DOCUMENT RESUME

ED 071 235 EC 050 510

AUTHOR	Bart, William M.; Krus, David J.
TITLE	An Ordering-Theoretic Method to Determine Hierarchies Among Items. Research Report #23.
INSTITUTION	Minnesota Univ., Minneapolis. Research, Development, and Demonstration Center in Education of Handicapped Children.
SPONS AGENCY	Bureau of Education for the Handicapped (DHEW/OE), Washington, D.C.
BUREAU NO	BR-332189
PUB DATE	Sep 71
GRANT	OEG-09-332189-4533(032)
NOTE	19p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Measurement Techniques; *Models; *Statistical Analysis; *Theories

ABSTRACT

Presented is a measurement method derived from ordering theory in which hierarchies among items are determined by processing item response patterns according to rules of symbolic logic. The method, based upon a boolean algebraic framework, is said to provide an alternative to classical measurement methods which assume that the trait to be measured is linearly ordered and can be measured with a simple additive model. An example of data analysis by the ordering theory method is provided for a rating scale of guilt in socially embarrassing situations. It is suggested that learning hierarchies, item hierarchies, behavioral sequences, and cognitive stage theories could all be more carefully studied by the proposed method. (Author/GW)

PROJ 332189
BEH
EC

RESEARCH REPORT #23

Project No. 332189
Grant No. OE-09-332189-4533 (032)

AN ORDERING-THEORETIC METHOD TO DETERMINE HIERARCHIES AMONG ITEMS

William M. Bart and David J. Krus
University of Minnesota
Research, Development and Demonstration
Center in Education of Handicapped Children
Minneapolis, Minnesota

September 1971

Department of Health, Education and Welfare
U. S. Office of Education
Bureau of Education for the Handicapped

ED 071235

Research Report #23

Project No. 332189
Grant No. OE-09-332189-4533 (032)

An Ordering-Theoretic Method to Determine Hierarchies Among Items

William M. Bart and David J. Krus
University of Minnesota

Research and Development Center in
Education of Handicapped Children
University of Minnesota
Minneapolis, Minnesota

September 1971

The research reported herein was performed pursuant to a grant from the Bureau of Education for the Handicapped, U.S. Office of Education, Department of Health, Education, and Welfare to the Center of Research and Development in Education of Handicapped Children, Department of Special Education, University of Minnesota. Contractors undertaking such projects under government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official position of the Bureau of Education for the Handicapped.

Department of Health, Education and Welfare

U. S. Office of Education

Bureau of Education for the Handicapped

Abstract

A measurement method derived from ordering theory is presented; ordering theory is an alternative measurement model with a boolean algebraic framework. With the use of this method, hierarchies among items can be determined. The item response patterns are processed according to rules of symbolic logic to describe the array of prerequisite relationships among the items. An example of actual data analysis by this method is given with attention to some antecedent formulations of the linear scaling problem.

An Ordering-Theoretic Method to Determine Hierarchies Among Items

William M. Bart and David J. Krus
University of Minnesota

Bart (in press) and Airasian and Bart (1970) introduced tree theory as an alternative measurement model with a boolean algebraic framework. Subsequent to those two statements tree theory was relabelled by Bart and Airasian so that the theory would more closely mirror its algebraic framework. Ordering theory has as its primary intent either the testing of hypothesized hierarchies among items or the determination of hierarchies among items. In ordering theory, response patterns for bivalued items are viewed as atoms in a boolean algebra with as many generators as there are items being considered. An ordering (formerly termed a tree) is the union of the obtained atoms and indicates the logical relationships among the items.

Ordering theory shares with classical models the item response matrix, but does not use summation across subject rows to express in the form of the correct responses of a subject his standing on the trait measured. This is the first departure of ordering theory from older models of measurement. These models invariably assume that the trait measured is linearly ordered and can be measured with a simple additive model - e.g., a summative score (Gulliksen, 1950). This assumption of linearity is rarely if ever tested and probably false (Bart, 1970). Instead of using summative scores as a starting point for further statistical analysis, ordering theory is used to determine logical relationships between items represented in the item response matrix and in this respect is more similar to pattern-analytic methods than to classical measurement models. Contemporary pattern-analytic methods share as a common feature the search for some parsimonious and manageable method of representing relationships among items and are coping with the basic difficulty that the number of response patterns increases exponentially with the number of items. Fine examples of this type of approach to measurement theory are papers by McQuitty (1955, 1956). Ordering theory, which draws heavily in its search for order among individual response patterns from the theory of boolean algebras (Halmos, 1963; Stoll, 1961), employs a logical-mathematical method instead of the otherwise mostly empirical search for relationships between items.

The prerequisite relationship

As stated, an ordering is a logical relationship among a set of items. Within that relationship various constituent relationships involving such logical connectives as "either...or" and "and" may be employed. For behavioral researchers the relationship "is a prerequisite to," which is the converse of "implies" or "if...then," is a useful relationship commonly indicated in many theoretical behavior frameworks (Gagne, 1965; Piaget, 1963). The salient qualities of the prerequisite relationship for behavioral research warrant its primary consideration in the designation of the logical relationships among items.

One requirement for the prerequisite relationship to be used to interrelate items is that the items be bivalued -- "1" is the item score for item i for subject k if subject k gives a correct response to item i and "0" is the item score for item i for subject k if subject k gives an incorrect response to item i. Then a defining statement for the prerequisite relationship may be cited based on logic theory (Kleene, 1950; Stoll, 1961): success on item i is a prerequisite to success on item j if and only if the response pattern (01) for items i and j respectively does not occur. The response patterns (00), (10), and (11) are called confirmatory and the response pattern (01) is called disconfirmatory with respect to the two-item ordering, item i is a prerequisite to item j.

For a set of n items such that a correct response for any item indicates a manifestation of the trait being measured one can construct a matrix which indicates the percentage of disconfirmatory response patterns for all of the two-item trees possible among the n items. In other words, a n x n matrix, indexed along the rows and columns by the item numbers, may be constructed such that the entry in the cell in the i-th row and j-th column is the percentage of the total response patterns that had a "0" for the i-th item and a "1" for the j-th column. With such a table one may consider a given tolerance level of disconfirmatory response patterns and then identify item pairs that are related by a prerequisite relationship. Thus if a cell entry for row 2 and column 5 in the matrix is .2 and one pre-established a tolerance level of 1 percent disconfirmatory response patterns then since .2 percent is within the 1 percent tolerance level one can say that item 2 is a prerequisite to item 5.

After various prerequisite-related item pairs are identified, one may construct a hierarchy among the items. Also for different tolerance levels one will construct different item hierarchies. It is recommended that low tolerance levels be used to determine the item hierarchies in order that the method be as compatible as possible with the logical definition provided for the prerequisite relationship. Also when indicating an item hierarchy, the corresponding tolerance level should be reported. Thus, as algebraic equations can be represented with algebraic expressions or with geometric forms, so an ordering can be represented algebraically with a boolean expression or graphically with a line graph that often has the form of a hierarchy.

Problem of statistical tests

To test the prerequisite relationship that an ite- i is a prerequisite to an item j, the hypothesis, that the probablility of a disconfirmatory response pattern (01) for the prequisite relationship for items i and j respectively is 0, needs to be tested. A direct way to conceive of the test for that hypothesis is to use the binomial distribution. One can consider two mutually exclusive but exhaustive events: a disconfirmatory response pattern (01) with hypothesized probability of 0 and a confirmatory response pattern (00,10, or 11) with hypothesized probability of 1. Thus, the problem reduces to testing for a binomial probability value (p) of 0. Unfortunately, this is insoluble for the variance for such a value is zero and nullifies any opportunity to use, for e ample, the normal distribution approximation to the binomial (e.g., Hoel, 1962).

Lending more credence to the insolubility of that test, there have been two recent attempts to articulate rigorous statistical test procedures for item hierarchies (or orderings). Proctor (1970) discussed a chi-square procedure to test whether a set of items form a Guttman scale or, in ordering-theoretic terms, a linear ordering. In the case of a two-item linear ordering which indicates a prerequisite relationship between two items, his procedure is not applicable for there would be no degrees of freedom being indicated for his chi-square test; at least three items must be considered in order for that technique to be used. Airasian

(1968) discussed a maximum likelihood procedure based on the multinomial distribution to test the fit of a hypothesized hierarchy for a set of three items. This method, though not as yet generalized to n items, may prove to be more general than that of Proctor in that the method cited by Proctor relates to linear hierarchies among items whereas the method cited by Airasian relates to linear hierarchies among items and branched hierarchies among items. However, both techniques provide neither a test for a linear ordering among two items nor statistical procedures to test the fit of each of the 2^{2^n} orderings for a set of n items.

In general, the problem of a statistical test for a prerequisite relationship between two items is unsolved. Thus, the tolerance level technique discussed in the previous section is the primary procedure to be used to indicate whether a hypothesized prerequisite relationship between two items is accepted or rejected.

Ordering theory and scalogram analysis

In the early forties Guttman (1944) formulated the requirements for a genuine scale, capable of legitimate measurement. The most salient features of his scale are that the scale should be virtually homogeneous and that persons receiving the same summative score as an index of their location on a unidimensional continuum should have responded in the same way to all items. As a matter of fact, Guttman himself never maintained that all of our measurements must strictly comply with this ideal form and his method later named scalogram analysis is really a procedure for evaluating to which degree our existing scales of measurement are approximating this ideal condition, the degree of approximation being expressed in his coefficient of reproducibility.

Despite the fact that virtually nobody could oppose the fact that the criteria of Guttman for legitimate measurement are highly desirable, his method was widely criticized (e.g., Festinger, 1947; Loevinger, 1948) mostly on empirical grounds, the most serious objection being that his criterion of scalability is rarely achieved even when total scores reach an acceptable level of reliability.

This objection is partially met in ordering theory which demands the Guttman criterion of scalability being met only in parts of a test (in separate branches of ordering) and not in the test as a whole, thus allowing spare channels into which the diversity in test behavior can be distributed. At this point of the discussion a real example will be discussed with a step by step description of item hierarchy construction by the ordering-theoretic method.

An example of the ordering-theoretic method

A simple rating scale of guilt in socially embarrassing situations as perceived by subjects was written. It is composed of 12 items and was administered to 15 students enrolled in a general psychology course at the University of Minnesota. Instructions were printed in caption on the rating scale: "Imagine that you find yourself in the situations described below. Rate how you would feel if it happened. Be frank."

The answer choices to each item were the following: 1) very bad; 2) a little bad; 3) not too bad; 4) don't care. The items were scored in a bivalued manner with "1" being given to either of the first two choices and "0" being given to either of the last two choices. Table 1 lists the items in the test in the order that they were presented to the subject.

Insert Table 1 about here

Table 2 indicates the item response matrix for the scale and marginal totals.

Insert Table 2 about here

With the information in the item response matrix cited in Table 2 the matrix of percentages of disconfirmatory response patterns for the two-item orderings indicating prerequisite relationships can be constructed. Table 3 indicates that matrix of percentages important in item hierarchy construction.

Insert Table 3 about here

If one allows for no tolerance level for disconfirmatory response patterns (or no error), then one can discern the item pairs that are related in a prerequisite manner from a consideration of the zero entries in the matrix of Table 3. Thus, item 1 is a prerequisite to item 6; item 2 is a prerequisite to items 4, 6, and 8; item 3 is a prerequisite to item 6; item 4 is a prerequisite to item 6; and so on to item 12 being a prerequisite to items 1-11. With the prerequisite relationships indicated in Table 3 one can construct the item hierarchy that is indicated in Figure 1a.

TABLE 1

Basic Items in the Scale of Guilt

1. drinking too much
2. shoplift
3. cheat on exams
4. have an extramarital affair
5. gossip
6. don't go to church on Sunday
7. have an homosexual experience
8. lie to parents
9. lie on income tax returns
10. being caught as a Peeping Tom
11. steal a book from the library
12. steal a book from a friend

TABLE 2

Item Response Matrix for Twelve-Item Scale

of Guilt for Sample of Fifteen Subjects

Item	1	2	3	4	5	6	7	8	9	10	11		row totals
Subject 1	1	1	1	1	0	0	1	1	1	1		1	10
Subject 2	1	1	0	1	1	0	1	0	1	1	1	1	9
Subject 3	1	1	1	1	0	0	1	1	1	1	1	1	10
Subject 4	0	1	0	0	0	0	1	0	0	1	0	1	4
Subject 5	0	0	1	0	1	0	1	0	0	1	1	1	6
Subject 6	1	1	1	1	1	0	1	1	1	1	1	1	11
Subject 7	0	1	1	1	1	0	1	0	1	1	1	1	9
Subject 8	0	1	0	1	1	0	0	1	0	0	0	1	5
Subject 9	0	1	1	0	0	0	1	1	0	0	0	1	5
Subject 10	1	0	0	0	0	0	1	0	1	1	0	1	5
Subject 11	1	1	0	1	0	0	1	0	0	1	0	1	6
Subject 12	0	1	0	1	1	0	0	0	1	1	1	1	7
Subject 13	0	1	1	1	1	0	1	1	0	1	0	1	8
Subject 14	1	1	0	1	0	0	1	0	1	1	1	1	8
Subject 15	0	1	1	1	0	0	1	0	1	1	1	1	8
column totals	7	13	8	11	7	0	13	6	9	13	9	15	

TABLE 3

M[illegible] Percentages of Disconfirmatory Response Patterns

for Two-item Prerequisite Relation Orderings for Scale of Guilt

Items	1	2	3	4	5	6	7	8	9	10	11	12
1	--	47	33	40	33	0	40	20	20	40	27	53
2	7	--	7	0	7	0	13	0	7	13	7	13
3	27	40	--	33	20	0	33	7	27	40	20	47
4	7	13	13	--	7	0	27	7	7	20	7	27
5	33	47	27	33	--	0	53	20	33	47	27	53
6	47	87	53	73	47	--	87	40	60	87	60	100
7	0	13	0	13	13	0	--	7	7	7	7	13
8	27	47	20	40	27	0	53	--	40	60	40	60
9	7	33	20	20	20	0	33	20	--	27	7	70
10	0	13	7	7	7	0	7	13	0	--	0	13
11	13	33	13	20	13	0	33	27	7	27	--	40
12	0	0	0	0	0	0	0	0	0	0	0	--

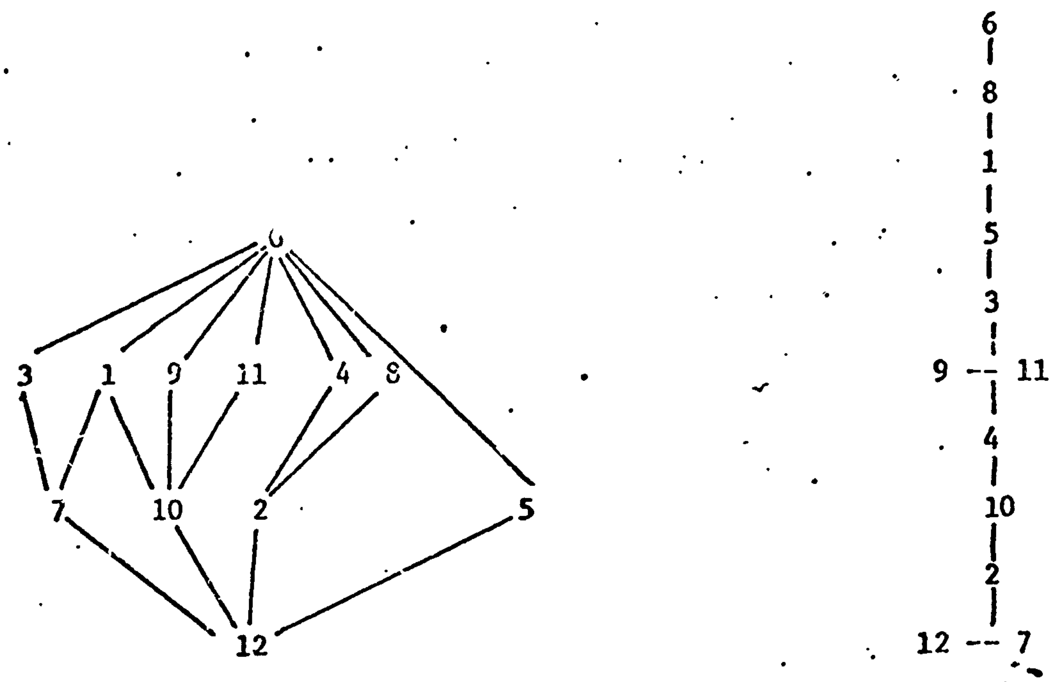
A close counterpart of this ordering-theoretic method of item hierarchy construction is the "Cornell technique" of Guttman (1947). For the data in Table 2 the coefficient of reproducibility is .86 and the ordering of items as in an ideal Guttman scale (G-scale) is indicated in Figure 1b. The G-scale using the "Cornell technique" was constructed from the matrix in Table 2 by setting cutting points in such a way as to minimize error (Edwards, 1957).

Insert Figure 1 about here

In comparing the Guttman procedure and the ordering-theoretic method as well as the two patterns in Figure 1, one can consider a variety of interesting properties. First, the Guttman procedure is used partly to determine the degree to which item data complies to a standard of a linearly ordered scale; with the ordering-theoretic method, that is not the case for an exact-fitting hierarchy can always be constructed for item data. Second, the ordering-theoretic method is able to provide one with all the two-item logical relationships such as is indicated in Fi, ure 1a item 10 being a prerequisite to item 9; whereas, the Guttman proced re can provide us with a markedly reduced number of inter-item relationships. Thus, for example, the relationships that feeling guilty for shoplifting (item 2) is a prerequisite to feeling guilty for lying to parents (item 8) and that feeling guilty for gossiping (item 5) is in a formal logical sense independent of feeling guilty for lying to parents (item 8) are specifically indicated in Figure 1a but not in Figure 1b. Third, the ordering-theoretic method provides valuable information as to the various

FIGURE 1

Ordering-theoretic item hierarchy (Figure 1a) and Guttman scale (Figure 1b) for items in the Scale of Guilt for a sample of fifteen subjects.



(a) (b)

sequences among the items in the guilt scale; whereas, the Guttman procedure cannot provide such information. With the ordering-theoretic method cited, the hierarchy for any set of items for any sample of subjects can be constructed through an ordering-theoretic processing of the item data.

Discussion

With the ordering-theoretic method of item hierarchy construction any array of bivalued item data can be analyzed to determine the item hierarchy structure. Learning hierarchies, item hierarchies, behavioral sequences, cognitive stage theories, etc., could all be more carefully studied with this method. There are many other uses to this ordering-theoretic method such as the following: 1) test data can be analyzed so that rich prescriptive, directive, and diagnostic information can be provided for the teacher and other test users; 2) the hierarchy of prerequisite skills necessary for reading could be determined with the use of this method. Also, research could be proposed on the stability of the item hierarchies over time and over samples of subjects. The ordering-theoretic method of item hierarchy construction is a tool for the behavioral researcher with a wide range of potentialities.

References

Airasian, Peter W. Formative evaluation instruments: a construction and validation of tests to evaluate learning over short time periods. Ph.D. dissertation, University of Chicago, 1969.

Airasian, Peter & Bart, William M. Tree-theory: a theory-generative measurement model. Paper read at the annual meeting of the American Educational Research Association, New York, February, 1971.

Bart, William M. A hierarchy among attitudes toward animals. Environmental Education, in press.

Bart, William M. Tree-theory and the order of a trait. Unpublished manuscript, 1970.

Edwards, Allen L. Techniques of attitude scale construction. New York: Appleton-Century-Crofts, 1957.

Festinger, Leon. The treatment of qualitative data by "scale analysis." Psychological Bulletin, 44, 1947, pp. 149-161.

Gagné, Robert. The conditions of learning. New York: Holt, Rinehart and Winston, Inc., 1965.

Gulliksen, Harold. Theory of mental tests. New York: John Wiley & Sons, Inc., 1950.

Guttman, Louis A. A basis for scaling qualitative data. American Sociological Review, 9, 1940, pp. 139-150.

Guttman, Louis A. The Cornell technique for scale and intensity analysis. Educational and Psychological Measurement, 7, 1947, pp. 247-280.

Hoel, Paul G. Introduction to mathematical statistics. New York: John Wiley & Sons, Inc., 1962.

Kleene, Stephan Cole. Introduction to metamathematics. New York: D. Van Nostrand Company, Inc., 1950.

Loevinger, Jane. The technic of homogeneous tests compared with some aspects of "scale analysis" and factor analysis, Psychological Bulletin, 45, 1948, pp. 507-529.

McQuitty, L. L. A method of pattern analysis for isolating typological and dimensional constructs. Research bulletin TN-55-62. San Antonio, Texas: Air Force Headquarters Personnel and Training Research Center, 1955.

McQuitty, L. L. Agreement analysis: classifying persons by predominant patterns of responses. British Journal of Statistical Psychology, 1956, 9, pp. 5-16.

Piaget, Jean. Psychology of intelligence. Paterson, New Jersey: Littlefield, Adams & Company, 1963.

Proctor, C. H. A probabilistic formulation and statistical analysis of Guttman scaling. Psychometrika, 1970, 35, 73-78.

Stoll, Robert R. Sets, logic, and axiomatic theories. San Francisco: W. H. Freeman, 1961.