hierarchy elements;

(3) the use of only one question per element to test the dependencies among elements;

(4) the absence of a test of hierarchical independence that takes error of measurement into account;

(5) the practice of testing for dependencies among hierarchy elements *after* instruction is completed on all elements, thus confounding forgetting with a lack of connection between elements;

and (6) lack of face validity for hierarchies which have been empirically evaluated and improved.

White's solutions to the first problem were to increase the size confirm the representativeness of thesample of subjects selected for learning hierarchy evaluations. Problem two was similar to Hively's objection and White suggested that this problem could be best solved through introspection during the generation of a learning hierarchy candidate. However, no clues were given on how to detect this problem during the evaluation phase of learning hierarchy research.

Problems three and four are related. The use of only one question per element does not allow the estimation of the error in measuring the dependency among elements. Moreover, none of the existing dependency indexes reviewed by Capie and Jones (1971) make the concept of error of measurement explicit. Also, White (1974) charged that indexes such as Gagne and Paradise's (1961) proportion of positive transfer as well as variants of it proposed by Walbesser and Eisenberg (1972), Guttman's coefficient of reproducibility (used in Resnick & Wang, 1969), and the four-fold contingency table and related phi correlation coefficient suggested by Capie and Jones (1971) share a common problem: each of these indexes can have values that indicate a hierarchical connection even when the skills are really independent. Certainly, the same charge could be directed at the Bart, *et al.*, ordering-theoretic method.

White and Clark (1973) introduced a statistical test of hierarchical dependency which they purport treats the problem of error of measurement and also allows the inference of hierarchical dependencies among objectives for some population of interest. White and Clark

carefully discussed the power function for this significance test in their *Psychometrika* paper. Unfortunately, White failed to discuss power issues in subsequent articles aimed at less technically sophisticated audiences. And it is precisely these audiences which repeatedly fail to recognize relationships between sample size and the power of significance tests in their research work (see Meehl, 1967). Consequently, educationally valid connections among elements in a hierarchy may, and probably will, be rejected if standard Neyman-Pearson hypothesis testing procedures are used in experiments with large samples of subjects that White also suggested. This point must be explained to prac· ·:al research workers.

To solve the fifth problem, White advocated the administration of test items at key points during a learning program designed to teach the subject matter in the hypothesized sequence under consideration. Then, White and Clark's significance test may be applied to determine whether dependencies exists among elements in a posited hierarchy. Of course, thi<r raises a fidelity issue: Does it make a difference that the hierarchy under consideration may not ultimately be used to sequence a learning program? Perhaps this question merits investigation.

The statement of White's sixth problem seems to be a reaction to blind, "dustbowl" empiricism in learning hierarchy research. White claimed that learning hierarchies modified on the basis of evaluative information should also be required to make logical, intuitive sense. If this is so, then what purpose does the evaluation serve and how are decisions to be made with an appropriate mix of logical and empirical informatin? On the other hand, unexamined learning hierarchies may have an intuitive appeal but may also seriously lack

empirical import. As has already been asserted in this paper, the explication of decisions and decision-making strategies for the improvement of learning hierarchies would be a noteworthy contribution to the literature.

Concluding Remarks

Summary

The need for an empirically defensible means of sequencing instruction appears to have been the primary motivator for research into learning hierarchies. Also, valid learning hierarchies could act as congealing forces in individualizing instructional systems by providing psychological roadmaps for diagnosing students' pre-instructional skills and for monitoring students' progress through instructional systems. Four methods for generating candidates for learning hierarchies were reviewed: introspection, formal analysis, observation, and statistical "fishing". Experimental transfer of training studies and statiscial studies have been used to evaluate posited hierarchies. White suggested significant modifications of these standard evaluative methods to improve the internal validity of research into learning hierarchies.

Should We Continue Research Into Learning Hierarchies?

The study of learning hierarchies has all the trappings of what Kuhn (1970; see also, Schwab, 1969) has called normal science: "Normal science means research firmly based upon one or more past scientific achievements, achievements that some particular scientific community acknowledges for a time as supplying the foundation for its further

practice [Kuhn, 1970, p.10]." There are three foci for normal science: gathering facts to generate theory, predicting new facts from theory, and reducing dissonance between the theory and the real world (Kuhn, 1970, p. 25-27). Kuhn uses the term paradigm to indicate the particular research tradition chosen to conduct normal science. The acquisition of a paradigm by an area of inquiry is a sign of the area's scientific maturity. Paradigms are also prescriptive since they govern the range of research problems acceptable to the scientific community. Moreover, they dictate the topics and modes of inquiry which students must master for acceptance and recognition in the scientific community. Studies in the sociology of science indicate that the scope and structure of paradigms are controlled by visible and revered statesmen in an area of inquiry who, in turn, influence the behavior of younger, less prestigious researchers struggling below them.

In terms of the sociology of psychology as a science, the importance of Robert Gagne to the current interest in learning hierarchy research is not trivial. Gagne certainly fits the elder stateman's role in psychology. Would the study of learning hierarchies have been important if Gagne had not chosen to become involved? Of course an equally valid question is whether learning hierarchy research would have been discovered or received any creative impetus without his involvement? More generally, are topics in instructional science researched because of the importance of their patrons or are they studied for their intrinsic importance? Unraveling the answers to these questions may serve to demystify the study of learning hierarchies and lead to the statement of several critical questions for the future

of learning hierarchy research.

Could we trim away our interest in learning hierarchies and, thereby, achieve a more parsimonious instructional science? Could we conduct the teaching/learning enterprise without learning hierarchies? Would any instructional sequence be better than none at least as good as one suggested by resource-consuming learning hierarchy research? Answers to these crucial questions, and others, may contribute evidence for determining the external validity of learning hierarchy studies. I would contend that these questions are just as important as the present exclusive and pervasive interest in the internal validity of learning hierarchy experiments which has been reviewed in this paper. But these are issues which must be expanded upon in subsequent papers.

References

Airasian, P. W. and Bart, W. M. Tree theory: A theory generative measurement model. Paper presented at the annual meeting of the American Educational Research Association, New York, 1971.

Airasian, P. W. and Bart, W. M. The use of ordering theory in the identification and specification of instructional sequences. Paper presented at the annual meeting of the American Educational Research Association, Chicago, 1974.

Ammerman, H. L. and Melching, W. H. Derivation, analysis, and classification of instructional objectives. Alexandria, Virginia: George Washington University, Human Resources Research Office, 1966.

Andreas, B. G. Experimental psychology. (Second edition) New York: Wiley, 1972.

Baker, F. B. Numerical taxonomy for educational researchers. Review of Educational Research, 1972, 42 (3), 345-358.

Bart, W. M. A hierarchy among attitudes towards animals. The Journal of Environmental Education, 1972, 3 (4), 4-6, (a).

Bart, W. M. A hierarchy among attitudes toward the environment. The Journal of Environmental Education. 1972, 4 (1), 10-14, (b).

Bart, W. M. and Airasian, P. W. The generation of item hierarchies by an ordering theoretic method on a Piagetian example. Paper presented at the annual meeting of the American Educational Research Association, Chicago, 1972.

Bart, W. M. and Kruz, D.J. An ordering-theoretic method to determine hierarchics among items. Educational and Psychological Measurement, 1973, 33, 291-300.

Briggs, L. Sequencing of instruction in relation to hierarchies of competence. Palo Alto, California: American Institutes for Research, 1967, ED 004034.

Callahan, R. E. Education and the cult of efficiency. Chicago: University of Chicago Press, 1962.

Capie, W. and Jones, H. L. An assessment of hierarchy validation techniques. Journal of Research in Science Teaching, 1971, 8 (2), 137-147.

Chenzoff, A. P. A review of the literature on task analysis methods. Port Washington, New York: U.S. Naval Training Device Center, 1964, (NAVTRADEVCEN 1218-3).

Coombs, C. H. and Satter, G. A. A factorial approach to job families. Psychometrika, 1949, 14, 33-42.

Cox, R. C. Evaluative aspects of criterion-referenced measures. In, Criterion-referenced measurement: Emerging issues. Symposium presented at the annual meeting of the American Educational Research Association, Minneapolis, 1970.

Cox, R. C. and Vargas, J. S. A comparison of item selection techniques for norm-referenced and criterion-referenced tests. Pittsburg: University of Pittsburg, Learning Research and Development Center, 1966.

Driver, H. E. Survey of numerical classification in anthropology. In D. H. Hymes, (Ed.), The use of computers in anthropology. The Hague: Mouton, 1963.

Durell, A. B. A computer simulation of measures for validating learning hierarchies. A paper presented at the annual meeting of the American Educational Research Association, Chicago, 1974.

Ferguson, R. L. The development of a computer-assisted branched test for a program of individually prescribed instruction. Unpublished doctoral dissertation, University of Pittsburg, 1969.

Folley, J. D. Guidelines for task analysis. Port Washington, New York: U. S. Naval Training Device Center, 1964, (NAVTRADEVCEN 1218-1).

Gagne, R. M. Learning hierarchies. Educational Psychologist, 1968, 6 (1), 1-9.

Gagne, R. M. The conditions of learning. (2nd edition) New York: Holt, Rinehart, and Winston, 1970.

Gagne, R. M. and Bassler, O. C. Study of retention of some topics in nonmetric geometry. Journal of Educational Psychology, 1963, 54, 123-131.

Gagne, R. M., Foster, H. and Crowley, M. E. The measurement of transfer of training. Psychological Bulletin, 1948, 45, 97-130.

Gagne, R. M., Mayor, J. L., Garstens, H. L. and Paradise, N. E. Factors in acquiring knowledge of a mathematical task. Psychological Monographs, 1963, 76 (7, Whole No. 526).

Gagne, R. M. and Paradise, N. E. Abilities and learning sets in knowledge acquisition. Psychological Monographs, 1961, 75, (14, Whole No. 518)

Gagne, R. M. and Staff, University of Maryland Mathematics Project. Some factors in learning nonmetric geometry. Monographs of Society for Research in Child Development, 1965, 30, 42-49.

Gregg, J. R. The language of taxonomy. New York: Columbia University Press, 1954.

Gronlund, N. E. Stating behavioral objectives for classroom instruction. New York, MacMillan, 1970.

Guttman, L. A basis for scaling qualitative data. American Sociological Review, 1944, 9, 139-150.

Hambleton, R. K. A review of testing and decision-making procedures for selected individualized instructional programs. Review of Educational Research, in press.

Hambleton, R. K. and Novick, M. Toward an integration of theory and method for criterion-referenced tests. Journal of Educational Measurement, 1973, 10, 159-170.

Hively, W. W. Informal comments made in University of Minnesota seminar in domain-referenced testing, 1972.

Hively, W. W., Patterson, H. L. and Page, S. H. A. A "universe-defined" system of arithmetic achievement tests. Journal of Educational Measurement, 1968, 5, 275-290.

Huntley, G. Empirical comparison of two methods for devising learning hierarchies. Paper presented at the annual meeting of the American Educational Research Association, Chicago, 1974.

Kaplan, J. D. An example of student-generated sequences in mathematics instruction. Mathematics Teacher, 1964, 57, (5), 298-302.

Kuhn, T. S. The structure of scientific revolutions. Chicago: University of Chicago Press, 1962.

Lawson, T. E. A performance-based instructional theory. Journal of Industrial Teacher Education, 1974, 11 (3), 52-64.

Lindvall, C. M. and Cox, R. C. The IPI evaluation program. AERA Monograph on curriculum evaluation. No. 5. Chicago: Rand-McNally, 1970.

Lingoes, J. C. Multiple sclaogram analysis: A set theoretic model for analyzing dichotomous items. Educational and Psychological Measurement, 1963, 23 (3), 501-524.

Lord, F. M. Some test theory for tailored testing. Princeton, New Jersey: Educational Testing Service, 1968, (RB-68-38).

McQuitty, L. L. Comprehensive hierarchical analysis. Educational and Psychological Measurement, 1960, 20, 805-816.

Mager, R. M. Preparing instructional objectives. Palo Alto, California: Fearon, 1962.

Meehl, P. Theory testing in psychology and physics: A methodological paradox. Philosophy of Science, 1967, 34, 103-115.

Morsh, J. E., Madden, J. M. and Christal, R. E. Job analysis in the United States air force. Lackland Air Force Base, Texas: Personnel Research Laboratory, 1966, (PRL-TR-66-6).

Moss, J. and Smith, B.B. Process and techniques of vocational curriculum development. Minneapolis, Minnesota: Minnesota Research Coordinating Unit for Vocational Education, 1970, ED 042 917.

Murdock, B. B. Transfer designs and formulas. Psychological Bulletin, 1957, 54, 313-326.

Palmer, E. J. and McCormick, E. J. A factor analysis of job activities. Journal of Applied Psychology, 1961, 45, 289-294.

Passmore, D. L., Asche, F. M., and O'Kelley, G. L. Formative evaluation of educational products in career education. Unpublished manuscript, University of Georgia/Athens, 1974.

Payton, O. D. Validation of Gagne's conditions of learning for a portion of a course in prosthetics. Unpublished doctoral dissertation, University of Maryland, 1971.

Philbrick, J. D. City school systems in the United States. Circular of Information No. 1, Washington, D.C.: U. S. Bureau of Education, 1885.

Phillips, R. Development of optimal instructional sequences. Paper presented at the annual meeting of the American Educational Research Association, Chicago, 1974.

Piaget, J. and Inhelder, B. The psychology of the child. New York: Basic Books, 1969.

Popham, W. J. and Husek, T. R. Implications for criterion-referenced measurement. Journal of Educational Measurement, 1969, 6, 1-9.

Resnick, L. (Ed.) Hierarchies in children's learning. Pittsburgh: University of Pittsburgh Learning Research and Development Center, 1971, ED 064 668.

Resnick, L. and Wang, M.C. Approaches to the validation of learning hierarchies. Preprint 50, Pittsburgh: Pittsburgh University Learning Research and Development Center, 1969, ED 035 943.

Schwab, J. J. The practical: A language for curriculum. School Review, 1969, 78, 1-23.

Smith, D. Determining learning sequences from a difficulty-scaling of test items. Paper presented at the annual meeting of the American Educational Research Association, Chicago, 1974.

Smith, E. L. Procedures for generating candidates for learning hierarchies. Paper presented at the annual meeting of the National Association for Research in Science Teaching, Chicago, 1972, ED 081 571.

Sneath P. H. A. and Sokal, R. R. Numerical taxonomy. Nature, 1962, 193, 855-860.

Sokal, R. R. and Sneath, P. H. A. Principles of numerical taxonomy. San Francisco: W. H. Freeman, 1963.

Spineti, J. P. and Hambleton, R. K. A computer simulation study of tailored testing strategies for objective based instructional programs. Laboratory of Psychometric and Evaluative Research Report No. 2, Amherst, Massachusetts: University of Massachusetts, School of Education, 1973.

Swaminathan, H., Hambleton, R. K. and Algina, J. Reliability of criterion-referenced tests: A decision-theoretic formulation. Journal of Educational Measurement, 1974, in press.

Taylor, M. Bart, W. and Rosehel, H. Applications of ordering theory analysis in the formative evaluation of instruction for retarded children. Paper presented at AERA,Chicago,1974.

Taylor, F. W. The principles of scientific management. New York: Harper and Row, 1911.

Towers, E., Lux, D., and Ray, W. A rationale and structure for industrial arts subject matter. Columbus, Ohio: The Ohio State University, 1966, ED 013 955.

Tryon, R. C. General dimensions of individual differences: Cluster analysis versus multiple factor analysis. Educational and Psychological Measurement, 1958, 18, 477-495.

Walbesser, H. H. and Eisenberg, T. A review of research on behavioral objectives and learning hierarchies. Columbus: ERIC Information Analysis Center for Science Education, 1972, ED 059 900.

White, R. T. Research into learning hierarchies. Review of Educational Research, 1973, 43 (3), 361-375.

White, R. T. A model for the validation of learning hierarchies. Journal of Research in Science Teaching, 1974, 11 (1), 1-3, (a).

White, R. T. Indexes used in testing the validity of learning hierarchies. Journal of Research on Science Teaching, 1974, 11 (1), 61-66, (b).

White, R. T. The validation of a learning hierarchy. American Educational Research Journal, 1974, 11 (2), 121-136, (c).

White, R. T. and Clark, r. M. A test of inclusion which allows for errors of measurement. Psychometrika, 1973, 38, 77-86.

Wood, R. Response-contingent testing. Review of Educational Research, 1973, 43 (4), 529-544.