DOCUMENT RESUME

ED 322 179 TM 015 327

AUTHOR	Olivarez, Arturo, Jr.; And Others
TITLE	Application of Categorization Methodology in Various Testing Situations.
PUB DATE	Apr 90
NOTE	77p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (Boston, MA, April 17-19, 1990).
PUB TYPE	Reports - Research/Technical (143) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC04 Plus Postage.
DESCRIPTORS	*Classification; *Cognitive Tests; Education Majors; Higher Education; *Mathematical Models; *Test Interpretation; Test Reliability; *Undergraduate Students
IDENTIFIERS	*Categorical Data; F Sort Categorization; Latent Partition Analysis

ABSTRACT

The purposes of the present investigation were to illustrate the applicability of categorization methodology for several empirical situations and to draw implications regarding the use of such methodology in examining categorical data. In using three tasks--two designed to measure cognitive dimensions (e.g., categorizing countries and categorizing action verbs) and one developed to tap personality differences (e.g., traits associated with assertiveness)--the study sought to understand how individuals group and structure stimuli. Specifically, the F-sort categorization technique and latent partition analysis (LPA) were used to evaluate the data collected. Fifty undergraduate preservice teachers performed the first two sorting tasks, and 45 education students performed the personality task. The first task consisted of sorting a random sample of 30 nations, the second sorting task consisted of a set of statements/items from a self-reporting questionnaire designed to measure the attributes needed to achieve personal success, and the last task consisted of a set of verbs. A deck of cards was constructed for each of the tasks: 100 decks for the 30-country sort, and 50 decks for sorting the 50 verbs and the 30 statements. The reliability indices provided insight as to how the stimuli, for each task, were clearly related to the corresponding latent partition. The use of theta reliability procedures in LPA appears to be justified given the nature of the tasks usually involved in this methodology. A 49-item list of references, 12 data tables, a list of stimuli description names for sorting tasks, instruction forms for the sorting tasks, and sample data sets for generating joint proportion matrices are provided. (RLC)

APPLICATION OF CATEGORIZATION METHODOLOGY IN VARIOUS TESTING SITUATIONS

Arturo Clivarez, Jr.
Victor L. Willson
Jonna M. Kulikovich
Texas A&M University

Paper presented at the National Council on Measurement in Education Annual Meeting, Boston, MA. 1990.

Abstract

The purpose of the present investigation was to illustrate the applicability of categorization methodology for a number of empirical situations and to draw implications regarding the use of such methodology in examining categorical data. In using three tasks, two designed to measure cognitive dimensions (e.g. categorizing countries and categorizing action verbs) and one developed to tap personality differences (e.g. traits associated with assertiveness), we sought to understand how individuals group and structure stimuli. Fifty undergraduate preservice teachers were sampled to performed the first two sorting tasks and 45 education students for the personality task. The application of this methodology was empirically justified in each of the sorting tasks. The reliability indices provided an insight as to how the stimuli, for each task, were clearly related to the corresponding latent partition.

APPLICATION OF CATEGORIZATION METHODOLOGY IN VARIOUS TESTING SITUATIONS

Introduction

Generally, when educational practitioners and researchers set out to understand academic abilities or personality traits, they administer a series of tests to their students. In the case of academic achievement and aptitude, these tests might take on a multiple-choice or a short answer format. For personality inventories, respondents may be asked to respond to a series of Likert-type statements. Whether intellectual or affective, these measures exemplify modes of assessment that are objective and researcher-controlled. When these instruments are reliable and valid, they have afforded investigators ease in administration and scoring. Furthermore, responses are typically easy to code into categories and therefore statistical analysis and interpretation is facilitated. However, do these standard means of evaluation present the most sensitive manner in which one can gain information regarding individual differences? That is, are these objective formats the most informative way of differentiating between experts and novices, gifted and nongifted learners, depressed and nondepressed individuals, etc.?

Schuell (1985) argued that perhaps the most informative means of understanding thought and behavior was through the use of interview techniques or verbal report data. These methodologies are extremely subjective in that they permit the respondent to answer in any fashion that he/she chooses. While one might gain a great number of insights regarding the cognitive processing of a student or the emotional trauma of a depressed child, interview techniques or verbal report data receive psychometric criticism because it is most difficult to obtain a reliability estimate on the data gathered. Furthermore, it is extremely difficult to code the data suitably for statistical analysis (White, 1985).

Thus, within the continuum of objective (e.g., multiple-choice, Likert scale) and subjective (e.g., interview techniques) assessment formats, is there a methodology that highlights more of the idiosyncratic natures of individuals' processing yet can be reliably and validly assessed? One alternative rests in the broad realm of categorization methodology. Most often these categorization techniques involve the sorting of stimuli cards. Unlike objective tests, the individual may group cards in any manner that makes sense to him/her rather than being directed to select a single response from a series of alternatives across a number of items. Unlike interview techniques, categorization methodology provides more well-defined parameters around which a researcher can code an analyze data (Schuell, 1985). For the respondent, categorization methodology may provide better guides as to what type of information is desired in an answer. As early as 1956, Bruner stated that categories "make the environment appear less complex" (p. 20), they "enable one to relate different classes of events" (p. 21). In essence, categorization for both researcher and respondent provide greater control in organizing and structuring information.

For example, categorization methodology has been extremely useful in the cognitive literature where attempts have been made to discern the processing differences between expert and novice performance in a variety of domains. Chi, Feltovich, and Glaser (1980), Shavelson and Stanton (1984) have effectively used categorization methodology to determine how physics experts and novices approach a variety of problems. Results indicate that experts most commonly sort problems according to principles such as the Law of Gravity or the properties of centrifugal force. Novices, on the other hand, often sort problems according to what are considered "surface features." That is, novices pay attention to the types of equipment used in certain physics problems (e.g., weights or pulleys) or how much mass or energy is presented in the given problem. While an expert and a novice could have obtained similar scores on an objective measure such as a multiple-choice test, the categorization methodology often presents the expert as possessing not only greater but also much better organized knowledge. The novice, on the other hand, may categorize information more loosely and in a more fragmented fashion.

It would seem that categorization methodology would also be extremely valuable to those studying affective aspects of individuals. While an individual may respond "strongly agree" to a certain Likert item, one might not know exactly to what extent this individual feels so strongly about the topic. Furthermore, is there one item to which this individual responds that more sensitively depicts an issue of great pertinence to this person than some other issue? As for interview techniques, the respondent may not be capable to describe fully how he/she feels (Schwell, 1985). Again, categorization methodology may be the best alternative of giving researcher and respondent a better framework of collecting and providing information.

Given that sorting techniques may suitably provide data about a reservoir of multifaceted processes, are there appropriate ways in which the data collected can be shown to be both reliable and valid? Such is the aim of the present investigation. In this study, we chose to examine how a series of categorization methods could be used to gather information about individual characteristics. For three tasks, two designed to measure cognitive dimensions and one developed to tap personality traits, we sought to understand how individuals would group together and structure stimuli. Specifically, we used the F-sort categorization technique and Latent Partition Analysis to evaluate the data that we collected. Before describing the methods and results of our study, we first provide an overview of the F-sort technique and latent partition analysis which are integral components of categorization methodology.

Categorization Methodology

As presented in the introduction, categorization methodology evolved to satisfy a need in the area of data collection and data reduction. The F-sort technique is used to collect data of the type required for analysis when using this methodology. The LPA technique is employed to summarize the data obtained from the F-sort task.

F-sort Technique

The F-sort technique is a method used for observing and recording categorical judgments manifested when subjects, acting as judges, conduct a series of sorting manipulations of a given set of objects. Typically, the subjects are provided with a deck of cards, which contain some type of stimuli that are supposed to be sorted into independent and exhaustive categories. Standard instructions are provided to each of the judges. These instructions contain pertinent information and a set of criteria about what is sought by the researcher and what is expected of the sorter. Basically, the judges are instructed to sort the stimuli, given on the cards, into disjoint categories of their own invention. There are no restrictions on the total number of categories or piles to be used, or on the number of objects that may be sorted in each category. Also, no a priori ordering of the category is assumed in the F-sort categorization task. The judge is free to decide on the homogeneous characteristic that the set of objects seem to possess.

This lack of restrictions in the F-sort technique implies also that there is no set standard in which the judges may base their observations (in this case their sorting of the stimuli). The F-sort technique is not supposed to be confused with the term Q-sort (Stephenson, 1953). The F-sort has been coined following the format of the Q-sort technique, but its methodology is dissimilar to Q-sort technique. The Q-sort technique involves assigning stimulus objects or items to fixed categories ordered along a predetermined dimension while F-sort is a free sorting technique, and the end result is a set of stimulus categories completely defined by the judge (sorter), (Miller et al., 1986).

Latent Partition Analysis

Latent partition analysis is the complementary technique employed in the categorization methodology to summarize the data obtained from the F-sort task. The LPA is a mathematical model for the sorting manipulations together with a computational algorithm for identifying the latent structure of the pooled sortings of several judges. LPA is, to some extent, similar in intent to various existing procedures, such as latent structure analysis (Lazarfeld & Henry, 1968); however, the mathematical models are not quite the same.

The latent partition model is briefly explained next, following its author's original development. As stated by the developer of the LPA model:

> "Latent partition analysis has been formulated to study the relationships between two or more partitions of the same set of items. A partition of a set of objects is a division of the set into independent, exhaustive categories. The data for latent partition analysis is a sequence of different partitions of the same set of items, and the basic structural hypothesis is that there is a latent partition which underlies the manifest partitions", Wiley (1967).

Latent partition analysis enables the examination of the modal population categorizations for a set of objects. One assumption made by Wiley is that for a given set of judges there exists a latent categorization or partitioning of objects presented. He assumed that the items are to be assigned to manifest categories according to independent, discrete probability distributions. The distributions are assumed to vary across items and across manifest partitions, but otherwise are invariant within a given manifest partition for items belonging to a given latent partition. A manifest partition is represented by the matrix Z_i with rows corresponding to categories and columns corresponding to objects. The (m,j)th entry is 1 if object j is included in category m, and 0 otherwise. An example of a manifest partition matrix is the following:

$$Z_i = \begin{bmatrix} 1&1&1&0&0&0&0&0&0&0&0&0 \\ 0&0&0&1&1&1&1&0&0&0&0&0 \\ 0&0&0&0&0&0&0&1&1&1&1&1 \end{bmatrix}$$

where the number of objects, K, is equal to 12; and the number of manifest categories, M_i, is equal to three.

The elements in the above matrix indicate that manifest category one consists of objects one, two, and three; manifest category two consists of objects four to seven inclusive; and manifest category three is made up of objects eight to twelve.

The latent partition, which is unobservable, is represented by Φ and also has a structure similar to Z_i. A major assumption is that the matrix is assumed to be constant across all the judges or sorters. The matrix Π_i consists of probabilities that relate manifest partition i to the latent partition Φ. The (m,u)th element of Π_i is the probability that any item from latent category u is included in manifest category m of partition i. The matrix Π_i is assumed to remain constant over independent partitionings of objects by any sorter or agent of partitioning as identified by Wiley (1967). Wiley further argued that the manifest categorization of any judge, i, may be derived and represented from the latent categorization in the following way:

$$E(Z_i) = \Pi_i \Phi \qquad (1)$$

where the expectation is over the transformation into the sample space of all possible partitions of the items, and Π_i determines the probability distribution of the partitions. The distribution for each of the items is assumed to be multinomial with probabilities determined by the column of Π_i corresponding to the latent category to which the object or stimuli belongs. The assignments are to be exhaustive and mutually independent. It was previously mentioned that the latent partition matrix is unobservable and the manifest partition matrix generated by each judge will be used to generate an approximate structure for the latent partition matrix. This is achieved by first obtaining the squared matrix S_i defined by

$$S_i = Z_i{}'Z_i \qquad (2)$$

where S_i is a matrix of item joint occurrence. The product of these matrices yields a square matrix indicating membership of an object to a particular category. In other words, S_i is a matrix which consists of submatrices with entries consisting of all ones. For the example provided above, the joint occurrence matrix will look as follows

$$S_i = \begin{bmatrix} 1&1&1&0&0&0&0&0&0&0&0&0 \\ 1&1&1&0&0&0&0&0&0&0&0&0 \\ 1&1&1&0&0&0&0&0&0&0&0&0 \\ 0&0&0&1&1&1&1&0&0&0&0&0 \\ 0&0&0&1&1&1&1&0&0&0&0&0 \\ 0&0&0&1&1&1&1&0&0&0&0&0 \\ 0&0&0&1&1&1&1&0&0&0&0&0 \\ 0&0&0&0&0&0&0&1&1&1&1&1 \\ 0&0&0&0&0&0&0&1&1&1&1&1 \\ 0&0&0&0&0&0&0&1&1&1&1&1 \\ 0&0&0&0&0&0&0&1&1&1&1&1 \\ 0&0&0&0&0&0&0&1&1&1&1&1 \end{bmatrix}$$

As observed from the example given there has been some arranging of the objects, this was done deliberately to illustrate the membership of each item to the particular category created by the sorter. But even if this nice ordering was not done, the diagonal entries of S_i are identically 1. We want to find the expectation of S_i which may be written as Σ_i. The elements in Σ_i are probabilities of joint occurrence for a pair of items. This leads to the definition of the LxL matrix

$$\Omega_i = \Pi_i' \Pi_i \tag{3}$$

where the (u,v)th entry of Ω_i is the probability of joint occurrence for any pair of distinct items from latent categories u and v. The square matrix $\Phi' \Omega \Phi$, which is similar in form to the square matrix S_i, is considered to describe the probability of membership for a pair of objects in the same latent category. The desired probabilities of joint occurrence are found in the matrix off-diagonal entries. These off-diagonal elements are similar to the ones in Σ_i. Therefore, the KxK matrix

$$\Delta^2_i = \Sigma_i - \Phi' \Omega_i \Phi \tag{4}$$

is a diagonal of "diversities". Thus, from (4) we have a representation of the expectation of S_i as follows

$$E(S_i) = \Sigma_i = \Phi' \Omega_i \Phi + \Delta^2_i, \tag{5}$$

where the expectation is conditional on the sorter.

In order to estimate Σ_i, the information obtained by each joint occurrence matrix S_i is used to form a KxK joint proportion matrix S defined as

$$S = N^{-1} \sum S_i \tag{6}$$

where the (j,k)th entry of S is the proportion of partitions in which items j and k were included in the same category. The derivation of the expectation of S, leads to the final representation of the expectation of S

$$\Sigma = \Phi'\Omega\Phi + \Delta^2. \qquad (7)$$

Wiley showed how Φ and Ω can be reconstructed by using concepts from factor analysis developed by Harris and Kaiser (1964) when it is known that the transpose of Φ has independent cluster structure. He further argues that the use of the quartimax procedure appears to be sufficient for the present purpose since the raw quartimax criterion value is maximized when independent cluster structure is achieved. Wiley also offers an algorithm for the estimation of Σ. Since Σ is not known, S must be used as an estimate of Σ, (see Wiley 1967 pp. 188-189). Using a more direct approach, Pruzek (1968) considered the parallels that exists between LPA and the common factor analysis model and a special case of conventional common factor analysis, viz., alpha factor analysis (Kaiser & Caffrey, 1965). Pruzek claims that using this algorithm leads to more positive conclusions than the ones obtained by Wiley's iterative algorithm. He emphasized that the "Kaiser-Caffrey rule for determining an appropriate number of alpha factors is justified in terms of maximizing common factor generalizability (internal consistency) as measured by Cronbach's (1951) coefficient alpha. It was found that Guttman's (1953) weaker lower bound for the number of common factors gives the number of alpha factors which have positive correlations (covariances) with a postulated universe of common factors", (Pruzek, 1968). According to both Wiley and Pruzek, Guttman's weaker lower bound corresponds to the number of roots in the initial correlations matrix R which are greater than unity leading to the approximate number of latent categories in LPA. The difference, of course, is that in latent partition analysis the initial matrix S is a matrix of cross-products or proportions and not a matrix of correlations, Pruzek (1968). The LPA technique has been expanded to include more than one latent partition (Evans, 1970) and its transformation procedure has been improved by Hofmann (1987).

Categorization Methodology in Applied Research

Categorization methodology has been applied in a variety of studies since its development. As already cited, Pruzek (1968) employed categorization methodology to examine mathematics achievement test items. He made theoretical and empirical comparisons between LPA and conventional factor analysis. He found that despite the apparent meaningfulness of derived item categories within each study, it was found that there is little correspondence between response data groups and derived LPA categories.

Two additional studies similar to Pruzek's, were those of Hartke (1978, 1979). Hartke used latent partition analysis to examine a conceptually homogeneous population of mathematics achievement items. He also used LPA in the development of conceptually independent sub-scales to reach a consensus of student's judgments.

Developmental stages of children's word meanings were investigated by Maguire, Patsula, and Evanechko (1975) while Conry (1973) studied the relationship between stimulus task characteristics to cognitive styles. Whitely (1976, 1977) analyzed the effects of task properties on the solution of verbal analogy items. As reported in Miller et al., (1986), Lane (1967) examined the kinds of classifications that counselors constructed for client statements. Willson and Palmer (1983) studied the categories that undergraduate students formed of attributions for examination performance. Colleta and Gable (1975) used categorization methodology to generate content categories from judgmental data gathered on Barth Scale items from twenty-three open education experts. While Coletta and Gable studied the views of teaching at the elementary level, Diamond (1980, 1983) examined these views at the secondary school level and Whitely and Doyle (1976, 1978) at the college level. Hambleton and Sheehan (1977) described the categorization methodology and applied it to one segment of an objective-based science program. Rideng and Schibeci (1984) used categorization methodology to examine the conceptual structure for an attitudinal test related to biology. Finally, Weitman (1986) investigated the hierarchies of effective early childhood teacher dimensions as perceived by directors, parents,and caregivers.

As it can be seen, the categorization methodology is not just used in one single area of study; it has been used on the development of new instruments and the improvement of existing ones. Categorization methodology can be utilized to improve curriculum systems in education. Its broad range of utility and flexibility makes it a desirable method to use for exploration and confirmation of underlying latent structures for sets of stimuli or phenomena. Despite the successful implementation of this methodology, there is still a lack of some kind of measure that may indicate the level of accuracy and consistency (agreement) that may exist among the sorters who create the several distinct categories. The following section deals briefly with the possible adaptation of such a measure.

Reliability of Composites

Reliability measurement of composite variables has been defined as a total score based on two or more subtests scores. Composite reliability has attracted a considerable amount of attention and interest among researchers in education, psychology, and sociology, Lord & Novick, (1968) and Armor, (1974). This procedures have been applied to current instruments which are made up of more than one component. For example, the WISC-R and the K-ABC are just two of countless number of instruments that are made of more than one component to measure intelligence. These tests have employed composite reliability procedures to determine the internal consistency of their different parts or components, Willson & Reynolds, (1984, 1985).

Alpha Reliability

In the present section, a brief review of the development of composite reliability is introduced and explained as well as the maximization of the alpha coefficient for internal consistency called theta by Armor (1974). Two other similar measures have been developed. Heise and Bohrnstedt's (1970) coefficient omega and Bentler's (1968) alpha coefficient. These coefficients will not considered here since their developmental approach is different from Armor's theta.

Due in part to the conceptual and computational simplicity of the split-half method, this coefficient was the most commonly used coefficient before Cronbach introduced the coefficient alpha which he proved to be the mean of all possible split-half reliabilities, Cronbach, (1951). Novick and Lewis (1967) proved that alpha was a lower bound to the true reliability which means that alpha is a conservative estimate of the reliability of a composite. In his review, Armor (1974) gave several computational forms for alpha and some instances of situations in which one form of alpha is more appropriate than the other. He adds that, since real data depart most of the cases from the parallel-item assumptions, it is necessary to decide on the adequacy of individual items. Under the parallel-assumptions, the more items the composite has the higher the reliability, except that when items do not correlate moderately high with other items in a composite, then reliability may be reduced and the items can be excluded. Armor defines the composite reliability alpha as

$$\alpha = [p/(p - 1)](C/S) \quad (8)$$

where S = the variance of the sum; C = the covariance between items i and j; and p = the number of items in the composite. That is, alpha can be derived as the proportion of scale variance due to item covariation adjusted to provide an upper limit of 1. He provides other computational forms for alpha. He claims that the different formulas for alpha are not just a matter of computational convenience and offers three considerations. He also pointed out some limitations for coefficient alpha when composites measure more than one independent dimension.

Theta Reliability

In order to maximize the alpha reliability of a composite, Armor (1974) suggests several steps that may be followed in order to construct a covariance scale or item analysis of a composite (see p. 24). He claims that the application of covariance scaling may enhance reliability. To some extent, the success of the application may depend on subjective judgments more than on analytic criteria. He lists some of the limitations that are inherent in alpha reliability and covariance scaling, one of them being that alpha depends on the assumption that all items in a composite are parallel items, which further implies that all the items measure a single underlying scale property equally.

As in the present study, in which the type of empirical data that were gathered through categorical methodology, the objects or items may measure two or more independent properties, for which alpha reliability will not be an appropriate measure to use in order to determine the internal consistency of the objects and further how consistent the sorters agree on these latent partitions or dimensions, as commonly labeled. Armor goes on to suggest a coefficient (theta) similar to alpha which uses factor analytic techniques to solve the problem of identifying dimensions in a set of data and relating specific items to each dimension. He claims that the key to these factor analytic techniques, in establishing optimally reliable scales, is the formal connection between reliability and scaling provided by principal component factor analysis. The results of a principal component analysis enables the researcher to compute an optimal reliability coefficient. Principal component analysis can be use to construct a set of factor scores, one set for each factor. A factor score is in effect a composite scale score based on a weighted sum of the individual items using factor loadings as the weight. Moreover each factor represents a statistically independent source of variation among the set of items. The solution of a component-analysis extracts these factors in an order corresponding to the magnitude of their variance contribution. In principal component analysis the amount of variance accounted for by a factor is called root and is denoted by λ_k. The kth root is simply the sum of the squared factor loadings for the kth factor. Armor (1974) provided the derivation of the theta reliability formulas for the case of a single-factor solution and for the case of multiple-factor solution. In the present study only the formulas will be presented to aid in the explanation and use of them.

For a given set of p items or objects and a single-factor solution with root λ_1, the reliability of the composite scores based on this factor is given by

$$\theta = [p/(p-1)][1 - (1/\lambda_1)] \quad (9)$$

where λ_1 is the first root of a principal component solution (Eentler 1968). Lord (1958) showed that weighting items according to their principal-component factor loadings theta is the maximum possible alpha. This same procedure has been used by Serlin and Kaiser (1978) to compute a principal component solution for scoring weights for all options in a multiple choice test. Willson (1982) extended the Serlin and Kaiser procedure by allowing several or all options of a given item to be correct. The assumption here is of a single underlying dimension being tapped by the composite. For the situation of multiple-factor solution with rotated factors, Armor (1974) gives a modified version of (9) and argues that multiple-solution with rotated factors is more complicated. Letting ϕ_{hk}^2 be the squared correlation between the original unrotated scores for factor h and the new factor k, and a given rotated m-factor solution with original eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_m$, the reliability of the k^{th} set of rotated factor scores is given by

$$\theta_k^* = (p/(p - 1))[\, 1 - \sum_{h=1}^{m} \phi_{hk}^2/\lambda_h] \qquad (10)$$

where p and λ_h are defined as before and ϕ_{hk} is the element in the h^{th} row and k^{th} column of the transformation matrix that maps the original factor loading into the rotated loadings. The formula only holds for orthogonal rotations like quartimax or varimax solution. The basic interpretation of this coefficient is similar to the one given to the alpha coefficient using classical true score theory and generalizability theory which is percentange of true score variance by the total score variance. It is this procedure and formulation that will be implemented to determine how each of the judges agree with each other in determining the dimensions of the set of objects for each to the sorting tasks since we are treating the different partitions for the total number of objects as a composite.

DESIGN OF THE STUDY

Methodology

The methodology used for each of the distinct tasks follows the exact procedures developed by Miller, et al. (1986). There are three numerical tasks used to illustrate the applicability of categorization methodology. The first task is that of the categorization of a set of countries. The second numerical example involves the sorting of behavior verbs. Finally, a set of statements dealing with what it takes for women to succeed is used as a final example. The sampled groups performed only two of the three task due to time constrains. Both groups performed the sorting of countries while the first group did the behavior verbs and the second did the personality trait statements. The countries sorting task was given as a training exercise.

Description of Samples

Sample I

The first sample of sorters consisted of 50 undergraduate preservice teachers enrolled in an educational psychology class. This particular undergraduate course was selected on the basis of the type of students enrolled in it. This type of sample specification is suggested in the categorization methodology. The course deals primarily with application of learning theory principles to problems of teaching. Of the 50 students selected, 42 declared to be education majors, the rest declared majors ranging from political science to theater arts. The sample was composed of 78% females and the mean age for the sample was 21.1 years of age.

Sample II

The second sample of sorters consisted of 45 undergraduate preservice teachers enrolled in a similar class to the group sample for the first task. The content of the course was identical to sample I. Of the 45 students selected, 42 declared education majors, while two declared majors in Scientific Nutrition. The other student was in Horticulture. This sample was made up of 76% females and the mean age for the sample was 20.9 years of age.

Description of Sorting Tasks

The Sorting of Countries

The first task consisted of sorting a random sample of 30 nations selected from a geography book by James & Webb, (1980): One World Divided. The primary utility of this example is to provide proper training and sufficient amount of instruction to the sorters, increasing the probability that the sorters or judges understood the task completely. Another reason for selecting this task was because of its potential for clear category classification. The classification of countries into common distinguishable categories lends itself to this type of sorting task. (See Appendix A for list of selected countries.)

The Sorting of Success Statements

The second sorting task consisted of set of statements or items from a self-reporting questionnaire designed to measure the attributes needed to achieve personal success (see Gardenswarts and Rowe, 1987). The sorting of these statements was used to determine the number of different categories that judges may derive from a list of 30 items. The main purpose for selecting this task is to compare how the categories created by the sorters differ from the categories the authors claim the instrument seems to measure. (See Appendix A for the list of success statements.)

The Sorting of Behavior Verbs

The last task consisted of a set of verbs used by Miller, Baker, Clasen, Conry, Conry, Pratt, Sheets, Wiley, & Wolfe (1967) in their study of teachers' views concerning facilitation of learning in the classroom. The main objective for using this task is the information it provided when it was first carried out by Miller et al. (1967). Their study reported the number and types of categories that teacher sorters produced for these fifty verbs of teachers' behavior in the classroom. This reported information has two uses. First, it will be useful for comparing the results of this study with their 1967 findings. Second, it will aid in creating the modal categories needed to make the contingency table for the agreement measure. (See Appendix A for list of verbs.)

Summary of Sorting Tasks

The aforementioned tasks were selected to cover just a small section of the wide applicability of categorization methodology. Categorization methodology has been applied in many research areas, and the selection of these three numerical tasks appear to cover some of the more common areas in which categorization methodology may be applied.

Materials

A deck of cards was constructed for each of the tasks. One hundred decks of cards were made for the sorting of the thirty countries. Fifty deck of cards for both the sorting of the 50 verbs and the 30 statements were also constructed. The decks of cards were color-coded to distinguish them and one thousand color-coded blank cards were also made available to the entire two samples doing the sorting task. These blank cards were used by the sorter to divide the different categories and to name the category they created. The decks of cards were wrapped with a rubber band. The same rubber band was used to wrap the deck of cards and the dividing blank cards when the sorter finished the task.

Also provided to the sorter was a set of written instructions for each of the tasks. This instructional handout was used also to gather demographic information about the sorter.

Procedures

An instructional handout and a deck of cards were distributed to each of the participants. The researcher explained the purpose and the procedure of the task and read the instruction to the group. The participants were reminded to consult the handout, or ask the researcher, should questions arise. The researcher provided help as needed. Appendix B contains the three different instructional handout forms used in the study.

The sorters were also provided with a second set of blank color-coded cards which were supposed to be used as title cards in the sorting task. There was no time limit set to perform each of the tasks. The response rate was 100 percent.

When everyone had finished the first task, the sets of cards were collected by the researcher, and the second instructional handout was distributed to each sorter. Similar steps were followed for the administration and completion of the second task. Once the sets of sorted decks of cards were collected, a data file was created for each of the tasks using a special recording form created by the researcher. From this form the computer data file was created, and the analyses were performed.

Analysis of Sorting Tasks

The computer data file was created so that each sorter had a matrix composed of just ones and zeroes. Thus, for the sorting of countries, ninety five rectangular matrices were created and readied for analysis. Similar computer data files were created for the other two tasks. Appendix

C contains a listing of these three data sets.

In order to analyze each data set, a FORTRAN program was written so that a joint proportion matrix, S, was created and analyzed in that form. Once this matrix was created principal-component factor analysis was performed on each of the three matrices outputted by the FORTRAN program. Three types of rotations were performed for all the data sets: quartimax, varimax, and oblique. In addition to using a principal component extraction procedure, an alpha factor analysis procedure was performed for all tasks. The use of more than one factor analytic method and rotation procedure was to make comparisons among them. The results reported will be those from the principal component solution and the varimax rotation.

In order to determine the interrelationships among the derived latent categories, the LISREL VI software was used. This program developed by Joreskog and Sorbom (1986) is one of the existing programs for estimating covariance structures. Wiley's model for latent partition follows this type of structure; therefore, the resulting latent categories from the latent partition analysis were used to specify the model needed to run LISREL.

To determine the theta reliability, a SAS program was written to obtain theta for each of the sorting tasks.

RESULTS

Latent Categories for All Tasks

The identification of latent structures from judgmental data requires the scoring of the sorting of objects into stimulus-similarity indices. These indices are the probabilities that two given objects or stimuli will be included in the same category by a judge. The cell entries in the matrix refer to percentage of judges who included the two distinct objects in the same manifest category regardless of what other items were placed in the category. The joint proportion matrices for each of the sorting tasks were used as correlation matrices using the option in SPSS-X (1987) for reading data in matrix form. Latent partition analysis was applied using a factoring procedure to the symmetrical joint proportion matrix to identify categories from the obtained manifest categories for each of the participants in the sorting task.

Latent Categories for Task I

Six clusters with eigenvalues over 1.0 were extracted from both the alpha factor analysis and the principal components analysis from the joint proportion matrix. All six clusters were deemed meaningful for further analysis but our purpose is to identify not to reduce no subsequent analyses were performed. Objects (countries) with the highest loadings, all over .5 for a particular cluster were retained. Table 1 presents the final rotated cluster pattern for this task. Content analysis of the clusters by the researcher yielded the following labels: European countries (F1), Latin American countries (F2), African countries (F3), Unknown countries (F4), Caribbean islands (F5), and Northernmost countries (F6). The label for cluster four came primarily from the labels provided by the sorters in their manifest categories. As it can be observed, the factor loadings were relatively high within each cluster and this allowed for easy labeling of the clusters. Nevertheless, a criterion of .40 was used to determine significant loadings in each category. The only country that did not meet this criterion was Morocco (T3). The following characteristic roots were obtained for these six factors: 6.649, 4.744, 3.992, 3.170, 1.788, 1.743. These eigenvalues were used to obtain the respective theta reliabilities for each of the factors in Task I.

Theta Reliability for Task I

The factor transformation square matrix which shows the correlation between the original unrotated scores for factor h and the new rotated factor k is provided in Table 2. These values were used to generate theta coefficients for each of the clusters produced by the factor analysis. This transformation matrix was obtained by using principal-component analysis, as recommended by Armor (1974). The obtained theta reliabilities for factors one through six for this task are provided in Table 3. The column labeled Lambda indicated the eigenvalues used for each latent partition and the column labeled Theta gives the reliabilities for each of the factors (composites).

Interrelationship of Latent Categories for Task I

Table 4 provides the interrelationships for the latent categories found in task I. The off-diagonal elements of these matrices provide an estimate of the probability that a random pair of concepts from two different latent categories will be sorted into the same observed category. These indices sometimes are called confusion indices. The diagonal elements provide an estimate of the probability that a random pair of elements from the *i*th latent category is being sorted into the same observed category. These indices provide information about the cohesiveness of the latent categories.

Latent Categories for Task II

The results of the latent categories for this task are found in Table 5. Six significant factors emerged for this set of thirty statements. Again, only those factors with eigenvalues over one were retained for further examination. The table presents the results of a principal component analysis and a final varimax rotation cluster for the task consisting of thirty statements. Content analysis of the individual cluster, as indicated by the inclusive underlined loadings, yield the following general labels: Excuses (F1), Magnificent Obsession (F2), Mega-Ambition (F3), Self-Motivator (F4), Discipline (F5), and Sharing with Others (F6). Using .40 as the criteria for significant loadings, the six latent categories contain 10, 5, 5, 4, 4, and 2 statements on F1 to F6 respectively. Some of the labels used here were similar to the ones used by the developers of the statements and by the sorters. A detailed analysis of this result, indicates that some statements appear to be loading high in more than one factor. For example, the statement labeled (T11) yielded two factor loadings higher than .50 on the factors labeled Magnificent Obsession and Self-Motivator. Other statements yielding high loadings on two factors, although not above 0.50, were T1, T3, T4, T13, T21, and T28.

Theta Reliability for Task II

Table 6 presents the transformation square matrix for this task. The elements of this matrix were used to calculate the respective theta coefficients for each of the composites. As before, the matrix was obtained by using principal component analysis. The obtained theta reliabilities are shown in Table 7. As previously noted the columns labeled Lambda and Theta indicate the eigenvalues and the actual reliability for the composites, respectively.

Interrelationship of Latent Categories for Task II

As mentioned before, this matrix provides information about how the sorters classified each of the statements in the proper latent categories (cohesiveness) and also how the sorters confused the latent categories. The results are shown in Table 8.

Latent Categories for Task III

Eight factors with eigenvalues greater than 1.0 were extracted by using principal component analysis on the joint proportion matrix for this task. The latent categories results are found in table 9. The table presents the final varimax rotation cluster for the fifty behavior verbs list used for this task. Examination of each individual factor yielded the following general cluster labels for this analysis. Aspects Conducive to Learning (F1), Aspects Detrimental to Learning (F2), Aspects that Reinforce Learning (F3), Aspects that make Teachers Efficient (F4), Aspects that deal with Evaluation of Learning (F5), Aspects that Ensure Learning (F6), Supervise (F7), and Reward (F8). The last two factors were single item factors. Using .40 as the criteria for significant loadings, the six multiple-item factors contained 15, 10, 11, 5, 3, and 4 verbs on F1 to F6. As observed in the previous two tasks, some of the verbs appeared to be loading on more than one factor. For example, let us look at the two single-verb factors. It is noted that the verb Supervise (T33) has two loading over 0.40 with one of the loadings in the factor dealing with aspects detrimental to learning (F2) and the other high loading by itself. Similarly with the verb rewards (T14), this verb has two loadings higher than .50. One high loading on the factor labeled Aspects Detrimental to Learning and the other forming a factor by itself. Other verbs yielding factor loadings higher than .40 on more than one factor were reminds (T27), advises (T36), assigns (T48), repeats (T23), and reviews (T24). The underlined factor loading scores indicate inclusion of the verbs in that column for that particular factor.

Theta Reliability for Task III

As before, each of the final solutions yielded a transformation matrix which is used to generate the theta reliabilities is found in Table 10. The transformation matrix for task III produced the interrelations that exist between the unrotated factor scores and rotated factor scores in the final solution.

The derived theta reliabilities for the eight factors were computed as presented in Table 11. Lambda and Theta columns in the same table are defined as before.

Interrelationship of Latent Categories for Task III

Table 13 provides the interrelationship square matrix for the eight latent categories obtained from the first analysis. The table is shown to indicate the level of confusion that the sorters had in sorting the verbs into the obtained latent categories and also the level or degree of cohesiveness for latent categories obtained. The rather small quantities obtained off the main diagonal indicate that a small percentage of sorters combined a verb from one latent category with a verb from another latent category. A diagonal entry from this matrix, as before, refers to a single latent category, say (F3), indicates the probability that a random pair of items (verbs) from the same latent category is sorted and put together.

DISCUSSION

Latent Catogories for Task I

The results for this task indicate clearly that the different clusters of objects appear to represent the underlying latent partition for the entire set of objects. The sorting of the 30 countries into these categories indicate that the latent partition for this task pertained to the general latent construct dealing with geographical locations. Of the 95 students performing the task, 20 failed to remain in the same dimension or construct while doing the sorting task. In other words, the sorters used more than one latent partition to classify or sort the countries. The rest of the sorters did remain within one latent dimension but this does not mean the selected dimension was only locational in nature.

The factor solution of the joint proportion matrix for this task led to the results obtained in Table 1. These results clearly showed that expected clustering of the individual countries in their respective factors led to the appropriate labeling of the resulting clusters or categories. There were some countries that contained high loadings and loaded on more than one factor. This indicates that the object or objects either belong to more than one category or the sorters simply perceived that object as having more than one common characteristic. For example, in the case of the countries forming the <u>Caribbean</u>

countries factor (F5) one of the countries had a high loading on another factor labeled Latin American countries (F2). The country (Dominican Republic) could belong also to the cluster of countries in the factor labeled Latin American countries without any loss of membership to either factor.

Theta Reliability for Task I

In order to make use of results from the principal component solution for task I, the information from the transformation matrix and the information gathered from the eigenvalues were used to generate the respective theta reliabilities for each of the categories in this composite. It was observed that the first three factors had the highest reliabilities ranging from .66 to .82. These factors contained countries that made up continents as opposed to the other factors which did not represent commonly known characteristics for those countries. Most of the other theta indices were low, particularly the factor dealing with Caribbean countries. This factor contained only three countries, and its factor loadings were not as high as the ones obtained for any of the other factors with higher reliability.

Armor (1974) claims that reliability of a composite can be improved by using factor scaling. This involves the use of factor scores for composite scales scores instead of the traditional unweighted sum of item scores. In other words, reliability will be improved when highest-loading items are kept in the factor. The reliability coefficient of the first factor in the composite for each of the analyses was lower than the second factor even though the eigenvalue for this first factor was the largest and should have yielded a larger theta reliability. However, when this first eigenvalue divided each of the quadratic terms in the transformation matrix, it yielded a larger term than expected and produced a lower index for the first factor of the composite. This is observed in each of the three theta analyses for each of the tasks. Once the eigenvalues tend gradually to level off for the other factors, the theta reliability indices tend to stabilize. Another aspect to be considered is that each of these factors with low theta reliabilities had two or three stimuli per factor. Thus, the fewer the number of objects or items in a composite the lower the reliability index.

The use of this coefficent (theta) to determine the internal consistency for each of the composites appears to indicate a more accurate assessment of the consistency that each of the objects appear to have for that particular factor whenever the case of a composite with multiple factors exists. It appears that if there are more objects within each of the factors or partitions, as identified here, then the reliability for each of the subscales will be improved.

Interrelationship of Latent Categories for Task I

In standard latent partition analysis two main matrices are of interest, the phi and omega matrices (Wiley, 1967). The matrix that describes the probability in which two items are combined is a func on of the latent categories to which the object or item belongs. In other words, the set of item combination probabilities is entirely dependent on a smaller probability set from the latent category combination. (Miller et al. 1986). In latent partition analysis, these probabilities are presented in the omega matrix (Table 4). As previously mentioned, the diagonal elements are an index of the degree of cohesiveness, and the off-diagonal elements provide an index of the degree of confusion among pairs of objects. From Table 4, it is clear that there was some confusion in some of the latent categories since two of the off-diagonal estimates are as large as one of the diagonal estimates. The category labeled Unknown countries had the largest off-diagonal estimate when paired with the latent category labeled Africa countries. The probability estimate for these two latent categories was 0.437. This result indicates that sorting concepts (countries) from two different latent categories (Unknown and African countries) were not being sorted into the same observed category. The same can be stated for the pair of latent categories labeled Latin American countries and Caribbean Islands. The probability estimate for this pair was .422. Similar argument can be made for these two latent categories. It is probable that the sorter will confuse pairs of countries from these categories since the Caribbean islands can also be considered as Latin America countries. The next largest estimate for the pair of latent categories labeled Europe and Northernmost countries was .290. Again, similar arguments can be stated for the results obtained for this pair of latent categories. The rest of the estimates indicated relatively little confusion among themselves. As indicated above, the diagonal entries for the omega matrix refer to a single latent category and are the probabilities that a given pair of items from that latent category will be sorted together. According to Miller et al. (1986), when the estimate is high (Miller provides no definition of "high") the concepts, or objects, and their latent category may be considered cohesive. The largest estimate in the diagonal was found on the African countries latent category. This estimate was .748 and indicated that any pair of countries for that latent category would have approximately a 75% chance of being sorted together. Miller et al. (1986) make a cautionary note with respect to this cohesiveness concept. They state that a latent category need not to be highly cohesive, but it must be homogoneous in its pattern of cohesion and confusion (p. 150).

Hambleton and Sheehan (1977) state that the average of the diagonal elements of the omega matrix may be used to reflect an overall agreement about the sorting or assignment of concepts into the latent categories. This particular mean for this task was .61. It is worth noting that there was no other source in the literature which provided more information about the theoretical basis and interpretation for this index than the one given by Hambleton and Sheehan.

Latent Categories for Task II

The analysis for this task are found in Table 5. The analyses for this data set contains success statements which yielded six latent categories for the 30 statement sorted by 45 students. The developers of these statements claimed that there are five major categories each with six statements. The appendix contains the actual subset of statements comprising each of the major categories as suggested by the researchers. In the present study, six major clusters or latent categories were identified by the principal component analysis. As clearly seen from Table 5, most of the statements within each latent category had very high loadings but it was also observed that some of these statements had high loadings in more than one latent category. This indicates certain degree of ambiguity for that statement as it was perceived by the sorters. A closer look at these statements and their respective latent category appears to indicate that the categories themselves were not totally unrelated. For example, consider the pair of latent categories labeled <u>Discipline</u> and <u>Self-Motivator</u>. There are four statements with high loadings in the Discipline cluster, and there were four in the Self-

Motivator cluster. Of these eight statements, there were four statements that showed moderate to high loadings in both latent categories.

Similar observations can be made for other pairs of latent categories as presented on Table 5. This overlapping seems to indicate that there are probably fewer than six distinct latent categories underlying this set of 30 statements.

Theta Reliability for Task II

In order to compute the theta reliability for this task, the results from the principal component solution were used to obtain the specific reliability index for each of the resulting latent categories. Specifically the factor transformation matrix and the eigenvalues for each of the latent categories as provided by the computed factor analysis results. Theta reliability computations were performed for both solutions. For the six latent category solution, the first two latent categories had the highest theta reliabilities. As was observed in the results in task I, the second latent category was larger in magnitude than the first latent category. The first latent category contained statements that pertained to negative or detrimental attributes and aspects while the second latent category contained statements that dealt with positive or desirable attributes and characteristics. These three latent categories by themselves accounted for 56 percent of the variance.

The next highest latent category was the one dealing with statements with connotations toward ambitious goals and dreams in order to achieve success. The rest of the theta reliabilities decreased in magnitude and the apparent reason for their low reliability index is attributed that each of these latent categories did not contain enough statements to form a true latent category.

Interrelationship of Latent Categories for Task II

As previously described, latent partition analysis generates two matrices which describe the sorting behavior of the sorters in a free categorization task. For task two, the omega matrix is presented in Table 8. The results of this analysis show that there was some confusion in several latent categories. The largest off-diagonal probability estimate was found with the pair of latent categories labeled Discipline and Self-Motivator. This probability estimate was 0.471 and indicates that pair of concepts from the two latent categories were not being categorized into the same observed or manifest category. Similar explanations can be made for the combination of latent categories labeled Mega-Ambition and Self-Motivator. It appears that sorters had more difficulty in sorting the statements which dealt with positive attributes than in sorting the concepts pertaining to negative statements from the latent category labeled Excuses. Notice the first column (F1) on the omega matrix, all the off-diagonal estimates were very low--meaning very little confusion among these type of statements. It is also observed that the first diagonal element in the matrix was the largest in relation to the rest of the diagonal probability estimates.

The main diagonal estimates for this task indicate that the latent categories do not seem to be highly cohesive among themselves. The average category cohesiveness value was .58, which is not very high in terms of the overall agreement that may have existed among the sorters.

Latent Categories for Task III

The results for the sorting task dealing with the teacher behaviors were presented in Table 9. Eight latent categories were extracted by the principal component procedure. This set of verbs depicting teachers' behaviors were the same employed in the original study by Miller et. al (1967). This very study led to the development of the categorization methodology which is composed of the F-sort technique and the latent partition technique. The primary use of this set of verbs was to make comparisons with the results obtained by the Miller et al. (1967) study. Their original study produced ten different latent categories for the set of verbs. The results of the study show that the there were two types of categorizations created by the sorters. According to Miller et al., the substance and structure of the categories suggest that the sorters grouped the set of verbs from different points of view. One of the types of categories created showed finer discriminations while the other category type reflects more evaluative ideas. They claimed that these two types of categorizations clearly portray the kinds of different perceptions that can be observed by the F-sort procedure.

The results for the present study showed six clearly identifiable clusters and two single item (verb) factors. These results seem to match the Miller experimental results primarily in the second type of categorization which dealt with aspects desirable and undesirable in teaching. It should be mentioned that the sample of sorters used in the original study was substantially larger than the sample size selected for the present study. Additionally, the original study selected both teachers and preservice teachers while in this study only preservice teachers were selected for the sorting task.

It is interesting to note that clustering of the individual verbs was very similar to those of the original study even though the sample size for the present study was relatively small when compared with the sample size of the original study. This seems to indicate that in order to establish categories of this type there is no need for sample sizes as large as the one used by Miller et al. (1967) since the results will not be that drastically different from each other. This may also reduce the probability that undetected subgroups within the sorter sample may perceive the same set of concepts in a different way yielding larger number of latent categories, which in turn may be difficult to describe and explain.

Theta Reliability for Task III

The theta reliabilities for the present task were reported in Table 11. The results for this analysis showed that the theta indices ranged from .30 to .88. Similar behavior was observed for the first and second latent categories. The first category, as previous tasks, yielded a smaller index as compared to the second latent category. The first five latent categories had reliabilities that were high in relation to the last three categories indicating the plausibility of an alternative model with a smaller number of latent categories explaining the same set of verbs. It is noted that both single verb categories had a higher theta reliability than the latent category dealing with aspects of teaching that ensure learning. This is explained by the high degree of ambiguity of the verbs belonging to this category.

Interrelationship of Latent Categories for Task III

The results of the interrelationships that exist among the latent categories were presented on Table 12. The results of the omega matrix were those from the eight factor solution. The off-diagonal elements in the matrix indicate relative amount of confusion for this set of latent categories. The largest off-diagonal probability estimate was .425, while the smallest probability estimate was .106. This large probability estimate indicates that pairs of concepts (verbs) from the latent categories labeled Aspects Conducive to Learning and Aspects that Ensure Learning were being sorted 43% of time. That is, a large number of of sorters did not differentiate the verbs of the two categories; therefore, verbs from the Aspects Conducive to Learning category were being combined with verbs from the category of Aspects that Ensure Learning. This shows that these two latent categories may not be entirely different from each other. The diagonal entries from this matrix shows low probability estimates for this set of latent categories. The largest values observed was 1.0 for the two categories with just one concept. These results were expected and are considered meaningless since it is desired that at least two concepts with common characteristics be sorted together. The largest estimate in this group was obtained for the latent category dealing with Aspects that make a Teacher more Efficient in the facilitation of learning. This index (.833) indicates that a given pair of items from that latent category had an 83% chance of being sorted together. The rest of the estimates were moderate implying moderate levels of cohesiveness.

SUMMARY

The central focus of this research was to illustrate the applicability of categorization methodology and to draw implications regarding the use of such methods developed and employed in other related fields and testing situations. The inclusion of the theta reliability index was considered of major importance to the enhancement of the existing and useful methodology of categorization developed by Miller et al. (1967). The selection of the three sorting tasks was based on the decision to employ different situations in which categorization methodology has been applied and may be applied by researchers. These three tasks were used to illustrate the use of these measures that were adapted for use in this methodology.

The use of theta reliability procedures in latent partition analysis appears to be justified given the nature of the tasks usually involved in this methodology. This methodology calls for the creation of more than one category for a given set of stimuli, therefore, conventional reliability procedures may not be applied to this type of multiple-category solutions as found in latent partition analysis. The indices of theta reliability for each of the illustrations provided an insight on how the stimuli in each partition were actually related to the latent category. This proposed index is endorsed in the present study as a means for the determination of the internal consistency of the subscales or subcategories that are derived from the latent partition analysis. This measure should be used to reflect the degree by which many sorters consistently sort sets of particular items into manifest categories. Due to the lack of existing measures such as theta, it is proposed that theta reliability analysis should be an integral part of the categorization methodology. The addition of this measure should only enhance the utility of this categorization methodology, regardless of the testing situation.

While the results of this study demonstrated the usefulness of categorization methodology by using factor analysis techniques, it became apparent that Armor's theta reliability coeffiecient greatly enhanced the present methodology. However, the coefficient requires further theoretical research. This research can be focused in determining theoretical distributional properties to determine significance and upper or lower confidence limits for sample statistics. This may be done through the use of Monte Carlo simulation studies. Additional studies can be geared toward determining differences between the empirical results obtained when methods other than principal component analyses are employed to determine the eigenvalues, which are used to calculate theta.

REFERENCES

Armor, D. J. (1974). Theta reliability and factor scaling. In H. Costner (Ed.), *Sociological methodology 1973-1974*. San Francisco: Jossey-Bass Pub.

Bentler, P. M. (1968). Alpha-maximized factor analysis and its relation to alpha and canonical factor analysis. *Psychometrika*, *33*, 335-346.

Bruner, J. S., Goodnow, J. J., & Austin, G. A. *A study of thinking*. New York: Wiley, 1956.

Chi, M., Feltovich, P, & Glaser, R. (1980). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, *5*, 121-152.

Colleta, A. J. & Gable, R. K. (1975). The content of construct validity of the Barth scale: Assumptions of open education. *Educational and Psychological Measurement*, *35*, 415-425.

Conry, R. F. (1973). Effects of cognitive style differences and stimulus factors on free-sorting behavior (Doctoral dissertation, University of Wisconsin-Madison). *Dissertation Abstracts International*, 34(3), 1123-A. (University Microfilms No. 73-19, 294).

Cronbach, L. J., (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*, 287-334.

Diamond, C. T. P. (1980). A note on measuring teachers' constructs. *Research in the Teaching of English*, *14(2)*, 161-164.

Diamond, C. T. P. (1983). Teacher Perspectives on the teaching of writing. The Alberta Journal of Research, 29(1), 15-30.

Evans, G. T. (1970). The analysis of categorizing behavior. Psychometrika, 35, 3, 367-392.

Gardenswart, L. & Rowe, A., (1987). What it takes. New York: Doubleday.

Gorodetsky, M. & Hoz, R. (1985). Change in the group cognitive structure of some chemical equilibrium concepts following a university course in general chemistry. Science Edu tion, 69, 2, 185-199.

Guttman, L., (1953). Image theory for the structure of quantitative variates. Psychometrika, 18, 277-296.

Hambleton, R. K. & Sheenan D. S. (1977). On the evaluation of higher-order science instructional objectives. Science Education, 61, 3, 307-315.

Hartke, A. R. (1978) The use of latent partition analysis to identify homogeneity of an item population. Journal of Educational Measurement 15, 43-47.

Hartke, A. R. (1979). The development of conceptually independent sub-scales in the measurement of attitudes. Educational and Psychological Measurement, 39, 585-592.

Harris, C. W. & Kaiser, H. F. (1964). Oblique factor analytic solutions by orthogonal transformations. Psychometrika, 29, 4, 347-362.

Heise, D. R. & Bohrnstedt, G. W. (1970). Validity, invalidity, and reliability. In E. F. Borgatta and G. W. Bohrnstedt (Eds.), Sociological Methodology 1970. San Francisco: Jossey-Bass.

Hofmann, R. J. (1987). Extended latent partition analysis. Educational and Psychological Measurement, 47, 39-44.

James, E. J. & Webb, K. (1980). One World Divided: A geographer look at the modern world, 3e. New York: Wiley.

Joreskog, K. G. & Sorbom, D., (1986). Lisrel VI. User's Guide. Chicago: Scientific Software Inc.

Kaiser, H. F. (1968). The varimax criterion for analytic rotation in factor analysis. Psychometrika, 23, 187-200.

Kaiser, H. F. & Caffrey, J., (1965). Alpha factor analysis. Psychometrika. 30, 1-14.

Lane, R. G. (1967). Counselors conceptualization of client statements (Doctoral dissertation, University of Wisconsin-Madison. Dissertation Abstracts International. 28(3), 966-A. (University Microfilms No. 66-13, 807).

Lazarsfeld, P. F. & Henry, N. W. (1968). Latent structure analysis. Boston: Houghton Mifflin.

Lord, F, M, & Novick, M. R. (1968). Statistical theories of mental test scores. Reading, MA: Addison-Wesley.

Lord, F. M. (1958). Some relations between Guttman's principal components of scale analysis and other psychometric theory. Psychometrika 23, 291-296.

Maguire, T. O. Patsula, R. B. & Evanechko, P. O. (1975) The development of word meaning discrimination in children. Alberta Journal of Educational Research, 21, 151-167.

Miller, D. M., Baker, J. F., Clases, R. E., Conry, R. F., Conry, A. D., Sheets, S. S., Wiley, D. E., & Wolfe, R. G. (1967). Elementary school teacher's viewpoints of classroom teaching and learning (Final Report, U.S. Office of Education Project 5-1015-2-12-1). Madison WI: University of Wisconsin-Madison), Instructional Research Laboratory, (ERIC Document Reproduction Service No. ED 046 904).

Miller, D. M., Wiley D. E., & Wolfe, R. G. (1986). Categorization methodology: An approach to the collection and analysis of certain classes of qualitative information. Multivariate Behavioral Research, 135-167.

Novick, M. R. & Lewis, G. (1967). Coefficient Alpha and the reliability of composite measurements. Psychometrika, 32, 1-13.

Pruzek, R. M. (1968). A comparison of two methods for studying test items (Doctoral dissertation, University of Wisconsin-Madison). Dissertation Abstracts International, 28(8), 3035-A. (University Microfilms No. 67-12, 464).

Rideng, I. M. & Schibeci, R. A. (1984). The development and Validation of a Test of Biology-related attitudes. Research in Science & Technological Education, 2, 1, 21-29.

Serlin, R. C. & Kaiser, H. F. (1978). A method for increasing the reliability for a short multiple-choice test. Educational and Psychological Measurement, 38, 2, 337-340.

Schuell, T. J. (1985). Knowledge representation, cognitive structure, and school learning: A historical perspective. In L. H. T. West & A. L. Pines (Eds.), Cognitive Structure and Conceptual Change. London: Academic Press.

Shavelson, R. J. & Stanton, G. C. (1975). Construct validation: Methodology and application of three measures of cognitive structure. Journal of Educational Measurement, 12(2), 67-85.

Stephenson, W. (1953). The study of behavior. Chicago: University of Chicago Press.

SPSS Inc. (1987). SPSS-X: User's Guide. New York, McGraw-Hill.

Weitman, C. J. (1986). Hierarchies of effective early childhood teacher dimensions as perceived by Texas directors, parents, and caregivers. Doctoral dissertation, Texas A&mM University, College Station, Texas.

White, R. T. (1985). Interview protocols and dimensions of cognitive structure. In L. H. T. West & A. L. Pines (Eds.), Cognitive Structure and Conceptual Change. London: Academic Press.

Whitely, S. E. (1976). Solving verbal analogies: Some cognitive components of intelligence test items. Journal of Educational Psychology, 68, 2, 234-242.

Whitely, S. E. (1977). Relationship in analogy items: A semantic component of a psychometric task. Educational and Psychological, 37, 725-739.

Whitely, S. E. & Doyle, K. O. (1976). Implicit theories in student ratings. American Educational Research Journal, 13(4), 241-253.

Whitely, S. E. & Doyle, K. O. (1978). Dimensions of effective teaching: Factors or artifacts. Educational and Psychological Measurement, 38, 1, 107-117.

Wiley, D. E. (1967). Latent partition analysis. Psychometrika, 32, 183-193.

Willson, V. L. (1982). Maximizing reliability in multiple choice questions. Educational and Psychological Measurement. 42, 69-72.

Willson, V. L. & Palmer, D. J. (1983). Latent partition analysis of attribution for actual achievement. American Educational Research Journal, 20, 4, 581-589.

Willson, V. L. & Reynolds, C. R. (1984). Regression effects on part scores based on whole-scores selected samples. Educational and Psychological Measurement, 44, 95-99.

Willson, V. L. & Reynolds, C. R. (1985). Constructing short forms from composite tests: Reliability and validity. Educational and Psychological Measurement, 45, 469-475.

TABLES AND APPENDICES

TABLE 1

Latent factor structure for task I: Six factor solution

Name of Stimulus	F1	F2	F3	F4	F5	F6
Belgium	.87798	.02379	.07483	.04888	.05782	.08099
Czechoslovaquia	.87667	.01773	.01717	-.03355	.05225	.04590
Luxembourg	.86790	-.01947	.00410	.14794	.16138	-.08395
Romania	.83514	.07213	.11101	.04087	.10073	-.00670
Denmark	.81962	.02473	.05064	.03524	-.04709	.34663
Spain	.80447	.20500	.07407	.00678	-.01666	.01444
Greece	.80383	.11696	.05103	.00070	-.03538	.06654
Liechtenstein	.79347	-.03288	-.01192	.25770	.15409	-.06504
Finland	.78619	.03271	.07585	.04658	-.08622	.36540
Peru	.08901	.91356	.07538	.06598	-.03162	.01181
Brazil	.08871	.89458	.06156	.03949	-.00889	.02958
Panama	.01498	.84995	.07425	.07445	.20507	.13169
Guatemala	.04397	.81080	.05254	.05758	.25137	.05669
Paraguay	.10802	.76471	.19205	.15858	.10468	-.06228
Mexico	.00028	.64600	.03848	.03562	.20428	.43498
Zimbabwe	.06445	.08977	.92291	.13602	.06119	.05204
Nigeria	.06430	.09692	.91425	.19731	.07088	.00120
Niger	.06362	.15263	.88817	.20296	.00528	.03145
Rwanda	.04668	.03188	.65762	.57763	.02746	.01713
Tanzania	.01309	.03093	.65716	.57819	.06777	.01081
Morocco	.27492	.14715	.46702	.22960	.28777	-.07346
Mauritania	.02221	.12252	.15994	.79236	.20107	.01112
Senegal	.06889	.09197	.23359	.75552	.17572	.05121
Cameroon	.01106	.04100	.41178	.71943	.19374	.08953
Nepal	.33420	.13560	.16088	.61995	-.13896	-.00666
Barbados	.112[illegible]	.36414	.02455	.16925	.73521	.12702
Trinidad Tobago	.09339	.17347	.25038	.26839	.65628	-.01237
Dominican Republic	.04204	.57727	.03215	.03023	.58730	.15747
Canada	.06243	.22708	-.00446	.04139	.08526	.82643
Iceland	.42321	.02561	.04908	.05455	.03569	.71125

Note. The underlined factor loadings for each factor indicate the membership of the objects to that factor.

F1 = European countries
F2 = Latin American countries
F3 = African countries
F4 = Unknown countries
F5 = Caribbean Islands
F6 = Northern countries

TABLE 2

Transformation matrix for task I: Six factor solution

Factors	F1	F2	F3	F4	F5	F6
F1	.63831	.43224	.41716	.38186	.23543	.17450
F2	-.75279	.38899	.39343	.30000	.18459	-.05593
F3	-.02628	.72755	-.51447	-.38934	.14399	.18161
F4	-.02835	-.18242	-.58023	.59288	.52348	-.06101
F5	-.15579	-.21265	.01692	.04102	.00205	.96361
F6	-.00971	.23244	-.26373	.50932	-.78469	.03434

Note. The elements of this matrix were used to calculate the respective theta coefficient for each of the categories in task I

F1 = European countries
F2 = Latin American countries
F3 = African countries
F4 = Unknown countries
F5 = Caribbean countries
F6 = Northernmost countries

TABLE 3

Theta reliabilities for task I: Six factor solution

	Latent partition	Lambda	Theta
1.	European countries	6.649	.788
2.	Latin American c.	4.744	.822
3.	African countries	3.992	.769
4.	Unknown countries	3.170	.664
5.	Caribbean countries	1.788	.469
6.	Northernmost countries	1.743	.563

Note. Lambda indicates the eigenvalues obtained after the rotation.

TABLE 4

Interrelationships of the latent categories

for task I: Six factor solution

Latent category	F1	F2	F3	F4	F5	F6
F1	0 682					
F2	0.108	0.669				
F3	0.122	0.177	0.748			
F4	0.136	0.169	0.437	0.553		
F5	0.131	0.422	0.195	0.261	0.533	
F6	0.290	0.186	0.079	0.107	0.184	0.459

F1 = European countries
F2 = Latin American countries
F3 = African countries
F4 = Unknown countries
F5 = Caribbean Islands
F6 = Northernmost countries

TABLE 5

Latent factor structure for task II: Six factor solution

Name of Stimulus	F1	F2	F3	F4	F5	F6
T2	<u>.94865</u>	-.00259	-.01516	.02112	-.02633	-.02893
T18	<u>.91441</u>	.01109	-.03264	.05537	-.03071	-.03418
T24	<u>.89777</u>	.04245	-.00808	.05914	-.01198	-.01666
T22	<u>.89250</u>	.00457	.03383	-.00140	.02003	-.01837
T14	<u>.87555</u>	.02196	-.00873	.04767	.01037	.02336
T20	<u>.80842</u>	.13532	-.01708	.03502	.06083	-.02755
T9	<u>.79655</u>	.18114	.09821	-.01444	-.01619	-.02108
T29	<u>.78335</u>	-.01010	.08163	-.03326	-.05313	.22426
T16	<u>.68489</u>	-.05291	.08693	-.11902	.26466	.15959
T26	<u>.52981</u>	.36421	-.00475	.09650	.18625	.00462
T8	.07404	<u>.80893</u>	.18868	.01009	.10921	.12797
T23	.02870	<u>.78500</u>	.39047	.11918	.04057	-.02567
T3	.01956	<u>.72609</u>	<u>.415[illegible]8</u>	.14701	.04839	.04435
T27	.16306	<u>.69720</u>	.05960	.20898	.09237	.18160
T11	.04635	<u>.56578</u>	-.00648	<u>.50730</u>	.03249	.29826
T17	.00837	.15056	<u>.79736</u>	.34896	.12087	.03492
T6	.01325	.09907	<u>.77362</u>	.25071	.22467	.04554
T19	.00836	.30705	<u>.71135</u>	.08812	.08578	.13950
T28	.01494	<u>.46412</u>	<u>.6362[illegible]</u>	.29700	.03153	-.04429
T13	.16309	<u>.49328</u>	<u>.59866</u>	-.10118	.07553	-.07717
T7	-.00242	.21716	.35889	<u>.76119</u>	.16955	.04301
T1	.00175	.10689	.17983	<u>.73062</u>	<u>.42269</u>	.15686
T12	.05068	.23734	.19754	<u>.70929</u>	.33017	-.00254
T4	.05202	-.04932	<u>.41934</u>	<u>.55476</u>	.22390	.32576
T10	.07137	.08330	.06488	.08735	<u>.820[illegible]7</u>	.16384
T25	.05655	.02348	.14728	.25628	<u>.81215</u>	.07566
T21	-.00261	.24633	.10063	<u>.41905</u>	<u>.64455</u>	.04413
T5	.01867	.05019	.31288	.39650	<u>.55470</u>	.15879
T15	.05764	.34006	.00890	.16870	.06036	<u>.82728</u>
T30	.05310	.04193	.11676	.11803	.34096	<u>.80323</u>

Note. The complete description of the stimulus name is provided in appendix B. The underlined factor loadings indicate membership of object to that factor.

F1 = Excuses
F2 = Magnificent obsession
F3 = Mega-ambition
F4 = Self-motivator
F5 = Discipline
F6 = Sharing with others

TABLE 6

Transformation matrix for task II: Six factor solution

Factors	F1	F2	F3	F4	F5	F6
F1	.48072	.47017	.45109	.42366	.35079	.20451
F2	.87246	-.18279	-.28117	-.28395	-.19506	-.08596
F3	.04327	-.61892	-.29048	.33019	.59940	.24977
F4	-.06788	.48247	-.61751	-.09858	.00413	.60956
F5	.00361	-.31896	.49587	-.36793	-.17701	.69688
F6	-.03505	.16720	.07670	-.69696	.66953	-.17579

Note. The elements of this matrix were used to calculate the respective theta coefficient for each of the categories in task II.

F1 = Excuses
F2 = Magnificent obsession
F3 = Mega-Ambition
F4 = Self-motivator
F5 = Discipline
F6 = Sharing with others

TABLE 7

Theta reliabilities for task II: Six factor solution

	Latent partition	Lambda	Theta
1.	Excuses	6.839	.744
2.	Magnificent obsession	3.683	.840
3.	Mega-ambition	3.404	.693
4.	Self-motivator	3.064	.632
5.	Discipline	2.756	.590
6.	Sharing with others	1.772	.674

Note. Lambda indicates the eigenvalues obtained after the rotation.

TABLE 8

Interrelationships of the latent categories

for task II: Six factor solution

Latent category	F1	F2	F3	F4	F5	F6
F1	0.737					
F2	0.087	0.606				
F3	0.050	0.428	0.577			
F4	0.042	0.320	0.408	0.624		
F5	0.052	0.220	0.283	0.471	0.547	
F6	0.066	0.257	0.177	0.301	0.303	0.620

F1 = Excuses
F2 = Magnificent obsession
F3 = Mega-ambition
F4 = Self-motivator
F5 = Discipline
F6 = Sharing with others

TABLE 9

Latent factor structure for task III: Eight factor solution

Name of Stimulus	F1	F2	F3	F4	F5	F6	F7	F8
Clarifies	.86816	.01043	.18232	.06711	.09432	.01565	.12334	.04921
Simplifies	.84433	.02618	.15323	.11639	-.00570	.05455	-.07257	.05058
Explains	.82466	.00173	.18268	.07704	.05640	.16856	.18605	-.02774
Discusses	.79147	.00695	.22097	.06126	-.01394	.20239	.14826	-.12410
Illustrates	.78330	.03198	.14529	.23045	-.01479	.23807	-.26099	.18247
Interprets	.78104	.03619	.19377	.10357	.20077	-.03526	.14055	-.07663
Introduces	.77380	.05947	.11035	.20797	.11310	.10964	.08458	-.02482
Demonstrates	.76422	.03302	.08939	.24500	-.00030	.26089	-.18874	.22169
Reasons	.71968	.02555	.31379	.08024	.11694	-.01394	.10913	-.20033
Answers	.71931	.02360	-.10358	.04944	.30046	.23264	.05484	-.02311
Displays	.65660	.02632	.18464	.34214	-.04961	.17426	-.25448	.28957
Exemplifies	.54769	.09780	.33376	.15454	.00507	.21215	.05672	.03636
Lectures	.48347	.18996	.09012	.12632	.11299	.36739	-.00351	-.39104
Questions	.47230	.03856	.19797	.10664	.31529	.35678	-.18495	.10643
Reminds	.42282	.07351	.40255	.10612	-.03942	.35911	.24887	.06810
Threatens	-.01111	.93868	.05386	-.00816	.00687	.06401	-.10031	.00109
Restricts	.01035	.91542	.06103	.03454	.04462	.01695	-.01391	-.03442
Reprimands	-.01536	.90482	.02796	.03149	.02314	.07491	-.04720	-.00309
Penalizes	.01770	.89324	.06783	.02697	.09017	.03225	-.07268	.06288
Controls	.03092	.81990	.05354	.17622	.06853	.06360	.27381	.00842
Enforces	.09877	.81815	.05744	.09304	.06625	.03525	.22319	.12329
Demands	.04080	.79432	.19391	-.03835	.05070	.05007	-.13658	-.06768
Regulates	.08290	.78852	.07403	.17672	.03869	-.00453	.30481	.01955
Judges	.04133	.59003	.12677	.00884	.35127	.05362	-.15997	-.08500
Permits	.03664	.42270	.36964	.18409	.06651	.13762	.30021	.12815

Note. The underlined factor loading for each factor indicate the membership of the objects to that factor.

F1 = Aspects conducive to learning
F2 = Aspects detrimental to learning
F3 = Aspects that reinforce learning
F4 = Aspects that make teachers efficient
F5 = Aspects that deal with evaluation of learning
F6 = Aspects that ensure learning
F7 = Supervise
F8 = Reward

TABLE 9

(Continued)

Name of Stimulus	F1	F2	F3	F4	F5	F6	F7	F8
Encourages	.17326	.03812	.86131	.11370	.01817	.04449	.06409	.15600
Inspires	.19370	.03660	.85920	.10256	.02103	.07343	.03014	.19287
Stimulates	.24825	.04297	.83363	.09953	.03632	.01685	-.0[illegible]172	.15621
Persuades	.24484	.06859	.82798	.07329	.05873	-.01438	-.00411	-.21442
Urges	.16930	.11638	.81337	.04339	-.01200	.10133	-.01527	-.11762
Impels	.00421	.38478	.69869	-.04178	.00215	.10422	-.07942	-.08009
Commends	.16947	.10346	.65411	.03538	.20743	-.02074	.11176	.43285
Convinces	.36769	.07477	.64790	.04206	.10300	.02128	-.00738	-.35742
Advises	.30995	.04238	.55371	.10035	.[illegible]0613	.10741	.40134	-.06799
Confirms	.38306	.05526	.43805	-.01240	.33075	.03806	.33799	.13514
Reinforces	.32786	.16007	.41859	-.03175	.03704	.36966	.21599	.15347
Plans	.18593	.09700	.08464	.9[illegible]49	.16395	.04833	.05872	-.02476
Organizes	.21672	.04390	.10256	.89[illegible]	.10315	.09538	.06135	.02125
Arranges	.25798	.07835	.12410	.86127	.05498	.04631	.02563	.04498
Schedules	.16316	.14371	.04098	.84265	.21972	.09048	-.04485	.01389
Assigns	.19642	.15537	.06431	.44799	.32685	.42457	.18109	-.12260
Grades	.05362	.19525	.04034	.23674	.83876	.16617	.00480	.03665
Evaluates	.22729	.08423	.14496	.13246	.80586	.01889	.18492	.04889
Tests	.14566	.16329	.93208	.22727	.77551	.26541	-.11489	.02230
Drills	.25696	.27785	.00328	.04163	.24509	.66899	-.14408	-.09932
Repeats	.52951	.09380	.08267	.03591	.08382	.60297	.11921	-.04126
Reviews	.47621	.00065	.07315	.14058	.21759	.59189	.03628	.12640
Tutors	.31660	-.02307	.19730	.26801	.09816	.50569	.14393	.04947
Supervises	.26541	.41603	.17857	.30609	.00171	.19576	.45762	.00354
Rewards	.15392	.15790	.50193	.09149	.27583	.07174	.01362	.56917

Note. The underlined factor loadings for each factor. indicate the membership of the objects to that factor

F1 = Aspects that are conducive to learning
F2 = Aspects that are detrimental to learning
F3 = Aspects that reinforce learning
F4 = Aspects that make teachers efficient
F5 = Aspects that deal with evaluation of learning
F6 = Aspects that ensure learning
F7 = Supervise
F8 = Reward

TABLE 10

Transformation matrix for task III: Eight factor solution

Factors	F1	F2	F3	F4	F5	F6	F7	F8
F1	.66571	.30341	.47782	.30073	.22964	.28410	.10123	.04831
F2	-.42255	.90096	-.00550	.00158	.07983	-.04209	.03897	-.00498
F3	.21036	.08599	-.82138	.40497	.22304	.22363	-.06443	-.07602
F4	-.47210	-.25694	.24634	.70983	.33176	-.10006	.05904	.15166
F5	-.13807	-.12015	.01265	-.47999	.80547	.29187	-.02838	.03428
F6	.04022	.06096	.00916	-.00231	-.04981	.02782	-.75718	.64652
F7	-.30063	-.06707	.10489	.05554	-.34134	.86679	-.07777	-.13095
F8	.01003	.01198	.15831	.09067	.11533	-.14217	-.63278	-.72976

Note. The elements of this matrix were used to calculate the respective theta coefficient for each of the categories in task III.

F1 = Aspects conducive to learning
F2 = Aspects detrimental to learning
F3 = Aspects that reinforce learning
F4 = Aspects that make teachers more efficient
F5 = Aspects that deal with evaluation of learning
F6 = Aspects that ensure learning
F7 = Supervise
F8 = Rewards

TABLE 11

Theta reliabilities for task III: Eight factor solution

	Latent Partition	Lambda	Theta
1.	Conducive to learning	9.395	.845
2.	Detrimental to learning	7.154	.881
3.	Reinforce learning	6.828	.830
4.	Teacher efficiency	4.144	.792
5.	Evaluation of learning	2.969	.703
6.	Ensure learning	2.713	.301
7.	Supervises	1.463	.668
8.	Rewards	1.343	.319

Note. Lambda indicates the eigenvalues obtained after rotation.

TABLE 12

Interrelationships of latent categories

for task III: Eight factor solution

Latent category	F1	F2	F3	F4	F5	F6	F7	F8
F1	0.558							
F2	0.106	0.618						
F3	0.328	0.185	0.480					
F4	0.306	0.165	0.210	0.833				
F5	0.240	0.215	0.209	0.328	0.720			
F6	0.425	0.148	0.250	0.278	0.315	0.477		
F7	0.329	0.400	0.298	0.355	0.240	0.295	1.000	
F8	0.268	0.220	0.413	0.225	0.320	0.230	0.260	1.000

F1 = Aspects conducive to learning
F2 = Aspects detrimental to learning
F3 = Aspects that reinforce learning
F4 = Aspects that make teachers efficient
F5 = Aspects that deal with evaluation of learning
F6 = Aspects that ensure learning
F7 = Supervise F8 = Reward

Note. The element i,j) or (j,i) for these matrices provides an estimate of the probability that pairs of concepts randomly drawn from two different latent categories will be sorted into the same observed category. The diagonal elements, (i,i), provide an estimate of that probability of pairs of stimuli randomly drawn from the ith latent category being sorted into the same observed category.

APPENDIX A

STIMULI DESCRIPTION NAME FOR SORTING TASKS

Description of stimulus name for task I

LIST OF COUNTRIES

T1.	TANZANIA	T2.	RWANDA
T3.	MORROCO	T4.	CAMEROON
T5.	MAURITANIA	T6.	NIGER
T7.	NIGERIA	T8.	ZIMBABWE
T9.	SENEGAL	T10.	NEPAL
T11.	DOMINICAN REPUBLIC	T12.	BARBADOS
T13.	TRINIDAD TOBAGO	T14.	CZECHOSLOVAQUIA
T15.	GREECE	T16.	DENMARK
T17.	BELGIUM	T18.	LUXEMBOURG
T19.	ROMANIA	T20.	SPAIN
T21.	LIECHTENSTEIN	T22.	FINLAND
T23.	ICELAND	T24.	CANADA
T25.	MEXICO	T26.	BRAZIL
T27.	PANAMA	T28.	GUATEMALA
T29.	PERU	T30.	PARAGUAY

Description of stimulus name for task III

MILLER ET AL. LIST OF VERBS

DEMONSTRATES	ILLUSTRATES
DISPLAYS	THREATENS
PENALIZES	REPRIMANDS
DEMANDS	RESTRICTS
IMPELS	INSPIRES
ENCOURAGES	COMMENDS
STIMULATES	REWARDS
GRADES	TESTS
EVALUATES	JUDGES
PLANS	ORGANIZES
ARRANGES	SCHEDULES
REPEATS	REVIEWS
DRILLS	REINFORCES
REMINDS	PERSUADES
CONVINCES	URGES
REGULATES	CONTROLS
SUPERVISES	PERMITS
ENFORCES	ADVISES
CLARIFIES	SIMPLIFIES
INTERPRETS	EXPLAINS
REASONS	CONFIRMS
ANSWERS	EXEMPLIFIES
LECTURES	DISCUSSES
QUESTIONS	ASSIGNS
INTRODUCES	TUTORS

Description of stimulus name for task II

GARDENSWARTZ AND ROWE LIST OF STATEMENTS

T1 I am energetic and enthusiastic about my life and work.

T2. I avoid new situations and challenges.

T3. I want to rank with the greats in my field.

T4. I accept the consequences of my choices.

T5. I parlay each of my experiences into many opportunities.

T6. The goals I am implementing are my own.

T7. I am a self-starter, and I'm quick to take action when I know what I want.

T8. I like to be a big fish in a big pond.

T9. My "wishbone" is stronger than by "backbone."

T10. Reading is a priority for me.

T11. I expect performance to be top-notch -both mine and others'.

T12. I commit no-holds-barred energy to tasks that are important to me.

T13. My goals and dreams are whale size.

T14. I am reluctant to continue when I'm told it can't be done.

T15. I present ideas so that people see what's in it for them.

T16. It's important to me to work no more than 40 hours a week.

T17. I continue to set challenging goals for myself.

T18. The feeling that "I can't" influences my behavior.

T19. Each person creates his or her own destiny.

Description of stimulus names for task II

T20. I resist new ideas once I've made up my mind.

T21. Work is a source of joy to me.

T22. I let things fall through the cracks.

T23. It's important to me to make my mark in the world.

T24. Discrimination about women holds me back.

T25. I am comfortable dealing with budgetary and financial matters.

T26. I get engrossed in my work and forget there's a world outside.

T27. I am happiest when I'm in charge.

T28. My motto is, "The sky's the limit."

T29. My responsibilities to others keep me from working on my own goals.

T30. I have a cadre of people I can count on for help and support.

APPENDIX B

INSTRUCTION FORMS FOR SORTING TASKS

INSTRUCTIONS

THE SORTING OF COUNTRIES INTO CATEGORIES

PURPOSE: THE OBJECTIVE OF THE TASK IS TO CREATE CATEGORIES (PILES) WHICH DESCRIBE THE COUNTRIES BY A COMMON FEATURE OR CHARACTERISTICS.

TASK DESCRIPTION:

1. START WITH THE FIRST COUNTRY CARD.
2. USE THE CARD TO START YOUR FIRST CATEGORY.
3. PICK UP THE NEXT COUNTRY CARD.
4. DECIDE WHETHER THIS COUNTRY CAN BE GROUPED WITH THE PREVIOUS ONE:

   IF YES, THEN PUT THE CARD IN THAT CATEGORY GROUP

   IF NOT, THEN CARD TO CREATE A NEW CATEGORY GROUP
5. PICK UP THE NEXT CARD AND REPEAT THE PROCESS.

IF YOU ARE OUT OF CARDS:

1. DECIDE WHETHER YOUR CATEGORIES ARE A SATISFACTORY REPRESENTATION OF YOUR SORTING.
2. IF NOT, MAKE CHANGES BY SWITCHING CARDS.
3. PUT A BLANK CARD ON TOP OF EACH CATEGORY AND WRITE A TITLE FOR THE GROUP ON THE BLANK CARD.
4. PUT THE CATEGORY GROUPS IN A SINGLE DECK.

DEMOGRAPHICS

ID # ______________________ (PLACE NUMBER ON EACH BLANK CARD USED FOR EACH CATEGORY)

COURSE ______________________

GENDER ______________________

MAJOR ______________________

CLASSIFICATION ______________

AGE ______________________

THANK YOU

INSTRUCTIONS

THE SORTING OF VERBS INTO CATEGORIES

PURPOSE: THE PURPOSE OF THIS TASK IS TO STUDY YOUR VIEWS ON CLASSROOM TEACHING BY USING A SET OF VERBS CHARACTERIZING TEACHING BEHAVIOR. YOUR VIEWS MAY REFLECT WHAT IS EXPECTED FROM A TEACHER IN ORDER TO FACILITATE LEARNING.
YOUR TASK IS TO MAKE A JUDGMENTAL DECISION FOR EACH STATEMENT AND CREATE THESE PILES THAT WILL POSSES A COMMON CHARACTERISTIC.

TASK DESCRIPTION:

1. START WITH THE FIRST VERB CARD.
2. CONSIDER THE VERB IN TERMS OF YOUR VIEWS FOR FACILITATING LEARNING.
3. USE THE FIRST CARD TO START YOUR FIRST CATEGORY.
4. PICK UP THE NEXT VERB CARD.
5. CONSIDER THE VERB IN TERMS OF YOUR VIEWS FOR FACILITATING LEARNING IN RELATION TO THE VERBS ALREADY SORTED.
6. DECIDE WHETHER YOUR IDEA FOR THIS VERB IS SIMILAR TO ONES ALREADY SORTED.

   IF YES, THEN PUT THE CARD IN THE CATEGORY GROUP YOU THINK IT BELONGS.

   IF NOT, THEN USE THE CARD TO START A NEW CATEGORY GROUP.
7. PICK UP THE NEXT CARD AND REPEAT THE PROCESS.

IF YOU ARE OUT OF CARDS:

1. DECIDE WHETHER YOUR CATEGORY GROUPS ARE A SATISFACTORY REPRESENTATION OF YOUR SORTING.
2. IF NOT, MAKE CHANGES BY SWITCHING CARDS.
3. PUT A BLANK CARD ON TOP OF EACH CATEGORY AND WRITE A TITLE FOR THAT GROUP ON THE BLANK CARD.

INSTRUCTIONS

THE SORTING OF STATEMENTS INTO CATEGORIES

PURPOSE: THE PURPOSE OF THIS TASK IS TO STUDY YOUR VIEWS ON WHAT IT TAKES FOR A WOMAN TO MAKE IT TO THE TOP HER CLASS. THE STATEMENTS MAY APPLY TO WORK, FAMILY, AND SCHOOL SETTINGS.
YOUR TASK IS TO MAKE A JUDGMENTAL DECISION FOR EACH STATEMENT AND CREATE THESE PILES THAT WILL POSSESS A COMMON CHARACTERISTIC.

TASK DESCRIPTION:

1. START WITH THE FIRST STATEMENT CARD.

2. CONSIDER THE STATEMENT IN TERMS OF YOUR VIEWS ABOUT WHAT IT TAKES FOR A WOMAN TO CONSIDER HERSELF SUCCESSFUL.

3. USE THE FIRST CARD TO START YOUR FIRST CATEGORY.

4. PICK UP THE NEXT STATEMENT CARD.

5. CONSIDER THE STATEMENT IN TERMS OF YOUR VIEWS ABOUT WHAT IT TAKES FOR A WOMAN TO BE TOP OF HER CLASS.(SAME IDEA AS IN 2)

6. DECIDE WHETHER THE CHARACTERISTIC OF THIS NEW CARD STATEMENT IS SIMILAR TO CARD STATEMENTS ALREADY SORTED.

   IF YES, THEN PUT THE CARD IN THE CATEGORY GROUP YOU THINK IT BELONGS TO.

   IF NOT, THEN USE THE CARD TO START A NEW CATEGORY GROUP.

7. PICK UP THE NEXT CARD AND REPEAT THE PROCESS

IF YOU ARE OUT OF CARDS:

1. DECIDE WHETHER YOUR CATEGORY ARE A SATISFACTORY REPRESENTATION OF YOUR SORTING.

2. IF NOT, MAKE CHANGES BY SWITCHING CARDS.

3. PUT A BLANK CARD ON TOP OF EACH CATEGORY AND WRITE A TENTATIVE TITLE FOR THE GROUP ON THE BLANK CARD.

APPENDIX C

SAMPLE OF CASES FROM DATA SETS TO GENERATE JOINT PROPORTION MATRICES

Data set for joint occurrence matrix for task I

```
0110000000000000000000010000010110
0120000000000100000000000001001000
0130000010000001100000000000100000
0141100011100000000000000000000000
0150000000000000100110010000000000
0160010000011000010000000000000000
0170001000000000001001001100000001

0210011001110010000011010011010001
0220000000000000001000010000000000
0231000100000000000000000000000100
0240000000000000100000000000000000
0250100010001000001100001100100100
0260000000000101000000000000000000

0310010000000111111101100100000000
0321101111111000000001001101111111

0411000100010000000000000000000000
0420000000000000000100000100000000
0430000000000010000000000000111111
0440000000000000001001110100000000
0450000000000000000000000011000000
0460111011100000000000000000000000
0470000000000100000000000000000000
0480000000000001100000000000000000
0490000000000000001000010000000000

0510011000000110000000000001111110
0521100111111001111111111110000001

0611101111110110000000000000111111
0620010000001000110111100000000000
0630000000000000000100001111000000

0710000000000000000000000011101000
0720000000000001101111111010001011
0731000011101001000000000000000001
0740111100010000000000000100000000

0810000000000000000000000011000000
0820000000000100011111111100000000
0830000000000011100000000000111111
0841111111111000000000000000000000
```

Data set for joint occurrence matrices for task II

```
011000111000000000001010001010000000
012101000100010001000000000000001001
013000000010001100000000000000010100
014000000000100000000000010000000000
015010000000100001010101010101000010

021010000000100001010101010100000010
022101111110111101010101010111111101

031100111100111000010101000100000000
032010000000100001010101010101011010
033001000010000100000000000010000100
034000000000000000100000000000000001

041101000010011001000000100000000001
042010000000100001010101010101011010
043000101100000010001010001000000100
044000001000010000000000000000100000

051100101100001000010000000000101000
052001000000001010000000000000010010000
053010000000000001010101010101000010
054000001000010000000000101000000000
055000000001100000100000000000000101

061010000000100001000101010101011000
062000001110011100000000001000100000
063100100000000000111000000000000011
064001000010000100000100010000100

071010000000000000101010101000000000
072100010000000100000000000000100001
073000000000000000000000010000001000
074001001011000100010100010000110
075000100100110001000000000001010000

081001000011010100010100010011100
082100111100101000100000110010000 1
083010000000000001000101000100000 0
084000000000000000100000000000000010

091010000000100001010001010100000010
092001000000000010000000010000000100
093000101100001000010000010001000
094100010000001000100100000000000000
095000000000010000000000100000110001
```