5-2000

A normative subtest taxonomy developed from the universal nonverbal intellegence test : implications and applications

Brian Eugene Wilhoit

To the Graduate Council:

I am submitting herewith a dissertation written by Brian Eugene Wilhoit entitled "A normative subtest taxonomy developed from the universal nonverbal intellegence test : implications and applications." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Education.

R. Steve McCallum, Major Professor

We have read this dissertation and recommend its acceptance:

Donald Dessart, Donald Dickinson, Thomas George

Accepted for the Council:

Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

To the Graduate Council:

I am submitting herewith a dissertation written by Brian Eugene Wilhoit entitled " A Normative Subtest Taxonomy Developed from the Universal Nonverbal Intelligence Test: Implications and Applications." I have examined the final copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Education I.

R Steve McCallum

R. Steve McCallum, Major Professor

We have read this dissertation
and recommend its acceptance:

Donald J. Dennis

Donald J. Dickinson

Thomas W. George

Accepted for the Council:

Cuminkel

Associate Vice Chancellor and
Dean of The Graduate School

A NORMATIVE SUBTEST TAXONOMY DEVELOPED FROM THE UNIVERSAL NONVERBAL INTELLIGENCE TEST: IMPLICATIONS AND APPLICATIONS

A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee

Brian Eugene Wilhoit
May, 2000

DEDICATION

This dissertation is dedicated to mom and dad for instilling the desire in me to learn and excel, and to Martha for her steadfast support and encouragement.

ACKNOWLEDGMENTS

There are many people who provided support and guided this research from genesis to end. First and foremost, R. Steve McCallum, who not only chaired my committee, but also co-authored the UNIT with Bruce Bracken, helped me wade through the minefield that is dissertation research. Donald J. Dickinson, Donald J. Dessart, and Thomas George, my other committee members, have all contributed to my graduate education immensely, as well as maintaining a sense of calm throughout the dissertation process. Also deserving a great deal of acknowledgement, Riverside Publishing, Inc., for allowing me to use the UNIT standardization data; and in particular, John Wasserman, the UNIT project director, and Kirk Becker, who provided initial statistical guidance. Also providing statistical support through tedious analyses was Mike O'Neil, a statistical consultant with the University of Tennessee. Along statistical analyses lines, Paul McDermott, professor at the University of Pennsylania, provided much of the SAS code necessary for the clustering analyses. Finally, the secretarial staff of the Educational Psychology department, April Phillips, Becky Bledsoe, and Betsy Johnson provided support and guidance through the completion of countless forms and procedures necessary for progression through the program.

ABSTRACT

The purposes of this research were threefold. The first goal was to develop and apply a normative typology using multivariate profile analysis of subtest scores of the Universal Nonverbal Intelligence Test (Bracken & McCallum, 1998; UNIT), taken from the national standardization sample; the second goal was to develop and apply a typology using multivariate subtest profile analysis of a subsample of identified children with Learning Disabilities from the national standardization sample; and the third goal was to provide practitioners a description of user friendly strategies needed to compare meaningful subtest profiles of individual examinees with commonly occuring normative profiles, as identified in goal one above.

To accomplish these goals, multistage cluster analyses were applied to the standardization sample for the UNIT Extended Battery, comprised of all six subtests; the UNIT Standard Battery, comprised of four subtests; and the UNIT Learning Disabled subsample, comprised of the four Standard Battery subtests. The results of these analyses yielded a seven profile cluster solution for the Extended Battery, a six profile cluster solution for the Standard Battery, and a four profile cluster solution for the Learning Disabled subsample. The psychometric properties of the respective analyses were impressive with extremely tight profile clusters that were separated extremely well from each other.

Demographic prevalence trends of the resultant clusters are similar to other studies, but help to describe the cluster composition. Additionally, the results lend support to the UNIT's underlying factor structure. To fulfill the third goal of this research, user-friendly methods of determining whether or not clinical profiles are unique when compared to the profiles identified in the standardization sample are discussed.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1

Introduction

The purposes of this research are threefold. The first goal is to develop and apply a normative typology using multivariate profile analysis of subtest scores of the Universal Nonverbal Intelligence Test (Bracken & McCallum, 1998; UNIT), taken from the national standardization sample; the second goal is to develop and apply a typology using multivariate subtest profile analysis of a subsample of identified children with Learning Disabilities from the national standardization sample; and the third goal is to provide practitioners a description of user friendly strategies needed to compare meaningful subtest profiles of individual examinees with commonly occurring normative profiles, as identified in goal one above.

A profile is the overall pattern of subtest scores on a particular test for a given individual. The technical elements of a profile consist of the shape, elevation, and scatter as set forth originally by Lee Cronbach and Goldine Gleser in 1953. According to Harvey Skinner (1978), "the shape is the actual pattern of ups and downs across variables in a profile, scatter describes how dispersed scores are from the average, while elevation is the mean score of the individual over all attribute measures in the profile" (p.297).

Profile analyses of intelligence tests have taken several directions since first being advocated by Wechsler (1949). There are two primary strategies. The first approach is intraindividual, and requires examination of the pattern of subtest or scale scores produced on an intelligence test to determine the relative strengths and weaknesses of a given individual. A second interpretive strategy, an interindividual approach, examines a profile pattern produced on a given scale for a given individual and compares it to the patterns of scores on the same scales from other individuals.

It is possible to use both approaches to profile analysis in a two-step process that includes both an interindividual (normative) and an intraindividual (ipsative) evaluation. The normative evaluation is characterized as an inter-individual profile analysis because the comparisons are between different persons, while the ipsative evaluation can be characterized as an intra-individual profile analysis since the comparisons are between the different subtest scores for an individual. The traditional normative analysis requires univariate statistical methods to compare the individual with the norm group, while the ipsative analysis examines the pattern of strengths and weaknesses of a given individual on that particular test.

Inter-Individual Profile Analysis

Inter-individual profile analysis is carried out using normative scores to aid in comparing how well one person performs in relation to another person. According to Cattell (1944), normative scores are obtained when " the subjects (persons) are placed in order relative to one another and assigned a standard score..." (p.293). It is at this normative level that some researchers and practitioners feel interpretation should stop (McDermott, Fantuzzo, and Glutting, 1990; McDermott, Fantuzzo, Glutting, Watkins, and Baggaley, 1992). All the major published intelligence tests' manuals report norming procedures to allow the among-individual comparisons that practitioners use. Usually, however, instead of examining the overall normative score profile through a multivariate analysis, traditionally practitioners have sequentially analyzed the scaled and standardized scores of the individual subtests, scales, and full-scale scores of intelligence tests. This sequential process is described as "univariate" because of the sequential comparisons designed to determine whether scores are different enough from the norming group to be statistically significant and/or exceptional (i.e., rare in the population).

The univariate analysis begins by examining the statistical significance, or lack thereof, between the obtained scaled score(s) and the calculated, tabled value(s) listed in the manual of the particular test being used. As suggested above, the comparisons are pairwise and sequential. Most comparisons use a significance level of $p < .05$ to rule out chance occurrence, but p values ranging from .01 to .15 have been commonly used (Kaufman, 1994; Sattler, 1988). Once the statistical significance comparisons between obtained global scores and the standardization sample is complete, the second univariate step of examining base rates begins.

Pairwise comparisons are considered next. By comparing a particular individual's subtest discrepancy scores to base rates, the evaluator can make a determination as to how common, or frequent, a pairwise subtest discrepancy is. Generally, as Kaufman (1994) advocates, this procedure begins by subtracting the lowest subtest score from the highest subtest score on the entire instrument or on the separate global scales (depending on the test). Once that discrepancy is obtained, it can be compared to the cumulative percentages for that particular discrepancy found in the standardization sample. If the discrepancy is large enough, it may then be determined to occur infrequently and be therefore interpretable (Kaufman, 1994). Both of these normative procedures are univariate, however, and have been shown to have statistical limitations due to the multiple pairwise, sequential comparisons, which may lead to the overidentification of strengths and weaknesses (Glutting, McDermott, Watkins, Kush, & Konold, 1997). Following this type of normative analysis, an ipsative or intra-individual, analysis of the subtest scores can occur.

Intra-Individual Profile Analysis

Intra-individual profile analysis is best known as ipsative interpretation of a psychological measure, usually a cognitive or personality scale. Raymond B. Cattell (1944) first defined ipsative scores as " scale units relative to other measurements on the person himself" (p. 294). Since the term's inception in 1944, over 200 articles have appeared in the psychological literature dealing with ipsative measures.

Typically, intelligence tests are composed of several subtests that yield raw scores for the individuals being assessed. These raw scores are then standardized through a derivation formula, then transformed to yield familiar standard scores. In order to ipsatize these scores then, one simply sums all the subtest scores for a given individual, takes the average of these scores, and compares the relative differences between the separate subtest scores with the calculated average of all the subtest scores (for that given individual) on a particular test battery, or scale within a test battery.

Ipsative, or intra-individual, profile analysis has been used extensively in the clinical assessment of intelligence. It is still taught as a viable, important approach to interpretation, and it is advocated by several influential texts in the field (Kaufman, 1979; Kaufman, 1994; Sattler, 1988). It continues to be a method of choice for several reasons. See Table 1 for an example of how this method is commonly used to identify strengths and weaknesses for a widely used test, the Wechsler Intelligence Scale for Children - 3rd Edition (Wechsler, 1991, WISC-III).

This intra-individual approach to profile analysis allows the clinician to concentrate on the individual's pattern of strengths and weaknesses while keeping the normative inter-individual comparisons to a minimum (Zachary, 1990).

Table 1.

Traditional Ipsative Approach to Subtest Profile Analysis with the WISC-III

Subtest[a]	Scaled Score	Difference[b]	Value for Significance[c]	Frequency[d]
PC	11	0.3	3.81	>25%
I	12	0.7	3.32	>25%
CD	10	1.3	3.77	>25%
S	9	2.3	3.48	25%
PA	9	2.3	3.91	>25%
A	11	0.3	3.74	>25%
BD	**16**	**4.7**	**3.04**	**<10%**
V	9	2.3	2.97	25%
OA	12	0.7	4.37	>25%
C	14	2.7	3.84	25%

[a] PC, Picture Completion; I, Information; CD, Coding; S, Similarities; PA, Picture Arrangement; A, Arithmetic; BD, Block Design; V, Vocabulary; OA, Object Assembly; C, Comprehension

[b] Absolute difference from the Scaled Score M, 11.3 (based on the 10 reported subtest scores)

[c] Value needed for statistical significance; values reported in Table B.3, WISC-III Manual (Wechsler, 1991); values in **Bold Typeface** are statistically significant

[d] Cumulative percentages of occurrence in standardization sample; values reported in Table B.3, WISC-III Manual (Wechsler, 1991)

Presumably, practitioners can use this information to design interventions that capitalize on the individual's strengths. In addition, some practitioners use this strategy to identify problem areas for remediating. For example, Table 1 shows that the abilities assessed by the Block Design subtest are considered strengths, and can be identified. Knowledge of these strengths may be used to help plan interventions.

Clinical hypotheses can be generated from this ipsative approach. These hypotheses can be used to guide diagnosis as well as intervention decisions (Zachary, 1990). The individual's strengths and weaknesses, as measured by the given intelligence scale, can be used conjointly with other assessment data in this decision-making process. In addition, a full assessment should include other measures of an individual's intellectual ability allowing the practitioner to use the data to triangulate the most accurate diagnosis.

While ipsative profile analysis is often used in practice, it is not without its problems. According to some, there is little empirical evidence to support its use (Hale, 1979; Hale & Landino, 1981; Hale & Saxe, 1983; Zachary, 1990). For example, according to McDermott, Fantuzzo, Glutting, Watkins, and Baggaley (1990), ipsatizing the subtest scores on the Wechsler Intelligence Scale for Children - Revised (WISC-R) automatically removes all the common variance associated with Spearman's g and almost 60% of the overall test's reliable variance. In essence then, by ipsatizing the scores in the traditional manner, McDermott et al. (1990) argue that the reliability of the scale is diminished drastically. Kaufman (1994) counters convincingly that the normative scaled scores are not changed in an ipsative analysis; only the profile mean is shifted. By doing so, no detrimental normative effects occur (e.g., there is no loss of systematic variability).

McDermott et al. (1990) examined the predictive validity associated with the WISC-R subtest scores both normatively and ipsatively. They found that normative test scores were able to account for nearly 40% of academic performance while ipsatized scores could account for less than 10% of academic performance. Therefore, they concluded that ipsatized scores were much less valid in predicting academic outcomes from the WISC-R.

Another drawback to using the ipsative approach to profile analysis according to McDermott et al. (1990) is that the practitioner is unable to compare one person's ipsatized profile to another's profile. Cattell's (1944) definition of ipsative scores makes this point. Ipsative scores can be compared only to other scores that a given individual receives, while normative scores may be compared between individuals. Although, McDermott et.al. (1990) argue that this is a weakness, Kaufman (1994) suggests that this is a strength. He points out that ipsative analyses are meant to supplement the group comparisons, not replace them (Kaufman, 1994).

Intra-individual profile analysis of intelligence scales has been supported over the years by several researchers (Delaney & Hopkins, 1987;

Kaufman, 1979, 1990, 1994; Kaufman & Kaufman, 1983; Sattler, 1988, 1992; Wechsler, 1974, 1991) while concurrently being discouraged by others (McDermott, Fantuzzo, and Glutting, 1990; McDermott, Fantuzzo, Glutting, Watkins, and Baggaley, 1992). With personality measures, this type of profile analysis is carried out quite commonly and is generally accepted as proper interpretation. However, because of the problems associated with intra-individual profile analysis, McDermott et.al.(1990; 1992) argue for another approach to subtest interpretation which makes use of a numerical taxonomy that utilizes a particular multivariate statistical approach.

Numerical taxonomy is the epistemological discipline that refers to the systematic distinguishing, ordering, and naming of types based on the analysis of numerical data using classification rules or models and has been referred to by Adams (1985) as "phenetic taxonomy". In essence, the objective of the (pheneticist) is to objectify and quantify the taxonomy by the use of empirical models to make reliable classifications of groups based on some form of measurement (Adams, 1985). Sneath and Sokal (1973) are credited as being the pioneers in the area of numerical taxonomy, especially in biological classification. In their work, Sneath and Sokal (1973) maintain that numerical taxonomy centers on the identification of latent relationships between objects at a given point in time. In order to accomplish these goals, a statistical approach known as cluster analysis has been relied upon heavily.

Cluster analysis refers to a set of multivariate statistical procedures, or algorithms, that allow a researcher to group subjects according to data gathered on several variables that describe the subjects. The resultant groups are composed of group members that are maximally similar to each other; concurrently then, each separate group is maximally dissimilar to each other. This classification or typological function of cluster analysis is by far the most commonly used by researchers (Aldenderfer & Blashfield, 1984).

In order to understand cluster analysis, it can be compared and contrasted to factor analysis. Most researchers are more familiar with factor analysis than cluster analysis; factor analysis can be described as a set of statistical procedures that produce groups of variables, called factors, based on data produced on the variables by the subjects. Factor analysis is a data reduction technique where the redundancy is removed from a set of correlated variables to produce a much smaller number of derived variables called factors (Kachigan, 1991). Whereas factor analysis reduces a large number of variables being measured into smaller groups of variables called factors, cluster analysis reduces a large number of subjects or objects into smaller groups of identifiable clusters based on their variable scores. Cattell (1944) called factor analysis an R-technique and cluster analysis a Q-technique.

Recently there has been a rapid increase in the use of cluster analysis, for two reasons (Aldenderfer & Blashfield, 1984). First, the advent of high speed computers has made the enormous data handling quite manageable;

second, the importance of classification as a scientific procedure, especially in the social sciences, has become more focal to research goals.

Aldenderfer and Blashfield (1984, p.12) propose that all solidly developed cluster analytic studies should be composed of the five following steps; a) select a sample that is to be clustered; b) define a set of variables that represent a measure of the entities, or subjects, in the sample; c) compute a measure of the similarities among the subjects; d) use a cluster analysis method to create groups of similar subjects; and e) validate the resulting cluster solution.

These are the basic steps of which every cluster analytic study should consist. While step one is fairly straightforward, steps two, three, four, and five will be discussed below.

After selecting a sample of subjects, a decision must be made as to what variables will be used to measure those subjects. For instance, researchers who are interested in personality types of a sample of alcoholics might choose to use the scales on a personality test such as the Minnesota Multiphasic Personality Inventory (Hathaway & McKinley, 1940; 1951) as the measurement variables as Goldstein and Linden (1969) did in their research. Those interested in the profiles of intelligence acquired by a sample of learning disabled children might choose the particular subtests on an intelligence test as measurement variables (Rourke, 1985). Once measurement variables have been selected, a similarities matrix must be computed among the subjects. A similarities matrix is an arrangement of similarity coefficients between cases in a row by column format often referred to as an N x N matrix, where N is a case.

There are several different computational methods for arriving at a similarities matrix. In the social sciences, the most common methods use either correlations or the distance in score units between subjects on the measurement variables. Each method has its associated strengths and weaknesses, depending on the data under consideration.

Profiles can be compared for similarity through shape, the pattern of low and high scores; elevation, the mean value of the case over all the variables; and scatter, the dispersion of scores about the mean (Cronbach & Gleser, 1953; Aldenderfer & Blashfield, 1984). With these characteristics in mind, the choice of a similarity measure becomes more apparent.

By using a correlation coefficient as the measure of similarity (a shape measurement), the researcher chooses a measure that is insensitive to the magnitude of the differences of the variables being compared. However, when dispersion and magnitude differences are not of critical importance, the correlation coefficient may become the similarity measure of choice (Aldenderfer & Blashfield, 1984). Additionally, of the similarity measures, the correlation coefficient is one of the easiest to calculate. The other most commonly used similarity measure is distance. As Aldenderfer and Blashfield (1984) point out, however, distance measures are more of a comparison of dissimilarity than similarity. That is, when the comparison of two variables is equal to zero, the variables are said to be the same, but as the comparison value increases, the variables become more dissimilar.

The most popular distance measure is Euclidean distance. In essence, it is literally the distance, in score units, between two points, and therein lies its appeal. Another multivariate distance measure is Mahalanobis D^2, or generalized distance (Mahalanobis, 1936). It is different from Euclidean distance in that it incorporates the variance-covariance matrix in its computational steps. When the correlation between variables is equal to zero, Mahalanobis D^2 and Euclidean distance are the same value (Aldenderfer & Blashfield, 1984).

The distance measures also have problems. Primarily, and intuitively, case similarities are affected by large or small variable magnitudes and amount of dispersion. That is, variables with large size and dispersion differences can mask variables with small size and dispersion differences (Aldenderfer & Blashfield, 1984). When using distance measures, most often researchers will first standardize the variables to overcome this problem.

In addition, Mahalanobis D^2 has historically been problematic in that it is tedious to compute (Mezzich & Solomon, 1980). However, with the advent of more powerful and faster computers, this problem has been somewhat alleviated. After selecting a similarity measure the next decision point then becomes the clustering method.

Clustering Methods

There are many clustering methods. Only the three major classes of clustering techniques used in the social and behavioral sciences will be discussed here (Aldenderfer & Blashfield, 1984). (For a more detailed description of the myriad of clustering techniques readers are referred to Aldenderfer and Blashfield, 1984; Anderberg, 1973; Everitt, 1980; Hartigan, 1975; Mezzich and Solomon, 1980; and Sneath and Sokal, 1973). The three classes of clustering techniques that are of primary importance to social and behavioral sciences are: (1) Hierarchical agglomerative, (2) Iterative partitioning, and (3) Factor analytic.

Hierarchical Agglomerative Methods

According to Aldenderfer and Blashfield (1984), there are four important points that distinguish hierarchical agglomerative clustering techniques. First, the methods in this class require a search of a similarity matrix and sequentially merge those cases most similar. Second, the merger of clusters can be visualized using a tree diagram. Third, at the first step of clustering, all cases are viewed as independent clusters; at the last step, all cases have been merged into one inclusive cluster. Lastly, the methods in this class are easy to conceptualize. Hierarchical agglomerative methods have some limitations. They require the calculation and storage of a huge similarity matrix. This requires a great deal of computer power and memory space. Second, these methods make only one pass through the data. Third, different solutions may be obtained by simply reordering the data in the similarity matrix. The four most popular clustering techniques in this class are: single linkage, average linkage, complete linkage, and Ward's method.

In single linkage, cases are joined to clusters if the case is similar to at least one case already in the cluster. Average linkage (Sokal & Michener, 1958) joins cases to clusters if the case is similar to the average of the cases already in the cluster. Complete linkage (Sokal & Michener, 1958) joins cases to clusters if the case is within a certain level of similarity to all the other cases in the cluster. Finally, Ward's method (1963) joins cases or groups of cases that result in the minimum increase in variance as represented by the error sum of squares ($ESS = x_i^2 - 1/n(Sx_i)^2$, where x_i is the score of the i^{th} case). When clustering begins with Ward's method, the ESS is equal to zero and as cases are combined the ESS is recalculated and used as the decision rule.

All of the discussed hierarchical agglomerative methods must be used in conjunction with some stopping rules in order to determine the point at which the data yield the most useful and relevant clusters. Obviously, one large all-inclusive cluster would not be any more useful than having each case in a sample represent separate clusters. Several stopping rules, or decision points, have been determined and tested for use with these methods (see Mojena, 1977; Mojena and Wishart, 1980; Wishart, 1982)

Iterative Partitioning Methods

Iterative partitioning methods allow the researcher to decide initially the number of clusters desired and where to partition the data (Anderberg, 1973). After the initial partition, each data point is assigned to the cluster having the nearest centroid. Once a data point is assigned, the cluster centroids are recomputed. The algorithms for iterative partitioning methods are designed to make a complete pass through the data each time a data point is assigned to a cluster. The algorithm repeats until no data points can be reassigned.

These methods have some important strengths over the hierarchical agglomerative methods. First, iterative partitioning methods can handle much larger data sets. Second, these methods make numerous passes through the data thereby compensating for a poor initial partition. Finally, these methods do not produce clusters within clusters, nor do the clusters overlap.

Although these methods have important strengths, they also suffer from one major limitation. The most optimal solution would require the data to be partitioned in all possible configurations. According to Aldenderfer and Blashfield (1984), an example with 15 cases and 3 clusters would require the examination of 217,945,728,000 unique partitions. This would be practically and computationally impossible, even for a data set as small as the example.

There are several iterative partitioning methods, but the most commonly used method in the social sciences is known as 'k-means' (MacQueen, 1967). In this procedure the researcher specifies k starting points as initial estimates of group centroids. Cases are then assigned to a cluster with the nearest centroid, and the centroids are recomputed after the addition of each case (Mezzich & Solomon, 1980).

Factor Analysis Methods

Variants of factor analysis have been used considerably in the field of psychology. Clustering methods based on factor analysis are commonly known as inverse factor analysis or Q-type factoring (Aldenderfer & Blashfield, 1984). Whereas traditional factor analysis groups variables based on the factor loadings, the clustering variants group cases based on the factor loadings. As in traditional factor analysis, clustering variants begin with a correlation matrix between entities as a measure of similarity from which the factors are extracted.

Clustering variants are not without their faults, however. These clustering variants have been cited as faulty for the use of a linear model across cases, the possible generation of multiple factor loadings, and the double centering of data (Aldenderfer & Blashfield, 1984; Everitt, 1980; Fleiss et al., 1971).

Once a clustering method is chosen and applied, the next step in a cluster analytic study is to validate the resultant solution. The three most commonly used validation techniques are; a) replication; b) significance tests on external variables; and c) tests of homogeneity.

Replication is straightforward and conceptually simple. The idea is to split the sample randomly into a derivation group and a validation group, apply clustering techniques to the derivation group to determine clusters, and finally, to apply the same clustering techniques to the validation group. Successful validation would result when clusters in the validation sample are compared to and found similar to those in the derivation sample. A word of caution is necessary, however. "The failure of a cluster solution to replicate is a reason for rejecting the solution, but a successful replication does not guarantee the validity of the solution" (Aldenderfer & Blashfield, 1984, p.65).

The second commonly used validation technique allows significance testing of external variables, i.e., the technique performs significance tests and compares the resultant clusters with variables not used in the original clustering procedures. For example, it is possible to compare a particular clustering solution with external demographic data or results obtained on other measures. The strength of this procedure is that the clusters are compared directly with relevant data (Aldenderfer & Blashfield, 1984).

The third most commonly used validation technique is the use of internal tests of homogeneity. Tryon and Bailey (1970) advocate the use of the

H statistic. The H statistic compares the variance for a particular cluster on a variable with the variance for that same variable across the entire data set (Blashfield & Aldenderfer, 1988). Another test of homogeneity used in this type of research has been Cattell's (1949) r_p statistic. Cattell's r_p compares the average similarity between the resultant clusters.

Other less commonly used validation techniques include the cophenetic correlation, significance tests on the variables used in the creation of clusters, and Monte Carlo procedures. Readers are referred to Aldenderfer and Blashfield (1984), Skinner and Blashfield (1982), and Blashfield and Aldenderfer (1988) for more information if interested in these less common techniques.

Application of Clustering Methods

The recent development of this plethora of clustering methods has led to an increase in taxonomic and typological research focusing on the study of intelligence. Various combinations of these clustering techniques have been used in the development of normative profile typologies of several individually administered intelligence tests. Typologies have been developed and described for the Wechsler Intelligence Scale for Children - Revised (Wechsler, 1974) by McDermott, Glutting, Jones, Watkins, and Kush (1989), the Wechsler Adult Intelligence Scale - Revised (Wechsler, 1981) by McDermott, Glutting, Jones, and Noonan (1989) and by Schinka and Vanderploeg (1997), the Wechsler Preschool and Primary Scale of Intelligence

(Wechsler, 1967) by Glutting and McDermott (1990a), the McCarthy Scales of Children's Abilities (McCarthy, 1972) by Glutting and McDermott (1990b), the Kaufman Assessment Battery for Children (Kaufman & Kaufman, 1983) by Glutting, McGrath, Kamphaus, and McDermott (1992), the Differential Ability Scales (Elliot, 1990) by Holland and McDermott (1996), and the Wechsler Intelligence Scale for Children - Third Edition (Wechsler, 1991) by Glutting, McDermott, and Konold (1997) and by Donders (1996). All of these typologies were developed and based on the national standardization samples for each respective instrument. Several of these authors have developed procedures to allow practitioners to compare a particular examinee's subtest profile to the "common" profiles of the respective typologies.

Additionally, cluster analytic procedures have been used in intelligence research that examines the incidence of profile types based on individually administered intelligence instruments for persons identified as learning disabled (Maller & McDermott, 1997; Shapiro, Buckhalt, & Herrod, 1995); persons who experienced traumatic-head-injury (Donders & Warschausky, 1997) and closed-head-injury (Crawford, Garthwaite, Johnson, Mychalkiw, & Moore, 1997); and children diagnosed as conduct disordered (Christian, Frick, Hill, Tyler, & Frazer, 1997). For example, DeLuca, Rourke, and Del Dotto (1991) applied cluster analytic procedures to the scaled scores of 4,000 students labeled learning disabled in arithmetic to determine possible subtypes. They used the scaled scores obtained on the Wechsler Intelligence Scale for

Children - Revised (WISC-R; Wechsler, 1974), the Wide Range Achievement Test (WRAT; Jastak & Jastak, 1978), the Personality Inventory for Children (PIC; Lachar & Gdowski, 1979), the Behavior Problem Checklist (Quay & Peterson, 1979), and the Activity Rating Scale (Werry, 1968) as clustering variables. A two-stage clustering procedure was used that incorporated a hierarchical procedure in the first stage and an iterative partitioning procedure in the final stage. Based on their analysis, they were able to determine four subtypes of arithmetic disabled students that needed distinctly different treatment regimens.

Statement of the Problem

It can be seen, then, that normative typologies based on clustering techniques can be both illuminating and useful when comparing individuals based on standardized testing results. These typologies have been generated using the most common intelligence and personality instruments. Practitioners now have a choice. They can use the older inter- and intra-individual interpretive strategies, or they can use the multivariate cluster-based interpretive strategies. However, practitioners using new multidimensional nonverbal tests of intelligence, as well as the newer special-focus tests, do not have the multivariate based analyses available. My research aims to apply the aforementioned clustering techniques to a new multidimensional nonverbal intelligence test, the Universal Nonverbal Intelligence Test (UNIT; Bracken & McCallum, 1998), for two populations. The first sample is representative of all school-age children in the United States; the second is representative of learning disabled (LD) children in the United States.

The UNIT (Bracken & McCallum, 1998) is a measure of general intelligence and cognitive abilities. It is designed to be used with children and adolescents between the ages of 5 years and 17 years who may be "disadvantaged by traditional verbal and language-loaded" (p.1) measures of intelligence. The UNIT measures a broad range of memory and reasoning abilities, including processes of verbal (symbolic) mediation and processes that do not rely on mediation (nonsymbolic). The UNIT memory subtests "measure recall of content, location, and sequence" (p.1), while the reasoning subtests "measure pattern processing, problem solving, understanding of relationships, and planning abilities" (p.1).

The UNIT is administered and completed without the use of language, making it a truly nonverbal test of intelligence. The UNIT consists of six subtests, Symbolic Memory, Spatial Memory, Object Memory, Cube Design, Analogic Reasoning, and Mazes. All of the subtests yield standard scores with a mean of 10 and a standard deviation of 3. The standard scores on these subtests can be combined to yield the following five scale quotients: Memory Quotient, Reasoning Quotient, Symbolic Quotient, Nonsymbolic Quotient, and Full Scale Intelligence Quotient. All of these scales use standard scores with a mean of 100 and a standard deviation of 15. Using the UNIT data from the standardization sample, two questions were developed.

Research Questions

1. What is the normative typology of the Universal Nonverbal Intelligence Test based on the obtained scaled scores of the six subtests, i.e., what are the "common" profiles and how many "common" profiles are available.
2. Are there distinctive normative subtypes of Learning Disabled students based on the UNIT standardization?

CHAPTER 2

Methods

Participants

The entire standardization sample of the Universal Nonverbal Intelligence Test (UNIT) was used. The sample included 2,100 children from age 5 years, 0 months through age 17 years, 11 months, 30 days. The sample was stratified according to the 1995 U.S. census data and incorporated current data from the U.S. Department of Education to target children from all major U.S. geographic regions, socioeconomic levels, ethnicities and races, community settings, classroom placements, special education services and parental education attainment. Representative samples of exceptional children such as learning disabled, deaf/hearing impaired, mentally retarded, and children learning English as a Second Language were included. The sample was divided into 12 groups of 175 children, categorized in age ranges of 1 year for all ages except for the 16 and 17 year olds; they were combined into one group. Because the sample was relatively equally distributed across months within years, norms were generated with four-month intervals.

Instrument

The UNIT is a technically sound measure of intelligence. Many reliability and validity studies were undertaken during its development.

In terms of reliability, internal consistency was measured using split-half correlations with Spearman-Brown corrections. Reliabilities of the four subscales ranged from .87 for the Symbolic Quotient to .91 for the Nonsymbolic Quotient with a .93 for the UNIT Full Scale scores. The reliabilities of the individual subtests ranged from .64 on Mazes to .91 on Cube Design with the majority of coefficients above .75. Additionally, reliabilities were examined over time with test-retest measures.

Test-retest reliabilities of the four subscales ranged from .78 on the Symbolic Quotient to .87 on the Reasoning Quotient with a .88 for the Full Scale scores. The reliabilities of the individual subtests ranged from .58 on Mazes to .85 on Cube Design. All reliability coefficients reported here were corrected for restriction or expansion in range; additionally, all composite coefficients were corrected for reliability of linear combinations (Bracken & McCallum, 1998).

Both internal and external evidence of validity for the UNIT was examined. Procedures were used to determine content validity and structural validity that included examinations of unidimensionality, intercorrelations, and both exploratory and confirmatory factor analysis.

Unidimensionality was achieved by selecting items with comparable item characteristic curves as measured by the Rasch model. Those original items with item characteristic curves not meeting best fit criteria were discarded. Intercorrelations among the four scales of the Standard and Extended batteries were consistently high with coefficients above .90. Confirmatory factor analytic conclusions consistently found evidence of four factors corresponding to the four global UNIT scales. Therefore, ample evidence of internal validity was shown. Evidence of external validity was also examined.

Correlational studies provided evidence of both concurrent and predictive validities for the UNIT. Correlations with the composite scores of Wechsler Intelligence Scale for Children - Third Edition, Woodcock-Johnson Tests of Cognitive Ability - Revised, Kaufman Brief Intelligence Test, and the Matrix Analogies Test yielded concurrent coefficients above .81. Additionally, the UNIT was compared with Raven's Standard Progressive Matrices and the Test of Nonverbal Intelligence - Second Edition yielding coefficients of .56 and .63 respectively.

The UNIT's predictive validity was also examined with comparisons between the UNIT Full Scale score and the Woodcock-Johnson Tests of Achievement - Revised Broad Scores (Reading, Mathematics, Knowledge, and Skills). The UNIT yielded correlations above .78 with all the achievement scores. Comparisons between the UNIT Full Scale score and the Wechsler Individual Achievement Test yielded correlations ranging from .44 to .74.

Additional validity studies with subsamples of the standardization sample who were diagnosed with an exceptionality (speech and language impaired, learning disabled, mentally retarded, intellectually gifted, and serious emotional disturbance). The external validity studies were used to aid in verifying the UNIT's technical validity properties.

Procedure

Cluster analysis procedures were used to sort the children's subtest profiles according to shape and level. The majority of procedures used for the initial part of this study will replicate those used by Glutting, McDermott, and Konold (1997); then procedures were used to develop a normative subtest taxonomy for the UNIT standardization sample.

The second part of the study used cluster analysis procedures to develop a subtest taxonomy for the subsample of learning disabled students taken from the overall standardization sample. Procedures used in the second part of this study replicated those in the initial part to ensure continuity.

The clustering strategy consisted of three stages, beginning with an agglomerative clustering procedure. Ward's (1963) hierarchical agglomerative procedure was used to initially cluster 12 initial partitions independently. The initial 12 partitions consisted of the groups divided by age levels. A proximity matrix of error sums of squares values was formed by pooling the clusters from the 12 independent analyses. This proximity matrix then underwent the second stage of clustering which also utilized Ward's hierarchical agglomerative procedure.

Ward's method utilizes error sum of squares values as a measure of similarity to produce the similarity matrix. This similarity matrix was subjected to the second stage of clustering, again using Ward's method to produce the clusters. From these clusters, the cluster centroids were be used as starting points for the third stage of clustering.

The third stage of clustering utilized an iterative partitioning procedure that uses a K-means passes algorithm. Generalized distance was used as the similarity measure with this iterative procedure.

Four stopping rules (Glutting, et al., 1997) were utilized during the first- and second-stage clustering with one stopping rule used with the final stage of clustering. The first four stopping rules are as follows:

1. Solutions must correspond to a hierarchical step occurring before an atypical change in the similarity measure.
2. Solutions must have a cluster variance to standardization sample variance ratio < 1.0.
3. Solutions must meet Mojena's (1977) first stopping rule. This mathematical rule requires stopping when:

$$a_{j+1} > a + ks_a ,$$

where a is the fusion coefficient; a_{j+1} is the value of the criterion at stage j + 1 of the clustering process; k is the standard deviate, a is the mean of the fusion coefficient, and s is the standard deviation of the fusion coefficient. The standard deviate, k, can be calculated at each stage of clustering where:

$$k_j = (a_{j+1} - a) / s_a$$

4. Solutions must fulfill the criteria set forth by Wishart's (1982) t-test.

The final stage of clustering (K-means) was to stop when subject relocations did not improve within cluster homogeneities or at no more than 100 iterations.

Profile Validation

Aldenderfer and Blashfield (1984) and Blashfield and Aldenderfer (1988) discuss several plausible methods for determining the validity of the clusters developed through the analyses. Three of those methods are:

1. Replication
2. Significance tests on independent variables
3. Tests of homogeneity

These methods tend to provide good indicators of cluster validity and have been used in similar research (McDermott, Glutting, Jones, Watkins, & Kush, 1989; McDermott, Glutting, Jones, & Noonan, 1989; Konold, Glutting, McDermott, Kush, & Watkins, 1999). Validation through replication is a straightforward procedure, simply split the sample in half, use one half to determine clusters, then cluster the other half as a check for similar results. If the second sample does not replicate the first then the validity of the clusters must be called into question (Aldenderfer & Blashfield, 1984; Blashfield & Aldenderfer, 1988). The second method for examining validity is to perform significance tests on independent variables.

Usually these significance tests are performed on external variables not used in the clustering procedures, such as demographics. Examples of demographics used in similar analyses are age, sex, race, and educational placement (Konold, Glutting, McDermott, Kush, and Watkins, 1999). Other pertinent variables could and should be used in this type of validation technique. The final validation technique is the use of internal tests of homogeneity.

Tryon and Bailey (1970) advocate the use of the H statistic as a test of homogeneity. The H statistic compares the variance for a particular cluster on a variable with the variance for that same variable across the entire data set (Blashfield & Aldenderfer, 1988). The value of the H statistic will approach 1.00 as the clusters become more homogenous. A cutoff value of $>$ .60 has been advocated (McDermott, Glutting, Jones, Watkins, & Kush, 1989; McDermott, Glutting, Jones, & Noonan, 1989; Konold, Glutting, McDermott, Kush, & Watkins, 1999). Another test of homogeneity used in this type of research has been Cattell's (1949) r_p statistic.

Cattell's r_p compares the average similarity between the resultant clusters. An r_p value of < .40 has also been suggested as a solid cutoff point (McDermott, Glutting, Jones, Watkins, & Kush, 1989; McDermott, Glutting, Jones, & Noonan, 1989; Konold, Glutting, McDermott, Kush, & Watkins, 1999). All three validation techniques were used to evaluate the resultant clusters produced by the UNIT standardization sample for the four subtest Standard Battery, the six subtest Extended Battery, and the four subtest Standard Battery for the Learning Disabled subsample.

Finally, a set of steps were created to help practitioners compare a particular UNIT profile to common UNIT profiles, as determined by the procedures described in this study. These steps are user-friendly and can be implemented by examiners in the field using simple mathematics.

CHAPTER 3

Results

The following results have been divided into three sections to correspond with the three overall analyses. First, the UNIT Standard Battery results are presented, including demographic prevalence trends, followed by results from the Extended Battery, and the Learning Disabled subsample.

Standard Battery

The MEG procedure was utilized for the clustering analyses (McDermott, 1998). The twelve initial age-partitioned blocks of the standardization sample (N = 2,100) were submitted to first stage clustering. These initial age-partitioned blocks yielded 93 profile clusters with an average of 7.75 clusters per age-partitioned block. Those resultant profile clusters were then merged into a 93 X 93 similarity matrix for subsequent second-stage clustering. The resultant second-stage cluster profiles were then analyzed through comparisons with the previously decided upon criteria. A six profile cluster solution met the decision criteria most clearly, and was therefore submitted to third-stage clustering to ensure accurate final profile clusters. The third-stage clustering relocates those profiles that were misclassified during second-stage clustering.

The six profile cluster solution produced extremely tight clusters ($\overline{H}$ = .981; range .978 to .985). Additionally, the six profile cluster solution was found to have excellent separation between each cluster ($\bar{r_p}$ = -.27; range -.81 to .34). The replication rate for the final-stage profile clusters across the twelve first-stage profile clusters was 88.89%. Replication was determined by assessing whether the final profile clusters existed in each of the first-stage profile clusters. Specific psychometric properties of each respective profile can be found in Table 2. Table 3 provides a summary of the subtest and Full Scale IQ scores for each profile while Table 4 and Figure 1 presents the subscale scores.

External variables were used in the description and to support the validity of the prototypic profile clusters. Full Scale IQ's, Subscale Quotients, and prevalence rates are presented with expectancy comparisons of gender, race, and parent education within each profile cluster reported. The following prevalence trends reported as higher or lower than expected were found to be statistically significant ($p < .05$ or less).

Core Standard Battery Profile Descriptions

1. Superior. (FSIQ = 122, MQ = 119, RQ = 120, SQ = 120, NSQ = 119; Prevalence = 16.8%). This profile type is composed of more Whites than expected; 92% versus the expected composition of 79%. Additionally, significantly fewer African Americans are represented; 3% versus the expected composition of 16%. More males (54%) are represented than females (46%). The majority of parents for these students have four or more years of college (52%).

Table 2

Psychometric Properties of UNIT Standard Battery Profiles

Profile Cluster	Within-Type Homogeneity (H)	Between-Type Similarity (r_p)	Replication: First-to-Last Cluster
1	.978	-.73	100%
2	.984	.21	100%
3	.981	-.45	50%
4	.985	.34	100%
5	.984	-.17	91.67%
6	.976	-.81	91.67%
Mean	.981	-.27	88.89%

Table 3

UNIT Standard Battery Subtest Scores[a] and Full-Scale IQs[b] for Respective Profiles[c]

Profile	SY	CD	SP	AR	FSIQ
1	13	14	13	13	122
2	9	13	11	10	105
3	12	9	12	10	105
4	10	9	9	12	99
5	8	8	9	8	88
6	5	6	5	5	71

[a]Subtest standard score Ms = 10 and SDs = 3

[b]FSIQ standard score M = 100 and SD = 15

[c]SY = Symbolic Memory, CD = Cube Design, SP = Spatial Memory, AR = Analogic Reasoning, FSIQ = Full-Scale IQ

Table 4

UNIT Standard Battery Subscale Scores[a] for Respective Profiles[b]

Profile	MQ	RQ	SQ	NSQ
1	119	120	120	119
2	101	109	100	110
3	113	97	106	104
4	95	103	106	93
5	91	89	88	91
6	73	75	73	74

[a]Subscale standard score Ms = 100 and Sds = 15

[b]MQ = Memory Quotient, RQ = Reasoning Quotient, SQ = Symbolic Quotient, NSQ = Nonsymbolic Quotient

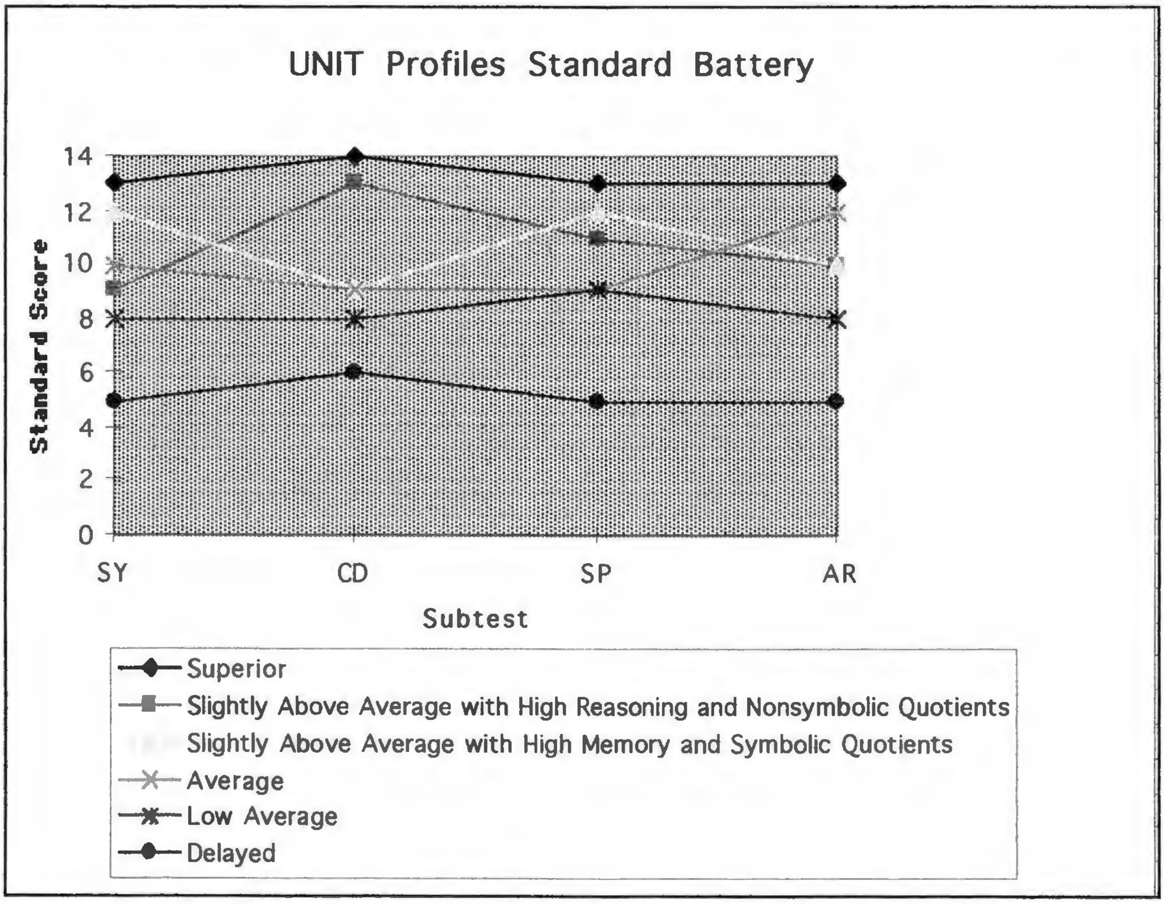

Figure 1. Mean UNIT Standard Battery subtest profiles

Note: SY = Symbolic Memory, CD = Cube Design, SP = Spatial Memory, AR = Analogic Reasoning

2. Slightly Above Average with High Reasoning and Nonsymbolic Quotients. (FSIQ = 105, MQ = 101, RQ = 109, SQ = 100, NSQ = 110; Prevalence = 17.5%). Again, Whites predominate at higher than expected levels within this type (84% versus 79%), while fewer than expected African Americans are present (8% versus 16%). Males predominate (61%) versus 44% female composition, and again the majority of parents have four or more years of college education (36%).

3. Slightly Above Average with High Memory and Symbolic Quotients. (FSIQ = 105, MQ = 113, RQ = 97, SQ = 106, NSQ = 104; Prevalence = 16.3%). More Whites than expected and less African Americans than expected compose this profile type (81% versus 11%), while females are disproportionately represented (58%; males - 42%). Parental education level is relatively evenly distributed among high school graduates (30%), those with up to three years of college (30%), and college graduates (29%).

4. Average. (FSIQ = 99, MQ = 95, RQ = 103, SQ = 106, NSQ = 93; Prevalence = 15.1%). The race composition of this profile approaches the overall population (Whites - 85%; African Americans - 13%). Gender composition also approaches the overall population with slightly more females than males (51% versus 49%). Parent education is again relatively evenly distributed with most completing high school (35%), 28 percent completing one to three years of college, 24 percent completing college, and 12 percent not finishing high school.

5. Low Average. (FSIQ = 88, MQ = 91, RQ = 89, SQ = 88, NSQ = 91; Prevalence = 25.7%). Slightly less Whites than expected (72% versus population expectancy of 79%) and significantly more African Americans than expected (32% versus population expectancy of 16%). In terms of gender, there are more females (52%) than males (48%). More parents have completed high school (38%; national proportion - 29%), with a large proportion also completing one to three years of college (26%).

6. Delayed. (FSIQ = 71, MQ = 73, RQ = 75, SQ = 73, NSQ = 74; Prevalence = 8.5%). More African Americans (32%) and less Whites (64%) than expected (versus 16% and 79%, respectively) comprise this profile. The largest proportion of parents within this profile have completed high school (37%) with some attending college (26%) and a large proportion not completing high school (20%). Sixteen percent of the parents have completed college. Significantly more females than males are present (54% and 46%, respectively) within this type.

<u>Extended Battery</u>

The MEG procedure was also utilized for the Extended Battery clustering analyses (McDermott, 1998). The twelve initial age-partitioned blocks of the standardization sample (N = 2,100) were submitted to first stage clustering. These initial age-partitioned blocks yielded 96 profile clusters with an average of 8 clusters per age-partitioned block. Those resultant profile clusters were then merged into a 96 X 96 similarity matrix for subsequent second-stage clustering. The resultant second-stage cluster profiles were then analyzed through comparisons with the previously decided upon criteria. A seven profile cluster solution met the decision criteria most clearly, and was therefore submitted to third-stage clustering to ensure accurate final profile clusters. The third-stage clustering relocates those profiles that were misclassified during second-stage clustering.

The seven profile cluster solution produced extremely tight clusters ($\overline{H}$ = .979; range .973 to .981). Additionally, the seven profile cluster solution was found to have extremely good separation between each cluster ($\overline{r_p}$ = -.09; range -.76 to .23). The replication rate for the final-stage profile clusters across the twelve first-stage profile clusters was 84.48%. Replication was determined by assessing whether the final profile clusters existed in each of the first-stage profile clusters. Specific psychometric properties of each respective profile can be found in Table 5. Table 6 provides a summary of the subtest and Full Scale IQ scores for each profile while Table 7 and Figure 2 presents the subscale scores.

Again, external variables were used in the description and to support the validity of the prototypic profile clusters. Full Scale IQ's, Subscale Quotients, and prevalence rates were presented with expectancy comparisons of gender, race, and parent education within each profile cluster reported. The following

Table 5

Psychometric Properties of UNIT Extended Battery Profiles

Profile Cluster	Within-Type Homogeneity (H)	Between-Type Similarity (r_p)	Replication: First-to-Last Cluster
1	.975	-.69	100%
2	.979	-.25	100%
3	.980	.11	66.67%
4	.981	.18	58%
5	.981	.05	91.67%
6	.981	..23	75%
7	..973	-.76	100%
Mean	.979	-.09	84.48%

Table 6

UNIT Extended Battery Subtest Scores[a] and Full-Scale IQs[b] for Respective Profiles[c]

Profile	SY	CD	SP	AR	OB	MZ	FSIQ
1	14	13	14	13	14	11	122
2	11	12	11	11	10	13	111
3	11	12	11	11	10	8	105
4	10	9	10	10	12	11	101
5	8	9	9	8	8	11	90
6	9	8	8	9	10	7	90
7	6	6	6	5	6	7	70

[a]Subtest standard score Ms = 10 and SDs = 3

[b]FSIQ standard score M = 100 and SD = 15

[c]SY = Symbolic Memory, CD = Cube Design, SP = Spatial Memory, AR = Analogic Reasoning, OB = Object Memory, MZ = Mazes, FSIQ = Full-Scale IQ

Table 7

UNIT Extended Battery Subscale Scores[a] for Respective Profiles[b]

Profile	MQ	RQ	SQ	NSQ
1	123	116	122	117
2	105	114	105	114
3	105	104	105	104
4	103	99	104	98
5	88	95	87	96
6	94	87	96	85
7	73	73	72	74

[a]Subscale standard score Ms = 100 and Sds = 15

[b]MQ = Memory Quotient, RQ = Reasoning Quotient, SQ = Symbolic Quotient, NSQ = Nonsymbolic Quotient

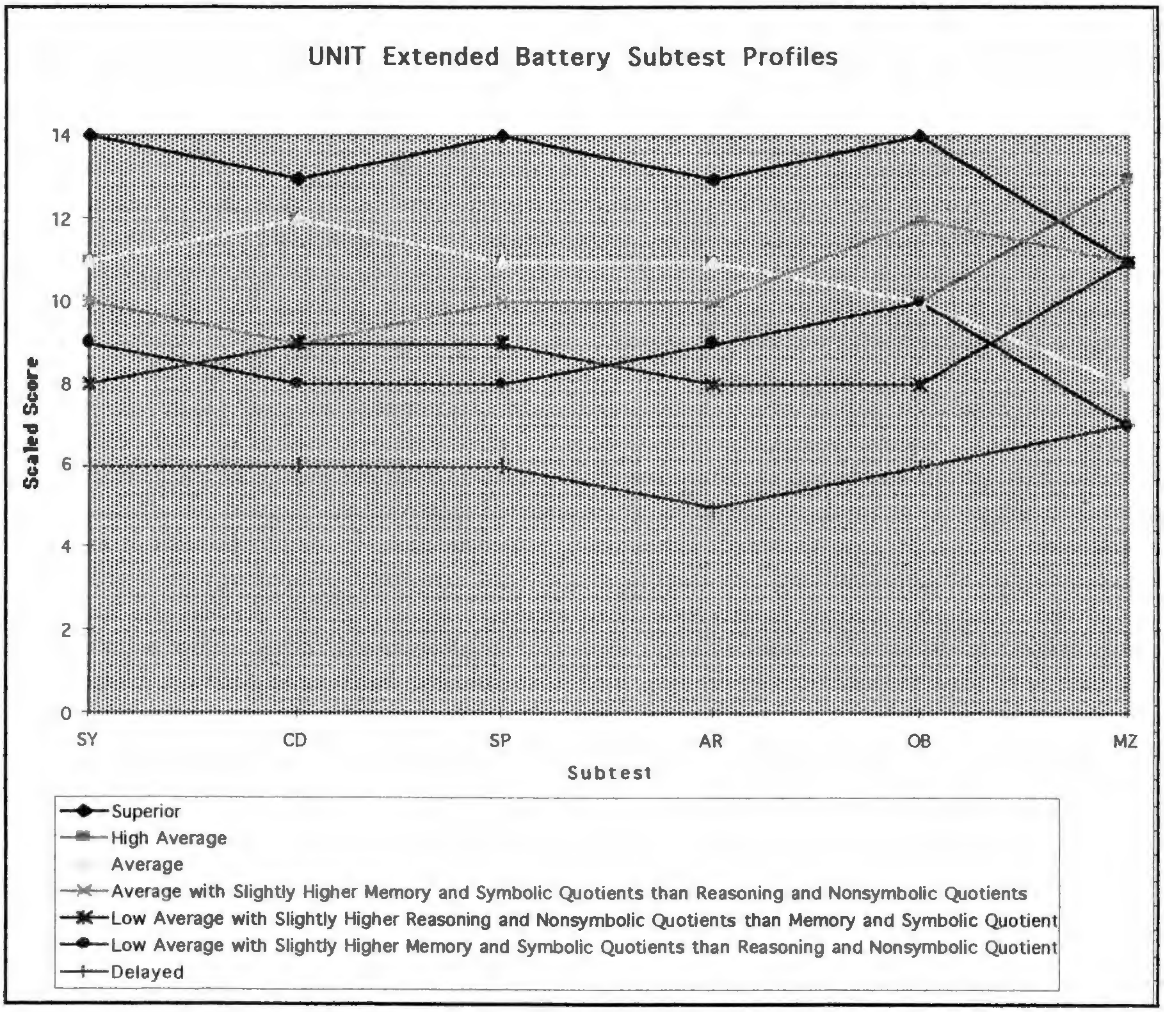

Figure 2. Mean UNIT Extended Battery subtest profiles

Note: SY = Symbolic Memory, CD = Cube Design, SP = Spatial Memory, AR = Analogic Reasoning, OB = Object Memory, MZ = Mazes prevalence trends reported as higher or lower than expected were found to be statistically significant ($p < .05$ or less).

Core Extended Battery Profile Descriptions

1. Superior. (FSIQ = 122, MQ = 123, RQ = 116, SQ = 122, NSQ = 117; Prevalence = 13.6%). This profile type is composed of more Whites than expected; 89% versus the expected composition of 79%. Additionally, significantly fewer African Americans are represented; 4% versus the expected composition of 16%. Males and females are equally represented. The majority of parents for these students have four or more years of college (55%).

2. High Average. (FSIQ = 111, MQ = 105, RQ = 114, SQ = 105, NSQ = 114; Prevalence = 15.6%). Again, Whites predominate at higher than expected levels within this type (89% versus 79%), while fewer than expected African Americans are present (5% versus 16%). There are more males than females (55% versus 44%), and again the majority of parents have four or more years of college education (39%).

3. Average. (FSIQ = 105, MQ = 105, RQ = 104, SQ = 105, NSQ = 104; Prevalence = 15.1%). More Whites than expected and less African Americans than expected compose this profile type (90% versus 6%), while males are disproportionately represented (56%; females - 44%). Most parents have four or more years of college (41%).

4. Average with Slightly Higher Memory and Symbolic Quotients than Reasoning and Nonsymbolic Quotients. (FSIQ = 101, MQ = 103, RQ = 99, SQ = 104, NSQ = 98; Prevalence = 14.6%). The race composition of this profile is very similar to the overall population (Whites - 76%; African Americans - 15%). More females than males are present (54% versus 46%). Parent education is relatively evenly distributed with most completing high school (35%).

5. Low Average with Slightly Higher Reasoning and Nonsymbolic Quotients than Memory and Symbolic Quotients. (FSIQ = 90, MQ = 88, RQ = 95, SQ = 87, NSQ = 96; Prevalence = 18.3%). The race of these students approaches expectancies (Whites - 81%; African Americans - 14%). In terms of gender, there are more males (53%) than females (47%). More parents have completed high school (36%; national proportion - 29%), with a large proportion also completing one to three years of college (31%; national proportion - 30%).

6. Low Average with Slightly Higher Memory and Symbolic Quotients than Reasoning and Nonsymbolic Quotients. (FSIQ = 90, MQ = 94, RQ = 87, SQ = 96, NSQ = 85; Prevalence = 14.2%). More African Americans (21%) and slightly less Whites (76%) than expected (versus 16% and 78%, respectively) comprise this profile. The largest proportion of parents within this profile have completed high school (39%) with some attending college (24%) and some graduating from college (22%). Fifteen percent of the parents have not

completed high school. Significantly more females than males are present (59% and 41%, respectively) within this type.

7. Delayed. (FSIQ = 70, MQ = 73, RQ = 73, SQ = 72, NSQ = 74; Prevalence = 8.7%). African Americans make up a disproportionally large group within this profile (37%) while Whites account for just under two thirds. Again, there are more females than males (52% and 48%). The largest proportion of parents have not completed high school (36%) with those completing high school following closely (32%).

Learning Disabled Subsample

The MEG procedure was again utilized for the clustering analyses (McDermott, 1998). Two randomly partitioned blocks of the Learning Disabled subsample of the standardization sample (N = 110) were submitted to first stage clustering. These initial partitioned blocks yielded 6 profile clusters with an average of 3 clusters per block. Those resultant profile clusters were then merged into a 6 X 6 similarity matrix for subsequent second-stage clustering. The resultant second-stage cluster profiles were then analyzed through comparisons with the previously decided upon criteria. A four profile cluster solution met the decision criteria most clearly, and was therefore submitted to third-stage clustering to ensure accurate final profile clusters. The third-stage clustering relocates those profiles that were misclassified during second-stage clustering.

The four profile cluster solution produced extremely tight clusters ($\overline{H}$ = .982; range .979 to .985). Additionally, the four profile cluster solution was found to have excellent separation between each cluster ($\overline{r_p}$ = -.49; range -.79 to .05). The replication rate for the final-stage profile clusters across the twelve first-stage profile clusters was 75%. Replication was determined by assessing whether the final profile clusters existed in each of the first-stage profile clusters. Specific psychometric properties of each respective profile can be found in Table 8. Table 9 provides a summary of the subtest and Full Scale IQ scores for each profile while Table 10 and Figure 3 present the subscale scores.

External variables were again used in the description and to support the validity of the prototypic profile clusters. Full Scale IQ's, Subscale Quotients, and prevalence rates were presented with expectancy comparisons of gender, race, and parent education within each profile cluster reported. The following prevalence trends reported as higher or lower than expected were found to be statistically significant ($p < .05$ or less).

Core Learning Disabled Subsample Profile Descriptions

1. Average with Higher Memory Quotient than Reasoning Quotient. (FSIQ = 106, MQ = 112, RQ = 100, SQ = 103, NSQ = 108; Prevalence = 17.3%). This profile type is composed predominantly of Whites; 84% versus the expected composition of 79%, while African Americans comprise only

Table 8

Psychometric Properties of UNIT Learning Disabled Profiles

Profile Cluster	Within-Type Homogeneity (H)	Between-Type Similarity (r_p)	Replication: First-to-Last Cluster
1	.985	-.76	50%
2	.981	-.45	100%
3	.979	.05	50%
4	.982	-.79	100%
Mean	.982	-.49	75%

Table 9

UNIT Subtest Scores[a] and Full-Scale IQs[b] for Respective Learning Disabled Profiles[c]

Profile	SY	CD	SP	AR	FSIQ
1	11	10	13	10	106
2	9	11	8	10	97
3	8	7	9	8	86
4	6	7	6	5	73

[a]Subtest standard score Ms = 10 and SDs = 3

[b]FSIQ standard score M = 100 and SD = 15

[c]SY = Symbolic Memory, CD = Cube Design, SP = Spatial Memory, AR = Analogic Reasoning, FSIQ = Full-Scale IQ

Table 10

UNIT Subscale Scores[a] for Respective Learning Disabled Profiles[b]

Profile	MQ	RQ	SQ	NSQ
1	112	100	103	108
2	92	103	100	94
3	89	86	86	88
4	74	77	73	78

[a]Subscale standard score Ms = 100 and Sds = 15

[b]MQ = Memory Quotient, RQ = Reasoning Quotient, SQ = Symbolic Quotient, NSQ = Nonsymbolic Quotient

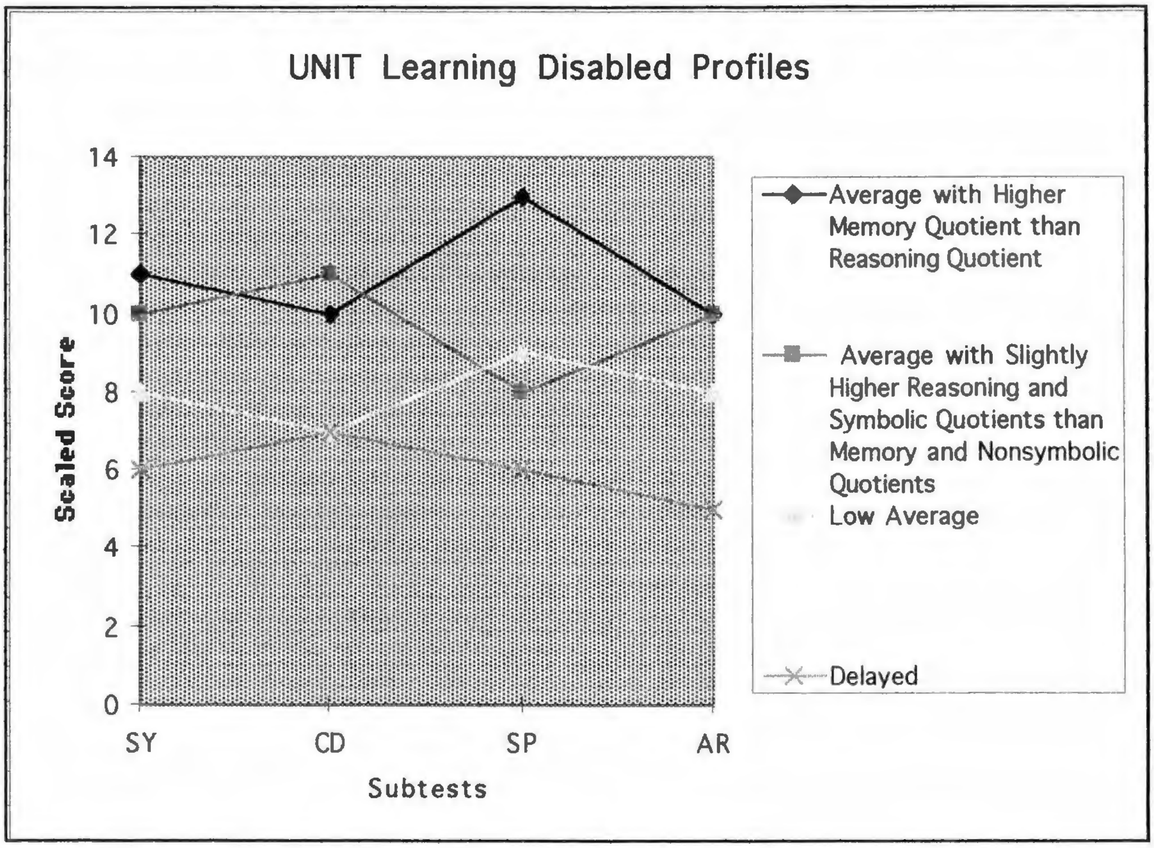

Figure 3. Mean UNIT Standard Battery subtest patterns for the Learning Disabled subsample

Note: SY = Symbolic Memory, CD = Cube Design, SP = Spatial Memory, AR = Analogic Reasoning

5 percent of this profile. More males (58%) are represented than females (42%). The majority of parents for these students have completed high school (41%), with sizable proportions of parents not completing high school (21%), and completing four or more years of college (26%).

2. Average with Slightly Higher Reasoning and Symbolic Quotients than Memory and Nonsymbolic Quotients. (FSIQ = 97, MQ = 92, RQ = 103, SQ = 100, NSQ = 94; Prevalence = 24.5%). Whites predominate this profile at higher than expected levels within this type (94% versus 79%), while no African Americans are present (versus the expectancy of 16%). Males predominate (70%) versus 30 percent female composition. The majority of parents have one to three years of college education (36%) with a large proportion completing college (30%).

3. Low Average. (FSIQ = 86, MQ = 89, RQ = 86, SQ = 86, NSQ = 88; Prevalence = 32.7%). The race composition of this profile approaches the overall population better than the other profile types (Whites - 86%; African Americans - 14%). More females than males are represented (56%; males - 44%). Parental education level is most evenly distributed within this profile with those who have not graduated high school (39%), those who have graduated high school (28%), those with one to three years of college (22%), and college graduates (11%).

4. Delayed. (FSIQ = 73, MQ = 74, RQ = 77, SQ = 73, NSQ = 78; Prevalence = 25.5%). The race composition of this profile also approaches the overall

population (Whites - 75% versus the expectancy of 79%; African Americans - 21% versus the expectancy of 16%). Gender composition also approaches the overall population with more males than females (54% versus 46%). Most parents have not completed high school (46%), 28 percent completing one to three years of college, 21 percent completing high school, and only 4 percent finishing college.

CHAPTER 4

Discussion

The results across the UNIT batteries are quite similar to the results of other similar studies using traditional verbally-laden tests. Additionally, parallels can be drawn between the results of the Learning Disabled subsample typology and other similar research. The implications and applications of the typologies are discussed below.

Standard Battery

Psychometrically, the UNIT Standard Battery produced extremely strong results with within-cluster homogeneity and between-cluster dissimilarity, $\overline{H}$ = .981 and $\overline{r_p}$ = -.27. The profiles comprising the clusters are considered more similar to each other as the $\overline{H}$ statistic approaches 1.0, and the farther the clusters are separated from each other the farther the $\overline{r_p}$ statistic departs from 1.0. These results then, can be considered somewhat stronger than similar studies of the KABC ($\overline{H}$ = .67, $\overline{r_p}$ = .19; Glutting, McGrath, Kamphaus, & McDermott, 1992), the WAIS-R ($\overline{H}$ = .74, $\overline{r_p}$ = .22; McDermott, Glutting, Jones, & Noonan, 1989), the WISC-R ($\overline{H}$ = .63, $\overline{r_p}$ = .33; McDermott, Glutting, Jones, Watkins, &Kush, 1989), the WISC-III ($\overline{H}$ = .67, $\overline{r_p}$ = .20; Konold, Glutting, McDermott, Kush, & Watkins, 1999), the WPPSI ($\overline{H}$ = .63, $\overline{r_p}$ = .33; Glutting & McDermott, 1990), and the DAS ($\overline{H}$ = .67, $\overline{r_p}$ = .31; Holland & McDermott, 1996).

The six profile cluster solution for the Standard Battery yielded not only the best overall psychometric properties, but was also the most parsimonious in terms of Subscale explication. Four of the profile clusters yielded would be considered relatively "flat" in terms of subtest variability. This lends good evidence to the factor analytic research (Reed & McCallum, 1995) indicating that the UNIT is a good measure of one overall factor, hypothesized as "g". Further, when considering the pattern of variability present among the Subscale scores of the prototypical profiles, more evidence is provided in support of the Primary and Secondary Scales described by Reed and McCallum (1995). That is, scores tended to cluster yielding patterns identifiable by the elevations and depressions of the Subscale scores, e.g., Memory versus Reasoning Quotients or Symbolic versus Nonsymbolic Quotients.

Similar to other typological research with intelligence scales (McDermott, Glutting, Jones, Watkins, & Kush, 1989; Holland & McDermott, 1996), the higher ability profiles were comprised of significantly more males than females while the lower ability profiles were comprised of more females than males. In terms of race, similar findings of higher than expected proportions of African Americans were found in the lower ability profiles while higher than expected proportions of Whites were present in the higher ability profiles (McDermott, Glutting, Jones, Watkins, & Kush, 1989; McDermott, Glutting, Jones, & Noonan, 1989; Holland & McDermott, 1996;

Konold, Glutting, McDermott, Kush, & Watkins, 1999). Interestingly though, the proportion of African Americans is higher than obtained in the WISC-III normative typology high ability profile (3% African Americans in UNIT Standard Battery Superior profile versus 0.1% African Americans in the WISC-III High Ability profile; Glutting, McDermott, & Konold, 1997, p. 362) and lower than obtained in the WISC-III normative typology Low Ability profile (32% African Americans in the UNIT Extended Battery Superior profile versus "more than two and one half times the national average" African Americans in the WISC-III Low Ability profile; Glutting, McDermott, & Konold, 1997, p. 364). Parental education levels for students in the lower and higher ability levels are also similar to previous typological studies (McDermott, Glutting, Jones, Watkins, & Kush, 1989; Holland & McDermott, 1996), with higher levels of education characterizing the higher ability profiles and lower levels of education characterizing the lower ability profiles.

Extended Battery

Psychometrically, the UNIT Extended Battery also produced extremely strong results with within-cluster homogeneity and between-cluster dissimilarity, $\overline{H} = .979$ and $\overline{r_p} = -.09$. The profiles comprising the clusters are considered more similar to each other as the $\overline{H}$ statistic approaches 1.0, and the farther the clusters are separated from each other the farther the $\overline{r_p}$ statistic departs from 1.0. These results are also stronger than the results of similar studies of the KABC ($\overline{H}$ = .67, $\bar{r}_p$ = .19; Glutting, McGrath, Kamphaus, & McDermott, 1992), the WAIS-R ($\overline{H}$ = .74, $\bar{r}_p$ = .22; McDermott, Glutting, Jones, & Noonan, 1989), the WISC-R ($\overline{H}$ = .63, $\bar{r}_p$ = .33; McDermott, Glutting, Jones, Watkins, &Kush, 1989), the WISC-III ($\overline{H}$ = .67, $\bar{r}_p$ = .20; Konold, Glutting, McDermott, Kush, & Watkins, 1999), the WPPSI ($\overline{H}$ = .63, $\bar{r}_p$ = .33; Glutting & McDermott, 1990), and the DAS ($\overline{H}$ = .67, $\bar{r}_p$ = .31; Holland & McDermott, 1996), but slightly weaker than those produced by the UNIT Standard Battery ($\overline{H}$ = .981 and $\bar{r}_p$ = -.27).

The seven profile cluster solution for the Extended Battery yielded the best overall psychometric properties, and as the Standard Battery profiles, was the most parsimonious in terms of Subscale explication. Six of the profile clusters yielded would, as with the Standard Battery profiles, be considered relatively "flat" in terms of subtest variability, again lending good evidence to the factor analytic research (Reed & McCallum, 1995) indicating that the UNIT is a good measure of "g". Also, as with the Standard Battery, when considering the pattern of variability present among the Subscale scores of the prototypical profiles, more evidence is provided in support of the Primary and Secondary Scales described by Reed and McCallum (1995). Again, scores tended to cluster yielding patterns identifiable by the elevations and depressions of the Subscale scores.

Similar to the UNIT Standard Battery profile clusters and other typological research with intelligence scales (McDermott, Glutting, Jones, Watkins, & Kush, 1989; Holland & McDermott, 1996), the higher ability profiles were comprised of significantly more males than females while the lower ability profiles were comprised of more females than males. Interestingly, however, profile five (Low Average with Slightly Higher Reasoning and Nonsymbolic Quotients than Memory and Symbolic Quotients) yielded a higher proportion of males. In terms of racial composition, similar findings of higher than expected proportions of African Americans were present in the lower ability profiles while higher than expected proportions of Whites were present in the higher ability profiles (McDermott, Glutting, Jones, Watkins, & Kush, 1989; McDermott, Glutting, Jones, & Noonan, 1989; Holland & McDermott, 1996; Konold, Glutting, McDermott, Kush, & Watkins, 1999). Again, as with the Standard Battery, the proportion is higher than obtained in the WISC-III normative typology high ability profile (4% African Americans in UNIT Extended Battery Superior profile versus 0.1% African Americans in the WISC-III High Ability profile; Glutting, McDermott, & Konold, 1997, p. 362) and lower than obtained in the WISC-III normative typology Low Ability profile (37% African Americans in the UNIT Extended Battery Superior profile versus "more than two and one half times the national average" African Americans in the WISC-III Low Ability profile; Glutting, McDermott, & Konold, 1997, p. 364).

Parental education levels for students in the lower and higher ability levels are also similar to previous typological studies (McDermott, Glutting, Jones, Watkins, & Kush, 1989; Holland & McDermott, 1996), with higher levels of education comprising the higher ability profiles and lower levels of education comprising the lower ability profiles.

Learning Disabled Subsample

The results of the Learning Disabled subsample must be analyzed cautiously due to a limited number of subjects. Missing data values forced the reduction in useable sample size to 110 individuals. Also, caution in interpretation is warranted due to the variability in Learning Disabled definitions used for identification for Special Education services across the different states from which the students were selected. Finally, the present analysis is limited to the UNIT cognitive measures of the Learning Disabled subsample. Further research will be needed to examine the UNIT and to link achievement measures in profile determination. However, these results may be useful as a starting point.

The UNIT Standard Battery Learning Disabled subsample produced a four profile cluster solution with extremely strong results with both excellent within-cluster homogeneity and between-cluster dissimilarity, $\overline{H} = .982$ and $\overline{r_p} = -.49$. Again, the profiles comprising the clusters are considered more similar to each other as the $\overline{H}$ statistic approaches 1.0, and the farther the clusters are separated from each other the farther the $\bar{r}_p$ statistic departs from 1.0. These results, as compared with the Standard Battery ($\bar{H}$ = .981 and $\bar{r}_p$ = -.27) and Extended Battery ($\bar{H}$ = .979 and $\bar{r}_p$ = -.09), would be considered slightly stronger than the UNIT normative typological results. Additionally, the replication rate of 75 percent across the resultant clusters from first stage to last stage was strong; however, the replication rate was not as strong as those produced with the Standard Battery (88.89%) or the Extended Battery (84.48%).

Consistent with other studies (Ward, Ward, Glutting, & Hatt, 1999), the UNIT Learning Disabled subsample produced a profile resembling slow learners, i.e., Profile 4, and another "profile with generally depressed abilities" (p.631), Profile 3. Interestingly, the two higher ability profiles produced the most variation in subscale scores with mirrored patterns between the two profiles, i.e., Profile 1 exhibited a higher Memory Quotient than Reasoning Quotient while Profile 2 exhibited a higher Reasoning Quotient than Memory Quotient. The same pattern was present with the Symbolic and Nonsymbolic Quotients, with Profile 1 having the higher Nonsymbolic Quotient and Profile 2 the higher Symbolic Quotient. These definitive patterns of strengths and weaknesses were not unexpected. Even though the population of Learning Disabled children is heterogeneous, patterns revealing increased variation have been found than have been found among their Non-Learning Disabled counterparts in neuropsychological research (Fisk & Rourke, 1983).

Clinical Comparisons, Limitations, and Future Directions

Konold, Glutting, McDermott, Kush, and Watkins (1999) have developed, and Holland and McDermott (1996) have advocated a method to compare subsequent clinical profiles with the normative profiles to determine uniqueness based upon the theoretical work of Osgood and Suci (1952). The method is easily conceptualized and user-friendly. The clinician simply compares the clinical subtest profile to the normative profiles that have the closest Full Scale IQ. The clinician then subtracts each of the respective subtest scores from the normative profile scores to yield difference scores. Those differences are then squared and summed. In terms of the UNIT Standard Battery, if the sum of the squared absolute differences of the four subtests is ≥ 272, the clinical profile can then be determined to be rare in the population. That is, a sum of squared differences ≥ 272 occurs less than five times out of one hundred in the general population. When making comparisons with the UNIT Extended Battery, a sum of the squared absolute differences of the six subtests ≥ 307 would be deemed rare in the population.

This method of comparing clinical profiles with the normative profiles provides a multivariate means of analysis that overcomes the statistical limitations associated with the univariate methods of comparison currently used with most intelligence batteries (Glutting, McDermott, Watkins, Kush,

& Konold, 1997). The univariate analyses rely on multiple stepwise, pairwise comparisons whereas this method allows for a multivariate comparison. Also, this method, while somewhat statistically complex in nature, is relatively easy to use in a clinical setting where complex statistical analysis software is either unavailable or cost prohibitive.

While this approach to the UNIT interpretation is useful, it does have limitations. First of all, for those clinicians who approach interpretation as a guide for developing treatments, this method's greatest strength is also its greatest weakness. The cluster analytic results are statistically superior to the more traditional approaches, but do not allow for the specific treatment generation like the traditional approaches. Although this approach may have limited utility currently, there are future directions worth exploring.

One possible direction for future research is to examine the five percent of clinical cases that are identified as not fitting one of the normative cluster profiles for possible explanations and implications. There may, in fact, be commonalities that exist among these "abnormal" profiles that could provide further information with clinical relevance. Another line of research that would be relevant and useful might involve the determination of what, if any, predictive validity the normative cluster profiles lend to educational outcomes.

REFERENCES

REFERENCES

Adams, K.M. (1985). Theoretical, methodological, and statistical issues. In B.P. Rourke (Ed.), Neuropsychology of learning disabilities: Essentials of subtype analysis (pp. 17-39). New York: Guilford Press.

Aldenderfer, M.S. & Blashfield, R.K. (1984). Cluster analysis. Beverly Hills, CA: Sage.

Anderberg, M. (1973). Cluster analysis for applications. New York: Academic Press.

Blashfield, R.K. (1980). Propositions regarding the use of cluster analysis in clinical research. Journal of Consulting and Clinical Psychology, 48(4), 456-459.

Blashfield, R.K. & Aldenderfer, M.S. (1988). The methods and problems of cluster analysis. In J.R. Nesselroade & R.B. Cattell (Eds.), Handbook of multivariate experimental psychology (2nd ed., pp. 447-473). New York: Plenum Press.

Bracken, B.A., & McCallum, R.S. (1998). Universal nonverbal intelligence test: Examiner's manual. Itasca, IL: Riverside Publishing

Cattell, R.B. (1944). Psychological measurement: Normative, ipsative, interactive. Psychological Review,51, 292-303.

Cattell, R.B. (1949). r_p and other coefficients of pattern similarity. Psychometrika, 14, 279-298.

Christian, R.E., Frick, P.J., Hill, N.L., Tyler, L., and Frazer, D.R. (1997). Psychopathy and conduct problems in children: II. Implications for subtyping children with conduct problems. Journal of the American Academy of Child and Adolescent Psychiatry, 36(2), 233-241.

Crawford, J.R., Garthwaite, P.H., Johnson, D.A., Mychalkiw, B., & Moore, J.W. (1997). WAIS-R subtest pattern clusters in closed-head-injured and healthy samples. Clinical Neuropsychologist, 11(3), 249-257.

Cronbach, L.J., & Gleser, G.C. (1953). Assessing similarity between profiles. Psychological Bulletin, 50(6), 456-473.

Delaney, E.A., & Hopkins, T.F. (1987). The Stanford-Binet Intelligence Scale: Fourth edition examiner's handbook. Chicago: Riverside.

DeLuca, J.W., Rourke, B.P., & Del Dotto, J.E. (1991). Subtypes of arithmetic-disabled children: cognitive and personality dimensions. In B.P. Rourke (Ed.), Neuropsychological validation of learning disability subtypes (pp. 180-219). New York: Guilford Press.

Donders, J. (1996). Cluster subtypes in the WISC-III standardization sample: Analysis of factor index scores. Psychological Assessment,8(3), 312-318.

Donders,J. & Warschausky, S. (1997). WISC-III factor index score pattern after traumatic head injury in children. Child Neuropsychology, 3(1), 71-78.

Everitt, B. (1980). Cluster Analysis. New York: Halstead.

Fisk, J.L. & Rourke, B.P. (1983). Neuropsychological subtyping of learning-disabled children: History, methods, implications. Journal of Learning Disabilities, 16, 529-531.

Fleiss, J., Lawler, W., Platman, S., and Fieve, R. (1971). On the use of inverted factor analysis for generating typologies. Journal of Abnormal Psychology, 77, 127-132.

Glutting, J.J., McDermott, P.A., & Konold, T.R. (1997). Ontology, structure, and diagnostic benefits of a normative subtest taxonomy from the WISC-III standardization sample. In D.P. Flanagan, J.L. Genshaft, & P.L. Harrison (Eds.), Contemporary intellectual assessment: Theories, tests, and issues (pp. 349-372). New York: Guilford.

Glutting, J.J., McDermott, P.A., Konold, T.R., Snelbaker, A.J., & Watkins, M.W.(In Press). More ups and downs of subtest analysis: Criterion validity of the DAS with an unselected cohort. School Psychology Review.

Glutting, J.J., McDermott, P.A., Watkins, M.M., Kush, J.C., & Konold, T.R (1997). The base rate problem and its consequences for interpreting children's ability profiles. School Psychology Review, 26(2), 176-188.

Glutting, J.J., McGrath, F.A., Kamphaus, R.W., & McDermott, P.A. (1992). Taxonomy and validity of subtest profiles on the Kaufman Assessment Battery for Children. Journal of Special Education, 26, 85-115.

Goldstein, S.G. & Linden, J.D. (1969). Multivariate classification of alcoholics by means of the MMPI. Journal of Abnormal Psychology, 74, 661-669.

Hartigan, J. (1975). Clustering algorithms. New York: John Wiley.

Hathaway, & McKinley, (1940). The MMPI manual. New York: Psychological Corporation

Holland, A.M., & McDermott, P.A. (1996). Discovering core profile types in the school-age standardization sample of the Differential Ability Scales. Journal of Psychoeducational Assessment, 14, 131-146.

Jastak, J.F., & Jastak, S.R. (1978). The Wide Range Achievement Test. Wilmington, DE: Jastak Associates.

Kachigan, S.K. (1991). Multivariate statistical analysis: A conceptual introduction (2nd ed.). New York: Radius Press.

Kaufman, A.S. (1979). Intelligent testing with the WISC-R. New York: Wiley.

Kaufman, A.S. (1990). Assessing adolescent and adult intelligence. Boston: Allyn & Bacon.

Kaufman, A.S. (1994). Intelligent testing with the WISC-III. New York: Wiley.

Kaufman, A.S., & Kaufman, N.L. (1983). Interpretive manual for the Kaufman Assessment Battery for Children. Circle Pines, MN: American Guidance Service.

Konold, T.R., Glutting, J.J., McDermott, P.A., Kush, J.C., & Watkins, M.M. (1999). Structure and diagnostic benefits of a normative subtest taxonomy developed from the WISC-III standardization sample. Journal of School Psychology, 37(1), 29-48.

Lachar, D., & Gdowski, C.L. (1979). Actuarial assessment of children and adolescent personality: An interpretive guide for the Personality Inventory for Children. Los Angeles: Western Psychological Services.

Mahalanobis, P. (1936). On the generalized distance in statistics. Proceedings of the National Institute of Science, 49-55.

Maller, S.J. & McDermott, P.A. (1997). WAIS-R profile analysis for college students with learning disabilities. School Psychology Review, 26(4), 575-585.

McDermott, P.A. (1998). MEG: Megacluster analytic strategy for multistage hierarchical grouping with relocations and replications. Educational and Psychological Measurement, 58(4), 677-686.

McDermott, P.A., & Glutting, J.J. (1997). Informing stylistic learning behavior, disposition, and achievement through ability subtests--or, more illusions of meaning?. School Psychology Review, 26(2), 163-175.

McDermott, P.A., Glutting, J.J., Jones, J.N.,& Noonan, J.V. (1989). Typology and prevailing composition of core profile types in the WAIS-R standardization sample. Psychological Assessment, 1, 118-125.

McDermott, P.A., Fantuzzo, J.W., & Glutting, J.J. (1990). Just say no to subtest analysis: A critique on Wechsler theory and practice. Journal of Psychoeducational Assessment, 8, 290-302.

McDermott, P.A., Fantuzzo, J.W., Glutting, J.J., Watkins, M.W., & Baggaley, A.R. (1992). Illusions of meaning in the ipsative assessment of children's ability. Journal of Special Education, 25(4), 504-526.

McDermott, P.A., Glutting, J.J., Jones, J.N., Watkins, M.M., & Kush, J. (1989). Identification and membership of core profile types in the WISC-R national standardization sample. Psychological Assessment, 1, 292-299.

Mezzich, J. & Solomon, H. (1980). Taxonomy and behavioral science: Comparative performance of grouping methods. New York: Academic Press.

Mojena, R. (1977). Hierarchical grouping methods and stopping rules: An evaluation. Computer Journal, 20, 359-363.

Mojena, R. & Wishart, D. (1980). Stopping rules for Ward's clustering method. Proceedings of COMPSTAT 1980, 426-432.

Quay, H.C., & Peterson, D.R. (1979). Manual for the Behavior Problem Checklist. Available from author.

Reed, M.T., & McCallum, R.S. (1995). Construct validity of the Universal Nonverbal Intelligence Test (UNIT). Psychology in the Schools,32, 277-290.

Rourke, B.P. (ed.). (1985). Neuropsychology of learning disabilities. New York: Guilford Press.

Sattler, J.M. (1988). Assessment of children (3rd ed.). San Diego, CA: J.M. Sattler.

Sattler, J.M. (1992). Assessment of children (3rd ed., rev.). San Diego, CA: J.M. Sattler.

Skinner, H.A. (1978). Differentiating the contribution of elevation, scatter, and shape in profile similarity. Educational and Psychological Measurement, 38, 297-308.

Skinner, H.A., & Blashfield, R.K. (1982). Increasing the impact of cluster analysis research: The case of psychiatric classification. Journal of Consulting and Clinical Psychology, 50(5), 727-735.

Sneath, P., & Sokal, R. (1973). Numerical taxonomy. San Francisco: Freeman.

Sokal, R. & Michener, C.D. (1958). A statistical method for evaluating systematic relationships. University of Kansas Scientific Bulletin, 38, 1409-1438.

Tatsuoka, M.M., & Lohnes, P.R. (1988). Multivariate analysis (3rd ed.). New York: Macmillan.

Tryon, R.C. & Bailey, D.E. (1970). Cluster analysis. New York: McGraw-Hill.

Ward, J.H., Jr. (1963). Hierarchical grouping to optimize an objective function. American Statistical Association Journal, 58, 236-244.

Ward, T.J., Ward, S.B., Glutting, J.J., & Hatt, C.V. (1999). Exceptional LD profile types for the WISC-III and WIAT. School Psychology Review, 28(4), 629-643.

Wechsler, D. (1974). Wechsler Intelligence Scale for Children - Revised. New York: Psychological Corporation.

Wechsler, D. (1981). Wechsler Adult Intelligence Scale - Revised. New York: Psychological Corporation.

Wechsler, D. (1991). Wechsler Intelligence Scale for Children - Third edition. San Antonio, TX: Psychological Corporation.

Werry, J.S. (1968). Developmental hyperactivity. Pediatric Clinics of North America, 15, 581-599.

Wishart, D. (1982). CLUSTAN users manual: Supplement (3rd ed). Edinburgh, Scotland: Program Library unit, Edinburgh University.

Zachary, R.A. (1990). Wechsler's intelligence scales: Theoretical and practical considerations. Journal of Psychoeducational Assessment, 8, 276-289.

VITA

The author was born and raised in the small rural community of Chuckey, Tennessee which is located in the heart of Appalachia. He attended public schools in Greene County, Tennessee and graduated from Chuckey-Doak High School located in Afton, Tennessee. After attending Tusculum College for one year, the author graduated with a Bachelor of Arts degree with Honors from the University of Tennessee - Knoxville with a major in Psychology. Subsequently, he attended the University of Tennessee - Knoxville and graduated with a Master of Science degree in Educational Psychology. He was awarded the Most Outstanding Master's Student of the College of Education for his graduating class. Finally, the Doctor of Philosophy degree was awarded from the University of Tennessee with a major in Education / School Psychology.