DOCUMENT RESUME

ED 174 623 TM 007 473

AUTHOR McGregor, John D.
TITLE A Comparison of White's and Walbesser's Learning Hierarchy Validation Techniques.
PUB DATE Mar 78
NOTE 16p.; Paper presented at the Annual Meeting of the Eastern Educational Research Association (Williamsburg, Virginia, March 1978)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Basic Skills; Curriculum Planning; Intermediate Grades; *Learning Modules; Learning Readiness; Learning Theories; *Models; *Organization; Programed Materials; *Sequential Learning; *Validity
IDENTIFIERS Gagne (Robert M); *Learning Hierarchies

ABSTRACT

An instructional program based on a learning hierarchy and involving the addition of fractions was presented to fourth grade students in a pretest - posttest design. The results of the tests were analyzed using both Walbesser's and White's techniques for validation. There was sufficient disagreement in the results to prevent a conclusion of equivalency of the two methods. Attempts to explain the disagreements did not provide sufficient evidence to disregard these differences. However, a brief comparison of one of Walbesser's ratios and White's statistics did show an interesting similarity, suggesting that a planned comparison might give statistical support to a cutoff score for the consistency ratio. Walbesser's procedure, with this statistical foundation, could provide superior information concerning both the hierarchical nature of the connection, and the quality of the learning program. (Author/RD)

# A Comparison of White's and Walbesser's Learning Hierarchy Validation Techniques

John D. McGregor
Christopher Newport College
Newport News, Virginia 23606



Paper Presented at the Spring 1978 Meeting
Eastern Educational Research Association
Williamsburg, Virginia

# A Comparison of White's and Walbesser's Learning Hierarchy Validation Techniques

John D. McGregor
Christopher Newport College
Newport News, VA 23606

In every curriculum construction effort, decisions must be made about how to formulate objectives, how to structure content, how to design and order instructional tasks, and how to adapt to unique student conditions.(Heimer,1973) All too often these decisions are based on the philosophical belief of those constructing the sequence rather than an empirically developed methodology.

In an attempt to provide a rational basis for the sequencing of content, Gagne, in the early sixties, initiated a program of research dealing with sequences of skills he termed learning hierarchies. Studies in the area have concentrated on three basic ideas: (1) theoretical concerns for the sequencing of content; (2) methods for generating the learning hierarchies; (3) methods for statistically supporting the generated hierarchies. Numerous statistical models have been proposed and supported by various philosophical and statistical arguements. Many of these methods have weaknesses ranging from slight to severe. Two of these techniques, White's and walbesser's, hold particular promise and are the subject of this study.

## Sequencing Content

With regard to content sequencing, Briggs stated that "for learning purposes, then, 'structure' means ... the description of the dependent and independent relationships among component competencies, arranged so as to imply when sequencing can be random or optional and when sequencing must be carefully planned, on the basis that transfers will be optional to build up from simple skills to more complex ones." (Briggs, 1968)

In the early sixties, Gagne took a common sense approach to the problem of structuring content. Beginning with a desired outcome, Gagne asked the question,"What capabilities would an individual have to possess if he were able to perform this task successfully given only instructions?" The answer to this question would be a new set of tasks, distinct from, but prerequisite to the original task. According to Gagne the derived tasks are "in some sense simpler and more general" than the original.(Gagne,1962 ) The process of questioning and deriving new tasks is applied to each of the tasks derived from the original task. Gagne found that by repaeting this process, he was defining a hierarchical structure, corresponding to Briggs' hierarchical structure, that was growing progressively simpler and more general. The hierarchical structure defined by this process is termed a learning hierarchy and represents a series of proposed dependencies among a set of intellectual skills.

Viewing the sequencing of content in the context of learning hierarchies, Ggane wrote," I am not sure that a learning hierarchy is supposed to represent a presentation sequence for instruction in an entirely uncomplicated way. Presumably, there should be some relation between an ordered set of intellectual skills and an ordering of a sequence of presentation ... in an instructional program."(Gagne, 1968 ) Others however have not been quite as hesitant as Gagne on the relationship between hierarchies and content sequencing. Okey said, "A hierarchy of tasks ... can serve as an instructional map or guide for a teacher because it is a list of which tasks are to be accomplished by students as well as the sequence in which they should be studied and taught."(Okey) Briggs in discussing the learning hierarchy presented in the Gagne and Paradise article stated that, "These subordinate competencies ... must be taught in a particular sequence (with some options within layers of the pyramid) rather than in a random sequence."(Briggs, 1968) Passmore in discussing objective-based instructional systems, listed three purposes for learning hierarchies:

(1) They describe the most efficient way to sequence instruction.

(2) They serve as a guide for the student's entry into a learning sequence.

(3) They serve as a tool used in criterion-referenced assessment plans for monitoring student progress through objective-based instructional systems. (Passmore)

Obviously, some school learning involves the memorization of independent facts, for such cases the sequence of presentation is of no consequence. Hopefully, however, much of the learning in school involves attainment of interrelated skills the learning of which facilitates the learning of other higher order skills. For this type of learning, a learning hierarchy probably does represent a valid sequence of instruction. Support for this statment can be found in a number of studies. Walbesser and Eisenberg, in an excellent summary of research, found that a majority of studies supported the following hypothesis: The aacquisition of a terminal behavior depends upon the attainment of a hierarchy of subordinate behaviors that mediate positive transfer from one set of behaviors to the next higher relevant behaviors in the sequence and eventually to the terminal behavior. (Walbesser, 1972)

## Generating Hierarchies

To generate hierarchies, suppositions must be made about what constitutes a valid element of the hierarchy and how they are to be ordered. Gagne used the term "learning hierarchy" to refer to a set of specified intellectual capabilities having an ordered relationship to each other. Gagne, in discussing these sets wrote that these capabilities are what some writers call "cognitive strategies.... What they are not is just as important. They are not entities of verbalizable knowledge. I have found that when deriving them one must carefully avoid statements of 'what an individual knows'." (Gagne, 1968 ) Smith gave three reasons for thinking of the elements of a learning hierarchy as performance requirements rather than as simply elements of verbalizable knowledge. First, since presentation of information does not imply assimilation of that information, some actual evidence must be specified to assure that the desired learning has taken place. Second, since recall of information is seldom justification for instruction, capacity for using that information to organize new information is a more useful outcome. Third, it is quite likely that more than presentation of verbal statements is required to develop capacity to utilize information.(Smith, 1972)

Once the character of elements which constitute the hierarchy is determined, a procedure must be chosen for generating the hierarchical structure. Passmore listed four methods for generating hierarchies: (1) observation, (2) formal analysis, (3) statistical fishing, and (4) introspection.(Passmore) Gagne's method of generating hierarchies, through questioning, is an example of introspection. The technique involves subject matter experts using their knowledge about the structure of their area and the nature of learning in their discipline. Using the process, the resulting hierarchy is somewhat dependent on the knowledge of the experts who developed the structure. To insure that a truely hierarchical ordering has been achieved, a validation procedure is needed.

## Validating Hierarchies

Theoretically, a 'valid' connection in a hierarchy implies that if the lower tasks are not mastered then the ones above them cannot be. A validated hierarchy then is one which consists of only valid connections between skills. Connections to be validated are examined as to the degree the performance of learners indicates compliance with the definition of a valid hierarchy.

Performance data can be gathered in essentially two ways. First, learners may be given a test consisting of questions on each of the skills in the hierarchy. The patterns of right/wrong answers can then be analyzed and the degree of agreement with the definition of hierarchical structure can be determined. Second, learners may be tested on their knowledge of each of the skills of the hierarchy during or immediately following instruction in the skills from the hierarchy. While practically all the validation techniques developed so far can use either type of data, various researchers have put forth arguements to support their use of one method or the other.

Although many different techniques have been developed for validation since Gagne initiated research into learning hierarchies, two are of particular importance as they represent basic modifications of Gagne's original efforts. Walbesser and Eisenberg expanded Gagne's original index of positive transfer into a series of 5 indexes.(Walbesser, 1972) Based on the performance data for both the higher and lower skill of a connection, the subjects are divided into four groups indicated by the following ordered pairs:

(0,0) which indicates failure on both the lower and higher level skills.

(0,1) which indicates failure on the higher level skill while passing the lower level skill.

(1,0) which indicates success on the higher level skill while failing the lower level skill.

(1,1) which indicates success at both levels.

The indexes were then defined in terms of the number of subjects in each group. In the ratios below f(1,1) represents the number of subjects in the (1,1) group. Of the five indexes, three are of immediate interest in hierarchy validation.

(1) Consistency Ratio - This ratio measures the strength of the implication that acquisition of the terminal behavior implies acquisition of all subordinate behaviors.

$$\frac{f(1,1)}{f(1,1) + f(1,0)}$$

(2) Adequacy Ratio - This ratio tests the strength of the implication that acquisition of all subordinate behaviors implies, with instruction, acquisition of the terminal behavior.

$$\frac{f(1,1)}{f(1,1) + f(0,1)}$$

(3) Completeness Ratio - This ratio estimates the percentage of individuals capable of traversing the hierarchical connection in question.

$$\frac{f(1,1)}{f(1,1) + f(0,0)}$$

A connection is considered validated if each of these 3 ratios has a value of 0.85 or above.

This is an important extension of Gagne's measure of validity. The ratios not only measure the negative side, that a learner cannot learn the higher task without the necessary lower order skills; it also measures the positive side, that having mastered the lower order skills,the student is more likely to achieve the higher level tasks. The consistency and adequacy ratios measure whether the connection is consistent with the hierarchical hypothesis while the completeness ratio measures whether a sufficient percentage of subjects were able to do either task and thus be included in the statistics.

White listed eight criticisms of the model for learning hierarchy validation used by Gagne.(White, 1974 ) One of White's objections was the absence of an index based on a standardized distribution. As part of a new model for hierarchy validation, White and Clark developed a statistic based on the binomial distribution.(White, 1973 ) Through the use of the statistical distribution, White's method takes into account errors of measurement associated with the testing of the skills. Although a complete description of White's method can be found in the previous reference, basically if two questions are being used to measure acquisition of each skill, the number of subjects who correctly answer both questions for the higher skill and miss both questions for the lower skill becomes the critical frequency. Using White's formulae, an expected frequency for this group is calculated. If the expected frequency exceeds the observed frequency then the connection is considered validated. While White's method is the more rigorous of the two methods discussed, it is much more difficult to use. Walbesser's model is relatively easy to use but lacks the statistical rigor of White's model. If a correspondence could be established between the two methods, then the use of the less complicated technique could be justified in terms of the the statistical validity of the other. The purpose of this study was to investigate the possibility of such a correspondence.

## Experimental Design

An examination of Walbesser's and White's models shows that the chief point at which the two techniques differ is the statistic used to either support or reject the hypothesized connections in the hierarchy. Thus, a learning program based on a learning hierarchy was presented to a single sample of students. The results of the associated tests were analyzed using both Walbesser's and White's techniques. Where the techniques led to different conclusions, attempts were made to explain these differences.

For the purposes od a study of validation techniques, Kane and others developed a proposed hierarchy on the addition of fractions. A programmed booklet was written to teach the skills of the hierarchy. Kane also used two pretests to measure the level of preparedness of the subjects and posttests to measure the achievement of the skills. The first pretest was used to insure that students included in the study had the background necessary to begin the hierarchy. The second pretest was used to eliminate students who already had mastered a majority of the material. For the purposes of this study, the programmed booklet was modified to include test frames for each skill. The posttest consisted of these test frames and no separate testing was done. Since it is important to know if a skill is available at the time it is needed to provide prerequisite knowledge, questions concerning a skill were placed immediately preceding the instruction for a skill for which the first skill was considered prerequisite.

The subjects in the study were fourth grade students who possessed the necessary prerequisite knowledges as evidenced by a score of 50% or better on Pretest I and who had not mastered a majority of the material as evidenced by a score of not more than 50% on Pretest II. The students spent approximately thirty minutes per day working through the booklet. The performance data was then extracted by the experimenter by grading each test frame.

Data Analysis

Two questions were used to measure the achievement of each skill. Thus performance data for each connection consisted of right/wrong responses from 2 questions for the lower skill and 2 questions for the higher skill. Subjects were then divided into nine groups based on 0,1,2 correct answers for the lower skill and 0,1,2 correct answers for the higher skill. The data for each connection was recorded in a 3 by 3 table.

| | | Higher Skill | | |
|---|---|---|---|---|
| | | 0 | 1 | 2 |
| Lower Skill | 0 | | | |
| | 1 | | | |
| | 2 | | | |

For White's method, the critical frequency was the number in the cell representing no correct responses on the lower skill and two correct responses on the higher skill. Using White's formulae and the data in the tables, an xpected frequency was calculated for each connection. If the expected frequency exceeded the observed frequency, the connection was validated. If it equals the observed frequency, then judgement about the connection is suspended. If the expected frequency is less than the observed then the connection is rejected. The results of these calculations are shown in Table I.

Table I

Analysis of Connections By White's Criteria

| Connection | Expected Value | Observed Value | Validated |
|---|---|---|---|
| Level I to Level IIb | 3 | 13 | No |
| Level I to Level IIc | 4 | 13 | No |
| Level IIa to Level Va | 41 | 76 | No |
| Level IIa to Level III | 31 | 72 | No |
| Level IIb to Level III | 2 | 16 | No |
| Level IIc to Level IV | 8 | 3 | Yes |
| Level IId to Level IV | 9 | 18 | No |
| Level III to Level IV | 5 | 5 | Suspend Judgement |
| Level IV to Level Vb | 7 | 5 | Yes |
| Level Va to Level VI | 4 | 3 | Yes |
| Level Vb to Level VI | 12 | 16 | No |
| Level VI to Level VII | 14 | 37 | No |

While White's technique can handle the case where a subject answers only one of the two questions correctly, Walbesser's method requires that a judgement be made as to whether or not a student has acquired the particular skill. For the purposes of this study, a student was judged to have acquired a skill if he answered at least one of the two questions correctly. The data in the 3 by 3 tables was reduced to 2 by 2 tables. For each connection, the consistency, adequacy and completeness ratios were calculated. The results of this analysis are shown in Table II. The results of the two procedures are summarized in Table III. Of the twelve connections, White's procedure validated three and suspended judgement on one. Walbesser's method failed to validate any of the connections.

## Conclusions

Prior to the start of this study it was agreed that the two methods would be considered equivalent if no more than two disagreements were found.

Table II

Analysis of Connections By Walbesser's Criteria

| Connection | Consistency Ratio | Adequacy Ratio | Completeness Ratio | Validated |
|---|---|---|---|---|
| Level I to Level IIb | .83 | .84 | .73 | No |
| Level I to Level IIc | .73 | .67 | .68 | No |
| Level IIa to Level Va | .23 | .74 | .46 | No |
| Level IIa to Level III | .19 | .78 | .29 | No |
| Level IIb to Level III | .79 | .77 | .74 | No |
| Level IIc to Level IV | .87 | .79 | .58 | No |
| Level IId to Level IV | .37 | .70 | .36 | No |
| Level III to Level IV | .87 | .69 | .64 | No |
| Level IV to Level Vb | .83 | .54 | .43 | No |
| Level Va to Level VI | .84 | .59 | .71 | No |
| Level Vb to Level VI | .51 | .75 | .41 | No |
| Level VI to Level VII | .57 | .72 | .69 | No |

Table III

Comparison of Evaluated Connections

| Connection | Results of White's Method | Results of Walbesser's Method | Agree/ Disagree |
|---|---|---|---|
| Level I to Level IIb | No | No | Agree |
| Level I to Level IIc | No | No | Agree |
| Level IIa to Level Va | No | No | Agree |
| Level IIa to Level III | No | No | Agree |
| Level IIb to Level III | No | No | Agree |
| Level IIc to Level IV | Yes | No | Disagree |
| Level IId to Level IV | No | No | Agree |
| Level III to Level IV | Suspend Judgement | No | Disagree |
| Level IV to Level Vb | Yes | No | Disagree |
| Level Va to Level VI | Yes | No | Disagree |
| Level Vb to Level VI | No | No | Agree |
| Level VI to Level VII | No | No | Agree |

There is sufficient disagreement in the data of this study to prevent a conclusion of equivalency. The results of using Walbesser's formulae with this hierarchy differ markedly from the results obtained by Kane. However, Kane restructured connections which were not validated and then recalculated the results to obtain an optimum ordering. Caution should be exercised in using this approach or a hierarchical relationship between two unrelated topics may result not unlike a high correlation between 2 unrelated measures.

Attempts to explain the disagreements do not provide sufficient evidence to disregard any of the differences. Therefore, the two methods should be examined to find any differences that may exist. A look at connection nine, between Level IV and Level Vb can illustrate some of these differences. For White's method, the expected frequency was 5 while the observed frequency was 7. Using Walbesser's method, the consistency ratio was .82, the adequacy ratio was .53 and the completeness ratio was .36.

Whether a hierarchy is validated using a learning program specifically designed to teach the skills of the hierarchy or using data based on prior learning, a hierarchy must be associated with some learning experience. A close look av the meanings of Walbesser's ratios and their numerical values for this connection gives valuable information about the relationship between the two skills and the related instructional sequence. The consistency ratio is designed to measure the degree to which the connection under scrutiny is consistent with the theory of hierarchical relationships. For this connection, the consistency ratio was .82. This indicates that 18% of the subjects exhibited behavior contrary to the hierarchical hypothesis. Although Walbesser set a cutoff of .85, the .82 ratio certainly gives support to the hierarchical hypothesis for this connection.

The adequacy ratio determines whether the elements of the hierarchy adequately define the skills necessary to achieve the higher skill. For connection nine, the adequacy ratio was .53. Rather than saying the relationship is not hierarchical, it indicates that 47% of the subjects who obtained the lower skill could not obtain the higher skill. This points to a need for either additional elements inbetween these two or to a need for improved instruction for the higher skill.

The completeness ratio measures the proportion of the available subjects who have achieved any of the skills involved in the connection. For connection nine, the completeness ratio was .36. This implies that 64% of the subjects achieved no skill in the connection. Thus any decision made on the validity of this connection is based on less than half the subjects.

White's method specifically tests the hypothesis that the proportion of persons with the superordinate skill only is zero. For connection nine, the expected frequency was 7 while the observed frequency was 5; and thus the connection was validated. Although the hierarchy can not be separated from the learning experience, White's formulae give no distinction between the validity of the hierarchical hypothesis and the effectiveness of the learning program.

It is interesting to note that this connection which was validated by White's method also had a consistency ratio of .82. In fact each of the connections validated by White's method had high consistency ratios. The calculation of the expected frequency for the critical cell involves the frequencies in all the cells of the table and thus the calculation is effected by the quality of the instructional program. While both methods are effected by the instructional sequence, Walbesser's method gives more detailed information which can be used to either revise the hierarchy or improve the instruction. While White's technique uses statistical rigor for validating the hierarchical nature of the connection, it gives no additional information which could be of value in revising the hierarchy. Because of the similarity between information from White's formulae and Walbesser's consistency ratio, the establishment of a cutoff value for the consistency ratio using the statistical basis of White's technique could be useful.

Although the data of the study did not support a claim of equivalency for the two methods, a brief comparison of one of Walbesser's ratios and White's statistics shows an interesting similarity. A planned comparison of these two might give statistical support to a cutoff score for the consistency ratio. Walbesser's procedure with this statistical backing would provide superior information concerning both the hierarchical nature of the connection and the quality of the learning program.

## Bibliography


Briggs, Leslie. Sequencing of Instruction in Relation to Hieracrhies of Competence. Pittsburgh American Institutes for Research, 1968.

Gagne, Robert M. "The Acquisition of Knowledge" Psychological Review 69, #4, 1962: 355-365.

Gagne, Robert M. "Presidential Address of Division 15: Learning Hierarchies" Educational Psychologist 6, #1, 1968.

Kane, Robert and others. Analyzing Learning Hierarchies Relative to Transfer Relationships Within Arithmetic, Purdue Research Foundation, 1971.

Okey, James and Gagne, Robert. "Revision of a Science Topic Using Evidence of Performance on Subordinate Skills" Journal of Research in Science Teaching 7, #4, 321 - 325.

Passmore, David L. Validation of Learning Hierarchies for Objective-Based Instructional Systems Massachusetts University (ERIC Document 99419)

Smith, Edward. Procedures for Generating Candidates for Learning Hierarchies, April, 1972.

Walbesser, Henry H. and Eisenberg, Theodore. A Review of Research in Behavioral Objectives and Learning Hierarchies Columbus, Ohio, 1972.

White, Richard. "A Model for Validation of Learning Hierarchies." Journal of Research in Science Teaching 11, #1, 1974.

White, Richard and Clark, Malcolm. "A Test of Inclusion Which Allows for Errors of Measurement." Psychometrika 38, #1, 1973.

Learning Hierarchy for Addition of Fractions
with like Denominators