Master's Theses   Theses and Dissertations

1984

WISC-R Performance in a Clinical Population Classified as Overcontrolled and Undercontrolled

Mary M. Gonzalez
Loyola University Chicago

$3\,\rho$

WISC-R PERFORMANCE IN A CLINICAL POPULATION

CLASSIFIED AS

OVERCONTROLLED AND UNDERCONTROLLED

by

Mary M. Gonzalez

A Thesis Submitted to the Faculty of the Graduate

School of Loyola University of Chicago in Partial

Fulfillment of the Requirements for the Degree of

Master of Arts

November

1984

ACKNOWLEDGMENTS

VITA

The author, Mary M. Gonzalez, was born on June 10, 1958 in Astoria, New York. She is the daughter of Ralph and Joan Gonzalez.

She obtained her secondary education at the Academy of the Sacred Heart in Bloomfield Hills, Michigan. She attended James Madison College at Michigan State University and in June, 1980 received the degree of Bachelor of Arts. Her majors were in psychology and political science.

The author entered Loyola University of Chicago's doctoral program in clinical psychology in the Fall of 1981. She began her graduate clinical training at the Charles I. Doyle Center and Day School of Loyola University in August of 1981 with a two year clerkship. In the Fall of 1983, the author was awarded a National Institute of Mental Health fellowship. The author currently holds a clinical externship position at the Illinois State Psychiatric Institute.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

Despite Wechsler's (1949) primary intention to design and organize a test of general knowledge, the broader assessment potentialities of the Wechsler Intelligence Scale for Children (WISC) were appreciated by creator and colleagues alike. In his review of the instrument, Wechsler concluded that the instrument lends itself to diagnostic assessment. As demonstration of faith in his conclusion, Wechsler observed that, adolescent sociopaths as a group, characteristically score higher on Performance IQ than Verbal IQ (1949). With the precedent established, others began to investigate the clinical usefulness of the WISC in formulating diagnostic statements.

More than three decades and one revision later, the Wechsler scales continue to be a major object of research. Wechsler (1974) in the manual of the Wechsler Intelligence Scale for Children-Revised (WISC-R), acknowledged the considerable research directed toward the identification of WISC-R syndrome patterns. It is thought that the robustness and stablility of the instrument accounts for its popularity (Shiek & Miller, 1978). Regardless of the reason, the search to evaluate the clinical utility of the WISC-R has generated numerous

1

studies involving a cross section of normal children, as well as atypical populations of emotionally disturbed, mentally impaired, and behavior disordered.

Others have chosen to investigate the factor structure of the WISC-R to ascertain whether there might be more clinically useful ways to organize WISC-R data. Reexaminations of the WISC-R standardization sample to document the existence of meaningful factors akin to the Verbal and Performance scales have proved fruitful. Kaufman (1975), in one of the first reexaminations of WISC-R normative sample, identified a three factor solution. His and others' efforts have not gone unnoticed by fellow researchers. More recent investigators (e.g., Hodges, Horwitz, Kline, & Brandt, 1982), have incorporated the three factor solutions to evaluate whether they provide novel information not found in the traditional two factor solution.

Even the most cursory exposure to the literature forces the reader to conclude that, despite extensive research efforts, the diagnostic utility of the WISC-R (as with the original) is, as yet, undetermined. The surrounding haze of conflicting results is no doubt, multidetermined and to single out one factor could be seen as simplistic. However, few would argue that the ambiguous, if not arbitrary, nature of classification system of emotional disorders is a likely candidate. Almost without exception, studies have shown a lack of agreement in the definition of emotional disturbance. The absence of

a sophisticated taxonomy of emotional disturbance, particularly among childhood disorders, confounds research efforts which demand clear differentiation among clinical groups.

The major classification system, the Diagnostic and Statistical Manual of Mental Disorder, is currently in its third edition (1980). Although the present edition is an improvement over its predecessors, it remains subject to criticism. Fault is found with its professed atheoretical foundation. DSM III is perhaps better described as a hybrid of theoretical systems, a hybrid which often leads to conflicting assumptions and conclusions. The system is also characterized by its narrative type descriptions of behaviors with no procedures for operationalizing them (Achenbach & Edelbrock, 1978). At best, the system is thought to provide mediocre reliability and validity. The present state of disarray in diagnostic assessment and terminology transcends research in WISC-R diagnostic utility since the classfication of emotional disturbance is a prerequiste in the evaluative efforts.

Another contributor to the prevalance of inconsistent research findings are the varying expectations that are elicited when one considers the diagnostic utility of an assessment instrument. Some expect that if the WISC-R is to lay claim as a diagnostic tool, it must have the precision of a classification metric. Others are less demanding, concluding that since many professionals use the WISC-R to

formulate diagnostic hypotheses, it is deserving of the description. Notwithstanding personal idiosyncracies in definitions, most would hope that WISC-R interpretive powers would legitimately include diagnostic usage. Perhaps, WISC-R diagnostic powers could become so exact as to be directly transferable into the process of differential diagnosis. This is certainly an area in need of refinement and the possibility that the WISC-R could contribute to the process cannot be prematurely rejected.

The present investigation examined the discriminatory utility of using WISC-R scores in differentiating between two clinical groups. The sample consists of 40 WISC-R protocols of male outpatients, who were labeled as Overcontrolled or Undercontrolled according to Davison and Neale's conceptualization (1982). The system is based on Achenbach's (1966) extensive factor analytic studies and subsequent identification of the Internalizer and Externalizer syndromes. It was postulated that if the two groups perform differentialy on the WISC-R, diagnostic formulations based on WISC-R measures would have an empirical basis.

CHAPTER II

REVIEW OF THE LITERATURE

Historical Perspective

Since the beginning of the testing movement, the investigation
of the effect of emotional factors on intelligence and performance on
IQ tests has interested psychologists. Attention of researchers has
been directed toward children, since childhood represents a critical
developmental period of intellectual growth (Piaget, 1952). Investi-
gations concerning the diagnostic usefulness of the Wechsler scales
have been numerous since its introduction. The general consensus is
that the Wechsler Intelligence Scale for Children-Revised (WISC-R) is
a stable and reliable instrument for evaluating children with emo-
tional disorders (Sattler, 1982). It is thought that the popularity
of the Wechsler scales is related to their potential use as a diagnos-
tic tool for emotional disturbance (Saccuzzo & Lewandowski, 1976).
Wechsler (1949) established the precedence of gathering diagnostic
information from the scaled IQ scores with his assertion that among
adolescent sociopaths, a Performance IQ greater than the Verbal IQ is
the characteristic pattern. However, further research has yielded
conflicting results concerning the legitimacy of treating the WISC-R

as a diagnostic aid (Hale & Landino, 1981; Hamm & Evans, 1978; Paget,1982; Schooler, Beebe, & Koepke, 1978)

The Wechsler Intelligence Scale for Children was originally published in 1949 and twenty five years later the revised edition was introduced. The revised editon has carried the tradition of application as a classifactory aid. Wechsler (1974) concluded that, "the scale (WISC-R) as a whole remains structurally and contextually the same" (p.iii). WISC-R content is quite similar to the WISC, with 78% taken directly from the WISC; 5.9% from the WISC but have undergone modification; and, 16.1% new items (Wechsler,1974). The WISC-R, as with the WISC, offers Verbal, Performance, and Full Scale IQs, with a mean of 100 and standard deviation of 15. For both instruments, the Verbal and Performance scales are comprised of six subtests, which yield scaled scores with a mean of 10 and a standard deviation of 3. Studies have indicated that the WISC-R has validity similar to the WISC (Wechsler, 1974).

Significant score differences between the WISC and WISC-R have been reported in the literature almost without exception. Studies indicate that the WISC-R provides lower scores of approximately 1/3 to 1/2 standard deviation for the three major scales. Doppelt and Kaufman (1977) reported that on the average, WISC-R IQs are four points lower than the WISC. Swerdlik (1977) states that the conflicting conclusions do not reflect negatively on the validity of the WISC-R but

rather reflect the influence of different normative groups. The new standardization sample, which includes 100 male and 100 female children at each of the age levels from 6-6 to 16-6, was designed to represent every racial, socio-economic stratum of the general population of normal children, and is considered perhaps the most significant improvement of the revision.

Swerdlik and Schweitzer (1978) suggest that since the content and structure of the WISC and the WISC-R are comparable, the factor structures are likely to be similar. In their study they requested 72 school psychologists in the tri-state area of Michigan, Illinois, and Ohio to administer the WISC and WISC-R to 164 black, white, and latino children in a counterbalanced order with a specific test retest interval of not less than one week nor more than a month. The children were referred to school psychologists because of concerns about their intellectual abilities. Swerdlik and Schweitzer's results indicated that the tests are quite similar in factor structure. In 1981 McMahon and Kunce examined the clinic records of children served at a midwestern university medical center. The first sample consisted of 120 male caucasians between the ages of 6 years-0 months and 9 years-11 months. who received Full Scale IQs greater than 85 on the WISC. A second sample included 67 children between the ages of 6-0 and 12-9 with Full Scale IQs greater than 85 on the WISC-R. The results from their analyses supported Swerdlik and Schweiter's earlier finding within groups of exceptional children with various psychoneurological diagnoses.

WISC-R Factor Analytic Studies

Factor analyses of the revised edition of the instrument have produced stable results with a wide variety of subject populations and statistical techniques (Wechsler, 1974). However, perhaps more critical were investigations into the internal stability and consistency of the factor structure of the standardization sample. Kaufman (1975) was among the first to factor analyze the standardization data of the WISC-R. Analyses produced three factors which he labelled Verbal Comprehension (VC) which includes Information, Similarities, Vocabulary, and Comprehension, Perceptual Organization (PO) which includes Picture Completion, Picture Arrangement, Block Design, and Object Assembly, and Freedom from Distractability (FD) which consists of Coding, Arithmetic, and Digit Span. Kaufman notes the close similarity between his factors and Wechsler's Verbal and Performance scales and concluded that this gives strong empirical support to the Wechsler dichotomy (1975). Kaufman states that there is substantial evidence in support of the Wechsler combination of the Verbal and Performance scaled scores to yield a Full Scale IQ. He found a large general factor accounting for 79 to 92% of the common factor variance. Furthermore, Kaufman found stronger factor analytic evidence for the construct validity of the WISC-R as compared to its predecessor (1975).

Wallbrown, Blaha, Wallbrown, and Engin (1975) also factor analyzed the WISC-R subtests in the standardization sample and found a

strong general factor and two subgeneral factors corresponding to a verbal and a perceptual scale. Reynolds and Gutkin (1980) observed a similar stability of the three factor solution across sex and age. They concluded that this supports the uniform interpretation of the WISC-R factor scores independent of the child's sex and age. In 1978 Shiek and Miller assessed the robustness of the WISC-R factor structure by comparing a sample of 126 children (62 males and 64 females with a mean age of 10.6 years) from lower and lower middle class homes. Eighty-seven of the children were white and 39 were black. A preliminary analysis revealed two basic differences between their sample and the standardization sample. First, the Verbal, Performance, and Full Scale IQs were significantly lower and the variances on the Performance and Full Scale variables were significantly restricted compared to the standardization sample. The factor structure, however, was highly consistent with the standardization sample.

The clinical usefulness of the WISC-R is enhanced by the generally agreed upon existence of a group of factors that coincides with the structure of the test (Schooler, Beebe, & Koepke, 1978). The research on factor structure cited above has been with normal subjects, others have assessed the generalizability of the factor structure with populations of exceptional children. The importance of these research efforts cannot be minimized if the WISC-R continues to serve as a diagnostic/ classifactory aid for psychologists.

In the search to ascertain whether the structure of intelligence is invariant thereby justifying the use of an intellectual assessment instrument to perform diagnostic functions, some researchers have chosen to study the population of the intellectually limited as the data base. Van Hagen and Kaufman (1975) factor analyzed WISC-R scores for a group of 80 children with Full Scale IQs ranging from 40 to 79. Their results yielded the same factors that had emerged with the standardization sample. Groff and Hubble (1982) examined the WISC-R records of 103 male and 82 female youths with Full Scale IQs in the mildly retarded and borderline intellectual range. The groups were then divided into a younger (mean age 10 years-4 months) and older group (mean age 15 years-2 months). The Verbal Comprehension and Perceptual Organization factors were indentified for both groups although the Freedom from Distractability factor was apparent only for the younger sample.

Hodges (1982) was interested in the generalizability of the factor structure in a psychiatric sample. Two hundred and forty children who received outpatient services at a community center served as the sample. Children with a diagnosis of mental retardation or a specific learning disability were excluded. Approximately half of the sample were diagnosed adjustment reaction. The 163 boys had a mean age of 11.6 years; the girls had a mean age of 11.9 years. The mean Full Scale IQ was 95.08 with a standard deviation of 12.03. All three Kaufman factors emerged with the Freedom from Distractability somewhat

weaker than the other two factors. Schooler, Beebe, and Koepke (1978) obtained factor loadings for the WISC-R scores for 799 children identified by school personnel as needing special education services. Children were classified by each school's legally mandated Educational Placement and Planning Committee. Categories of classification used were Educable Mentally Impaired (EMI), Learning Disabled (LD), Emotionally Impaired (EI), an Other category for those children needing special services but not meeting the requirements for an EMI, LD, or EI diagnoses, and, finally, a None category for those children not needing special services. Of the 799 chldren, 275 were classified as LD, 127 as EMI, 69 as EI, 59 Other, and 209 children were labeled None. Schooler et al. concluded that the factor structure is remarkably similar for all clinical groups.

Peterson and Hart (1979) examined the stability of the factor structure for a clinic referred population using the factor labels described by Kaufman. Six hundred and fifty-four second through sixth grade children who had been identified because of learning and/or behavioral problems comprised the sample. The sample was divided into groups labeled: learning disabled, mentally retarded, emotionally handicapped, slow learner, culturally disadvantaged, and no significant problem. Factor analyses showed the Verbal Comprehension and Perceptual Organization factors clearly in evidence. The third factor, Freedom from Distractability, was much less stable. Peterson and Hart (1979) conclude that there might be some genuine differences on this third factor between normal and clinical populations.

DeHorn and Klinge's (1978) research supported these earlier findings. They administered the WISC-R to 100 adolescents (52 male and 48 female) with either inpatient or outpatient status. The factors that emerged were similar to the earlier findings and they concluded that the WISC-R scores of an adolescent psychiatric sample can be compared meaningfully to those of retarded or normal children. The WISC-R protocols of 100 children referred for behavior problems in school were data in the Finch, Kendall, Spirito, Enfin, Montgomery, and Schwartz study (1979). Factor analysis revealed two factors, Verbal Comprehension and Perceptual Organization which corresponded to the subtest dichotomy. The Freedom from Distractability factor was absent. Hodges (1982) suggested that the results may have been confounded because the mean IQ for the sample was in the borderline intellectual range.

The overwhelming consensus in the literature is in favor of the presence of a least two meaningful factors in the WISC-R which correspond to the Verbal and Performance scale IQs. This has been found for both normal populations on a wide variety of demographics as well as the more recent research focused on clinical populations. Reynolds and Gutkin (1980;1981) conclude that the factors of the WISC-R are essentially invariant across sex and age with regard to the pattern of factor loadings. The three factors of VC, PO, and FD account for a similar percentage of the total variance across groups, the magnitude of the factor loadings are also similar. Some authors have suggested

the use of factorally pure measures as substitutes for the Wechsler Verbal and Performance IQs (Gutkin, 1982).

Interpreting WISC-R Score Discrepancies

Peterson and Hart (1979) observe that it is essential that there be comparable qualitative attributes measured by the instrument if qualitative and quantitative comparisons are to made. The presence of a stable factor structure across clinical groups and normal populations suggests the possibility of an exploration as to the clinical usefulness of factor scores in differential diagnosis. Typically analyzed differences among diagnostic groups are Verbal minus Performance discrepancy scores, subtest scatter, and subtest patterning. It is speculated that all categories of psychopathology have a higher intertest scatter and lower performance than one would expect from a normal population (Dean, 1977; Wechsler, 1949). The variability in performance is assumed to be the result of emotional factors or maladaptive life styles (Dean, 1977).

Evidence regarding the diagnostic utility of the WISC-R is inconsistent. Given the existence of the one diagnostic pattern of adolescent sociopaths (Wechsler, 1949) the possibility exists for others, hence the search continues. Protocols of 80 adolescents who had been tested at the Center for the Study of Crime, Law Enforcement, and Corrections at Tennesse State University served as the sample in the Saccuzzo and Lewandowski study (1976). The adolescents were from pre-

dominantly lower socio-economic level; ages ranged from 13 to 16 years with a mean of 14.95 years; Full Scale IQ was between 80-89 with a mean of 85.19. The results showed a highly significant difference in favor of Performance IQ. Dean (1977) examined the WISC-R profiles of 41 male Caucasian adolescents described as conduct disordered. The age range of the sample was 13 to 15 years and the Full Scale IQ ranged from 80 to 105. Dean observed a general depression of verbal functions. Furthermore, he noted a tendency for the verbal and performance subtest scores to scatter more widely from their respective means than in nonclinical populations.

In a later study, Dean (1978) utilized a stepwise discriminant analysis to evaluate the subtest scores of 60 learning disabled youths and a matched sample of emotionally disturbed children. Forty-eight caucasian males and 12 caucasian females comprised each sample. The results indicated that the Emotionally Disturbed (ED) children had significantly higher performance means than that obtained on the verbal scale. The differences between the Verbal IQ (VIQ) and Performance IQ (PIQ) for the Learning Disabled (LD) group were not significant. There were four subtests that were found to differentiate significantly between the two samples. Morris, Evans, and Pearson (1978) reviewed the WISC-Rs of 113 youths (88 male and 25 female; 71 white and 42 black) ranging in age 6 years-11 months to 13 years-8 months. All had been classified as severely emotionally disturbed. Their profiles showed significantly smaller scaled score means on all ten subtests in comparison with the standardization sample.

Other attempts to analyze the diagnostic utility of the WISC-R have not met with such successs. Sattler (1982) concludes that there are no WISC-R patterns that have been found that can reliably distinguish between various groups of emotionally disturbed children, although there may be greater variabilty of scores in some ED children. A major conceptual difference between researchers supportive of the gathering of diagnostic information from the WISC-R and those disapproving of such usage, often centers around the degree of specificity that is demanded from the instrument. The criterion on which the alleged diagnostic utility of the WISC-R is evaluated ranges from an instrument which is expected to make post facto dichotomous distinctions to more sophisticated maneuvers requiring prediction and categorization into diagnostic groups.

A noteworthy effort founded upon a generous amount of faith in the WISC-R is the 1967 Fernald and Wisser study. Their study pursued Wechsler's observation regarding the higher Performance IQ scale than Verbal IQ. The hypothesis was that given a group of juvenile offenders, the amount by which the PIQ exceeds VIQ would indicate the degree of acting out. The data were the records of 72 male caucasian juvenile offenders ranging in age from 12 to 15 years and referred to a detention center for clinical evaluation. Excluded from the study were youths with diagnoses of organic damage, psychosis, or mental deficiency. The results failed to support little if any justification for using the Verbal minus Performance discrepancy score as a predictor of acting out.

Many would consider the Fernald and Wisser (1967) study premature if not unrealistic. Perhaps first among the contenders would be Sattler who questions the very premise of this and other studies which hope to assess the alleged diagnostic utility of the WISC-R. Sattler (1974) asserts that in general, nonpsychotic conditions do not seriously affect the overall level of intellectual performance, although he acknowledges Wechsler's observation of delinquent youths, but suggests that the finding is unique. Further, according to Sattler, there is no evidence to support the assumption that pathology and scatter are necessary linked (1982).

Others share Sattler's negative evaluation of the WISC-R diagnostic utility. Schoonover and Hertel (1970) concluded from their study that diagnostic categories are not readily differentiated by WISC scores alone. They had analyzed the WISC scores of 351 children from nine diagnostic categories, in terms of Verbal minus Performance differences, subtest scatter, and subtest patterning. The WISC scores seem to differentiate between two groupings of categories rather than among the categories themselves. McMahon and Kunce (1981) found the relationship between WISC and WISC-R and interdisciplinary diagnoses although consistent with clinical expectations, to be weak and therefore, concluded that it "raises questions about the adequacy of the Wechsler scales as aids in the diagnostic process" (p.410).

More recently, Hale and Landino (1981) used 100 caucasian males aged 7 to 16 years as the subjects in their study. Conduct disorders, withdrawn disorders, mixed, and nonproblem were the criterion groups formed on the basis of ratings from the Behavior Problem Checklist (Quay & Peterson, 1979). Their first hypothesis tested was to determine if the WISC-R could discriminate among the criterion groups. This was supported. However their second hypothesis was not supported; namely the WISC-R was unable to classify subjects at a rate significantly greater than chance. The obtained accuracy rate was only 66%. Hale and Landino (1981) concluded that, "although the WISC-R subtest analysis may not be of value for the discrimination of one diagnostic group from another, it may still have clinical utility for generating hypotheses about the intellectual functioning of individual children" (p.94).

Vance, Singer, Kitson, and Brennner (1983) are quite pessimistic concerning the final outcome decision regarding the diagnostic utility of the WISC-R. They claim that the continual search for a specific diagnostic pattern that will discriminate LD and ED children is a relatively fruitless task and a single clear-cut pattern that is characteristic of these children is not expected to emerge. Their criticism is harsh. They state that intellectual patterning in itself should not be a basis for assigning children to classes for brain damaged or for behavioral disorders. "The continual use of diagnosing LD and ED children solely on the WISC-R subtest scatter or verbal performance

discrepancy score, in light of all the results against these practices might be considered a case of malpractice" (Vance et al., 1983, p.130). Schoonover and Hertel (1970) and Schooler, Beebe, and Koepke (1978) observed a similar lack of distinct WISC-R patterns of subtest scores for children lumped in the heterogeneous emotionally impaired group. Schooler, Beebe, and Koepke (1978) suggest that the WISC-R is not suitable for distinguishing among groups. However, the results of their study were later questioned because of their failure to include the Digit Span scores into their computations (Hodges,1982).

Thompson (1980) assessed whether groups of children determined to be clinically distinct through an interdisciplinary evaluation differed significantly among themselves and in comparison with the standardization sample on various WISC-R measures. The traditional summary scores of VIQ, PIQ, and FSIQ reflected differences with the standardization sample having the highest score, followed by the emotional disordered group, then the learning disabled, and finally those diagnosed as mentally retarded. The VIQ minus PIQ score did not differ significantly among the clinical groups or between any of the clinical groups and the standardization sample. While there were demonstrated differences among the composite scores, Thompson (1980) concluded, given the general scarcity of significant difference especially between children with psychological disorders and those with learning disabilities, caution should be exercised with regard to utilizing the WISC-R as a diagnostic tool.

WISC-R <u>Normative Sample</u>: <u>Implications for Profile Analysis</u>

The assumption of WISC-R profile patterns as having diagnostic and remedial implications remains questionable. One of the difficulties has been the derivation of clinically meaningful and scientifically justified methods of interpreting individual WISC-R profiles. Sattler (1982) states that the cardinal rule for the use of profile analysis is dependent upon the presence of statistically significant differences between Verbal and Performance scale IQs and between subtests scaled scores. The use of profile analysis in itself presents problems, according to Sattler, because the subtests are not as reliable as the three IQs obtained on the test. The primary methods of profile analysis are the comparison of PIQs and VIQs, each Verbal subtest scaled score to the mean Verbal scaled score, each Performance subtest scaled score to the mean Performance scaled score, each subtest scaled score to the mean subtest scaled score, and sets of individual subtest scores.

While scatter is not fortuitous, Sattler cautions that other factors separate from emotional variables could be accountable such as, age, sex, racial or ethnic membership, socio-economic-status, parental level of education, social or physical environment, family background, and parental occupation. Of course, the possibility always exists that scatter is simply a reflection of the unreliability of the individual subtest scores, examiner variables, situational

variables, rather than even a reflection of cognitive strengths and weaknesses (Sattler, 1982). Guertin (1966) similarly stressed the importance of systematically considering these relevant variables, and that score configurations are a function of a variety of factors. Sattler (1982) reminds his readers that the goal of profile analysis is not to classify or categorize children, but rather to enhance the examiner's understanding of children's abilities.

Perhaps in response to the growing trend toward holistic interpretation of the WISC-R (Dean, 1977), several researchers have reexamined the normative sample with the idea that one needs to know the characteristics of the normative sample before one can make comparative statements. While this appears to be a reasonable conclusion it has been subject to frequent oversight. A likely contributing cause has been a bit of confusion regarding the nature of factor analysis. Blaha and Wallbrown (1984) address this dilemma. They state that, "the very nature of factor analysis is such that research findings cannot be applied appropriately to individuals without the intervention of clinical judgment. It can suggest a series of likely hypotheses but only a sensitive clinician can determine which, if any, of these hypothesis are applicable to an individual " (p.566).

There is a general consensus that factor analysis while it can contribute to the research endeavors regarding the diagnostic utility of the WISC-R, it alone can not provide the answer. Since diagnostic

and clinical meaning are frequently attributed to VIQ-PIQ discrepancy, several researchers have taken discrepancy scores to be a natural starting ground for their investigations. Clampit, Adair, and Strenio (1983) present a table of frequencies with which a discrepancy is manifest on the basis that the frequency may be the more meaningful statistic for clinicians. The table is designed not only to identify those cases where a discrepancy is rare but also to minimize the over-interpretation of minor discrepancies that, although may be statistically significant, may well be clinically unimportant. Kaufman (1976) in his reexamination of the normative sample, found that 45% of the children had discrepancies of nine or more points, 34% had 12 or more point discrepancies, and 25% of the sample had a 15 or more point VIQ minus PIQ score. He computed that the mean VIQ-PIQ discrepancy was 9.7 (SD=7.6), demonstrating that the average child could have a significant (p<.15) VIQ-PIQ discrepancy.

In further analysis of the normative sample, Kaufman (1976) found that the average child had about one subtest score that deviated significantly (greater than or equal to three scaled score points) from their own VIQ or PIQ mean and about two subtest scores that deviated significantly from their own Full Scale mean. Therefore, he concluded that considerable scatter, in terms of significant subtest differences is common and the "flat profile" often associated with normal children is a myth. The fact, Kaufman (1976) states, that the discrepancy is normal in its magnitude, suggests that it may be unrelated to

the ultimate diagnosis. Kaufman proposes computing the means of the various subtest grouping or indices and comparing them to the overall mean of the respective scale with deviations of three points significant for individual subjects. Kaufman states that test interpreters are fortunate that the factor analysis of the WISC-R generally support specific interpretations since there is sufficient test specificity (1979). He urges that if a hypothesis is generated from one scale, that the interpreter examine the subtests from the other scale which tap the same ability for a verification or rejection. Overall, Kaufman (1979) offers his support for the clinicians interpretation of the child's profile of scaled scores.

Hodges, Horwitz, Kline, and Brandt (1982) comment on the hesitancy to explore other summary scores in studying the WISC-R. The purpose of their study was to determine whether the three factors (Verbal Comprehension, Perceptual Organization, and Freedom from Distractability) identified by Kaufman differ in apparent clinical utility from the traditional summary scores. The subjects were 177 male and 83 female psychiatric outpatients from a community mental health clinic. The independent variables were sex and DSM II diagnosis. Children with a diagnosis of mental retardation or with a FSIQ less than 70 were excluded. Adjustment reactions of childhood or adolescence were common, describing half of the subjects. Thirty-one subjects were classified as overanxious, 18 as hyperkinetic, 23 as conduct disorder, 40 as Learning Disabled (LD) without Behavior Disorder

(BD), and the remaining 19 were diagnosed as LD with BD. Significant main effects for diagnosis for the verbal summary scores (VIQ and VC) and the FD factor. There was no difference between the two summary scores for performance, the main effect for diagnosis was nonsignificant for PIQ as well as PO. The Verbal-Performance discrepancy score also lacked significance.

Hodges et al.(1982) conclude, based on their findings, that the Kaufman scores yield useful information about WISC-R performance that is not made available by the traditional summary scores. The Kaufman factors conveyed novel information about several diagnostic groups specifically antisocial, hyperactive, and LD. Delinquent children performed more poorly than the others on the VC factor; hyperactive children had the highest VC score and a relatively low score on the FD factor; LD children showed deficits on both verbal summary scales and on the FD factor. These scores are consistent with the clinical impressions of these groups. In contrast the traditional scores yielded a difference only for VIQ. Hodges et al.(1982) concluded, "important information that was contained in the 11 subtest scores would have been overlooked if reference had been made only to the traditional summary scores" (p.837).

Taxonomies of Childhood Disorders

There remains those researchers who are doubtful of the alleged diagnostic usefulness of the instrument (Vance, Singer, Kitson, & Brennner,1983). Thompson (1980) suggests that one of the difficulties that immediately confront researchers is the problem in arriving at precisely defined categories of childhood disorders, particularly among Behavior Disordered and various Learning Disabled children. Achenbach and Edelbrook (1978) conclude that the "study of psychopathology in children has long lacked a coherent taxonomic framework within which training, treatment, epidemiology, and research could be integrated" (p.1275).

Others have called attention to the confounding effects of imprecise categorization of childhood disorders on the assessment process of the alleged diagnostic utility of the WISC-R (Davison & Neale, 1982; McDermott, 1980). The absence of well established clinical categories against which to validate profile types prevents systematic attempts at categoric prediction (Achenbach, 1978; Dean, 1977). Typically, childhood disorders are classified according to the Diagnostic and Statistical Manual (DSM). The third edition (American Psychiatric Association, 1980) attempted to correct some of the weaknesses of the earlier publication, such as the narrative style with an absence of explicit criterion and paucity of possible diagnoses of childhood disorders. DSM III boasts of a wider variety of possible diagnoses

(40 specific diagnoses organized under nine general groups) that are specific to the developmental period and not simply scaled down versions of adult disorders. Furthermore, there is an increased emphasis on operationally defining the disorders.

Davison and Neale (1982) note that the major influence on the classification of childhood disorders has been factor analytic studies of symptoms. A most notable investigation was conducted by Achenbach (1966) in his analysis of the case histories of 300 psychiatric inpatients and outpatients. He concluded that the dichotomy of Internalizer/Externalizer aptly included most subjects. Encouraged by these findings, Achenbach (1978) later published The Child Behavior Profile with the goal of developing a, "descriptive classification system that could be used to group children for research and clinical purposes, to reflect adaptive competencies as well as behavior change" (p.478). In 1978 Achenbach published an extensive review and analysis of empirical efforts in the classification of child pathology. He demonstrated consistent justification for the dichotomous categorization of syndromes into internalizing and externalizing.

Davison and Neale (1982), in their presentation of childhood disorder refer to Achenbach's broadband undercontrolled and overcontrolled syndromes. The undercontrolled or externalizers are characterized by behavior excesses and the overcontrollers or internalizers show behavior deficits. The key to the distinction, "lies in whether

the child's way of reacting, such as disobedience, creates more of a problem for others or, such as anxiety, affects the self" (Davison & Neale, 1982, p.457). The authors suggest that the DSM III diagnoses which would qualify as undercontrolled are the conduct disorders, attention deficit disorders, and diagnoses of hyperactivity. The diagnoses of anxiety, avoidance, withdrawal disorders, somatic symptoms, and childhood fears would be best described as overcontrolled. The rationale for Davison and Neale's categorization system seemingly rests on the ease of matching the detailed criterion for the DSM III diagnosis with elaborate description of the syndromes in the factor analytic studies as well as its intuitive appeal.

A classification system in order to be useful must offer a minimal amount of ambiguity to those applying its standards (Davison & Neale, 1982). Reliablity becomes the primary prerequisite for evaluating a classification system. However the paradox is recognized if one considers the typical means of assessing reliablity, i.e., whether diagnosticians agree (Davison & Neale, 1982). There are few who would argue that this method is anything but infallible. According to McDermott (1980), errors of inconsistency in the diagnostic process can result from misapplication of criteria in rendering diagnosis or misadoption or mixed adoption of theoretical schemas from which to evaluate.

The difficulty of clinical diagnoses is apparent when one considers that the criteria of nosological entities are of doubtful validity and moreover, the reliable relationship between test performance, selective impairment or personality characteristics is at best uncertain. Achenbach and Edelbroock (1978) describe the state of the art as having a "bootstrapping" quality, that is, "investigators have attempted to lift themselves up by their own bootstraps by establishing relationships among measures all of which are known to be imperfect" (p.1290). Albeit the dismal atmosphere, researchers continue in their efforts perhaps motivated by the acceptance that increments of knowledge can be gleaned from even the most ambiguous study.

The present study has not only borrowed Achenbach's terminology but also his critical attitude. Rather than utilizing narrow band syndromes which have inconsistent support in the literature, the broad band syndromes of overcontrolled and undercontrolled will act at the diagnostic categories because of their clinical usefulness and strong empirical foundation. Davison and Neale's (1982) conceptualization of the Undercontrolled and Overcontrolled syndromes with respect to DSM III diagnoses will be applied since it appears sufficiently grounded in theory and empirical analysis. Furthermore, the Kaufman factors as well as the traditional summary scores on the WISC-R will be computed and analyzed since they have demonstrated clinical usefulness (Hodges, 1982; Hodges, Horwitz, Kline, & Brandt 1983; Thompson, 1980).

Hypotheses

In the present study it was postulated that the two clinical groups would be differentiated by WISC-R scores in the following manner:

1. For the Undercontrolled clinical group, the Performance IQ will be greater than the Verbal IQ.

2. The Undercontrolled clinical group will reflect a larger Performance IQ compared to the Overcontrolled clinical group.

3. For the Undercontrolled clinical group, the Perceptual Organization factor score will be greater than the Verbal Comprehension factor score.

4. The Undercontrolled clinical group will exhibit a larger Perceptual Organization factor score compared to the Overcontrolled clinical group.

5. The scatter and range scores of the two clinical groups will be similar, and will be significantly greater than the normative sample.

Although no specific hypotheses are formulated with regard to Kaufman factors, these factors will be evaluated to determine whether they provide clinical data not found with the traditional summary scores.

CHAPTER III

METHOD

Subjects

Data for 40 male children were obtained from an outpatient, com-
munity mental health center servicing a specific geographical region
of a large urban area. The center, affilitated with a major midwest-
ern university, also functions as a training institution for graduate
psychology and social work students. The center offers multidiscipli-
nary assessment and treatment services to children and their families.
For the most part the sample consisted of Caucasian children from mid-
dle to lower middle class backgrounds (36 Caucasian, 4 Hispanics).
All of the children were male and their ages ranged from 6 years, 3
months to 15 years, 2 months.

The psychological testing records of the 40 subjects who were
seen as outpatients at the center provided the data for this study.
The WISC-R is used as part of the center's standard psychodiagnostic
battery. The tests are administered by graduate clinical psychology
students under the supervision of doctoral level clinical psycholo-
gists. All children in this study had received a thorough multidisci-

plinary diagnostic evaluation. Primary diagnoses coded in the classi-
factory system of the Diagnostic and Statistical Manual (2nd & 3rd
editions) were recorded from the clinical files. In order to maximize
sample homogeneity the following exclusionary criteria were imple-
mented: a full scale IQ of less than 75, known history of thought dis-
order, good evidence of organic impairment, or a diagnosis of learn-
ing disability with no secondary diagnosis.

Instrument

The revised WISC, like its predecessor, was designed and organ-
ized as a test of general intelligence (Wechsler, 1974). It maintains
the original subdivison of the Scale into Verbal and Performance meas-
ures as well as the technique of weighting each of the component tests
equally to obtain the subject's IQ. Wechsler (1974) describes the
process as assortative in contrast to hierachical, the implication
being that each test is considered necessary for the full appraisal of
intelligence. The WISC-R was standardized on 2,200 white and nonwhite
American (e.g., Blacks, American Indians, Orientals, Puerto Ricans and
Mexican Americans) chosen to be representative of the population in
the same proportions reported in the 1970 United States census data.
Two hundred children in each of the eleven different age groups (6-6
to 16-6) comprised the sample.

Reliability The WISC-R has outstanding reliability (Anatasi,
1976; Kaufman, 1979; Sattler, 1982; Wechsler, 1974). Split-half and

retest reliability coefficients were computed for the WISC-R subtests and for the Verbal, Performance and Full Scale IQs. Each of the three IQ scales has a reliability coefficient of at least .89 in the standardization group over the entire age range. The average reliability coefficients are .94 for the Verbal Scale IQ, .90 for the Performance Scale IQ, .96 and for the Full Scale IQ, (Wechsler,1974). Albeit less, the average reliability for the individual Verbal tests with a range of .77 to .86, is still considered satisfactory. Reliability coefficients range from .70 to .85 for the individual Performance subtests.

In comparison to the original version, the WISC-R offers increased reliability of the Information, Comprehension, Arithmetic, Similarities, and Picture Completion tests at the younger age levels as a result of the incorporation of additional items of appropriate difficulty. Digit Span has more reliability on the WISC-R than its predecessor. This is thought to be due to the change in administration, which requires that both trials of an item be given, even if the first trial is passed (Kaufman, 1975). Considered separately, for each age group, the reliability coefficients range from a low of .57 for Mazes at the 16-6 age level to a high of .92 for Vocabulary at the 16-6 age level.

Standard Errors of Measurement The standard errors of measurements based on the mean performance of the eleven age groups, are 3.60

for the Verbal Scale, 4.66 for the Performance Scale, and 3.19 for the Full Scale. The Verbal Scale subtests tend to have smaller standard errors of measurement (ranges from 1.15 to 1.44) than the Performance Scale subtests which ranges from 1.17 to 1.70. The smallest standard errors of measurement within the Verbal Scale are the Information and Vocabulary subtests, while Block Design and Picture Completion have the smallest within the Performance Scale.

Stability The stability coefficients of the subtests and IQ scales were assessed by retesting a group of 303 children from six age groups in the standardization sample after a one month interval (Wechsler, 1974). Sattler (1982) reported stability coefficients of .95, .93, and .90 for the Full Scale, Verbal Scale, and Performance Scale IQs respectively. The median coefficient for the twelve subtests is .78 ranging from Mazes with a coefficient of .65 to Information with a coefficient value of .88. The gains in IQs values from the first to the second testing were approximately 3 1/2 points on the Verbal Scale, 9 1/2 points on the Performance Scale, and 7 points on the Full Scale. These increases were attributed to a practice effect. Wechsler (1974) concludes that the stability coefficients are of sufficient magnitude to lend further support to the overall reliability of the WISC-R.

Validity Anatasi (1976) commented on the absence in the WISC-R manual of a discussion on the validity of the instrument. She notes

that, "to be sure, normative tables of standard score equivalents for each subtest provide evidence of age differentiation, but no evaluation of the data in terms of the criteria is given" (p.259). Wechsler does address this issue, although perhaps in an indirect manner. He states that the WISC-R is not predicated on any one definition of intelligence. Wechsler describes the construct as, "the overall capacity of an individual to understand and cope with the world around him" (1974, p.5). Wechsler underscores the distinctiveness of his definiton both in his conception of intelligence as a multifaceted construct rather than a independent trait, as well as his reluctance to single out a particular ability as crucial.

Wechsler (1974) states that the the primary intent and function of intelligence tests is, "not to evaluate...a subject's cognitive abilities; nor is its purpose ... to appraise his educational, vocational, or other competencies" (p.1). However he concedes that they are inevitably used so. The information is relevant only to the extent that it establishes and reflects whatever it is that one defines as overall capacity for intelligent behavior. Wechsler is firm in his belief that a general intelligence exists and that it is possible to measure it objectively (1974). Furthermore, he asserts that the concept of an "intelligence quotient" is a scientifically sound and useful measure.
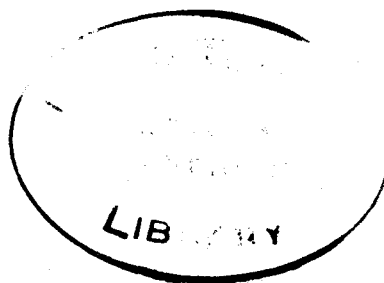
Wechsler (1974) concluded on the basis of available research, that the correlations between the three scales on the WISC-R and WPPSI were .80 for both the Verbal and Performance Scales and .82 for the Full Scale. Differences between the mean IQs on the Verbal, Performance, and Full Scale IQs were 1.5, 2.8, and 2.5 IQ points respectively with higher IQs on the WPPSI. Correlations between the WISC-R and the WAIS are even more impressive with .96 for the Verbal Scales, .83 for the Performance Scales, and .95 for the Full Scales (Wechsler, 1974). Differences between the mean IQs obtained on the two tests were less satisfactory, 5.3 for the Verbal Scale, 5.2 for the Performance Scale, and 6.2 for the Full Scale, with higher IQs on the WAIS than on the WISC-R.

Sattler (1982) discusses the criterion validity of the instrument in his presentation of studies correlating the WISC-R with other intelligence tests, and measures of achievement and school grades. Based on this research, Sattler concluded that when intelligence tests, receptive vocabulary tests, achievement tests, and school grades are used as criteria, the WISC-R has satisfactory concurrent validity. For example, the median correlations of the WISC-R to the Stanford-Binet is .75 for the Verbal Scale, .68 for Performance Scale, and .82 for the Full Scale; a correlation of .60 was found with the Peabody Individual Achievement Test; and a correlation of .68 was computed with the McCarthy Scales of Children's Abilities (Sattler, 1982).

Procedure

The sample was gathered by numerically working through the available psychological testing records at the center. Records selected included those with complete WISC-Rs except Mazes which is not routinely administered at the center. Using the WISC-R record forms, the following measures were recorded. The Full Scale IQ (FSIQ), Verbal IQ (VIQ), and Performance IQ (PIQ) scores were derived according to the procedure outlined in the WISC-R manual (Wechsler, 1974). The scales that constitute VIQ are: Information, Similarities, Arithmetic, Vocabulary, and Comprehension. The scales of Picture Completion, Picture Arrangement, Block Design Object Assembly, and Coding sum to the PIQ scale. The present study included the discrepancy score defined by Seashore (1951) of Verbal IQ minus Performance IQ (VIQ-PIQ).

A score for each Kaufman scale was obtained by summing the scaled scores of subtests that comprise each factor and converting them to deviation quotients. Verbal Comprehension (VC) is comprised of Information, Similarities, Vocabulary, and Comprehension. The Perceptual Organization Factor (PO) consists of the subscales Picture Completion, Picture Arrangement, Block Design, and Object Assembly. The factor Freedom from Distractability (FD) is composed of Arithmetic, Digit Span, and Coding. The Kaufman factors were combined to form the following discrepancy scores: VC-PO, VC-FD, and PO-FD.

Scores were obtained for the verbal scale range and performance scale range by subtracting the lowest from the highest scaled score respectively. Intertest scatter, according to Kaufman (1979) is best understood by quantifying the number of subtests that deviate significantly from each scales respective mean. Kaufman defines a significant deviation as one which is plus or minus three points from the scale's own mean. In this study, WISC-R data were computed by Apple-II program, "WISC-R Scoring and Interpretative Report" (Honaker & Harrell, 1982), therefore making available actual probabilities. In later analyses, these actual probabilities were used rather than the Kaufman guidelines. Scatter scores for each subject were calculated by summing the number of subtests that deviated significantly.

The clinical groupings used in this study were based on the primary diagnoses on file for 32 of the subjects. For the remaining eight subjects, diagnostic impressions were based on the description of presenting symptomology in the clinical summaries recorded during the intake procedure. Each individual then was classified according to the system outlined by Davison and Neale (1982). For the eight protocols which were classified based on diagnostic impressions, rather than primary diagnoses on file, examiner reliability was ascertained by the use of an independent judge. There was a 100% agreement between the examiner's and the judge's categorization decisions on the eight protocols. After extensive review, Davison and Neale (1982) propose that the following DSMII and DSM III diagnoses fall within the Overcontrolled category:

1. Withdrawing reaction (DSM II)

2. Overanxious reaction (DSM II)

3. Separation anxiety disorder (DSM III)

4. Avoidant disorder (DSM III)

5. Overanxious disorder (DSM III)

6. Schizoid disorder (DSM III)

7. Elective mutism (DSM III)

8. Oppositional disorder (DSM III)

9. Identity disorder (DSM III)

Diagnoses classified as Undercontrolled are as follows:

1. Hyperkinetic reaction (DSM II)

2. Unsocialized aggressive reaction (DSM II)

3. Group delinquent reaction (DSM II)

4. Attention deficit disorder with or without hyperactivity and residual type (DSM III)

5. Conduct disorders, undersocialized and aggressive, under- socialized and nonaggressive, socialized and aggressive, socialized and nonaggressive, and atypical (DSM III)

Analyses To ascertain the existence and, therefore, influence of extraneous variation in the dependent variable, analysis of covariance was computed on the principle factors with race and age as the desig- nated covariates. With minimal covariate effects, analyses of vari- ance were performed. Analyses of variance were computed with diagno- sis as the factor. Oneway analyses of variance were performed on the

Verbal IQ and Performance IQ, Verbal Comprehension, Perceptual Organization, and Freedom from Distractability factors, the Verbal and Performance scatter and range. Oneway analyses were also computed for the discrepancy scores between Verbal IQ and Performance IQ, Verbal Comprehension and Perceptual Organization, Verbal Comprehension and Freedom from Distractability, and lastly Perceptual Organization and Freedom from Distractability. In addition, oneway analyses were performed for each of the eleven subtests.

Although it was the intention of the author to compute a chi square to determine the frequency of direction with respect to strength (i.e., highest to lowest scaled score) in the Kaufman factors, it was impossible to perform since there were several cells which failed to meet the minimum required cell number. Finally, t-tests were performed on the Full Scale IQ and the range scores. A minimum alpha level of $p < .05$ was required for significance on all tests with a $p < .10$ considered indicative of a trend.

CHAPTER IV

RESULTS

Since WISC-R subtest reliabilities are a complex function of age (Wechsler, 1974), and the present sample varied across nine years, potential age differences were analyzed first. With age and diagnosis as the variables of interest, a 9x2 ANOVA indicated that there were no significant age differences between the two diagnostic criterion groups for any WISC-R summary scores. The F values are as follows: for the Full Scale IQ with $F(1,38)=1.58$, $p<.22$; Verbal IQ $F(1,38)=2.72$, $p<.08$; Performance IQ, $F(1,38)=.28$, $p<.76$; Verbal Comprehension, $F(1,38)=3.73$, $p<.03$; Perceptual Organization $F(1,38)=.85$, $p<.43$; and Freedom from Distractability, $F(1,38)=.15$, $p<.86$. Diagnostic groups were collapsed across ages for the subsequent analyses.

To test hypotheses regarding the potential role of WISC-R scores in discriminating diagnostic groups, the scores were evaluated by means of the Statistical Package for the Social Sciences, 2nd edition. The two diagnostic groups were compared on their mean Full Scale IQ (FSIQ), Verbal IQ (VIQ), Performance IQ (PIQ), Verbal Comprehension Factor (VC), Perceptual Organization Factor (PO), Freedom from Distractability Factor (FD), Verbal IQ minus Performance IQ score

(VIQ-PIQ), Verbal Comprehension minus Perceptual Oraganization score (VC-PO), Verbal Comprehension minus Freedom from Distractability score (VC-FD), Perceptual Organization minus Freedom from Distractability score (PO-FD), Verbal range, Performance range, Verbal scatter, Performance scatter, direction, and each of the eleven subtests. In each case one-way ANOVA's were used.

Traditional Summary Scores

The first and second hypotheses proposed that the Undercontrolled diagnostic group would manifest a Performance IQ greater than their Verbal IQ, and similarly their Perceptual Organization factor score would be larger than their Verbal Comprehension score. The means and standard deviations of the traditional indices, along with the F values are presented in Table 1. The conclusion reached following an examination of the Table is not only a lack of statistical significance between the criterion groups on the FSIQ, VIQ, and PIQ, but a marked similarity between the two groups.

Kaufman Scales

The third hypothesis proposed that the Undercontrolled group would score higher on the Perceptual Organization factor than Verbal Comprehension. The between group hypothesis stated that the Undercontrolled would exhibit a larger Perceptual Organization factor score in comparison to the Overcontrolled group. The means, standard devia-

TABLE 1

Means and Standard Deviations of Diagnostic Groups

on Traditional Indices

	Mean	SD	F value

VERBAL IQ

| Overcontrolled | 103.50 | 17.25 | |
| Undercontrolled | 102.75 | 14.47 | 1.42 |

PERFORMANCE IQ

| Overcontrolled | 106.45 | 12.80 | |
| Undercontrolled | 106.55 | 12.55 | 1.04 |

FULL SCALE IQ

| Overcontrolled | 105.45 | 15.23 | |
| Undercontrolled | 104.70 | 13.33 | 1.31 |

----

tions, and F values for the three scales based on Kaufman's (1975) factors: Verbal Comprehension, Perceptual Organization, and Freedom from Distractability are reported in Table 2. One way analyses of variance on the VC, PO, and FD factors indicate lack of significance.

The question whether the diagnostic groups differ in their relative strengths and weaknesses on the Kaufman scales was investigated by computing the frequency of the various configurations (e.g. VC>PO>FD, PO>VC>FD, PO>FD>VC FD>PO>VC, VC>FD>PO, PO>FD>VC, FD>VC>PO, PO>VC=FD, VC=PO>FD, and VC>PO=FD). A chi-square was not performed since the minimal number per cell required for analysis was not achieved in all cases. However, examination of the data shows that the VC>PO>FD configuration describes ten of the subjects; seven Overcontrolled subjects and three Undercontrolled subjects. The PO>VC>FD organization configuration was found in ten of the Overcontrolled group and nine of the Undercontrolled. The remaining 11 protocols were distributed among the other eight configuration types. Although a formal analysis was not computed, the absence of a significant pattern was apparent. In addition, the Kaufman factors failed to provide further clinical information not found in the traditional scores.

Discrepancy Scores

Although no specific hypotheses were articulated in this area, it was expected that given the postulated strength of the Undercontrolled group in performance summary scores, (Performance IQ and Per-

TABLE 2

Means and Standard Deviations of Diagnostic Groups

on Three Kaufman Factors

	Mean	SD	F value
VERBAL COMPREHENSION			
Overcontrolled	104.85	18.18	
			1.66
Undercontrolled	103.10	14.11	
PERCEPTUAL ORGANIZATION			
Overcontrolled	109.35	10.47	
			1.62
Undercontrolled	108.70	13.32	
FREEDOM FROM DISTRACTABILITY			
Overcontrolled	90.55	17.51	
			1.39
Undercontrolled	96.40	14.83	

ceptual Organization) compared to the Verbal summary scores, (Verbal IQ and Verbal Comprehension), the discrepancy scores of this group would be of a lesser value than the Overcontrolled group who had postulated strengths in the opposite direction. One way analyses of variance of the mean VIQ-PIQ, VC-PO, VC-FD, and PO-FD were performed. The VIQ-PIQ difference of $F(1,38)=1.10, p<.84$; VC-PO difference of $F(1,38)=1.37, p<.79$; VC-FD difference of $F(1,38)=1.97$ $p<.17$ failed to meet the level of significance $(p<.05)$ in each case assumed in this study. However, in the PO-FD discrepancy, significance was approached $F(1,38)=1.16, p<.10$ reflecting a larger PO-FD difference (M=19.3) for the Overcontrolled diagnostic group than the Undercontrolled diagnostic group (M=12.3). The reader is referred to Table 3 for specific information regarding means, standard deviations, and F values.

Subtest Scatter

As a fifth hypothsis, it was postulated that both dignostic groups would exhibit a similar level of scatter. Moreover, the amount of scatter and range would be significant when compared to the normative sample. The comparisons were made through referencing the pertinent tables provided by Kaufman (1979). Verbal and Performance range scores were obtained by subtracting the lowest scaled score from the highest within the respective scales. A one way analysis of variance of the groups' average range scores suggest a trend $F(1,38)=1.58, p<.06$ in the Overcontrolled group. The Overcontrolled criterion group

TABLE 3

Means and Standard Deviations of Diagnostic Groups

on Discrepancy Scores

	Mean	SD	F value
VERBAL IQ - PERFORMANCE IQ			
Overcontrolled	-2.96	13.92	
			1.10
Undercontrolled	-3.80	13.25	
VERBAL COMPREHENSION-PERCEPTUAL ORGANIZATION			
Overcontrolled	-4.50	13.90	
			1.37
Undercontrolled	-5.60	11.87	
VERBAL COMPREHENSION-FREEDOM FROM DISTRACTABILITY			
Overcontrolled	14.30	19.81	
			1.97
Undercontrolled	6.70	14.12	
PERCEPTUAL ORGANIZATION-FREEDOM FROM DISTRACTABILITY			
Overcontrolled	19.30	12.34	
			1.16
Undercontrolled	12.30	13.26	

exhibited a greater scatter within the Verbal Scale,(M=6.55) therefore producing a larger verbal range score than the Undercontrolled group (M=5.40).

A second approach was adopted to investigate scatter. For each protocol, the number of subtests that deviated significantly from the child's own verbal and performance mean respectively were recorded. Means for each diagnostic group were obtained. The Overcontrolled group showed a greater scatter of verbal subtest score scatter (M=1.25) than the Undercontrolled group (M=.50). Oneway analysis of variance yielded a significant finding $F(1,38)=3.08, p<.03$. There were no significant differences between the groups on the performance scale (Overcontolled M=.85; Undercontrolled M=.70, $F(1,38)=1.23, p<.55$). The results of the analyses on the verbal subtests are presented in Table 4.

Subtest Patterning

To further investigate the significant findings with the Over-controlled group on verbal scatter (and the strong trend found in the Verbal range), oneway analyses of variance was performed on each of the eleven subtests. No significant differences were found. The results of the analyses are summarized in Table 5 and Table 6.

Finally, as a post hoc analysis, the subtests which exhibited the greatest amount of difference between the two diagnostic groups

TABLE 4

Means and Standard Deviations of Diagnostic Groups

on Verbal Range and Verbal Scatter Scores'

	Mean	SD	F value
VERBAL RANGE			
Overcontrolled	6.55	2.01	
			1.58**
Undercontrolled	5.40	1.60	
PERFORMANCE RANGE			
Overcontrolled	6.75	2.29	
			1.23
Undercontrolled	6.45	2.06	
VERBAL SCATTER			
Overcontrolled	1.25	1.21	
			3.08*
Undercontrolled	0.50	0.69	
PERFORMANCE SCATTER			
Overcontrolled	0.85	0.81	
			1.23
Undercontrolled	0.70	0.73	

*p<.05
*p<.10

were identified. An examination of Table 5 and Table 6 shows that the

Overcontrolled and Undercontrolled groups differ most on the Informa-

tion subtest, with means of 10.50 and 9.45 and on Picture Arrangement

with means of 11.30 and 10.35, respectively. The scores of these two

subtests were summed, yielding a total score. The means of the two

diagnostic groups were strikingly similar. The Overcontrolled had a

mean of 21.80 and the Undercontrolled had a mean of 19.80. Certainly,

no significance is indicated.

TABLE 5

Means and Standard Deviations of Diagnostic Groups

on the Verbal Subtests

	Mean	SD	F value
INFORMATION			
Overcontrolled	10.50	3.70	
			1.89
Undercontrolled	9.45	2.69	
SIMILARITIES			
Overcontrolled	11.30	4.31	
			3.26
Undercontrolled	11.65	2.39	
ARITHMETIC			
Overcontrolled	9.25	2.92	
			1.06
Undercontrolled	10.00	3.01	
VOCABULARY			
Overcontrolled	10.60	3.25	
			1.51
Undercontrolled	10.15	2.64	
COMPREHENSION			
Overcontrolled	11.10	3.42	
			1.03
Undercontrolled	11.25	3.37	
DIGIT SPAN			
Overcontrolled	8.15	3.00	
			1.50
Undercontrolled	9.35	3.68	

TABLE 6

Means and Standard Deviations of Diagnostic Groups

on the Performance Subtests

	Mean	SD	F value
PICTURE COMPLETION			
Overcontrolled	12.55	2.69	
			1.74
Undercontrolled	12.40	3.55	
PICTURE ARRANGEMENT			
Overcontrolled	11.30	2.98	
			1.58
Undercontrolled	10.35	2.39	
BLOCK DESIGN			
Overcontrolled	11.40	2.56	
			1.28
Undercontrolled	11.70	2.90	
OBJECT ASSEMBLY			
Overcontrolled	10.85	2.01	
			1.90
Undercontrolled	11.25	2.77	
CODING			
Overcontrolled	8.55	3.65	
			2.25
Undercontrolled	9.15	2.43	

CHAPTER V

DISCUSSION

The purpose of this study was to evaluate the clinical utility
of various WISC-R summary scores by assessing whether two groups of
children judged to be clinically distinct based on diagnoses formu-
lated in an multidisciplinary assessment context, differed signifi-
cantly in their performance on the WISC-R. It was posulated that if
the two groups manifest differential performance, diagnostic interpre-
tations based on WISC-R measures would have some empirical basis.
Performance was evaluated on a number of indices to approximate the
holistic style of interpretation recommended in recent literature
(Blaha & Wallbrown, 1984; Kaufman, 1979).

The premise of this study, that the two clinical groups could be
differentiated by their WISC-R performance, received partial support.
The qualifer, 'partial', must be underscored since while the groups
were differentiated by WISC-R measures, it was not in the direction
predicted.   Moreover, few statistically significant differences
between groups were found. They were as follows:

    1. The Overcontrolled group demonstrated significantly

        greater scatter on the verbal scale than the Undercont-

        rolled group.

2.   The Overcontrolled group exhibited a trend for a larger
     range of Verbal subtest scores in comparison to the
     Undercontrolled group

3.   The Overcontrolled group also tended toward a larger Per-
     ceptual Organization minus Freedom from Distractability
     score than the Undercontrolled Group.

Given the clinical status of the groups and clinical literature
linking pathology and test score scatter, it was postulated that the
two criterion groups would exhibit more scatter and range than the
normative sample of the WISC-R. An examination of the respective
table of norms of range and scatter scores in the standardization sam-
ple (Kaufman, 1979), fails to support this hypothesis. Kaufman (1979)
lists the average verbal range scores as four to five and the average
verbal scatter score as zero to one depending on the age of the chil-
dren which are figures equivalent to the scatter and range scores
found in both criterion groups.

Albeit the significance was not in the direction postulated, the
range and scatter scores formed the significant findings in this
study. Of surprise and interest was the finding that the Overcont-
rolled group, not the Undercontrolled group, manifested a significant
level of scatter and range on the Verbal subtests. Intuitively, one
might expect this finding to be more descriptive of the Undercont-

rolled group because it consists of diagnoses of conduct disorders and attention deficit disorders. These disorders are characterized by erratic behavior, and yet this was not manifested on their WISC-R performance. The greater variance was found in the Overcontrolled group which consists of anxiety disorders of avoidant and overanxious behaviors. Consider the assumptions that a "flat profile" is indicative of emotional health, and that pathology and scatter are associated. If there is validity to these assumptions, then it appears that the characteristic inner psychic distress experienced by the Overcontrolled group is more distruptive than the external chaos evoked by the Undercontrollers. The third finding that the Overcontrolled group Perceptual Organization minus Freedom from Distractability (PO-FD) factor score approached significance is difficult to interpret. One of the problems lies in the failure of either factor (PO or FD) in isolation or in combination with the Verbal Comprehension factor to be of significance. What makes the discrepancy, Perceptual Organziation minus Freedom from Distractability unique for the Overcontrolled group is not immediately apparent.

The results of this study are discouraging in their failure to identify a single clear-cut pattern characteristic of either of the two diagnostic groups. Although significant discriminant results were obtained, little of the variability was accounted for in later analyses, although ANOVAs and t-tests were performed on each of the eleven subtests. It is, however, important to note that the significant

findings emerged exclusively on the Verbal scores. So while a charac-
teristic pattern seemed to be emerging in the Verbal Scale, further
analysis was not successful in providing clarification. Results simi-
lar to those of the present study have encouraged some (e.g., Vance,
Singer, Kitson, & Brennner, 1983) to conclude that the search for a
specific diagnostic pattern that will discriminate LD and ED children
from normal children, is a fruitless task. However, before the pres-
ent study is treated as a confirmation of this conclusion, the author
would like to take the opportunity to critically analyze the project.

A major structural flaw in this study is the small sample size.
An investigation limited to 40 testing protocols undermines external
validity and thereby restricts the generalizability of results. Fur-
thermore, the small N also reduces internal validity since it allows
minor deviations to substantially affect the results. There were
other methodological problems with the study. The exclusionary cri-
terion, although intended to sufficiently establish the parameters for
subject selection (e.g., Full Scale IQ of less than 75, known history
of thought disorder, good evidence of organic impairment or a diagno-
sis of learning disablility with no secondary diagnosis) is likely to
have been too broad.

In the selection of the protocols, careful attention was paid to
the possible influence of the race and age of the children. Thiry-six
of the 40 subjects were Caucasian, the remaining four were Hispanic.

The rationale for the inclusion of the Hispanic subjects was provided by Kaufman's (1979) comments regarding ethnic and racial membership as it influences test performance and assessment. Kaufman (1979) had concluded that while an examiner should be familiar with the characteristics of the pertinent group, group differences are not meaningful for individual test interpretation. In addition, a review of the four Hispanic WISC-R protocols indicated that these subjects did not differ substantially from the other subjects. The age variable could also be considered problematic. The age span of this study approached nine years, a necessity to meet a minimum sample size. The age factor was found not to have differential significance between the two criterion groups, but one cannot ignore the implications of such heterogeneity in the age span.

Finally the nature of this study prevented matching subjects on such demographic variables as family composition, socio-economic status, parental occupation, educational level of parents. It was hoped that the combined effects of a broad catchment area and random selection within the inclusionary criteria would prevent systematic biases. The possiblity of a bias between the two groups, however, that confounded the findings must be considered.

This discussion has focused on the structural flaws of the study that are correctable by using a large enough population needed to match protocols on all relevant subject variables. In considering

another type of flaw in this study, the outcome is a bit more precari-
ous. This type of flaw is not particular to the design of the present
study but is seemingly inherent in the field of psychological assess-
ment and classification. This concerns the question of pathology.
The expertise of the field, as both a science and a profession rests
on its ability to make reliable and valid distinctions between what
constitutes normality and pathology. The field has been struggling
with this issue since its inception and has yet to attain a completely
satisfactory conceptual schema.

The criticism of the arbitrary nature of nosological categories
has had a long tradition. Achenbach (1979) noted that the ambiguous
and conflicting statements regarding diagnostic attributes, particu-
larly among behavior and learning disorders, has frustrated research
and training efforts. Achenbach (1978) dealt with the field's inade-
quacy by constructing his own instrument to assess childhood pathol-
ogy. Others, such as Davison and Neale (1982) have attempted to make
what is already available, more manageable by classifying the array of
diagnoses presented in the Diagnostic and Statistical Manual into a
dichotomy of syndromes. Not to detract from the appeal of the system,
it may not have been relevant for this research endeavor. The cat-
egorization lacked the specificity needed to make reliable distinc-
tions. It seems likely that the gross level of significance obtained
in this study which resisted further analysis is related to the over-
general diagnostic categories utilized in this study. A more sophis-

ticated classificatory system may increase the frequency and stability of differences between diagnostic categories.

Of course, the alternative explanation is that diagnostic categories, per se, have no relevance to the WISC-R. Consider Wechsler's (1974) statements that, a good part of a diagnostician's skill in appraising test performance depends on his or her ability to detect and interpret unusual and aberrant test response and, moreover, there is no general rule for making such interpretations since much depends on what one considers deviant. The ambiguity and subjectivity of Wechsler's conclusion are readily apparent. It is statements of this sort that arose apprehension in those reticent to grant WISC-R's claims to diagnostic use. However, progress has been made in determining the parameters of the instrument's diagnostic powers and this movement needs to be recognized. Profile analyses, factorally pure measures, and frequency tabulations have been incorporated with success in attempts to ascertain WISC-R interpretive bounds. Kaufman (1979) presents perhaps, the best conceptualization. He recommends treating the three IQs and 12 subtest scores as raw material. Profile analysis in addition to the traditional three factor solution adds the flexibility and sophistication essential for breaking an examiner's overdependency on the three IQ scores.

Currently the question of the diagnostic utility of the WISC-R is formulated around the degree of its clinical usefulness, research-

ers are no longer satisfied with simple affirmation or rejection as an
answer. Certainly few would disagree with the Vance et al.(1983) dra-
matic statement, that diagnosing LD and ED children solely on the
WISC-R subtest scatter or V-P difference is a case of malpractice.
But again few are recommending the type of "armchair analysis" of
which they are so critical. For serious researchers the evaluation of
the clinical utility of the WISC-R is pursued with the understanding
that the WISC-R can serve as a framework to generate clinical hypoth-
eses about a child's performance.

REFERENCES

Achenbach, T.M. (1966). The classification of children's psychiatric symptoms: A factor analytic study. Psychological Monographs, 80(7, Whole No.615).

Achenbach, T.M. (1978). The child behavior profile:I.boys aged 6-11. Journal of Consulting and Clinical Psychology, 46, 478-488.

Achenbach, T.M. (1979). The child behavior profile: An empirically based system for assessing children's behavioral problems and competencies. International Journal of Mental Health, 7, 24-42.

Achenbach, T.M., & Edelbrock, C.S. (1978). The classification of child psychopathology: A review and analysis of empirical efforts. Psychological Bulletin, 85, 1275-1301.

Achenbach, T.M., & Edelbrock, C.S. (1979). The child behavior profile:II. boys aged 12-16 and girls aged 6-11 and 12-16. Journal of Consulting and Clinical Psychology, 47, 223-233.

American Psychiatric Association. (1980). Diagnostic and statistical manual of mental disorders (3rd ed.). Washington, D.C.: Authors

Anastasi,A. (1976). Psychological Testing (4th ed.). Macmillan Publishing Co., Inc. New York.

Blaha, J., & Wallbrown, F.H. (1984). Hierarchical analysis of the WISC and WISC-R synthesis and clinical implications. Journal of Clinical Psychology, 40, 556-571.

Clampit, M.K., Adair, J., & Strenio, J. (1983). Frequency of discrepancies between deviation quotients on the WISC-R: a table for clinicians. Journal of Consulting and Clinical Psychology, 51, 795-796.

Davison G.C., & Neale, J.M. (1982). Abnormal Psychology (3rd ed.). New York: Wiley and Sons.

Dean, R.S. (1977). Patterns of emotional disturbance on the WISC-R. Journal of Clinical Psychology, 33, 486-490.

Dean, R.S. (1978). Distinguishing learning-disabled and emotionally disturbed children on the WISC-R. Journal of Consulting and Clinical Psychology, 46, 381-382.

DeHorn, A., & Klinge, V. (1978). Correlations and factor analysis of the WISC-R and the Peabody Picture Vocabulary Test for an adolescent psychiatric sample. Journal of Consulting and Clinical Psychology, 46, 1160-1161.

Doppelt, J.E., & Kaufman, A.S. (1977). Estimation of the differences between WISC-R and WISC IQs. Educational and Psychological Measurement, 37, 417-424.

Edelbrock C.S., & Achenbach T.M. (1980). A typoplogy of Child Behavior Profile patterns: Distribution and correlates for disturbed children aged 6-16. Journal of Abnormal Child Psychology, 8, 441-470.

Fernald, P.S., & Wisser, R.E. (1967). Using WISC-R verbal-performance discrepancy to predict degree of acting out. Journal of Clinical Psychology, 23, 92-93.

Finch, A.J., Kendall, P.C., Spirito, A., Enfin A., Montgomery, L.E., & Schwartz, D.J. (1979). Short form and factor analytic studies of the WISC-R and behavior problem children. Journal of Abnormal Child Psychology, 7, 337-344.

Finch, A.J., Spirito, A., & Brophy C.J. (1982). Reflection-impulsivity and WISC-R performance in behavior-problem children. The Journal of Psychology, 111, 217-221.

Groff, M., & Hubble L. (1982). WISC-R factor structures of younger and older youth with low IQs. Journal of Consulting and Clinical Psychology, 50, 148-149.

Guertin, W.H. (1966). The search for recurring patterns among individual profiles. Educational and Psychological Measurement, 26, 151-165.

Gutkin, T.B. (1978). Some useful statistics for the interpretation of the WISC-R. Journal of Consulting and Clinical Psychology, 46, 1561-1563.

Gutkin, T.B. (1982). WISC-R deviation quotients vs. traditional IQ's: An examination of the standardization sample and some implications for score interpretation. Journal of Clinical Psychology, 38, 179-182.

Gutkin, T.B., & Reynolds, C. R. (1981). Factorial similarity of the WISC-R for White & Black children from the standardization sample. Journal of Educational Psychology, 73, 227-231.

Hale, R.L. (1979). The utility of the WISC-R subtest scores in discriminating among adequate and underachieving children. Multivariate Behavioral Research, 14, 245-253.

Hale, R.L., & Landino, S.A. (1981). Utility of WISC-R subtest analysis in discriminating among groups of conduct problem, withdrawn, mixed, and nonproblem boys. Journal of Consulting and Clinical Psychology, 49, 91-95.

Hamm, H.A., & Evans, J.G. (1978). WISC-R subtest patterns of severely emotionally disturbed students. Psychology in the Schools, 15, 188-190.

Hodges, K. (1982). Factor structure of the WISC-R for a psychiatric sample. Journal of Consulting and Clinical Psychology, 50, 141-142.

Hodges, K., Horwitz, E., Kline,J., & Brandt, D. (1982). Comparison of various WISC-R summary scores for a psychiatric sample, Journal of Clinical Psychology, 38, 830-837.

Honaker, L.M., & Harrell, T.H. (1982). WISC-R scoring and interpretative report. Indialantic, Fla.: Psychologistics, Inc.

Kaufman, A.S. (1975). Factor analysis of the WISC-R at age levels between 6 1/2 and 16 1/2 years. Journal of Consulting and Clinical Psychology, 43, 135-147.

Kaufman, A.S. (1976). Verbal-Performance IQ discrepancies on the WISC-R. Journal of Consulting and Clinical Psychology, 44, 739-744.

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

McDermott, P.A. (1980). Congruence and typology of diagnoses in school psychology: An empirical study. Psychology in the Schools, 17, 12-24.

McMahon, R.C., & Kunce, J.T. (1981). A comparison of the factor structure of the WISC and WISC-R in normal and exceptional groups. Journal of Clinical Psychology, 37, 408-410.

Morris, J.D., Evans, J.G., & Pearson, D.R. (1978). The WISC-R subtest profile of a sample of severely emotionally disturbed children. Psychological Reports, 42, 319-325.

Paget, K.D. (1982). Intellectual patterns of conduct problem children on the WISC-R. Psychology in the Schools, 19, 439-445.

Peterson, C.R., & Hart, D.H. (1979). Factor structure of the WISC-R for a clinic-referred population and specific subgroups. Journal of Consulting and Clinical Psychology, 47, 643-645.

Piaget, J. (1952). The Origins of intelligence in children. New York: International Universities Press.

Piotrowski, R.J. (1978). Abnormality of subtest score differences on the WISC-R. Journal of Consulting and Clinical Psychology, 46, 569-570.

Quay, H.C., & Peterson, D.R. (1967). Manual for the behavior problem checklist. Champaign, Ill.: Children's Research Center, University of Illinois.

Reynolds, C.R. (1979). Interpreting the index of abnormality when the distribution of score differences is known: Comment on Piotrowski. Journal of Consulting and Clinical Psychology, 47, 401-402.

Reynolds, C.R., & Gutkin, T.B. (1980). Stability of the WISC-R factor structure across sex at two age levels. Journal of Clinical Psychology, 36, 775-777.

Saccuzzo, D.P., & Lewandowski, D.G. (1976). The WISC as a diagnostic tool. Journal of Clinical Psychology, 32, 115-124.

Sattler, J.M. (1974). Assessment of children's intelligence rev.ed., Saunders, Philadelphia.

Sattler, J.M. (1982). Assessment of children's intelligence and special abilities (2nd ed.). Boston: Allyn and Bacon Inc.

Schooler, D.L., Beebe, M.C., & Koepke, T. (1978). Factor analysis of WISC-R scores for children identified as learning disabled, educable mentally impaired, and emotionally impaired. Psychology in the Schools, 15, 478-485.

Schoonover, S.M., & Hertel, R.K. (1970). Diagnostic implications of WISC scores. Psychological Reports, 26, 967-973.

Seashore, H.G. (1951). Differences between verbal and performance IQs on the Wechsler Intelligence Scale for Children. Journal of Consulting Psychology, 15, 62-67.

Shiek, D.A., & Miller, J.E. (1978). Validity generalizations of the WISC-R factor structure with 10 1/2-year-old children. _Journal of Consulting and Clinical Psychology, 46_, 583.

Swerdlik, M.E. (1977). The question of the comparability of the WISC and WISC-R: Review of the research and implications for school psychologists. _Psychology in the Schools, 14_, 260-270.

Swerdlik, M.E., & Schweitzer, J. (1978). A comparison of factor structures of the WISC and WISC-R. _Psychology in the Schools, 15_, 166-172.

Thompson, R.J. (1980). The diagnostic utility of the WISC-R measures with children referred to a developmental evaluation center. _Journal of Consulting and Clinical Psychology, 48_, 440-447.

Vance, B., Singer, M.G., Kitson, D.L., & Brennner, O.C. (1983). WISC-R profile analysis in differentiating LD from ED children. _Journal of Clinical Psychology, 39_, 125-131.

Van Hagen, J., & Kaufman, A.S. (1975). Factor analysis of the WISC-R for a group of mentally retarded children and adolescents. _Journal of Consulting and Clinical Psychology, 43_, 661-667.

Wallbrown, F.H., Blaha, J., Wallbrown, J., & Engin, A. (1975). The hierarchical factor structure of the Wechsler intelligence test for children-revised. _Journal of Psychology, 89_, 223-235.

Wechsler, D. (1949). _Manual for the Wechsler intelligence scale for children-revised_. New York: Psychological Corporation.

Wechsler, D. (1974). _Manual for the Wechsler intelligence scale for children-revised_. New York: Psychological Corporation.

APPROVAL SHEET

The thesis submitted by Mary M. Gonzalez has been read and
approved by the following Committee:

> Dr. Thomas Petzel, Director
> Professor, Psychology, Loyola University
>
> Dr. John M. Paolella
> Assistant Professor, Psychology, Loyola University

The final copies have been examined by the Director of this
thesis and the signature which appears below verifies the fact
that the thesis is now given final approval by the Committee
with reference to content and form.

The thesis is therefore accepted in partial fulfillment of the
requirements for the degree of Masters of Arts.

11/27/84
Date

Director's Signature