AN EXPLORATION OF PARAMETRIC VERSUS NONPARAMETRIC STATISTICS IN OCCUPATIONAL THERAPY CLINICAL RESEARCH

by

Charlotte Brasic Royeen

Dissertation submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Educational Research and Evaluation

APPROVED:

M. Lichtman, Chairperson

M. G. Cline

W. L. Seaver

J. C. Fortune

C. Sears

R. McKeen

March, 1986

Blacksburg, Virginia

AN EXPLORATION OF PARAMETRIC VERSUS NONPARAMETRIC STATISTICS IN OCCUPATIONAL THERAPY CLINICAL RESEARCH

by

Charlotte Brasic Royeen

Committee Chairperson: Marilyn Lichtman
Educational Research and Evaluation

(ABSTRACT)

Data sets from research in clinical practice professions often do not meet assumptions necessary for appropriate use of parametric statistics (Lezak and Gray, 1984). When assumptions underlying parametric tests are violated or cannot be documented, the power of the parametric test may be invalidated and consequently, the significance levels inaccurate (Gibbons, 1976). Much research has investigated the relative merits of parametric versus nonparametric procedures using simulation studies, but little has been done using actual data sets from a particular discipline. This study compared the application of parametric and nonparametric statistics using a body of literature in clinical occupational therapy. The most common parametric procedures in occupational therapy research literature from 1980 - 1984 were identified using methodology adapted from Goodwin and Goodwin (1985). Five small sample size data sets from published occupational

therapy research articles typifying the most commonly used univariate parametric procedures were obtained, and subjected to exploratory data analyses (Tukey, 1977) in order to evaluate whether or not assumptions underlying appropriate use of the respective parametric procedures had been met. Subsequently, the nonparametric analogue test was identified and computed.

Results revealed that in three of the five cases (paired t-test, one factor ANOVA and Pearson Correlation Coefficient) assumptions underlying the use of the parametric test were not met. In one case (independent t-test) the assumptions were met with a minor qualification. In only one case (simple linear regression) were assumptions clearly met. It was also found that in each of the two cases where parametric assumptions were met, no significant differences in p values between the parametric and the nonparametric tests were found. And conversely, in each of the three cases where parametric assumptions were not met, significant differences between the parametric and nonparametric results were found. These findings indicate that if cases were considered as a whole, there was a one hundred percent agreement between whether or not parametric assumptions were violated and whether or not differences were discovered regarding parametric versus nonparametric results.

Other findings regarding (a) non-normality, (b) outliers, (c) multiple violation of assumptions for a given procedure, and (d) research designs employed are discussed and implications identified. Suggestions for future research are put forth.

ACKNOWLEDGEMENTS

Appreciation is extended to the American Occupational Therapy Foundation for the financial support of a Doctoral Fellowship.

The support of my husband, is gratefully acknowledged. Also, appreciated is extended to my dissertation committee members.

Finally, the clinical researchers who shared their data for the investigation are heartily thanked. They shared their data in the true spirit of science and represent commitment to the scientific process.

TABLE OF CONTENTS

List of Tables

List of Figures

Chapter One

Introduction and Background of the Study

Within the field of occupational therapy, clinical research is recognized as a crucial element in the continued development of the profession (Gillette, 1982; Rogers, 1981; West, 1981; Williamson, 1982; Yerxa, 1984). In addition, the advent of cost indexing formulas, implemented by the Department of Human Health Services for determining hospital reimbursement based upon a client's classification into diagnostic related groups (DRG's) versus actual cost of medical care, will have a significant impact upon health care in general (Christiansen, 1983). Christiansen states:

> It seems clear that as administration and policymakers render decisions about how health care providers are used and reimbursed, those disciplines with objective evidence of their effectiveness and efficiency will have a commanding advantage. Such evidence is generally provided via research. (p. 197)

Thus, for continued development of the profession and for a competitive edge in the health care market, it appears that more clinical research is necessary in occupational therapy.

Clinical research may be considered to be a specialized form of applied research, having some unique characteristics. First, clinical research employs subjects identified as deviating from normal in some way, even if a comparison or control group is composed of normal subjects. Second, clinical research centers on research questions which, when answered, will influence either evaluation or treatment of individuals characterized by a particular deviation from normal. Third and finally, clinical research is usually based upon relatively small samples since acquisition of subjects with identified and specified deviations from the norm is often difficult.

In spite of the increase in occupational therapy clinical research over the last thirty years (Ethridge & McSweeney, 1971; Ottenbacher, 1982a), the profession still needs more and better clinical research for the reasons specified previously. Suggestions for improving the quality of occupational therapy clinical research have been put forth in the literature. Kielhofner (1982) and Schmidt (1981) suggest use of qualitative research methods and Ottenbacher (1983) suggests that 'critical attention' be paid to the uses and abuses of research procedures in occupational therapy. For example, Ottenbacher proposes the use of ANOVA with appropriate post hoc analysis instead of employing multiple t-tests as found in five out of fourteen articles he sampled in the occupational therapy literature.

Data analysis procedures for clinical research in occupational therapy are primarily centered upon the use of statistical procedures designed for experimental research such as the aforementioned t-test and other parametric statistical procedures. Parametric statistics are based upon a normal distribution model. Certain assumptions or suppositions underlie the normal curve model (Bucklaw, 1983) and these assumptions relate to the distributions of the variables being studied. In general, these assumptions are as follows.

1) The observations are independent (Bucklaw, 1983).

2) The variables are distributed normally (Bucklaw, 1983; Seaver, 1979).

3) Homogeneity of variance exists among variables for two or more independent samples (Bucklaw, 1983; Hinkle, Wirsma, & Jurs, 1979).

4) The measurement scale of the variable is at least interval level (Bucklaw, 1983; Hinkle, Weirsma & Jurs, 1979).

The assumption of a normal distribution is frequently justified using the central limit theorem (Edgington, 1969). The central limit theorem states that when randomly sampled, any population with finite variance will yield sample means that approach normality (Conover, 1980). Furthermore, the larger the sample size, the more the distribution of sample means will approach normality. However, the appropriateness of the central limit theorem for use with finite populations, as is often encountered in clinical research, is questionable. To illustrate, Plane and Gordon (1982) state, "...when sampling from a finite population, increased sample size will not always bring the shape of the sampling distribution of the sample mean closer to normality" (p. 175). In fact, they found that sometimes the "...sampling distribution departs from normality as the sample size approaches the finite population" (p. 175). Thus, assuming the central limit theorem applies to all samples may not be prudent, and may lead to violation of one or more of the assumptions underlying use of parametric statistical tests.

Violation of one or more of the assumptions underlying the use of parametric statistics and the effect such violations have upon the results of the statistical analyses have been the topics of considerable research and theoretical analysis throughout the last decades (Boneau, 1960; Bradley, 1978; Glass, Peckman & Sanders, 1972; Lezak & Gray, 1984; Singer, 1979; Tupper, 1984). Two concepts are crucial when considering the body of literature regarding statistics and violation of the assumptions underlying their appropriate use. These concepts are robustness and power.

Robustness refers to the capacity of a statistical test to come to the correct conclusion when one or more of the underlying assumptions have been violated (Gardner, 1975). Boneau (1960) defines robustness more simply. He maintains robustness refers to violations of assumptions producing minimal effects (Boneau, 1960).

Related to the notion of robustness is the concept of power. According to Ottenbacher (1982b), power pertains to the probability of making a Type II error, that is, the probability of retaining a false, null hypothesis (Welkowitz, Ewen, & Cohen, 1971). Power, therefore, is the probability of a statistical test coming to the correct decisions regarding rejecting or accepting the null hypothesis. Gibbons (1976) defines it as "...the ability of a test procedure to use sample evidence to detect whether or not the true situation differs from the hypothesized situation"

(p. 26). Clarke (1971) defines it as "... the ability to lead to a decision to reject the null if the null is false" (p. 297).

Due to the reasons related to power and robustness parametric statistical procedures dominate data analysis in occupational therapy as well as elsewhere. For, generally speaking, parametric procedures are considered to be the most powerful class of data analysis procedures (Welkowitz et al., 1971). Moreover, parametric procedures are generally considered to be robust to violations of assumptions (Bradley, 1978).

Yet, the prevalent use of parametric procedures in occupational therapy clinical research may be questionable. Robustness with clinical data may be questionable due to distributional characteristics of the data sets (Hill & Dixon, 1982). For example, in a related discipline similarly based upon clinical research, clinical neuropsychology, it is suggested by Lezak and Gray (1984) that "...the research data in clinical neuropsychology do not conform to the requirements for parametric statistical analysis" (p. 101). Furthermore, Lezak and Gray cite three primary reasons why parametric procedures are problematic in clinical neuropsychology.

First, data sets in clinical neuropsychology are often fraught with outliers. Second, since the research studies are primarily clinically based, large samples are difficult to obtain due, in part, to mortality and morbidity. Third, there is an excessive amount of empty cells due to the wide range of variability in behavioral dimensions, and that given an array of tests, not all subjects can perform all tests and, consequently, there is considerable missing data. All of these problems often occur in occupational therapy clinical research.

For example, research raises the possibility that distributional characteristics or factors related to oral motor function may be skewed (Ottenbacher, Dauk, Givelinger, Grahm & Hasset, 1985). Thus, the assumption of normal distribution of variables may not fit variables of interest in occupational therapy clinical research. As a solution to these problems, Lezak and Gray propose the increased use of nonparametric statistical procedures. Since the problems specified in clinical neuropsychological research are similar to those in occupational therapy clinical research, nonparametric methods may also hold great promise for increasing the quality and quantity of clinical research in occupational therapy. The quality may be enhanced by improving power of the analysis and therefore the accuracy and validity. The quantity may be improved by allowing for more research due to nonparametric analyses applicability with small samples, the ability to compute variables at less than interval level, and the relative ease of computation. To illustrate, nonparametric statistical procedures allow for comparison of distributions when little is known about the form of the distribution (Mosteller & Rourke, 1973). And, very little is known about the distributional characteristics of the variables studied in occupational therapy clinical research. Thus, nonparametric statistics may better allow for comparisons between such variables since, computationally speaking, nonparametric statistical methods are based upon nondistributional factors such as the "...signs of differences, ranks of measurements, and counts of objects or events falling into categories and for this reason they are called nonparametric" (Mosteller & Rourke, 1973, p. 1).

Relative to the definition of nonparametric statistics, Mosteller and Rourke state:

> The adjective 'nonparametric' is somewhat misleading, because nonparametric statistics do in fact deal with parameters such as the median of a distribution or the probability of success (p) in a binomial. Indeed, the word nonparametric as commonly used does not lend itself to a precise definition. (p. 1)

Marascuilo and McSweeney (1977) suggest that perhaps nonparametric and distribution free procedures are best categorized as assumption freer. Furthermore, McSweeney and Katz (1978) state that "since the nonparametric-distribution free distributions are situation specific and frequently blurred, these terms are often used interchangably" (p. 1024). However, Bradley (1968) makes a distinction between nonparametric and distribution free as follows:

> ...a nonparametric test is one which makes no hypothesis about the value of a parameter in a statistical density function, whereas the distribution free test is one which makes no assumption about the precise form of the sampled population. (p. 15)

But, even Bradley concedes that common use of the terms has made them equal. Thus, based upon Bradley (1968), Marascuillo and McSweeney (1977), and McSweeney and Katz (1978), the terms nonparametric and distribution free will be used interchangably in the current investigation.

There is evidence that many scholars consider nonparametric statistics as the preferred set of statistical analyses for clinical and applied research. Borg (1984) believes that nonparametric procedures are more appropriate for much of the data generated in educational research. For example, according to Siegel (1956) nonparametric statistics are well suited to behavioral science methodology for three reasons. First, the tests do not assume that the scores or data are drawn from a normal population. In occupational therapy very little is known about the population of clients constituting specific disability groups. However, it is unlikely the populations fit the normal curve model nor that they are infinite. Second, nonparametrics can use scores which are not exact but are in less than interval level form. Thus, even if the population is distibuted normally, nonparametrics should be used if the data is in the nominal, categorical or ratio level of measurement (Bucklaw, 1983). Much occupational therapy data which could be employed in clinical research is measured at these levels. Third, nonparametrics are useful with a small sample. In fact, as far back as 1959 Gaito stated that for samples less than six, nonparametric statistics are the only choice (Gaito, 1959). And, frequently populations or samples investigated in occupational therapy clinical research are small in number.

As previously stated, nonparametric statistical methods may enhance occupational therapy clinical research. They could do so by increasing the quality of clinical research by improving the validity of data analysis procedures (matching statistical tests with assumptions). They could also increase the quantity of occupational therapy research by allowing for the use of small samples and problematic data sets (outliers, non-normal and data in ordinal or rank order) to be amenable to quantitative data analysis. These later examples of problematic data sets are often seen in clinical samples. McSweeney and Katz (1978) present an excellent summary of reasons for using nonparametric statistics which can also serve as a further explanation for why nonparametrics can enhance the quality and quantity of occupational therapy research. These reasons are (a) nonparametrics are assumption freer; (b) the data may not appropriately be put into metric form, i.e., unordered qualitative variables; (c) the data may be rank ordered; (d) the data may be from small samples; (e) there may be a non-normal distribution of the variable(s); (f) heterogeneity of variance between groups may exist; and (g) outliers may be present.

Gibbons (1976) argues four main points for the advantage of nonparametric statistics which are germane to the improved quality and quantity of occupational therapy clinical research based upon nonparametric statistics. First, for the nonstatistician, there is a lesser chance of misusing statistics by violating assumptions. Secondly, the cost of compiling and collecting the data may be less since interval level of measurement for the data isn't necessary and small samples can be employed. Third, nonparametric methods are easy to understand. Fourth, nonparametrics have a wide scope of application.

Why are parametric tests so prevalent if, indeed, nonparametric methods offer more valid procedures when assumptions for parametric statistics are violated? One reason relates to power. Given a comparison between an equivalent parametric and nonparametric test run under conditions of normality, a parametric test has greater power (Siegel, 1956). Therefore, it is often assumed that a parametric test is better.

But, how powerful is a parametric test when the assumptions upon which it is based are violated? Gibbons (1976) infers that when assumptions underlying parametric tests are violated, or if one cannot document that the assumptions underlying parametric statistics have been met, the concept of power is irrelevant since the tests may not be valid and the significance levels not accurate.

Another reason parametric statistics are more prevalent than nonparametric statistics has to do with the concept of robustness. Many parametric procedures, especially the F and t tests, are thought to be robust to violations of assumptions under certain conditions (Boneau, 1960, 1962; Box, 1953; Gardner, 1975; Minium, 1978). However, the claim of robustness to violation of assumptions appears to have been overgeneralized beyond highly specified situations (usually delineated from simulation studies) to almost all real life situations and data sets. Thus, almost all violation of assumptions, often ill defined, are tolerated in the name of robustness (Bradley, 1978; Singer, 1979). Furthermore, Bradley (1978) argues that there is no quantitative definition of robustness and that it is therefore a subjective claim and highly variable in form as it is commonly used.

Robustness is, in fact, a specifically defined condition existing only under specified situations. For example, Bradley (1978) states that basic statistical tests deal with robustness to nonnormality or heterogeneity of variance, but do not elucidate how interactions between dual violations of assumptions in a given data set can affect robustness. Also, Bradley (1978) postulates that if assumptions are compromised, the corresponding change in alpha depends upon "...a complex interaction involving many factors" (p. 196). These are:

1) size of alpha,
2) location of the region of rejection (left, right or two tailed),
3) size and shape of the sample, and
4) size, shape and variance of the population(s)

Thus, in order to study the possible effects of violation of assumptions using parametric procedures, the current study will be an exploratory investigation into the use of parametric versus nonparametric procedures using a multiple case study design. That is, actual data sets from occupational therapy clinical research will serve as units of analysis embedded in each "case" or published research.

Thus, the current investigation proposes to study parametric versus nonparametric statistics in a way that is "authentic" (Sprott, 1978). That is, the multiple case study approach "... investigates a ... phenomenon within its real-life context" (p. 23). To illustrate, most work in theoretical statistics centers on simulation studies. However, such research is limited by artifacts of how data sets are generated (Carroll, 1976) and by the degree to which the data configurations validly reflect what would be found in real life situations. Conversely, use of existing data sets from published occupational therapy clinical research allow for study of statistics in real life situations.

In addition, there appears a recent trend to use authentic data, i.e., data sets derived from actual investigations, as a basis for statistical research (Krauth, 1980; Duan, 1983; Katz & McSweeney, 1980; Lezak & Gray, 1974; Marascuilo & Daginas, 1982; Sprott, 1978). Thus, the current investigation will build upon this innovative methodology using a multiple case study design that is systematic and comprehensive for a given discipline, occupational therapy.

Statement of the Problem

Parametric statistical procedures are employed in data analysis of occupational therapy clincial research when, in all likelihood, one or more of the assumptions underlying the appropriate use of parametric procedures has been violated. The central issues to this problem are (a) whether or not assumptions have indeed been violated, and (b) do violations of the assumptions make a difference to the substantive meaning of the statistical analysis. The problem has been stated by Gardner (1975) as the need for current research to focus on the degree to which an inappropriate statistic may lead to a deviate conclusion. Lezak and Gray (1984) have illustrated that the innapropriate use of parametric procedures in clinical neuropsychological analysis often misses statistical significance by a "hair's breath", therefore a Type II error is committed. However, use of parametric analyses when assumptions are violated can also result in inappropriate significant findings, or a Type I error.

The problem can be restated as the possibility of deriving false or inappropriate conclusions in occupational therapy clinical research when employing parametric statistical procedures in questionable circumstances.

Purpose of the Study

The purpose of the study is twofold. First, it is the purpose of the current study to investigate data sets which are typical of the most commonly employed parametric procedures in occupational therapy research when the procedure is employed with a small sample size. The assumptions underlying use of the the parametric data analysis procedures will also be evaluated. Second, the current investigation will compare the findings of the parametric analysis to those of a comparable analysis using nonparametric or distribution free procedures in order to determine the similarity or differences between results under conditions identified by exploration of the data sets.

Research Questions

The problem under study will be approached by investigation of a series of research questions using a multiple case study methodology (Yin, 1984; Yin, 1985). These are specified as follows.

Research Question One

In samples of clinical occupational therapy research data sets in which parametric statistics have been employed, have the assumptions underlying use of parametric procedures been met?

Research Question Two

If assumptions regarding the use of parametric procedures in examples of clinical occupational therapy research data sets are violated, how crucial or important are those violations according to the statistical and methodological literature?

Research Question Three

In cases of clinical occupational therapy research, how similar or different are the substantive findings of the data analysis and the p values of the tests if the findings of the parametric procedures are compared to those of the findings of equivalent nonparametric procedures?

Limitations of the Study

The current investigation will be limited to the study of occupational therapy literature from 1980 to 1984. Moreover, the study of clinical occupational therapy research data sets will be limited to those data sets upon which an occupational therapy research paper has been published during that time, and which can be accessed by the investigator.

In order to limit the breadth of the current study, it is necessary to limit the scope of it. Therefore, the study will center on comparing data analysis procedures and will not address the issues of research design, or the lack of random samples often encountered in clinical research. That is to say, it is deemed beyond the scope of the current study to identify alternative design methodologies and consequent statistical analyses from those employed by the original author. In addition, comparison of parametric versus nonparametric procedures is limited to those procedures identified as the most commonly used in occupational therapy clinical research. Thus, analyses are limited to univariate as compared to multivariate analyses.

Though the scope of the study is limited to occupational therapy, the investigation may have application to other professions or disciplines which base theory and practice on clincial research grounded in statistical theory.

Significance of Study

There are three major points germane to the significance of the current study. First, McSweeney and Katz (1978) maintain that much of the previous work regarding parametric versus nonparametric free methods has been equivocal. They suggested that future work should center upon the systematic investigation of comparable efficiency of parametric versus nonparametric procedures with small data sets. Indeed, that type of investigation is precisely what this study proposes, and thus it is significant in terms of directing research to an important issue identified by major theorists in the field.

Second, the current investigation proposes to study the problem in a way that is "authentic" (Sprott, 1978) by using actual data sets from occupational therapy clinical research as "cases" for analysis. Thus, the applicability of the findings to clinical research in occupational therapy and statistical theory underlying decision rules for data analysis procedures will be enhanced and will be a valid representation of real life situations.

In summary, the current investigation will provide the means for further study of parametric versus nonparametric statistics which will enhance the quality and the quantity of clinical research in occupational therapy by empirical investigation of parametric versus nonparametric statistical procedures.

Chapter Two

Review of the Related Literature

This chapter will present a summary of literature pertinent to the following topics: assumptions underlying power and robustness of selected parametric tests; parametric versus nonparametric tests; and, a historical overview of nonparametric statistics.

Assumptions and Robustness of Selected Parametric Procedures

Bradley (1978) states that robustness exists only under highly specified situations for any given statistical procedure. Thus, the current section will highlight the conditions under which a given statistical procedure will likely be robust as revealed by simulation studies. Simulation studies are investigations into theoretical statistics which usually employ Monte Carlo methods for computer generation of data sets and thus, simulation studies employ artificially developed data sets. As a concomitant to this, the assumptions underlying use of each procedure will also be specified.

Pearson Product Moment Correlation

The assumptions of the Pearson Product Moment Correlation are specified as follow.

1) There is bivariate normality of the underlying

distribution (Welkowitz, Ewen & Cohen, 1971).

2) There is linearity in the relationship between two variables (Welkowitz et al., 1971).

3) Scale of measurement of the variables is interval or ratio level (Hinkle et al., 1979).

Generally, the Pearson r appears robust when sample size is large, that is, greater than 25 or 30 (Welkowitz et al., 1971). However, if the relationship between variables is non-linear, r will likely be an underestimate of the relationship (Furfey, 1958). Moreover, Haldane (1949), states that r is robust to skewness, but that it is not robust to kurtosis differences between distributions which may affect the precision of r. Thus, certain types of non-normality differentially influence the Pearson Product Moment Correlation.

Student's t-test (Single sample, independent samples and paired or matched samples)

The assumptions underlying use of different forms of the t test follow.

1) The variable(s) are distributed normally (Welkowitz et al., 1971).

2) For the two sample case, the variances of the two populations are equal (Welkowitz et al., 1971).

3) There is random selection of the subjects under study (Hinkle et al., 1979).

4) There is interval or ratio level of measurement.

It is often assumed that one and two sample t-tests are robust and accurate due to the properties of the central limit theorem (Pocock, 1982). For example, Welkowitz et al. (1971) maintain that the t-test is robust even "if the assumption of normality is not met" (p. 141). Furthermore, they maintain that the assumption of equality of variance (for the two sample case) is also irrelevant if the sample sizes are equal. However, it is reported that the test is non-robust if one sample is 1.5 times larger than the other in size, and if the respective variances are unequal (Welkowitz et al., 1971). The issue of robustness of the two sample t test becomes complex and highly situationally specific if one considers that Gans (1981) reports that the t-test is robust for equal samples even if variances are unequal, unless the distribution is skewed, but that the t-test may be compromised if sample sizes are unequal.

Furthermore, the failure of the central limit theorem to assure accuracy of the t-test computed with samples as large as 500 has been documented (Pocock, 1982). Other researchers argue that the student t is invalid if the distribution is non-normal (Benjamin, 1983; Tikie & Singh, 1981), particularly if the tails are long and heavy (Blair, 1981; Gross, 1976; Hettmamsperger & McKean, 1978; Horn, 1983; Lehman, 1975; Irvine & Ramey, 1984). In spite of Boneau's findings (1962) to the contrary, many researchers say not to use the two sample t-test for pooled variances if variances are unequal (Ramsey, 1980; Seaver, 1977; Bhattacharjee, 1968). It is evident that controversy exists regarding robustness of the t test.

ANOVA

The following are assumptions underlying use of analysis of variance (ANOVA) specified by Hinkel et al. (1979).

1) Observations are random and independent samples from a population.

2) Interval or ratio level of measurement is employed.

3) Samples are derived from normally distributed populations.

4) The variances of populations are equal.

For the fixed effects model, Gardner (1975) maintains that the robustness of ANOVA to non-normality and heterogeneity of variance is well accepted, but does not exist under all conditions. Consistant with this, Minium (1978) reports that ANOVA is robust to heterogeneity of variance only if the sample size is very large. However, Feir-Walsh and Toolhaker (1974) found that a one factor ANOVA was not robust when the underlying distributions were not normal even if variances were homogeneous.

Bishop and Dudeqicz (1978) report unequal variance can seriously effect inferences about means, especially if the cell sizes are not equal. Somewhat similarly, as early as 1947, Geary maintained that a small departure from normality can seriously affect the ANOVA. In addition, for the random effects model, non-normality can be more serious and heterogeneity of variance can be detrimental even if sample sizes between groups are equal (Neter & Wasserman, 1974). Conflicting views are presented by Box (1953) and Boneau (1962), who report that ANOVA is robust to non-normality and heterogeneity of variance while Hollington and Smith (1979) provide a highly situation specific description of the robustness of the ANOVA.

REGRESSION

The following are the assumptions underlying use of regression analysis (Montgomery & Peck, 1982).

1) The errors are uncorrelated.

2) The errors are normally distributed.

3) There is constant variance of the error term, and error has a mean of zero.

4) There is a linear relationship between dependent and independent variables.

Sen (1968) states that the least square estimate of a regression coefficient is vulnerable to gross errors.

Parametric Versus Nonparametric and Distribution Free Statistics

As was stated previously in chapter one, parametric statistical procedures are those that have underlying assumptions based upon normal curve theory and interval or ratio level of measurement. However, there is controversy as to whether or not it is appropriate, in spite of stated assumptions, to employ parametric procedures with data measured at less than interval level which usually characterizes data derived from scaling instruments. Some theorists state it is permissable to do so (Johnson & Heyer, 1980) and that it is a psychometric, not a statistical problem. Others, though, state that use of less than interval level data does indeed violate an assumption of parametric statistics and should not be done (Siegel, 1956).

A study by Rowney and Zenisek (1980) investigated the liklihood of approval or rejection of manuscripts in Canadian psychological journals based upon parametric procedures used with ordinal level data. They found sixteen percent of those reviewers interviewed for the study were likely to reject manuscripts on the basis of using parametric procedures with ordinal level data.

However, in spite of the controversy regarding use of parametrics with scaling data and data measured at less than the ratio or interval level, parametric procedures are prevalent in the published literature. The following reasons are put forth as a summary explanation for the preferred use of parametric procedures on part of most individuals:

1) They are thought to be more powerful (Kirk, 1968).

2) They are commonly assumed to be robust under almost

all conditions.

3) They are easily accessible in computer packages.

4) They are the most familiar, i.e., parametric procedures predominate in most textbooks and training in statistics and research.

5) Until recently, only parametric statistical procedures were available for the more complicated research designs (Singer, 1979).

6) If converted into ranks or counts, loss of information from the data occurs (Boneau, 1960).

The development of nonparametric and distribution free procedures has been seen as a release from the normal curve model underlying use of parametric procedures and, thus, as necessary tools in many fields and applicable in most all disciplines (Clarke, 1971; Greco, 1979; Savage, 1957). Currently, there appears to be a renewed interest in the application of nonparametric statistics in disciplines having primarily clinical populations. For example, Conover and Iman (1981) advocate the use of nonparametric procedures as a solution to fitting real world problems into statistical theory with data sets that are not normal. In counseling Kirkpatrick (1981) proposed an increased use of nonparametrics for counselor research. In psychology (Bucklaw, 1983) and clinical neuropsychology (Lezak & Gray, 1984) the increased use of nonparametric procedures is advocated as a method of improving research. Even in non-clinically based research, nonparametrics are being used since some investigators deem them to be more robust and more congruent with valid data analysis (Carmelli & Jorde, 1982).

In spite of these arguments for the increased use of nonparametric statistics, there is much misunderstanding and many misconceptions regarding them (Singer, 1979). Some cited by Singer are:

1) They are not always the quickest or easier form of data analysis.

2) They are not absolutely assumption free. Sometimes, there is an assumption of symmetry or similarity of scale.

3) They are available only for the simplest designs.

4) They are inferior to classical methods (parametrics).

5) They are very recent in origin.

In summary, it is apparent that in spite of the long standing preference for use of parametric procedures in data analysis, there is an interest and commitment to expansion of the role nonparametric procedures should play in data analysis for disciplines based upon clinical research. The interest in nonparmetrics is not new, and an historical overview of nonparametrics in the next section of this chapter

Historical Overview of Nonparametric and Distribution Free Tests

Singer (1979) presents a brief overview of the history of nonparametric procedures. In his review, he cites Arbuthnot as the originator of what is currently termed the sign test, which was not only the first nonparametric test, but also the first statistical test of hypothesis ever devised and employed. Arbuthnot used a binomial expansion to demonstrate that the chance of more male births versus female births for the preceding 82 years was so outside the realm of probability that it could not be an accident, but rather the function of "Divine providence". He published his findings in the article "An Argument for Divine Providence, Taken from the Constant Regularity Observed in the Births of Both Sexes" (Noether, 1984).

Singer (1979) further credits Fechner, who worked in the 1800's, as the father of modern experimental psychology and also the originator of rank transformation for data analysis. Fechner pioneered the use of a method similar to Kendall's Tau and worked with the median as compared to the mean.

Galton furthered the work in nonparametrics based upon ranks. He also criticized the excessive use of the assumption of normal distribution (Singer, 1979). Thus, it is clear that nonparametrics are not "new" procedures, a misconception that Singer (1979) has tried to dispel.

Wolfowitz is credited with originating the term nonparametrics in 1942 (McSweeney & Katz, 1978; Noether, 1984). At that time, only a few rank order tests, goodness of fit tests, and correlation tests existed (McSweeney & Katz, 1978). Until the 1930's, nonparametric techniques were unusual in methodological and statistical texts and articles (McSweeney & Katz, 1978). However, in the 1950's nonparametrics became very popular and gained favor as "quick and dirty" analysis procedures (McSweeney & Katz, 1978).

But the commonplace use of nonparametric procedures has not kept pace with the level of theoretical development. For, "distribution free statistics are still not receiving the attention they deserve" (Penfield, 1971 p. 8). The following reasons are put forth as explanations.

First, there is an over emphasis on Type I error control but not Type II error control (McSweeney & Katz, 1978). Second, textbooks on data analysis devote less than ten percent to nonparametric procedures (Smith, 1976). Third, there is a lack of conceptual and consistant connections betweeen nonparametric procedures and research design (McSweeney & Katz, 1978). And fourth, there is a prejudice against parametric procedures based upon their early reputation as "ersatz and make-do" (Noether, 1984 p. 176).

Chapter two has included a review of the assumptions and robustness of selected parametric procedures, a discussion of parametric versus nonparametric and distribution free tests, and an historical overview of nonparametric tests. The next chapter will focus upon the design of the current investigation.

Chapter Three

Methods and Procedures

Overview

The case study process of investigation as outlined by Yin (1984) serves as the methodology for the current investigation. Case study methods employing a set of "pre-specified strategies" are often employed if "how" type research questions underlie the investigation (Yin, 1984). Since this investigation is based upon "how" questions, a case study methodology appropriately structures the investigation in order to ". . . understand events as they occur in their natural context" (Greene & David, 1984 p. 73). And, use of multiple cases allows for research comparable to that of replicating investigations over time or across sites (Yin, 1984).

Issues of generalizability and construct validity pertaining to case study methods need to be addressed. This current research design insures generalizability to theoretical propositions (analytic generalizability) driving or underlying the investigation but not to populations (statistical generalizability per se (Yin, 1984). Thus, final interpretation of the investigation can be made in terms of statistical theory pertaining to selection and application of parametric and nonparametric statistics. Construct validity has been built into the investigation by the use of multiple sources of evidence, i.e., multiple cases, coupled with literature search.

Chapter three specifically describes the case study design employed in the current investigation and is composed of two main sections. Section one, data collection, has two subsections which deal with data identification and acquisition. Section two, design of the case study, has three subsections. These deal with (a) data exploration, (b) replication of parametric analysis and selection and calculation of the nonparametric analysis and (c) case study guidelines.

Data Collection

Data Identification

Articles published in the occupational therapy literature over a five year period were indexed according to whether or not the article was statistically based. A statistically based article was operationally defined as an article indicating that one or more statistical procedures were executed for data analysis in the article under review. A nonstatistical article was operationally defined as an article wherein one or more statistical procedures were not executed for use and reporting in the article under review.

All major articles published in the _American Journal of Occupational Therapy_, _Occupational Therapy Journal of Research_, _Occupational Therapy in Mental Health_, _Physical_ and _Occupational Therapy in Geriatrics_, and _Physical and Occupational Therapy in Pediatrics_ from 1980, or the onset of the journal, through 1984, inclusive, were indexed for tabulation. Other types of articles, such as brief reports, were indexed for a journal if, overall, they were statistically oriented. For example, brief reports were indexed for _Occupational Therapy Journal of Research_ but not in _Physical and Occupational Therapy in Geriatrics_. Also, theme issues for any given issue of a journal were not indexed since they were not statistically oriented.

Table 1 presents the overall figures for statistical versus nonstatistical articles for all journals combined by year.

Table 1

Frequencies Across Years (1980-1984), _Totals and Percentage of Statistical Procedures for All Journals Combined_

	Frequency by Year						
Category	1980	1981	1982	1983	1984	Total	Percentage
Statistical Articles	33	35	47	52	51	218	42.08
Non Statistical Articles	64	36	74	72	54	300	57.92
Totals	97	71	121	124	105	518	100.00

Table 1 reveals that, overall, there are more nonstatistical articles than statistical articles in occupational therapy literature during this time period. Table 1 also reveals that there has been an increase in the total number of statistical articles over the five year period, with a corresponding decrease in nonstatistical articles. This finding is similar to that of Ottenbacher and Peterson (1985) who found an increase in the percentage of statistical articles in the American Journal of Occupational Therapy when all issues of 1983 were compared to those of 1973. In the 1983 volume of the American Journal of Occupational Therapy Ottenbacher and Peterson found that 31% of the articles were nonstatistical and 20% employed descriptive statistics. In the current investigation descriptive statistical articles were categorized as nonstatistical. Thus, combining these values for the study by Ottenbacher and Peterson reveals that 51% of the articles were nonstatistically based. Similarly, in the current investigation which combined all journals for 1983, the number of nonstatistically based articles was 53%.

All of the statistically based articles were further reviewed for types of statistical procedure employed. Such empirical investigation into the statistical procedure used in occupational therapy clinical research is limited except for the previously cited investigation by Ottenbacher and Peterson (1985).

A brief explanation of the criteria and method of categorization of statistical procedures follows and is based, in part, upon a categorization scheme used by Goodwin and Goodwin (1985) for a similar indexing of educational research literature in Educational Researcher. If a particular technique was used more than once in a statistically based article, it was counted only once for that article. However, if two different statistical techniques were employed within a certain article, they were both counted. Thus, there are more statistical procedures than total number of statistical articles. As previously stated, descriptive statistics were specified as non-statistical. If it could not be determined what statistical technique was employed, the technique was classified as "nonspecified".

Thus, all articles were reviewed to determine if they were statistically or nonstatistically based, and, if they were statistically based, the articles were additionally reviewed for statistical procedure(s) employed. Inter-rater reliability was computed in a manner similar to that employed by Ottenbacher and Peterson (1985) and Goodwin and Goodwin (1985). A rater was asked to index all statistically oriented articles from the 1984 issues of the American Journal of Occupational Therapy. The rater independently indexed these 25 articles. The number of agreements between the principal investigator and the independent rater was divided by the number of agreements added to the number of disagreements (# agree / # agree + # disagree) for a percentage of agreement score. In the current investigation the percentage of agreement score was 90%. this is comparable to that reported by Ottenbacher and Petersen (1985) and Goodwin and Goodwin (1985).

The results of indexing the occupational therapy literature for each journal individually are included in tables found in Appendixes A to E. Summary information is found in Appendix F.

Table 2 presents the summary data for types of statistical procedure employed by categories of parametric and nonparametric procedure.

Table 2

Frequencies Across Years (1980-1984), Totals and Percentages of Parametric, Nonparametric and Nonspecified Procedures For All Journals Combined

	1980	1981	1982	1983	1984	Total	Percent
Parametric Procedures	43	63	59	76	92	333	77.44
Non-Parametric Procedures	4	16	25	11	22	78	18.14
Non Specified Tests	3	3	4	2	7	19	4.42
Totals	50	82	88	89	121	430	100.00

Note. Non-specified tests refer to correlations or tests of difference which were not specified by the author in the respective article.

Note. Articles were indexed according to types of statistical procedures employed. Thus, one article could have more than one statistical procedure and accounts for a greater number of statistical procedures compared to the number of statistical articles.

Table 2 indicates that parametric procedures outnumber nonparametric procedures at a ratio of approximately four to one. Table 3 presents the frequency and rank order of the particular parametric procedures found in the occupational therapy literature.

Table 3

Summed Across Years (1980-1984) and Rank Order of Parametric Statistical Procedures for All Journals Combined

Procedures	Total Frequency	Rank Order
Pearson Correlation	74	1
Chi Square	49	2
Student t	4	12
Paired t	26	7
Independent t	46	3
One Way ANOVA	35	5
Factorial ANOVA	44	4
Planned Contrasts	3	13
Post-hoc MC	27	6
ANCOVA	10	9
Regression	14	8
Discriminant	5	11
Canonical Correlation	0	0
Factor Analysis	4	12
MANOVA	1	15
Meta Analysis	2	14
Other	9	10

Note. The "other" category includes infrequently indexed procedures such as weighted kappa, path analysis, post hoc power analysis, multiple correlation analysis, Chronbach's Alpha and confidence intervals.

Table 3 shows that the ten most commonly employed parametric techniques in the occupational therapy literature are the Pearson Product Moment Correlation, the Chi Square, the independent t-test, factorial ANOVA, one way ANOVA, post-hoc multiple comparisons used with ANOVA, paired t-test, regression, analysis of covariance, and the category of other.

Goodwin and Goodwin (1985) categorized the Pearson Product Moment Correlation, Chi Square, forms of the t-test, and the one way ANOVA as basic techniques. Other procedures such as regression and factorial ANOVA were classified as advanced procedures. Based upon an adaptation of their scheme, recategorization of the frequency of statistical procedures in occupational therapy literature into basic and advanced procedures is presented in Table 4.

Table 4

Basic Versus Advanced Techniques in Occupational Therapy Literature

Category	Frequency	Percentage
Basic	234	68.02
Advanced	110	31.97
Total	344	100.00

Note. These calculations do not include the category of other.

Table 4 reveals that basic techniques are employed approximately 66% of the time, about twice as often as advanced techniques. Since the indexing methodology was essentially similar, comparisons can be made with the results presented by Goodwin and Goodwin (1985) when they indexed the literature of the _American Educational Research Journal_ (AERJ) from 1979 to 1983. According to their definition, they found basic techniques comprised about thirty-three percent of the overall number of procedures used. Thus, occupational therapy research literature reflects the use of basic procedures moreso than literature in educational research at a ratio of approximately two to one.

Table 4 further reveals that advanced procedures were employed in about 31% of the cases for all journals over this five year period. Ottenbacher and Peterson (1985) similarly found that advanced procedures as they defined them were used in 31.97% of the cases for articles in the _American Journal of Occupational Therapy_ for 1983. Thus, the findings of these two separate investigations are similar.

Goodwin and Goodwin (1985) found the total percentage of nonparametric statistical techniqes to be ten percent. The percentage of nonparametric techniques in the current review of occupational therapy literature is much higher (refer to Table 2) at nearly 20%. Table 5 presents the frequency and rank order of particular nonparametric statistical procedures.

Table 5

Frequencies and Rank Order of Nonparametric Statistical Procedures for All Journals Combined by Year 1980-1984

Procedures	Total Frequency	Rank Order
Sign Test	1	9
Spearman's Rho	11	2
Kendall Tau	8	3
Point Biserial	1	9
Cramer's V	2	8
Fisher's Exact Test	6	5
Mann Whitney U-test	14	1
Wilcoxon Matched Pair Signed-Rank Test	5	6
Kruskall-Wallis ANOVA	7	4
Randomized Permutation Test (ANOVA)	1	9
Split Middle Test of Trend	4	7

Review of Table 5 reveals that the five most commonly employed nonparametric procedures in the occupational therapy literature were the Mann Whitney U-test, Spearman's Rho, Kendall Tau, Kruskall Wallis ANOVA, and Fisher's Exact Test. The first four of these procedures deal with correlation and testing for differences between groups. Thus, the most commonly used nonparametric procedures are comparable to the most commonly used parametric procedures in occupational therapy literature, for those also dealt primarily with correlation and testing for differences between groups. One might surmise the bulk of the nonparametric procedures in occupational therapy literature could also be categorized as basic versus advanced techniques.

In summary, the review of the occupational therapy literature revealed that parametric statistical techniques are indeed the most commonly used group of statistical procedures and the review also indicated which of those procedures are most common. The current investigation will focus upon some of the parametric procedures previously identified as most commonly used. In order to appropriately limit the scope of the investigation, only univariate statistical procedures identified as the most commonly employed procedures in occupational therapy literature will be studied. Thus, the following procedures will be the basis of the current investigation: (a) Pearson Product Moment Correlation Procedure, (b) Paired t-test, (c) Independent t-test, (d) Regression, (e) and One-factor ANOVA.

Data Acquisition

Sample selection for case study investigations often utilizes purposive sampling. That is, cases are not selected at random but are selected purposefully according to some theoretical rationale (Yin, 1985). Since Minium (1978) stated that nonparametric procedures are most appropriate to use when the sample size is small, data sets from published occupational therapy research were identified as possible cases for the current investigation when (a) one of the previously mentioned parametric procedures was employed, and (b) when the sample size was small (less than fifty subjects). Thus, sampling was set up in order to assure a collection of extreme cases (small sample size for investigation of parametric versus nonparametric procedures). Accordingly, four to five articles per statistical procedure were identified for possible use. The category of statistical procedure in each article is the case, and the unit of analyses are the parametric and corresponding nonparametric procedures.

Of those articles so identified for each category of parametric procedure, the two with the smallest sample size were selected for possible use in the current study. Once such articles were identified, the authors were contacted by letter followed by telephone contact. Initially, an attempt was made to contact fourteen authors. All contact occurred between September, 1984 and May, 1985. The purpose of the current study was explained and the authors' data sets were requested for use. Not all authors were successfully contacted and some no longer had the data sets. Since some authors responded that they could not provide data sets, and in certain instances it was not possible to make contact with an author, an additional five articles were selected for possible use. These articles were selected from later years so that the authors' would be more likely to still have the data sets. Attempts were made to contact the authors of this set of articles.

In all, five authors could not be located and contacted. One author refused to participate in the study, and three authors were willing to participate but could not locate their data sets. Ten authors sent their data sets. Of these, three were duplicates, i.e., two data sets for one category of procedure. In such cases, the data set with the smallest sample was selected from among the duplicate sets. Thus, seven data sets were available for use in the investigation. (Note that at this time ANCOVA and factorial ANOVA were part of the investigation but were subsequently dropped as part of limiting the scope of the investigation).

Data sets were inputted into the computer as data bases. Values of data points were checked and confirmed before commencing with the analysis. The level at which variables were measured was evaluated according to criteria set forth by Leedy (1974), Minium (1978), and Siegel (1956).

Data Analysis

Data Exploration

Tukey (1977) coined the phrase "data exploration" to describe a step-by-step process of employing descriptive analysis in order to really understand the characteristics of a data set. This is commmonly referred to as exploratory data analysis or EDA. Other authors (Looney & Gullidge, 1985) similarly recommend probability plots and correlational analyses to check distributions and characteristics of data sets. Thus, each of the five data sets were subjected to data exploration (Fienberg, 1979; McGill, Tukey & Larson, 1978; Beniger & Robyn, 1978). However, a minimum criteria of a sample size of at least five per group was set in order to conduct exploratory data analysis (Hoaglin, Mosteller & Tukey, 1983). Thus, those data sets or subsets with less than five were not subjected to EDA.

One or more of the following procedures was conducted with each data set in order to test the assumptions underlying use of the respective parametric analysis employed. All tests were performed on SAS (Statistical Analysis System User's Guide, 1982) unless otherwise specified. The procedures used were:

1) stem-and-leaf plots;
2) manually executed boxplots;
3) the Shapiro Wilks test or W statistic;

4) the F test was used to test for equality of variance in two sample situations (Statistical Analysis System, 1982);

5) the Bartlett-Box F test of homogeneity of variance was used for three or more samples but was run using another statistical package, i.e.,Statistical Package for the Social Sciences (1983);

6) the Squared Ranks test for distribution free testing of variance for k groups was computed by hand for appropriate cases (Conover, 1980);

7) correlations of one variable with another, and;

8) scattergrams.

The Shapiro-Wilk test was chosen since it is an excellent test when small samples are employed (Shapiro & Wilk, 1965). Procedures for the manually executed boxplots are included in Appendix G.

Note that the above procedures refer to univariate tests for normality. In the bivariate case, the Pearson Product Moment Correlation, procedures for testing bivariate normality were based upon Filliben (1975).

If exploratory data analysis could not be conducted due to small sample size, certain assumptions about the data set were made. First, it was assumed that the central limit theorem did not necessarily apply. And, since evidence suggests that extremely small samples often do not follow a normal distribution (Lezak and Gray, 1984) it was assumed that distributional characteristics of the data set were questionable as regards normality.

Replication of Parametric Analysis and Selection and Calculation of Nonparametric Analysis

Each of the parametric analyses were replicated. Subsequently, the results of the exploratory data analysis were used in order to determine the appropriate nonparametric statistic. The nonparametric statistic was also selected based upon the research question the original author wished to investigate.

Considering the results of the replication of the parametric analyses and the analyses using nonparametric statistics, the probability values (p values) of the significance tests from the two sets of procedures were compared in tabular form. An efficiency comparison of selected parametric versus nonparametric procedures under conditions of normality is presented in Table 6. This table served as a reference for evaluation of the parametric and nonparametric p values obtained in the current investigation. Note that since comparisons are made under conditions of normality, a real advantage for the power for the parametric test is built into the figures presented (Bradley, 1968).

Specific guidelines for procedures and interpretation regarding the comparison of parametric versus nonparametric statistics follow in the next and last section of this chapter.

Table 6

Comparable Efficiency of Selected Parametric Versus Nonparametric Tests

Parametric Test	Nonparametric Test	ARE*	Reference
Independent t-test	Wilcoxon Rank-Sum	.955	Bradley (1968)
Paired t-test	Wilcoxon Matched Pairs Sign Test	.637	Gibbons (1976)
	Wilcoxon Signed Rank Test	.955	Gibbons (1976)
Pearson r	Spearman's r	.91	Siegel (1956)
	Kendall Tau	.91	Gibbons (1976)
One Facor ANOVA	Kruskal-Wallis	.955	Gibbons (1976)
Regression	Thiel (slope only)	.955	Hollander & Wolfe (1983)

Note. ARE = asymptotic relative efficiency for normal distributions.

Case Study Guidelines

Yin (1984) identified the delineation of the criteria for interpretation as a necessary step in case study methodology. Therefore, the following guidelines have been developed in order to structure the interpretation process of comparing parametric versus nonparametric tests. This process of interpretation is a complicated one based upon iterative analysis of the data sets (Chambers, Cleveland, Kleiner, and Tukey, 1983) and factors related to the statistical tests and assumptions underlying their use.

A limitation to the interpretation guidelines needs to be presented. In describing exploratory data analysis, Velleman and Hoaglin (1981) state:

> Exploratory methods . . . call for frequent application of the analysts' judgement, and the judgement cannot readily be cast in simple rules . . ." (p. XVI).

Indeed, this statement is applicable to the entire process of developing guidelines for the comparison of parametric versus nonparametric tests because situation specific difficulties can arise which may alter guidelines and demand judgmental decisions. However, considering this possible limitation to the guidelines, they are presented in six main categories: preliminary procedures, analysis of characteristics of the data set, analysis of characteristics of the statistical tests, comparison of significance levels, supporting data, and synthesis. The following are guidelines for the preliminary procedures.

Guidelines for Preliminary Procedures.

I. Research Design and Statistical Test

A. What is the research question expressed by the original author?

1. Specify the hypothesis.

2. State alternative hypothesis or outcomes.

B. What is the research design?

1. Identify number of groups in the study.

2. Identify characteristics of the groups regarding validity and reliability of measurement.

C. Confirm the parametric statistical test used by the author.

D. Does the test fit with the research design?

1. If yes, continue to E.

2. If no, evaluate suitability of data set for use in the study.

E. What are the conclusions of the author?

1. What is the statistical conclusion?

2. What is the substantive conclusion?

F. At what level of measurement are the variables?

1. Use criteria of Leedy (1974).

2. Use criteria of Siegel (1956).

3. Use criteria of Minium (1978).

H. Is the level of measurement consistant with the statistical test employed?

1. If yes, continue to II.

2. If no, evaluate suitability of the data set for use in the current study.

II. Replication of the Parametric Analysis

A. Execute the parametric analysis.

B. Has the parametric analysis of the original author been replicated?

1. If yes, continue to C.

2. If no, investigate why. Contact original author to clarify situation as needed.

C. Review the assumptions underlying the use of the parametric test.

The following are the case analysis guidelines.

Case Analysis Guidelines.

I. Analysis of the Characteristics of the Data Set

A. What is the size of the data set?

1. If the data set or subset thereof has a sample size of five or more, that data set or subset thereof will be subjected to exploratory data analysis.

2. If the data set or subset thereof has a sample size of less than five, it will not be subjected to exploratory data analysis since the sample size is too small to render judgments about EDA.

B. What is the shape of the data set?

1. This will be revealed by a boxplot and a stem and leaf plot.

2. Is the data set symmetric?

3. Is it skewed? Skewness is determined by an absolute value greater than .5 (Walburg, Stykowski & Hung 1984). Mildly skewed would be a score of .5 to .99. Moderately skewed would be a score of 1.0 to 1.99. Profoundly skewed would be a score of 2.0 or greater.

4. Are there long tails present?

C. Is the distribution normal?

1. Univariate tests of normality such as the Shapiro-Wilkes will address this question. The normal probability plot will also provide information.

2. Bivariate normality will be determined by the process of computing the generalized distance squared and chi square values of each two variables upon which correlations are based. Consequently, bivariate normality will be determined considering the following categories:

2.1 Percentage of values falling within the ellipsoid at the .5 level. At least 50% of the scores must be at this level.

2.2 Percentage of values falling within the ellipsoid at the .9 level. At least 90% of the scores must be at this level.

2.3 Linearity of the scattergram of the chi square value by the generalized distance squared.

2.4 Degree of correlation between the generalized distance squared to the chi square value. An acceptable correlation is .90 or greater.

2.5 In judging whether or not the correlation was bivariate normal, the following criteria were used. First, it was determined if there were irregularities in three of the four categories (considering outliers and linearity simultaneously). Presence of three irregularities suggested non-bivariate normality. Second, if less than three categories were irregular, but the generalized distance squared percentage was less than .45 for 50% or less than .85 for 90%, the correlation was considered marginally bivariate normal or non-bivariate normal.

D. What is the range of the data?

1. The range will be revealed by stem and leaf displays, standard deviations, and min/max determinations.

2. Gaps in the data will be determined by the stem and leaf plot.

3. Repeated values will be revealed by the stem and leaf.

4. Single outliers will be revealed by the box plot.

5. Two or more data points acting together as an outlier will be revealed by the scattergram.

E. Is the relationship between variables linear?

1. Results revealed by scattergram.

2. If yes, continue to F.

3. If no, identify the nature of the relationship.

F. Using appropriate evidence drawn from the exploratory data analysis, conclusions about the distributional characteristics of the data sets will be summarized and justified.

II. Evaluation of Assumptions Underlying Use of the Parametric Test

A. Determination will be made of which assumptions were tested during the exploratory data analysis. Any assumptions, such as normal distribution of the error term in regression analysis, which were not tested will be tested at this point.

B. Based upon the EDA, each assumption underlying use of the parametric test will be either accepted or rejected, or classfied as questionable.

C. For those assumptions which are rejected or questionable, relate the severity of the assumption as documented in the literature.

III. Nonparametric Analysis

A. Select the nonparametric test according to the following criteria.

1. EDA

2. Guidelines in Royeen and Seaver (1986)

3. Level of Measurement

4. Sample Size

5. Recommendations in the literature

B. Have the assumptions of the nonparametric test been met?

C. Execute the nonparametric test.

D. Chart the results (critical values and probability levels) of the nonparametric, replicated parametric, as well as the orginial parametric analysis. Difference in p values will be operationally defined in two ways. First, the same or different outcomes in reference to a p value of .05 will be used (Eisenhart and Swed, 1940). Second, a p value difference equal to or greater than .118 will be used as this was the median difference identified by Gans (1984) when he investigated five different tests run on seven different samples varying by six distributions.

1. If the results between the parametric versus the nonparametric test are ambiguous, run a robust procedure.

   1.1 Select a robust procedure.

   1.2 Execute it.

   1.3 Analyze it.

   1.4 Go to interpretation.

2. If there was a clear difference between the results of the parametric and nonparametric analyses, but unknown distributional characteristics of the data set (N is less than

five) run a robust procedure if feasible. If not, assume the ambiguity surrounding assumptions of the parametric procedure weighs in favor of the nonparametric procedure, and go to interpretation.

3. If there is a clear difference in the parametric versus the nonparametric test under known distributional assumptions and documented violations of assumptions, go to interpretation.

IV. Guidelines for Interpretation

A. Compare the probability levels considering ARE under normal conditions identified in Table 6 and the EDA. Also, refer to literature on power conditions of nonparametric tests under conditions of non-normality, as available.

B. Compare the probability levels considering which, if any, assumptions have been violated and relate it to the literature.

C. Consider the substantive results between the sets of procedures in light of subjects and measurement characteristics.

D. Compare probability levels for those sets with robust procedures.

E. Consider any supporting or clarifying data from the literature.

VI. Synthesis

Based upon all interpretation, an overall synthesis will be made relative to which procedure, if either, was more appropriate in which case and why.

Summary

Chapter three has presented the research design employed in the current investigation. Since a case study methodology was employed, the research design included specification of the steps of the process using guidelines for data analysis and interpretation. The next chapter will present the results of the investigation.

Chapter Four

Results

Introduction

The current chapter will present results of each case study regarding the following categories of statistical analyses; paired t-test, independent t-test, one factor ANOVA, correlation and regression. For each category, a description of the data set is presented. Also, results of the exploratory data analyses are presented. Subsequently, results of the data analysis are put forth as is a concluding summary paragraph for each category of procedure.

Paired t-test

Description of the Data Set.

The paired t-test analysis was conducted using data, included in Appendix H, from a study by Nelson, Weidensaul, Anderson and Shih (1984). Their investigation sought to determine if there is a significant difference between the duration of nystagmus (a rhythmical back - and - forth eye movement which is an automatic response to spinning) as measured by the Southern California Postrotary Nystagmus Test (SCPNT) during bright light and dim light conditions. Null and alternative hypotheses were tested.

HO: There is no significant difference.

HA: There is a significant difference.

The research design was composed of a test and retest situation using one sample. Subjects, aged eighteen to thirty-five, were tested under both bright and dim lighting conditions. The conditions were randomly counterbalanced but the subjects were not randomly sampled. Ten females and eight males participated in the study.

A paired t-test for dependent samples was employed by the authors for the analysis. Interval level data were used so that the level of measurement was appropriate for the statistical procedure employed. Since the null hypothesis was not rejected, it was concluded that lighting conditions did not influence nystagmus duration.

Exploratory Data Analysis.

The data (scores of the difference between the nystagmus duration between bright and dim light conditions) were subjected to exploratory data analysis. A boxplot of the difference scores is presented in Figure 1.

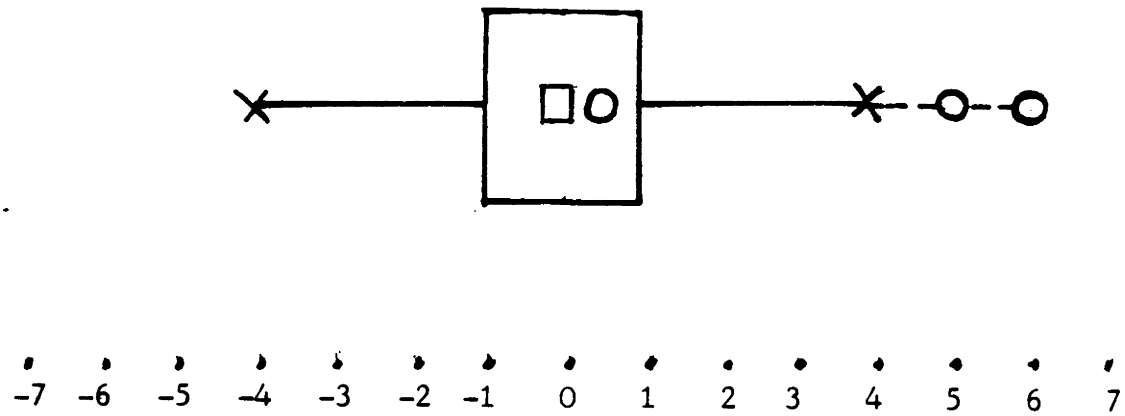


FIGURE 1. Boxplot of the Difference Scores

The boxplot of data from the difference scores presented in Figure 1 reveals that this distribution may be slightly skewed since the mean and median are not coincident. Also, there are two mild outliers associated with the right tail. Considering the presence of two outliers in the right tail, the distribution of the data set is mildly asymmetrical.

Table 7 presented results of more formal tests used for exploratory data analysis.

Table 7

Distributional Testing for Difference Scores Between Bright and Dim Light Conditions

Test	Value
Normality	W=.912587 p=.095
Skewness	.845092

The distributional testing presented in Table 7 indicates that the data set tests as normally distributed, but just marginally so (W = .912587, p = .095). And, the data set is mildly skewed beyond the absolute value criterion of .5.

In summary the data set tests as marginally normal but is skewed and has two mild outliers.

Results of the Data Analyses.

The Sign Test and the Wilcoxon signed Rank Test were identified as analogs to the paired t-test for data analyses. However, the Wilcoxon Signed Rank Test assumes symmetry and this assumption may not be tenable since the data is mildly skewed and has two outliers. Thus, the Sign Test was employed considering that no assumption of symmetry is required by this procedure. Table 8 presents the results of the original parametric analysis, the replicated parametric analysis and the Sign Test analysis.

Table 8

Results of the Originial, Replicated and Nonparametric Analog for the Paired t-test (two tailed)

Analyses	Test Statistic		p Value
Original	$t=-0.68$	$df=17$	.5061
Replicated	$t=-0.68$	$df=17$	.5061
Sign Test	$s=8$ $s=9$	$n=18$ $n=18$	$.8146 < p < 1.00$ @

@ Solved for most and least advantageous to rejection of null hypothesis (range of probability) considering ties or zero differences (Gibbons, 1985).

Table 8 reveals that the Sign Test produces a much higher range for the p value ($.81456 < p < 1.00$) than does the paired t-test (.5061).

Application of the case study interpretation criteria to evaluate similarity and difference of the findings reveals that in terms of substantive meaning, the findings are similar. To illustrate, each test accepts the null hypothesis at alpha = .05. However, consideration of the meaningful difference in p values, i.e., equal to or greater than a difference of .118, suggests that the findings are different in terms of probability levels. The p value of the Sign Test solved most advantageously regarding ties minus the p value of the replicated paired t-test (.8146 - .5061 = 0.3085) is greater than the set criteria of .118. The discrepancy in p values is even greater if solved using the Sign Test solved most advantageously considering ties.

If one reviews the raw data presented in Appendix H, it is readily apparent that the two data sets are essentially similar. Considering the difference in p values between the parametric and nonparametric tests, one may surmise that the Sign Test indicates a greater similarity between data sets than does the paired t-test (.8146 to 1.00 versus .5061). Considering that under certain conditions of non-normality the Sign Test can be up to two times more powerful than the t-test (Gibbons, 1985), and, in light of the presence of irregularities in the current data set (marginal normality, skewness and outliers), it appears that the Sign Test more powerfully reveals the similarly of the bright and dim data sets, and that the nonparametric test more validly reflects the degree of similarly of the data sets by yielding the higher p values. Thus, for the first case it appears that the nonparametric test does as well or better than the parametric test.

In summary, it appears that there is a difference in p values between the parametric and nonparametric tests, and that the nonparametric p value more powerfully reflects the degree of similarity between the data sets from the bright light and dim light conditions, when, in this case, skewness, outliers, and marginal normality were present in the data sets consisting of difference scores.

Independent t-test

Description of the Data Sets.

The independent t-test was analyzed using data from an investigation by Slavik (1982) and the data sets are included in Appendix I. The investigation studied duration of postrotary nystagmus, defined in the previous section, as measured by the Southern California Postrotary Nystagmus Test (SCPNT) in normal and strabismic children. Null and alternative hypotheses were tested.

HO: There is not a significant difference in duration of nystagmus between normal and strabismic children.

HA: There is a significant difference in duration of nystagmus between normal and strabismic children.

The subjects were from two different settings. Ten normal children were from a Montessori Day School whereas the five strabismic children were from an optometrist's clinic. Random sampling was not employed and the subjects were administered the SCPNT one time.

A two tailed, two sample t-test for common dispersion was used for the original analysis. The data were measured at the interval level so the statistical test was appropriate to the level of measurement.

The null hypothesis was rejected, so it was concluded that strabismic children differed from normal children in terms of the duration of postrotary nystagmus.

Exploratory Data Analysis.

Figure 2 presents the boxplots of the data sets from normal and strabismic children.

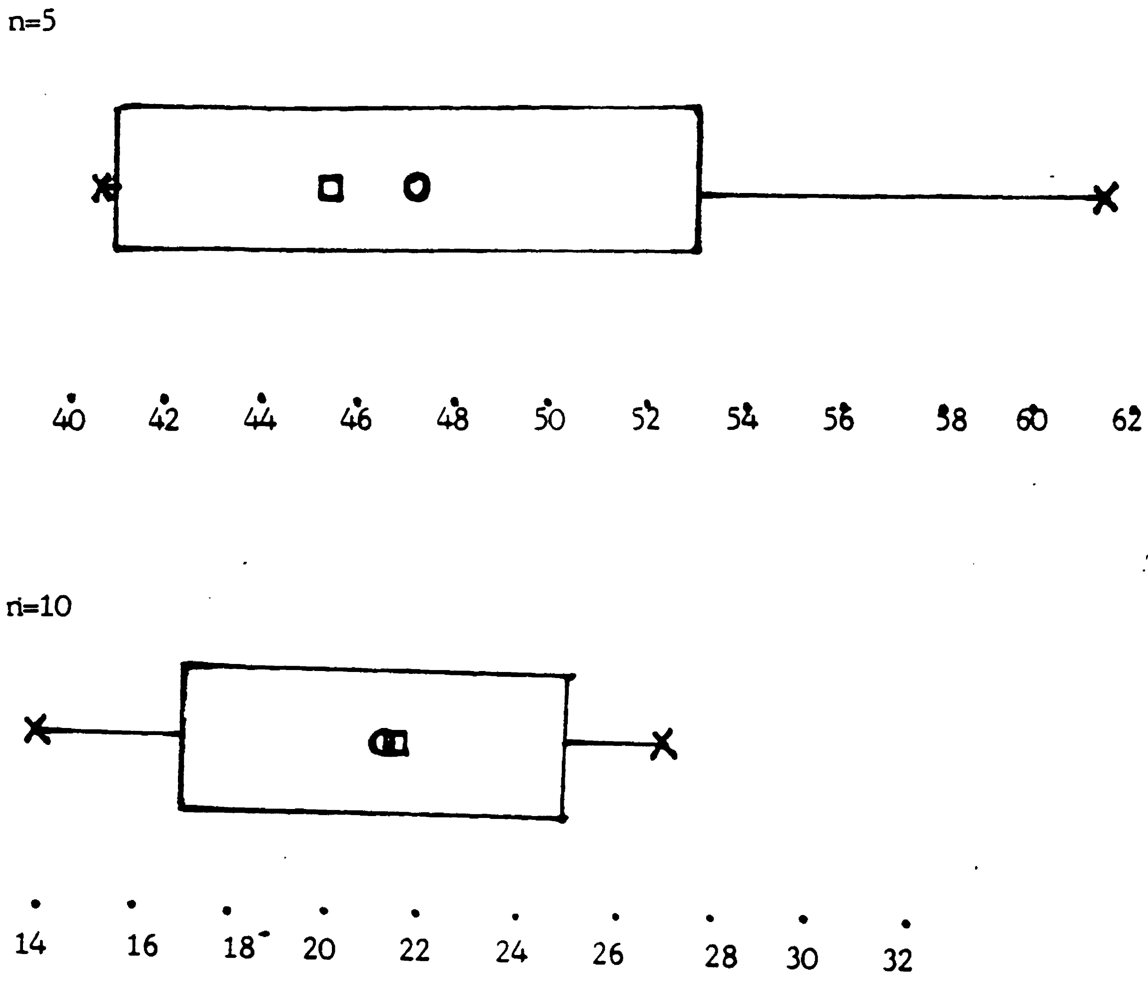


FIGURE 2. Boxplots of Data from Normal and Strabismic Children

Figure 2 reveals that the data set from the normal group (n=10) appears symmetric. However, the data set from the strabismic group (n=5) is positively skewed; furthermore, the left tail is almost nonexistent and the mean and median are divergent.

Table 9 presents the result of formal testing for distributional characteristics.

Table 9

Distributional Testing for Data from Normal and Strabismic Children

	Data by Group	
Test	Normal Children	Strabismic Children
Normality	W=.914503 p=.363	W=.892229 p=.385
Skewness	-.454793	.891195

Results in Table 9 indicate that the two data sets test as normally distributed. The data set from the normal children is not skewed. But, the data set having to do with strabismic children exceeds the absolute value of .5 in terms of degree of skewness and is therefore a mildly skewed distribution, as was indicated by the boxplot.

In addition, since equality of variance between the two groups is an assumption underlying the two sample t-test, testing for variances was done. The results of an F test were equivocal ($F=3.67$, $df=4/9$, $p=.0962$). Due to the small sample sizes involved (5 and 10) and due to the skewness of the strabismic sample, a distribution free test of variances, the Two-Sample Squared Ranks Test of Variances, was calculated (Conover, 1980). Conover and Iman (1981) report the asymptotic relative efficiency of this test to the F test to be .76 for a normal distribution. The Squared Ranks test did not support the assumption of homogeneity of variances ($T=3.2265996$, $p < .01$). Considering that the results of the F test indicated marginal acceptance of the assumption of homogeneity of variances and the results of the Squared Ranks test rejected the assumption, replication of the t-test employed testing under conditions of common as well as uncommon dispersion. This was similarly done for the nonparametric analysis.

Results of the Data Analysis.

For the common dispersion analysis, the Mann Whitney Wilcoxon test was used; and for the uncommon dispersion analysis, a modification of Mood's median test for the generalized Behren's-Fisher problem (Fligner & Rust, 1982) was used since it is a distribution free test. Table 10 presents the results of the analyses.

Table 10

<u>Original, Replicated and Nonparametric Analog for Common and Uncommon Dispersion Tests of Location for Two Samples</u>

	Dispersion	
Test	Common	Uncommon
Original	t=7.77 p<.0001	
Parametric	t=7.8518 p=.0001	t=6.3513 p=.0013
Nonparametric	m=5,n=10,SumR=65 p=.000	T=10.271404 @ .005<p<.001

@ The test statistic was in between two critical values. Thus, the corresponding upper and lower values of p are presented.

Table 10 reveals that two sample testing assuming common dispersion yields nearly identical p values. Furthermore, the nonparametric test yielded an equivalent or slightly lesser p value (depending upon how ties were resolved) in the case of uncommon dispersion depending upon where the p value fell within the possible range.

Application of the criteria for evaluation of the similarity or differences between the findings reveals that the substantive conclusions are the same for the parametric and nonparametric tests for common as well as uncommon dispersion. Similarly, there are no differences among any of the p values using the criterion of a meaningful difference of at least .118. Thus, for the second case it was conluded that the nonparametric test performed as well as the parametric test.

In summary, it is concluded that for the current analysis there is no meaningful or substantive difference regarding p values for the parametric and nonparametric analysis when one data set (the strabismic one) was slightly skewed and homogeneity of variance was marginal.

One Factor ANOVA

Description of the Data Set.

An analysis of a one factor ANOVA was conducted using data from an investigation by Cunningham and Trickey (1983) in which they ran three separate ANOVAs. The data are included in Appendix J. They studied personality characteristics of occupational therapy students, as measured by the Learning Styles Inventory (LSI) in order to determine if groups of students with an identified learning style (accommodator, assimilator, divergent, and convergent), differed by performance on grade point average, psychiatric internship, or physical dysfunction internship. For each condition of performance a separate one factor analysis of variance was calculated. Thus, three one way ANOVAs were computed by Cunninghan and Trickey (1982). The null and alternative hypothesis, applicable to each one factor ANOVA, follows:

HO: There are no significant differences between groups regarding performance.

HA: At least one group differs significantly from the others regarding performance.

The subjects were thirteen senior level occupational therapy students, eleven females and two males. Each was administered the LSI in order to determine learning style and corresponding group membership.

The null hypothesis was accepted in all three data analyses; there were no significant differences among groups.

Level of measurement employed is problematic. Even though grade point scales for GPA and the internships are, theoretically, interval level of measurement, inspection of the data sets reveals that in actuality the scores appear to function more on the ordinal scale of measurement due to the lack of range in the scores and numerous ties. Thus, there is evidence that the parametric one factor ANOVA, based upon level of measurement, does not match the research design. Additionally, the extremely small sample sizes of each group (N=6,2,3,2, respectively) also raises questions about the match between the statistical procedure and research design.

Exploratory Data Analysis.

Exploratory data analysis can only be conducted with one group, the accommodators, since that is the only group with a sample size of five or greater. And of the three performance categories, formal exploratory data analysis can only be done for GPA since that is the only data set of the three performance categories that has some degree of variance. The other two performance categories have the same score for all subjects in the accommodator group.

The boxplot for the accommodator group for performance on GPA is presented in Figure 3.

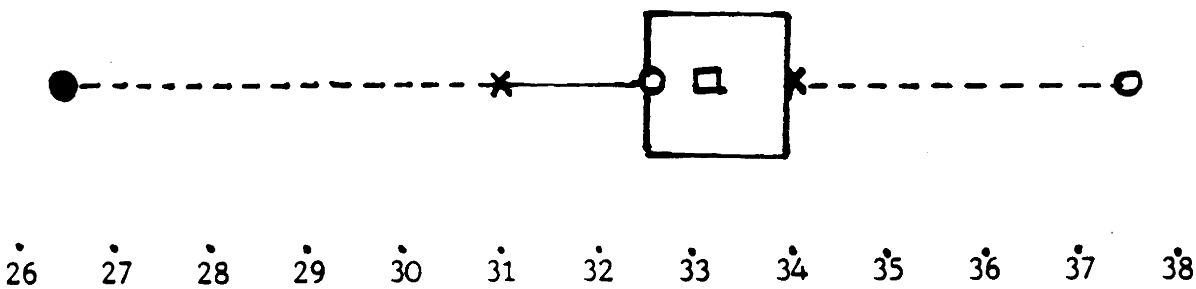


FIGURE 3. Boxplot of Data from Accommodator Group's Performance on GPA

Figure 3 reveals the following regarding the data set. One mild outlier and one extreme outlier is present. And, the median and mean are not coincident.

Table 11 presents the results of more formal tests of exploratory data analysis as well as logical analysis of the accomodator group for all three performance categories.

Table 11

<u>Distributional Testing and Logical Analysis of Accommodator Group for the Three Performance Conditions</u>

Condition	Normality	Skewness	Variance
GPA	W=.956731 p=.753	-.641909	.1360002
Psychiatric Internship	*	*	0.0
Physical Dysfunction Internship	*	*	0.0

<u>Note</u>. "*" indicates that SAS was unable to compute the value due to the lack of variance in the data sets.

Table 11 reveals that for accommodators in the GPA performance category, the distribution tested as normal. However, the data set is mildly skewed. In addition, it can be concluded that the accommodator group in the performance categories of psychiatric and physical dysfunction internships can be considered non-normal since all the data values were the same and lack any variance which is, by definition, non-normal.

Since homogeneity of variances between groups is an assumption of a one factor ANOVA, testing for variances was done. A distribution free test, k sample Squared Ranks Test for homogeneity of variances (Conover, 1980) was chosen since the F test can be sensitive to non-normality. The results are presented in Table 12.

Table 12

Results of k Sample Squared Ranks Test for Homogenity of Variances

Condition	Test Statistic	df	p Value
GPA	T=1.6947459	3	$p > .25$
Psychiatric Internship	T=11.089295	3	$.025 < p < .01$ @
Physical Dysfunction Internship	T=9.1868899	3	$.05 < p < .025$ @

Note. @ The test statistic was in between two critical values. Thus, the corresponding upper and lower values of p are presented.

Table 12 indicates that for GPA the assumption of homogenity of variances has been supported. However, for the psychiatric and physical dysfunction internship performances the assumption has been violated.

For data from psychiatric and physical dysfunction internships the conditions of heterogenity of variances exist as does non-normality. For the GPA performance category, the accomodator group tests out as distributed normally, has one mild and one extreme outlier, is mildly skewed, and homogeneity of variance across groups is present.

<u>Results of the Data Analysis</u>.

A one factor ANOVA using the F statistic was applied to the data transformed into ranks (SAS, 1982). This is equivalent to the Kruskal-Wallis using the chi square distribution (Hollander & Wolfe, 1973), but is easier for generation of p values and is often more powerful (SAS, 1982). Furthermore, when alpha levels are in a range less than .10, and when groups are small, the chi square distribution is not exact for the Kruskall-Wallis test, but is conservative (Conover and Iman, 1980). Consequently, use of the Kruskal-Wallis when three of the fours groups has a sample size less than five is questionable. Thus, the F statistic applied to ranks was chosen for the analysis not as the optimal data analysis procedure for this problematic data set, but rather as the "least compromised" procedure. The average rank method of compensating for ties was used (SAS, 1982).

Table 13

Results of the Orginial, Replicated and Nonparametric One Factor ANOVA Analyses by Performance Condition

Condition	Original	Parametric	Nonparametric
GPA	@ p>.05	F=0.49 p=0.7001	F=0.64 p=0.6059
Psychiatric Internship	@ p>.05	F=1.39 p=0.3074	F=1.29 p=0.3348
Physical Dysfunction Internship	@ p=0.056	F=3.69 p=0.0557	F=4.76 p=0.0297

Note. @ The F value was not supplied by the author.

Note. The Kruskal-Wallis procedure with the test statistic adjusted for ties was computed for comparison. The results (GPA, H=2.12, df=3, p=.5477; psychiatric internship, H=3.69, df=3, p=.3060; physical dysfunction internship, H=7.36, df=3, p=.0612) confirm the finding of Conover and Iman (1980) that the chi square distribution for small samples (less than five per group) when the alpha is less than .10 is not exact and is conservative. The F statistics applied to ranks did produce a more "powerful" result as indicated would be the case in SAS (1982).

Table 13 reveals that using the .118 criteria for meaningful differences in p values between procedures, there are no differences in the p values generated (GPA, 0.7001 - 0.6059 = 0.1042 < 0.118; Psychiatry, 0.3074 - 0.3348 = 0.0274 < 0.118; Physical Dysfunction, 0.0557- 0.0297 = 0.0349 < 0.118). However, for the GPA performance category the p value difference is close to the set criterion.

Table 13 further reveals differences regarding substantive conclusions when evaluating the p values at the 0.05 level of significance. For performance conditions of GPA and psychiatric internship, the substantive conclusions are the same. However, for the physical dysfunction internship, the substantive conclusion changes and the null hypothesis would be rejected at the 0.05 level of significance if using the nonparametric analysis.

In summary, there are no substantive or meaningful differences between the parametric and the nonparametric tests regarding the GPA performance category when homogeneity of variance existed. The accommodator group of the data set was skewed and had one mild and one extreme outlier. The psychiatric internship one factor ANOVA had no meaningful or substantive difference in p values when non-normality and heterogeneity were present. Thus, for these units within the case there were no differences between the parametric and nonparametric analyses.

The one factor ANOVA with the physical dysfunction internship had a substantive difference in p values. At the .05 level of significance one would reject the null hypothesis using the nonparametric analog, but would accept the null using the parametric analysis. In this case, the accommodator group of the data set was not distributed normally, and heterogeneity of variance existed. Bradley (1978) identified that the efficiency of the Kruskal-Wallis can be 1.00 under conditions of non-normality. Thus, given that the normality assumption was clearly violated and based upon Bradley's findings, it appears the relative efficiency of the nonparametric procedure is greater than the parametric test for this unit within the case. Considering that homogeneity of variance assumptions may also have been violated, it was concluded that the nonparametric ANOVA was more powerful than the parametric ANOVA and that the nonparametric test performed "better" than the parametric test in this situation.

Pearson Product Moment Correlation

Description of the Data Set.

The analysis using a Pearson Product Moment Correlation was executed using data, included in Appendix K, from an investigation by Koomar and Cermack (1981). The purpose of their investigation was to establish test reliability of the consonant vowel (CV) and digit (D) formats of dichotic listening with normal and learning disabled children aged seven to ten using different formulas for calculation. Dichotic listening is a central nervous system integration process measured by the counting of sounds reported to be heard in one ear compared to the other ear (Koomar & Cermack, 1981). Thus, the variables under study are fourteen methods of calculating dichotic listening performance under two conditions (test and retest). The variables are identified by the formula used to derive them. For ease of presentation, the formulae for computing the variables for the test and the retest conditions are converted to alphabetical designations. The specific formula corresponding to the alphabetical designations A through N are presented in Appendix L with 1 representing the test condition and 2 representing the retest condition.

Thirty children, fifteen normal children (four girls and eleven boys) and fifteen learning disabled children (four girls and eleven boys) were tested and then retested one week later at the same time of day. Within the scope of common practice for reliability analysis, the research design and the statistical procedure are consistent. Also, the level of measurement of the variables, ratio, is consistent with the statistical procedure.

The substantive conclusions were that normal children had a higher reliability than learning disabled children, and that both groups had higher reliability using CV versus the D format.

Exploratory Data Analysis.

The seven procedures (for seven different formulas or ways of calculating reliability indices for dichotic listening) for two conditions, CV and D for each group was subjected to exploratory data analysis.

In the previous analyses, boxplots were hand calculated according to procedures specified in Appendix G. Since the correlational analyses had fourteen analyses per subject category (learning disabled and normal children) and employed many computed variables, the boxplots for the correlational analyses were executed using SAS plotting procedures. Table 14 presents summary boxplot information and information regarding univariate normality as well as linearity of the plot of each set of two variables for the normal children.

Table 14

Distributional Characteristics of Variables for Normal Children Using Univariate Analyses for Test (1) and Retest (2)

Variable (n=15)	Skewness	Outliers	Univariate Normality	Linearity
A1	-.31669	2	P=.703	Yes
A2	-.189574		p=.801	
B1	.123293		p=.909	Yes
B2	.617911	1	p=.320	
C1	-1.3405		p=.033	Yes
C2	.125934		p=.648	
D1	.416653		p=.750	Yes
D2	.759126		p=.164	
E1	-.380832		p=.581	Yes
E2	-.572251		p=.594	
F1	-.93264		p=.174	Yes
F2	-1.40133	1	p=.055	
G1	-.210002		p=.297	Yes
G2	-.3794552		p=.490	
H1	-.564318		p=.455	Yes
H2	-.127675		p=.479	
I1	2.30216	1*	p<.01	Yes
I2	1.58521	1	p<.01	
J1	-1.01779	1	p=.358	No
J2	2.17702	2,1*	p<.01	
K1	.214265	1	p=.910	Yes
K2	.269949		p=.933	
L1	-1.546286	1	p=.032	No
L2	.701565	1,1*	p=.412	
M1	.362703	1	p=.841	Yes
M2	.471803		p=.8781	
N1	-1.3112	1	p=.036	No
N2	.543336		p=.664	

Note. * Indicates an extreme, versus a mild, outlier.

Table 14 reveals that for normal children the following variables were either skewed, had outliers, were non-normal or lacked linearity with the paired test variable: A1, B2, C1, D2, E2, F2, H1, I1, I2, J1, J2, L1, L2, and N1.

Table 15 presents analysis of bivariate normality for the normal subgroup.

Table 15

Bivariate Normality Assessment for Normal Subjects (n=15)

Variables	Generalized Distance Squared .5 (a)	Generalized Distance Squared .9 (b)	Scattergram Linear	Outlier	Pearson	Bivarite Normality
A	.5666	.80	Gaps	2	.91474	Yes
B	.6666	.9333	No	2	.91937	Yes
C	.6	.9333	Gaps	1	.94446	Yes
D	.6	.8666	No	3	.90729	Yes
E	.5333	1.000	Yes	1	.98112	Yes
F	.5333	.9333	Partial	2	.87329	Yes
G	.40	.9333	Gaps	1	.93879	No
H	.5333	.9333	Partial	2	.85543	Yes
I	.6663	.933	Partial	1	.80204	Yes
J	.7333	.8666	No	2-3	.81858	No
K	.60	.933	Gaps	1	.97353	Yes
L	.7333	.80	Yes	3-4	.86293	No
M	.4666	.8666	Gaps	2	.96984	No
N	.7333	.80	No	2	.866882	No

(a) Refers to the percentage of values within the criteria of .50.

(b) Refers to the percentage of values within the criteria of .90.

Note. Refer to guidelines in chapter three for review of determination of bivariate normality.

Table 15 reveals the following sets of variables lack bivariate normality: B, G, I, J, L, M, and N.

Table 16 presents results of univariate exploratory data analysis for learning disabled subjects as a subgroup.

Table 16

Distributional Characteristics of Variables for Learning Disabled Children for Test (1) and Retest (2) (n=15)

Variable	Skewness	Outliers	Univariate Normality	Linearity
A1	.00708951		P=.346	Yes
A2	-.2255944	1,1*	p=.084	
B1	1.033457		p=.076	Yes
B2	.875769		p=.292	
C1	-1.0279		p=.019	Yes
C2	-.428709		p=.910	
D1	1.0499		p=.245	Yes
D2	1.43077	1	p=.037	
E1	-.940414		p=.162	Yes
E2	-.189489		p=.955	
F1	-.514549		p=.675	Yes
F2	-.932907		p=.459	
G1	-.0619533		p=.916	Yes
G2	-.2953142		p=.989	
H1	.369166		p=.855	Yes
H2	-.443554		p=.784	
I1	2.80876	1*	p<.01	Yes
I2	3.31239	1*	p<.01	
J1	2.04945	1,1*	p<.01	Yes
J2	2.2061	1*	p<.01	
K1	.626904		p=.644	Yes
K2	.327453	1	p=.285	
L1	1.156326	1	p=.189	Yes
L2	.136077		p=.619	
M1	.937305		p=.357	Yes
M2	.860352	1	p=.099	
N1	.841489	1	p=.640	Yes
N2	-.009425		p=.610	

Note. * Indicates an extreme, versus a mild, outlier.

Table 16 reveals that considering skewness, outliers, non-normality and lack of linearity, the following variables were affected: A1, B1, C1, D1, D2, E1, F1, F2, I1, I2, J1, J2, K1, L1, M1, M2, and N1.

Table 19 presents bivariate normality assessment for learning disabled children as a subgroup.

Table 17

Bivariate Normality Assessment for Learning Disabled Subjects (n=15)

Vari-ables	Generalized Distance Squared .5 (a)	Generalized Distance Squared .9 (b)	Scattergram Linear	Out-lier	Pearson	Bivariate Normality
A	.6	.866	No	2	.90555	Yes
B	.6666	.8666	Partial	2	.82590	No
C	.5333	.9333	Yes	–	.96767	Yes
D	.80	.866	No	3	.81754	No
E	.6	.866	No	3	.89570	No
F	.5333	.9333	Partial	1–3	.89939	Yes
G	.4	.9333	Yes	–	.98105	Marginal
H	.4666	.9333	Yes	–	.97792	Yes
I	.8666	.866	No	2	.73966	No
J	.8666	.866	No	2	.73217	No
K	.6666	.866	No	3	.86227	No
L	.7333	.866	No	3	.80222	No
M	.666	.80	No	3	.82303	No
N	.666	.866	No	3–4	.86049	No

(a) Refers to the percentage of values within the criteria of .50.

(b) Refers to the percentage of values within the criteria of .90.

Note. Refer to guidelines in chapter three for review of determination of bivariate normality.

Table 17 reveals that four of the pairs of the pairs of variables (A, C, F, H) are bivariate normal. One other, G is marginally bivariate normal and the others are all non-normally distributed.

Results of the Data Analysis

Kendall's Tau was identified as an alternative, nonparametric statistical procedure for association analysis. Tables 18 and 19 present the results of the original, parametric and nonparametric analysis for normal and learning disabled children, respectively.

Table 18

Original, Replicated and Kendall Tau Association Analysis for Normal Children on the Dichotic Listening Test (n=15)

Condition	Original	Replication	p value	Kendall Tau	p value	Difference
A	.82	.80303	.003	.55556	.0056	.0053
B	.80	.82614	.0001	.64660	.0013	.0012
C	.61	.60705	.0164	.42653	.0343	.0179
D	.82	.82199	.0002	.67013	.0008	.0006
E	.83	.86744	.0001	.67002	.0007	.0006
F	.87	.82890	.0001	.59807	.0023	.0022
G	.92	.92240	.0001	.79243	.0001	.000
H	.32	.32296	.2403	.32536	.0921	.1482 *
I	.89	.88872	.0001	.65714	.0006	.0005
J	.59	.59470	.0194	.44020	.0227	.0033
K	.84	.84009	.0001	.65714	.0006	.0005
L	.68	.67832	.0054	.44020	.0227	.0173
M	.85	.85240	.0001	.65701	.0007	.0006
N	.60	.59600	.0190	.45633	.0194	.0004

Note. * Indicates a meaningful difference in p values greater than .118.

Table 18 reveals that there are no substantive differences in p values. However, one set of variables (H), yielded a meaningful difference in p values of .1482. The p value for the Pearson r is .2403 whereas the p value for the Kendall Tau is .0921, thus it is clear that the nonparametric test produces a much lesser p which is considerably closer to a set level of significance. Variable H is bivariate normal, yet has two outliers which may have affected the parpametric test. In certain cases of non-normality such as the uniform and double exponential distributions, Bradley identified the efficiency of the Kendall's Tau as 1.00 and 1.266, respectively. It may be the Kendal Rau was more efficient in this case when outliers were present.

Ten of the twenty-eight variables had univariate outliers. Eleven of the fourteen pairs of variables had bivariate outliers.

Table 19 presents analysis using parametric and nonparametric procedures for learning disabled children (n=15).

Table 19

Original, Replicated and Kendall Tau Association Analysis for Learning Disabled Children on Dichotic Listening Test (n=15)

Condition	Original	Replication	p value	Kendall Tau	p value	Difference
A	.39	.39203	.1484	.05181	.7996	.6512 @*
B	.64	.63757	.0106	.39409	.0457	.0351
C	.85	.84836	.0001	.72370	.0003	.0002
D	.58	.58309	.0225	.19512	.3194	.2969 @*
E	.56	.55663	.0311	.33171	.0905	.0594 *
F	.68	.67772	.0055	.54640	.0053	.0002
G	.89	.89124	.0001	.69906	.0003	.0002
H	.66	.66484	.0068	.38648	.0469	.0401
I	.88	.87546	.0001	.20000	.2987	.2986 @*
J	.58	.58280	.0226	.06678	.7290	.7064 @*
K	.41	.40155	.1379	.20008	.2987	.1608 @
L	.26	.26123	.3470	.06678	.7290	.382 @
M	.66	.507901	.0537	.20389	.2963	.2426 @
N	.41	.412204	.1270	.17392	.3713	.2443 @

Note. @ Meaningful difference in p values.
Note. * Substantive difference in p values.

Table 19 reveals that of the fourteen correlations for learning disabled children, eight had meaningful differences in p values. Additionally, five of the correlations, all but one of which had the formerly cited meaningful difference in p values, had differences regarding substantive conclusions at the .05 level of significance. Thus, considered simultaneously, nine of the correlations had very different results using parametric versus nonparametric procedures. These were A, D, E, I, J, K, L, M, and N. Univariate abnormalities had been noted in A2, D1, D2, E1, I1, I2, J1, J2, K1, K2, L1, M1, M2, and N1. Bivariate normality was lacking in D, E, I, J, K, L, M, and N. Additionally, bivariate normality was lacking in B but no meaningful or substantive differences in p values were noted for B. Thus, for this group of "non-nomral" subjects, differences between parametric and nonparametric tests were found in all but one case where bivariate normality was lacking. Such a pattern of discrepancy between the parametric and nonparametric tests corresponding to non-normality was not seen in the "normal" subjects.

In summary, the correlational analysis comparing parametric and nonparametric procedures yielded one substantive difference for normal children as a subgroup (n=15) when bivariate normality was lacking and the Pearson appears to have underestimated the significance of the correlation. And, nine meaningful or substantive - or both - differences for learning disabled children as a subgroup (n=15) were found, and in each case bivariate normality was lacking and the significance of the correlation was overestimated by the parametric procedure.

In cases of non-normality such as uniform and double exponential distributions Bradley (1978) identified the efficiency of the Kendal Tau to the Pearson r to be 1.00 and 1.266, respectively. Thus, it was concluded that in cases of non-normality the nonparametric procedure was more efficient than the parametric in appropriately revealing probability levels. And, when assumptions were violated the nonparametric test performed comparable to the parametric test.

Regression

Description of the Data Sets.

Regression analysis was conducted using data from an investigation by Fisher and Bundy (1982) in which they studied equilibrium reactions in children as a function of age and group membership (normal girls = NG, normal boys = NB, boys with sensory integrative dysfunction = SID). Specifically, they investigated whether age for each group can be predicted by an equilibrium score. Equilibrium scores consisted of seven measures of joint angles. Fisher and Bundy used stepwise multiple regression for their analyses. Since the current investigation is limited in scope to univariate analysis, their analyses were not replicated exactly. Rather, simple linear regression was computed based upon those single variables with the most significant correlation with age.

Ten normal girls, fourteen normal boys, and nine boys with sensory integrative dysfunction participated in the study. The subjects participated in a one time testing session during which four equilibrium tests yielding seven different angle scores were administered.

Regression is an appropriate technique for prediction and the variables are measured at the interval level. The substantive results of the investigation by Fisher and Bundy were that childrens' ages can be appropriately predicted by an equilibrium measure.

In order to conduct the analysis, certain preliminary calculations were done. Parametric and nonparametric correlations of subjects' ages with the seven angle scores were computed for each of the three groups. These calculations are presented in Appendix M. For each group, the parametric and the nonparametric correlational analysis revealed a single variable to be most significantly correlated with age. For the normal girls it was the angle score represented by TBRA; for the normal boys it was the angle score represented by FBPT; and, for the boys with sensory integrative dysfunction it was the angle score represented by TBRA. Thus, the respective variable for each group was used in the simple linear regression calculations. The data sets for each group, including age and angle scores by subject, are in Appendix N.

Distributional characteristics of these variables was conducted primarily to identify if outliers were present.

Exploratory Data Analysis.

Boxplots on age and a single equilibrium measure for each group (normal boys, normal girls and boys with sensory integrative dysfunction) are presented in the following figures 4, 5, and 6.

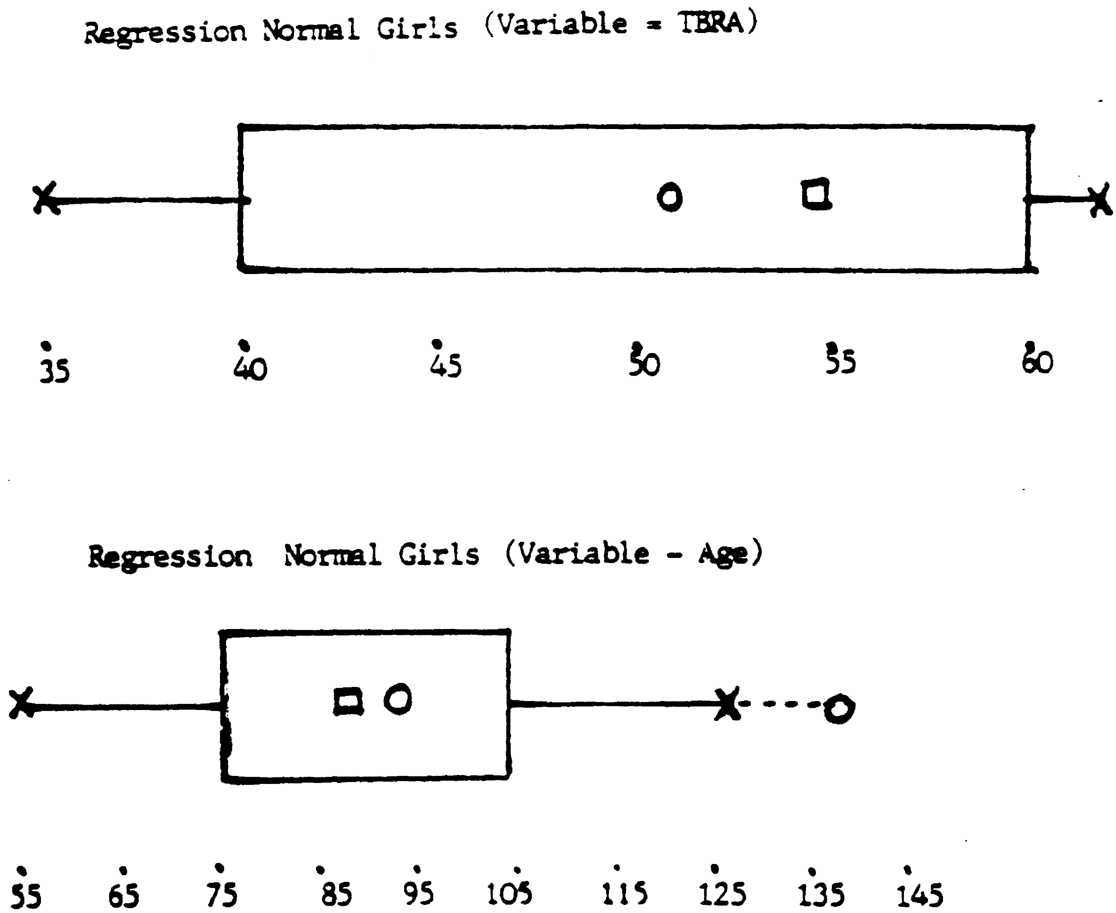
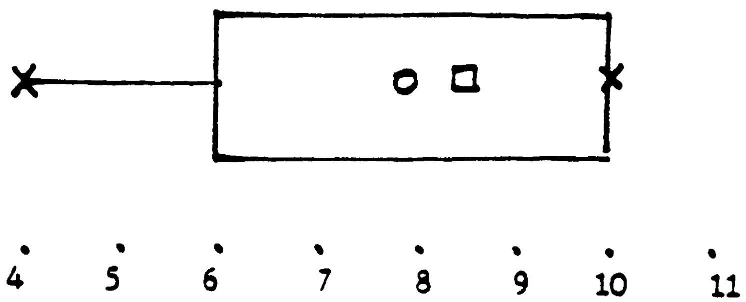


FIGURE 4. Boxplots of TBRA and Age for Normal Girls

Regression Normal Boys (Variable - FBPT)



Regression Normal Boys (Variable = Age)

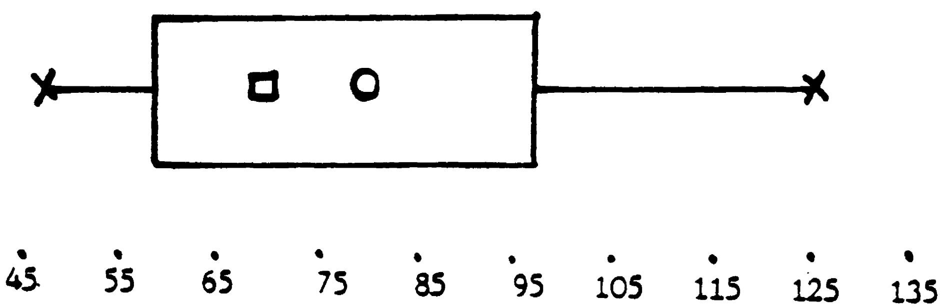
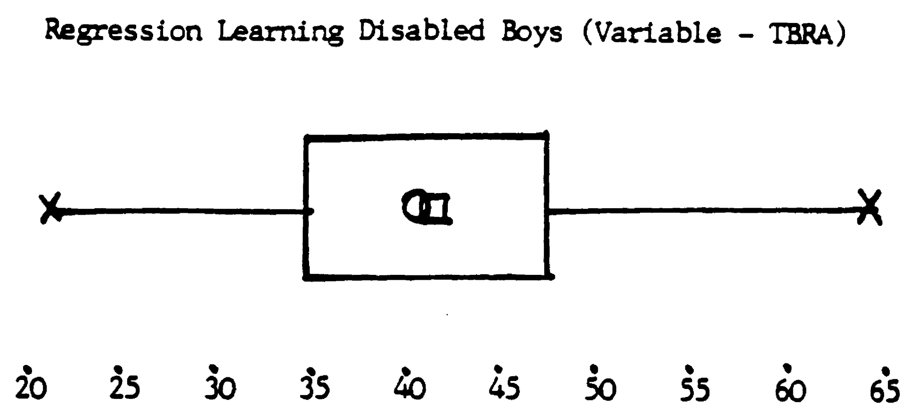


FIGURE 5. Boxplots of FBPT and Age for Normal Boys



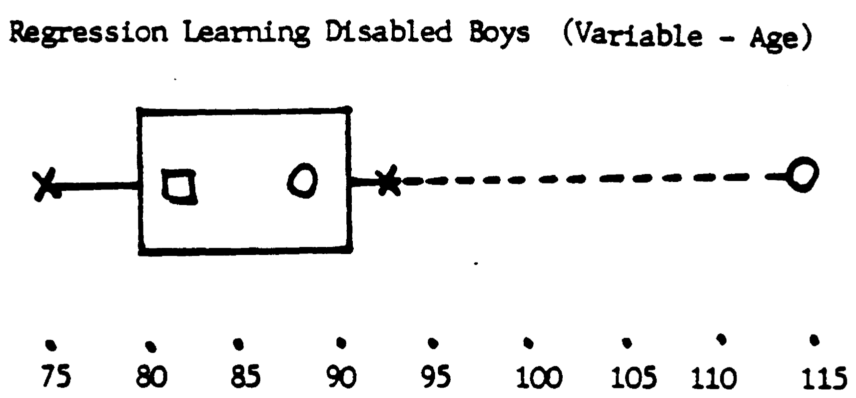


FIGURE 6. Boxplots of TBRA and Age for Boys With Sensory Integrative Dysfunction

Figure 4 reveals a discrepancy between the median and the mean of the variable TBRA and the variable age has a mild outlier.

Figure 5 reveals that for normal boys, the variable FBPT is slightly negatively skewed whereas the variable age does not have an approximately equivalent mean and median.

Figure 6 reveals that for learning disabled boys, the variable TBRA has a symmetrical distribution, but the variable age is discrepant in terms of the mean and the median and has a mild outlier.

Table 20

Distributional Analysis of Equilibrium Scores and Age of Normal Girls, Normal Boys and Boys With Sensory Integrative Dysfunction

Group	Variable	Skewness	Univariate Normality
Normal Girls	TBRA	-.661307	p=.088
	Age	.3511	p=.776
Normal Boys	FBPT	-.431142	p=.058
	Age	.476157	p=.224
Boys with Sensory Integrative Dysfunction	TBRA	.232321	p=.785
	Age	.994059	p=.202

Table 20 reveals that the variable TBRA is slightly skewed and tests marginally normal (p=.088). In spite of a mild outlier, the variable age tests as normal. For normal boys the variable FBPT tests as marginally normal (.058). Finally, the variable age for boys with sensory integrative dysfunction is skewed.

<u>Results of the Data Analysis</u>.

Simple linear regression was computed for each group by regressing age on the specified equilbrium measure. Subsequently, a nonparametric regression for each group was computed. These are presented in Table 21.

Table 21

Results of the Parametric and Nonparametric Simple Linear Regression and Tests of Signifiance

Group	Procedure	Equation	Test of Slope	Test of Slope and Intercept
Normal Girls	Parametric	Y=5.596882+ 1.737830(X)	$t=2.2735$ $p=.0256$	$F=120.81$ $p=.0001$
	Non- parametric	Y=6.6613+ 1.717(X)	$T=.5111$ $T=.4667$ $.036<p<.023$	$C=10.51218$ $C=10.19996$ $.0047<p<.0011$
	Difference		None	None
Normal Boys	Parametric	Y=26.919527+ 7.074556(X)	$t=2.845$ $p=.0147$	$F=130.20$ $p=.0001$
	Non- parametric	Y=33.32+ 6.25(X)	$T=.3186813$ $.1<p<.05$	$C=12.367$ $C=13.166$ $.0048<p<.0009$
	Difference		None	None
SID Boys	Parametric	Y=54.297229+ .833753(X)	$t=3.009$ $p=.0197$	$F=411.40$ $p=.0001$
	Non- parametric	Y=58+ .75095(X)	$T-=.6667$ $T+=.5556$ $09546<p<.005475$	$C=9.39132$ $C=.917388$ $.0096<p<.0055$
	Difference		None	None

Table 21 reveals that the parametric and nonparametric procedures yield no meaningful or substantive differences between p values.

Residual analysis from the parametric and the nonparametric equations was conducted and the results are presented in Table 22.

Table 22

Residual Analysis of Parametric and Nonparametric Regression Equations

Group	Condition	Skewness	Normality	Outliers
NG	Parametric	-.11958	p=.560	none
	Nonparametric	-.102461	p=.538	none
NB	Parametric	-.32737	p=.310	one
	Nonparametric	.925449	p=.334	none
SIB	Parametric	.0691171	p=.325	none
	Nonparametric	.0989474	p=.372	none

Table 22 reveals that the residual analysis is very similar between the parametric and nonparametric regression analyses. For normal boys, the nonparametric residuals are slightly skewed and for the parametric regression residuals there is one mild outlier in the normal boys regression sets. However, evaluation of outlier diagnostics suggested no outliers were present. In addition, all test as distributed normally.

In summary, overall there are no difference between the parametric and nonparametric regression analyses. It should be noted that Hollander and Wolfe (1973) report the efficiency of Thiel's method for slope, compared to the parametric counterpart, to be 1.172 for a double exponential distribution and .919 for a uniform distribution. For this case the relative efficiency appears to be a ratio of one to one for the nonparmatric performed comparable to the parametric test.

Summary

This chapter has presented the results of six case studies divided into categories of statistical analyses using existing data sets. Chapter five will present an integration of the cases, a multiple case study, and a discussion of the findings.

Chapter Five

Discussion

Chapter four presented the results of the individual case studies investigating parametric versus nonparametric statistics with existing data sets from occupational therapy literature. This chapter will present a summary table of the results of the individual cases. Subsequently, the multiple case investigation will be discussed in terms of the research questions posed, issues arising from the investigation, and areas for future research.

Table 23 presents the results of the individual cases in terms of a multiple case study.

Table 23

Multiple Case Study Summary Table of Results of Parametric versus Nonparametric Statistics Using Data Sets from Occupational Therapy Literature

Parametric	Nonparametric	Conclusion
		CASE ONE
Paired t-test	Sign Test	Meaningful difference in p values. The nonparametric p value was much larger and better reflected the degree of similarity between the data sets. Skewness, two mild outliers and marginal normality were found in the data set of difference scores.
		CASE TWO
Independent t-test		
Common Dispersion	Mann Whitney Wilcoxon	No meaningful or substantive differences in p values for either comparison. One data set (strabismic sample) was mildly, positively skewed.
Uncommon Dispersion	Modified Mood's Median Test	
		CASE THREE
ANOVA Psychiatric	F test on Ranks	No meaningful or substantive differences in p values. Data of one group was a non-normal, uniform distribution and homogeneity of variance was marginal.
ANOVA Grade Point Average	F test on Ranks	No substantive differences in p values. A meaningful difference was not present using set criteria but results were close to a meaningful difference. Accommodator group of the data sets was mildly, negatively skewed with one mild and one extreme outlier.

Table 23 (continued)

Parametric	Nonparametric	Conclusion
ANOVA Physical Dysfunction	F test on Ranks	Substantive difference in p values. Reject null using nonparametric test, accept null using parametric test. Accommodator group of data sets was not normally distributed and heterogeneity of variance existed.
		CASE FOUR
Pearson r	Kendall Tau	No substantive difference in p values. Meaningful difference with one set of paired variables (H). H set of variables was not bivariate normal and had two bivariate outliers.
Normal Children (n=15)		The nonparmetric p value is much less than the parametric p value.
Pearson r	Kendall Tau	Nine of the 14 correlations had substantive or meaningful (or both) differences in p values. In each
Learning Disabled (n=15)		instance the Kendall Tau was more conservative, yielding a larger p value. Univariate and bivariate abnormalities had been noted in all single and paired variables having p value differences. All of the nine cases of differences had bivariate outliers and lacked bivariate normality.
		CASE FIVE
Normal Boys		
Parametric Regression	Nonparametric Regression	No meaningful or substantive difference in p values for slope or overall equation. Variable age had one mild outlier. Residual analysis for the normal boys data set yielded one outlier for the parametric analysis and mild skewness for the nonparametric analysis.

Table 23 (continued)

Parametric	Nonparametric	Conclusion
Normal Girls		
Parametric Regression	Nonparametric Regression	No meaningful or substantive difference in p values for slope or overall equation. Neither residual analysis yielded abnormalities.
Boys with Sensory Integrative Dysfunction		
Parametric Regression	Nonparametric Regression	No meaningful or substantive differences in p values for slope or overall equation. The variable age had one mild outlier. The results of the residual analysis yielded no abnormalities.

Table 23 reveals that in certain cases, and in certain instances within certain cases, there were clear differences in results between the parametric and the nonparametric statistics. However, in other cases there were no differences in results between the parametric and nonparametric statistics. These findings will be analyzed by progressing through an analysis based upon each research question and then integrating the analyses into a synthesis or summary.

Research question one focused upon whether or not assumptions underlying the use of parametric statistics were met. Table 24 presents a summary analysis of assumptions underlying each case.

Table 24

Analysis of Assumptions Underlying Parametric Test in Each Case

Case	Procedure	Was Underlying Assumption Met?
1	Paired	No. Data of difference scores tested as marginally normal, had two mild outliers, and was mildly skewed.
2	Independent t-test	Yes, with a minor qualification. Data sets tested as normal but strabismic data set was mildly, positively skewed.
3	One factor ANOVA	No. True interval or ratio level of measurement was not employed. Samples did not appear to be from a normally distributed population. Heterogeneity of variance existed.
4	Pearson r	No. Bivariate normality lacking in nine (normal subgroup), and thirteen (learning disabled subgroup) of the paired variables.
5	Regression	Yes.

In only one case (regression analysis) were assumptions unequivocally and clearly met. In one case (independent t-test), assumptions were met with a minor qualification and in three cases assumptions clearly were not met. Thus, research question one can be answered that in three out of five cases assumptions underlying use of parametric procedures were not met during a multiple case investigation of small sample size data sets.

Thus, the belief that there is no need to test for assumptions when using small data sets can be questioned due to the prevalence of violated assumptions with three of the five cases, purposefully selected due to small sample size.

Research question two addressed whether or not the assumption violated is important in terms of the statistical or methodological literature. This question is addressed using information from chapter two in summary form in Table 25.

Table 25

Importance of Assumptions Violated According to Literature

Case	Procedure	Violated Assumption	Important
1	Paired t-test	Related to normality	No, as long as there are no heavy tails.
2	Independent t-test	None.	NA
3	ANOVA	Normality Homogeneity Level of measurement	Yes, since cell sizes not equal.
4	Pearson r	Bivariate normality	Yes, if sample less than 30.
5	Regression	None	NA

Table 25 reveals that three of the cases (one, three and four) have important or significant violations of assumptions as revealed by the literature cited in chapter two.

The remaining research question three, i.e., how similar or different are findings of parametric versus nonparametric tests, determined (a) whether violations of assumptions in these cases employing small data sets made a difference, and (b) whether the importance of the violated assumptions as predicted from the literature was confirmed empirically.

Table 26 is presented as a summary of interpretations pertaining to research question three.

Table 26

Parametric versus Nonparametric Findings for Individual Cases

Case	Procedure	Findings	Difference Defined
1	Paired t-test	Difference	Parametric p value is .5061. Nonparametric p value is .8146 to 1.0
2	Independent t-test	None	NA
3	ANOVA	Difference in one of three instances	Parametric p value is .0557. Nonparametric p value is .0297
4	Pearson r	One out of 14 is different for normal subgroup	Nonparametric p value is smaller than parametric (.09 vs. .24)
		9 out of 14 are different for learning disabled	Nonparametric p value is more conservative in every case
5	Regression	None	NA

Table 26 reveals that in cases one, three and four, the violated assumptions were indeed important as suggested in Table 25's summary of the literature reviewed. Furthermore, Table 26 reveals that the marginal normality and outliers present in the difference data set of case one apparently did affect the parametric test, and the outliers may have functioned like a 'heavy tail' which compromised the robustness of the paired t-test as had been indicated in the literature.

Table 27 presents a concise summary of these findings regarding differences found related to whether or not parametric assumptions were met.

Table 27

Summary of Assumptions met Related to Differences Found By Case

Case	Procedure	Assumptions Met	Differences Found
1	Paired t-test	No	Yes
2	Independent t-test	Yes	No
3	One Factor ANOVA	No	Yes
4	Pearson r	No	Yes
5	Regression	Yes	No

Review of Table 27 reveals that in each of the two cases where parametric assumptions were met, no differences between parametric versus nonparametric procedures were found. And conversely, in each of the three cases where parametric assumptions were not met, differences between the parametric and nonparametric procedures were found. This finding indicates that if cases are considered as a whole, there is a one hundred percent agreement between whether or not parametric assumptions were violated and whether or not differences were discovered regarding parametric versus nonparametric tests. This is support for the notion that assumptions need to be tested and selection of parametric or nonparametric data analysis procedures conducted accordingly.

Further analysis regarding research question three - how similar or different are findings - is based upon the concept of power. The use of parametric analysis in these five cases did not support the notion that when conditions underlying appropriate use of the parametric procedures are met, parametric tests are more powerful. Siegel (1956) and Fergusen (1976) had found parametric and nonparametric procedures to be within five percent range of comparable efficiency or power. Two cases within this multiple case study were found to be within an eleven percent comparable efficiency or power range: the independent t-test (case two) and regression with the subgroups normal girls, normal boys and boys with sensory integrative dysfunction (case three). For these cases, the nonparametric procedures appeared as powerful as the parametric procedure.

Thus, the claims that nonparametric procedures throw out information (Ager & Jacobson, 1980) and that they lack power (Boneau, 1960) are questioned since nonparametric procedures rendered equivalent findings to those from parametric procedures in these cases. In addition, David and

Perez (1960) propose that the argument for using a nonparametric procedure is strengthened by knowing it will usually yield the same answer as a parametric test. The current multiple case lends support to this concept since, generally speaking, if assumptions weren't violated, the nonparametric test performed comparable to the parametric test.

In certain instances when assumptions were violated, the nonparametric test was more powerful that the parametric test. This is consistant with Blair (1980) who stated that nonparametric tests can be more powerful when assumptions are violated. In the case of the physical dysfunction ANOVA (assumptions violated were normality and homogeneity of variance) the nonparametric p value was much smaller than the parametric value, suggesting the nonparametric test was more powerful to reject the null hypothesis in these circumstances. Use of the nonparametric test in this instance theoretically could have prevented a Type II error.

In another instance, use of the nonparametric test yielded a p value much closer to a level of significance than did the parametric procedure (Pearson r=.24, Kendall Tau=.09) when bivariate normality was lacking and there were two bivariate outliers. Again, the nonparametric test appeared more "powerful" in this case.

However, use of parametric tests when assumptions were violated also appeared to lead to Type I errors, that is innappropriate findings of significance. Thus, in these cases the nonparametric tests appeared to be more appropriate, not less powerful. For example, this happened in the case of the Pearson r with the subgroup of learning disabled (case four) for nine of the fourteen correlations.

Based upon the findings regarding Type I and Type II errors, it may be most appropriate not to define nonparametric procedures as more or less powerful related to parametric procedures, but as possibly more validly employed with greater confidence in the Type I or Type II error rate probability.

Though not addressed by the research questions posited for the study, three issues are in need of discussion based upon the findings. These are (a) interactions of violations of assumptions, (b) distributional characteristics of the variables studied in occupational therapy clinical research, and (c) research design/methods in occupational therapy clinical research. Each of these will be discussed in turn.

Bradley (1978) noted that most simulation studies into robustness and effects of the violation of assumptions on parametric tests investigated the effects of single violations in isolation from each other. It is worth noting that in these authentic data sets all violations occurred in multiples of two or three. That is, it was never a single violation influencing the robustness of a particular test but rather the combination of two or three violated assumptions. These preliminary findings may have implications for continued research into the robustness of procedures when dual or multiple violations of assumptions are introduced into computer generated data sets for simulation studies. Furthermore, the concept of robustness as generalized from simulation research to practice may need re-evaluation in terms of the complexity of interactions between violated assumptions and the parametric tests as occurs in actual practice.

The distributional characteristics of the variables investigated by the studies comprising this multiple case investigation are noteworthy regarding two dimensions; non-normality and outliers. In three of the five cases (paired t-test, one factor ANOVA and Pearson r) instances of univariate or bivariate non-normality were documented. These findings lend support to the notion proposed by Pocock (1982) that the central limit theorem may not apply in all sampling situations and that an automatic assumption of normality based upon the central limit theorem may not be prudent.

Furthermore, in four of the five cases (paired t-test, one factor ANOVA, Pearson r, and regression) there were outliers either in the original data set, the derived data set or calculated residuals. Lezak and Gray (1984) purport that data sets in clinical neuropsychology are often fraught with outliers, and it appears that - at least in these cases - the same is true for data sets in clinical occupational therapy research. Given the prevalence of outliers, the issue of "why" is paramount. Often, outliers are considered to be "errors" in a data set. Comrey (1985) states:

> These "bad" observations can result from many different causes, e.g., errors in test scoring, recording, and/or keypunching; errors in reading data into the computer, such as incorrect formatting, extra or missing cards or lines in the data deck or data set; computer errors; errors in observation or recording; deliberate falsification of responses by subjects for any one of a variety of reasons, and so on. (p. 273)

At this point, it is difficult to assess whether (a) the prevalence of outliers is an artifact of small data sets such as employed in this multiple case investigation, or (b) outliers represent valid manifestations of variables of interest in occupational therapy clinical research. The outlier problem is not unique to occupational therapy. For example, Tupper and Rosenblood (1984) have discussed how characteristics of attribute variables can confound neurological research. Recently, the problem of outliers and non-normality was raised by Mooijaart (1985) regarding factor analytic procedures. He stated that, as regards non-normality and factor analysis:

> Unfortunately there are few psychological theories about how variables are distributed. It seems reasonable to apply models in which weak distributional assumptions are made. (p. 324)

Mooijaart's solution to the problem of non-normality and factor analysis seems reasonable to apply to occupational therapy clinical research. That is, until more research is done on distributional characteristics of variables of concern in occupational therapy clinical research, and unless variables under investigation are documented to be distributed normally, it appears prudent to apply data analysis procedures in which weak distributional assumptions are made, i.e., nonparametric procedures.

The third issue, occupational therapy research design and methods, is brought to light considering the state of the art of the application of design/methods in occupational therapy research over the period 1980 through 1984. For example, indexing of the occupational therapy literature from that time revealed that 66 percent of the data analysis procedures employed could be considered "basic" whereas only 31 percent were "advanced". The fact that the bulk of the data analysis procedures dealt with basic statistics suggests that occupational therapy clinical research is in an emergence period and not yet fully developed in terms of design/methodological/statistical expertise. The field does not yet have a strong and vigorous methodology tradition.

The state of the art of the practice of research/design methods as illustrated by the purposive sample of procedures within this multiple case study further confirms the emergence of clinical research in occupational therapy, but highlights the lack of a rigorous tradition of methodology. To illustrate, two of the five cases within the multiple case study employed data analysis procedures which lack a reflection of advanced practice in research design/methods.

Thus, a confounding variable affecting the study of parametric versus nonparametric statistics using authentic data sets is design of the investigation from which data sets are acquired. For, it is difficult to evaluate similarities or differences between parametric verus nonparametric procedures when the research design is problematic. For example, the correlational investigation into test-retest reliability (case five) may have been more appropriately investigated using a research design and analysis procedures based upon analysis of variance (Cronback, 1947; Jackson & Messick, 1967). Similarly, the three one-factor ANOVA's executed (case three) may have been more appropriately executed as a factorial ANOVA with subsequent multiple comparisons (Box, Hunter & Hunter, 1978).

Consequently, it must be noted that the process of improving the quality and quantity of clinical research in occupational therapy will not be accomplished by a singular approach along one dimension, i.e., parametric versus nonparametric procedures. As an illustration, Baum, Boyle and Edward (1984) state "The most difficult phase of clinical research is the "design . . .phase" (p. 267). And, it is the domain of research design/methods (of which parametric versus nonparametric procedures is just a part) which needs to be addressed. In this vein, Box (1984) discusses the general importance of practice regarding development of statistical procedures. In occupational therapy clinical research, the relationship between practice, research design and statistical procedures is crucial and underlies the future development of a tradition of research methods and practice unique to occupational therapy clinical research. The adequate development of such may be predicated upon fostering and generating research methods within the profession of occupational therapy, which in turn may be predicated upon the level of training (bachelor's or master's level) for entry into the profession (Royeen, in press-a). For, it is impossible to facilitate the growth of a tradition of methodology in occupational therapy without support for graduate level education in methodology related to occupational therapy (master's and doctoral level).

Conclusions and Future Research

The current investigation highlights the importance of exploratory data analysis when analyzing clinical data. In each case investigated herein, meaningful patterns or characteristics of the data emerged which could aid interpretation regardless of class of data analysis procedure (parametric versus nonparametric) employed. Furthermore, information on distributional characteristics of variables as revealed by exploratory data analysis compiled from research investigations over time could serve to define characteristics of variables studied in occupational therapy clinical research. Thus, it is recommended that occupational therapy clinical research include exploratory data analysis as a component for published research.

Research into parametric versus nonparametric procedures is, as stated by McSweeney and Katz (1978), "equivocal". Since the current multiple case investigation can be applied to nonparametric theory directing application and applied only as individual cases resemble the cases presented herein, continued research which could be generalizable is warranted. Such future research could be structured as follows.

First, based upon design and methodology developed for the current investigation, conduct of multiple case investigations into single categories of procedures using data sets randomly selected from the occupational therapy literature from 1980 through 1986 can build upon the current investigation. The most commonly employed procedures in occupational therapy should be investigated, and these would be (in order of frequency in the literature): (a) Pearson r, (b) independent t-test, (c) factorial ANOVA and post hoc analyses, (d) one factor ANOVA, (e) pair t-test, (f) regression, and (g) ANCOVA. Second, future research could also address design/methodological issues and employ alternate designs (such as factorial ANOVA versus multiple one factor ANOVA's and ANOVA based reliability analyses versus reliability analyses based upon correlations).

LITERATURE CITED

Ager, C. L.. & Jacobson, B. R. (1980). Pure versus practical application of research methods: The robustness of the t and F. Letter to the Editor, American Journal of Occupational Therapy, 34(6), 406.

Baum, C.M., Boyle, M.A., & Edwards, D.F. (1984). Initiating occupational therapy clinical research. Amercian Journal of Occupationl Therapy, 38(4), 267-269.

Beniger, J. R., & Robyn, D. L. (1978). Quantitative graphics in statistics: A brief history. American Statistician, 32 (1), 1-11.

Benjamin, Y. (1983). Is the t-test really conservative when the parent distribution is long tailed: Journal of the American Statistical Association, 78(383), 645-654.

Bhattacharjee, G. P. (1968). Non-normality and heterogenity in a two sample test. Annals of the Institute of Statistical Mathematics, 20, 239-254.

Bishop, T. A., & Dudeqicz, E. J. (1978). Exact sample analyses of variance with and without equality of variance: Test probability and tables. Technometrics, 20(4), 419-430.

Blair, R. C. (1981). A reaction to "Consequences of failure to meet assumptions underlying fixed effects analysis of variance and covariance". Review of Educational Research,51(4), 499-507.

Boneau, C. A. (1960). The effects of violations of the assumptions underlying the t test. Psychological Bulletin, 57(1), 49-64.

Boneau, C. A. (1962). A comparison of the power of the U and t tests. Psychological Review, 69, 246-256.

Borg, W. R. (1984). Some important changes in educational research methods over the past twenty years. Paper presentation the the Annual Meeting of the American Educational Research Association, April (ERIC Document Reproduction Service No. 242797).

Box, G. E. P. (1984). The importance of practice in the development of statistics. Technometrics, 26(1), 1-8.

Box, G. E. P., Hunter, W. G., & Hunter, J. S. (1978). Statistics for experimenters. New York: John Wiley and Sons.

Box, G. E. P. (1953). Non-normality and tests of variance. Biometrika, 40, 318-335.

Bradley, J. V. (1978). Robustness? British Journal of Mathematical and Statistical Psychology, 31, 144-152.

Bradley, J. V. (1968). Distribution-free statistical tests. Englewood Cliffs, New Jersey: Prentice Hall.

Bucklaw, L. W. (1983). Nonparametric and psychology: A revitalized alliance. Perceptual and Motor Skills, 57(2), 447-450.

Carmelli, D., & Jorde, L. B. (1982). A nonparametric distance analysis of biochemical genetic data from the Aland islands, Finland. American Journal of Physical Anthropology, 57, 331-340.

Chambers, J. M., Cleveland, W. S., Kleiner, B., & Tukey, J. W. (1983). Graphical methods for data analysis. Boston: Duxbury Press.

Christiansen, C. H. (1983). Research: An economic imperative. Journal of Occupational Therapy Research, 3(4), 195-198.

Clarke, L. E. (1971). Down with the mean. Mathematical Gazette, 55(393), 286-298.

Comrey, A. L. 91985). a method for removing outliers to improve factor analytic results. Multivariate Behavioral Research, 20(3), 273-280.

Conover, W. J. (1980). Practical Nonparametric Statistics. New York: Wiley and Sons.

Conover, W. J., & Iman, R. L. (1981). Rank transformations as a bridge between parametric and nonparametric statistics. The American Statistician, 35(3),124-129.

Cronbach, L. (1947). Test reliability: It's meaning and determination. Psychometrika, 12, 1-16.

Cunninghan, S., & Trickey, B. (1983). The correlation of learning styles and student performance in academic and clinical coursework. Occupational Therapy Journal of Research, 3(1), 54-55.

David, H. A., & Perez, C. A. (1960). On comparing different tests of the same hypotheses. Biometrika, 47(3+4), 297-306.

Edgington, E. S. (1969). _Statistical inference: The distribution free approach_. New York: McGraw Hill.

Eisenhart, C., & Swed, F. S. (1960). On certain criteria for testing the homogeneity of variance. _Annals of Mathematical Statistics_, _11_, 111-112.

Ethridge, D. A., & McSweeney, M. (1971). _Research in occupational therapy_. Dubuque, IA: Kendall/Hunt.

Feir-Walsh, B. J., & Toolhaker, L. E. (1974). An empirical comparison of the ANOVA F-test, normal scores test and Kruskal Wallis Test under violation of assumptions. _Educational and Psychological Measurement_, _34_(4), 789-799.

Ferguson, F. A. (1976). _Statistical analysis in psychology and education_. New York: McGraw Hill.

Fienberg, S. E. (1979). Graphical methods in statistics. _American Statistician_, _33_(4), 165-178.

Filliben, J. J. (1975). The probability plot correlation coefficient test for normality. _Technometrics_, _17_, 1.

Fisher, A. G., & Bundy, A. C. (1982). Equilibrium reactions in normal children and boys with sensory integrative dysfunction. _Occupational Therapy Journal of Research_, _2_(3), 171-183.

Fligner, M. A., Rust, S. W. (1982). a modification of Mood's median test for the generalized Behren_Fisher problem. _Biometrika_, _69_, 221-226.

Furfey, P. H. (1958). Comment on "The needless assumptions on normality in Pearson's r". _American Psychologist_, _13_, 545-546.

Gaito, J. (1959). Nonparametric methods in psychological research. _Psychological Reports_, _5_, 115-125.

Gans, D. J. (1981). Use of a preliminary test in comparing two sample means. _Communication in Statistics - Simulation and Computation_, _B10_(2), 164-174.

Gans, D. J. (1984). The search for significance: different tests on the same data. _Journal of Statistics: Computer Simulation_, _19_,1-21.

Gardner, P. L. (1975). Scales and statistics. _Reveiw of Educational Research_. _45_(1), 43-57.

Geary, R. C. (1947). Testing for normality. _Biometrika_, 34, 209-242.

Gibbons, J. D. (1985). _Statistical inference_. New York: Marcel Decker, Inc.

Gibbons, J. D. (1976). _Nonparametric methods for quantitative analysis_. Columbus, Ohio: American Sciences Press, Inc.

Gillette, N. P. (1982). A data base for occupational therapy: Documentation through research. _American Journal of Occupatonal Therapy_, _36_(8), 499-501.

Glass, G. U., Peckman, P. P., & Sanders, J. R. (1972). Consequences of failure to meet assumptions underlying the fixed effects analysis of variance and covariance. _Review of Educational Research_, _42_(3), 237-287.

Goodwin, L. D., & Goodwin, W. L. (1985). Statistical techniques in American Educational Resarch Journal Articles, 1979-1983: The preparation of graduate students to read the educational research literature. _Educational Research_, February, 5-11.

Greco, J. L. (1979). Nonparametric data analysis. _Journal of College Science Teaching_, _8_(3), 159-161.

Gross, A. M. (1976). Confidence interval robustness with long tailed symmetrical distributions. _Journal of the American Statistical Association_, _71_, 409-416.

Greene, D., & David, D. L. (1984). A research design for generalizing from multiple case studies. _Evaluation and Program Planning_, Vol. 7, 73-85.

Haldane, J. B. S. (1949). A note on non-normal correlations. _Biometrika_, _36_, 467-468.

Hettmansperger, T. P., & McKean, J. W. (1978). Statistical inferences based upon ranks. _Psychometrika_, _43_,(1), 69-79.

Hill, M. A., & Dixon, W. J. (1982). Robustness in real life: A study of clinical laboratory data. _Biometrics_, _38_, 377-386.

Hinkle, D. E., Weirsma, W., & Jurs, S. G. (1979). _Applied statistics for the behavioral sciences_. New York: Houghton Mifflin Co.

Hollander, M., & Wolfe, D. A. (1973). _Nonparametric statistical methods_. New York: John Wiley and Sons.

Hollington, T. L., & Smith, P. J. (1979). Distribution of the normal scores statistic for nonparametric one-way ANOVA. _Journal of the American Statistical Association_,_74_(367), 715-722.

Horn, P. S. (1983). Some easy t statistics. Journal of the American Statistical Association, 78(384), 930-935.

Hussein, S. S., & Sprent, P. (1983). Nonparametric regression. Journal of the Royal Statistical Society, A, Part 2, (146), 182791.

Iman, R. L., & Conover, W. J. (1978). Approximation of the critical region for Spearman's Rho with and without ties present. Communication in Statistics, Series B(7), 269-282.

Irvine, J. M., Ramey, D. B., (1984). T-statistics: A short tale of long tales. Paper presentation at the Joint Statistical Meeting, Philadelphia, PA.

Jackson, D. N., & Messick, S. (1967). Problems in human assessment. New York, McGraw Hill.

Johnson, R. W., & Heyer, K. D. (1980) On the enduring untruth about measurement and parametric statistics. Canadian Psychology, 21(3), 134-135.

Katz, B. M., & McSweeney, M. (1980). A multivariate Kruskal-Wallis test with post hoc procedures. Multivariate Behavioral Research, 15, 281-297.

Kielhofner, F. (1982). Qualitative research: Part one - paradigmatic grounds and issues of reliability and validity. Occupational Therapy Journal of Research, 2(1), 67-79.

Kirk, R. E. (1968). Experimental design: Procedures for the behavioral sciences. Belmont, CA: Wadsworth Publishing Co.

Kirkpatrick, J. S. (1981). Nonparametric statistics: Useful tools for counselors. Personnel and Guidance Journal,59(10), 619-651.

Koomar, J., & Cermack, S. A. (1981). Reliability of dichotic listening using two stimulus formats with normal and learning disabled children. American Journal of Occupational Therapy, 37(7), 456-463.

Kratochwill, T. R. (1978). Single subject research. New York: Academic Press.

Krauth, J. (1980). Nonparametric analysis of response curves. Journal of Neuroscience Methods, 2(2), 239-252.

Leedy, P. D. (1974). Practical research planning and design. New York: MacMillan and Co.

Lehman, E. L. (1975). Nonparametrics. San Francisco: Holden-Day and McGraw Hill.

Lezak, M. D., Gray, D. K. (1984). Sampling problems and nonparametric solutions in clinical neuropsychological research. Journal of Clinical Neuropsychology, 6(1), 101-109.

Looney, S. W., & Gullidge, T. R. (1985). Use of corrrelation coefficients with normal probability plots. American Statistician, 39(1), 75-79.

Marascuilo, L. A., & Daginais, F. (1982). Planned and post hoc comparisons for tests of homogeneity where the dependent variable is categorical and ordered. Educational and Psychological Measurement, 42(3), 777-781.

Marascuilo, L. A., & McSweeney, M. (1977). Nonparametric and distribution free methods for the social sciences. Montery, CA: Brooks-Cole Publishing Co.

McGill, R., Tukey, J. W., & Larsen, W. A. (1978). Variations of box plots. American Statistics Teacher, 32(1), 12-22.

McSweeny, M., & Katz, B. M. (1978). Nonparametric statistics: Use and nonuse. Perceptual and Motor Skills, 4(3), 1023-1032.

Minium, E. W. (1978). Statistical reasoning in psychology and education. New York: John Wiley and Sons.

Montgomery, D. C., & Peck, F. A. (1982). Introduction to linear regression analysis. New York: John Wiley

Mooijaart, A. (1985). Factor analysis for non-normal variables. Psychometrika, 50(3), 323-342.

Mosteller, F., & Rourke, R. E. F. (1973). Sturdy statistics: Nonparametrics and order statistics. Reading, MA: Addison-Wesley.

Nelson, D. L., Weidensaul, N. K., Anderson, V. G., Shing_Ru Shih (1984). The Southern California Postrotary Nystagmus Test land electronystagmography under differing conditions of visual input. American Journal of Occupational Therapy, 38(8), 535-540.

Neter, J., & Wasserman, W. (1974). Applied linear statistical models. Homewood, Ill.: Richard D. Irwin, Inc.

Noether, N. E. (1980). The role of nonparametric in introductory statistics courses. American

Statistician, 34(1),22-23.

Ottenbacher, K. (1982a). Publication trends in occupational therapy. Journal of Occupational Therapy Research, 2(2), 80-88.

Ottenbacher, K. (1982b). Statistical power and research in occupational therapy. Journal of Occupational Therapy Research, 2(1), 13-25.

Ottenbacher, K. (1983). A "Tempest" over t-tests. American Journal of Occupational Therapy, 37, (10), 700-701.

Ottenbacher, K. & Peterson, P. (1985). Quantitative trends in occupational therapy research: Implications for practice and education. American Journal of Occupational Therapy, 39(4),240-246.

Ottenbacher, K., Dauck, B. S., Givelinger, M., Grahm, V., & Hasset, C. (1985). Reliability of the behavioral assessment scale of oral functions in feeding. American Journal of Occupational Therapy, 39 (7), 436-441.

Pearson, E. S., & Please, N. W. (1975). Relation between the shape of population distribution and the robustness of four simple test statistics. Biometrika, 62(2), 223-241.

Penfield, D. A. (1971). AERA Conference Summaries III. Educational Statistics. ERIC Clearinghouse on Tests, Measurement and Evaluation. Princeton, NJ: Mead.

Plane, D. R., & Gordon, K. R. (1982). A simple proof of the nonapplicability of the central limit theorem to finite populations. American Statistician, 36(3), 175-176.

Pocock, S. J. (1982). When not to rely on the central limit theorem - an example from absenteeism data.

Ramsey, P. H. (1980). Exact type I error rates for robustness of student's t test with unequal variances. Journal of Educational Statistics, 5(4), 337-349.

Rogers, J. C. (1981). Guest editorial: Educating the inquisitive practitioner. Journal of Occupational Therapy Research, 2(1), 3-11.

Royeen, C. B. (in press-a). Entry level education in occupational therapy. American Journal of Occupational Therapy

Royeen, C. B. (in press-b). The boxplot: A screening test for research data. American Journal of Occupational Therapy

Royeen, C. B., & Seaver, W. F. (1986). Promise in nonparametrics. American Journal of Occupational Therapy, 40(3), 191-193.

Rowney, J. A., & Zenisik, L. (1980). Manuscript characteristics influencing reviewers decisions. Canadian Psychology, 2(1), 17-21.

Savage, R. (1957). Nonparametric statistics. Journal of the American Statistical Association, 52, 331-344.

Schmidt, H. (1981). Qualitative research and occupational therapy. American Journal of Occupational Therapy, 35(2), 105-106.

Seaver, W. J. (1977). The right test but the wrong occasions. National Association Business Teachers Educational Review, p. 15-17.

Sen, P. K. (1968). estimates of the regression coefficient based upon the Kendall Tau. Journal of the American Statistical Association, 63,1379-1389.

Shapiro, S. S., & Wilk, M. B., (1965). An analysis of variance test. Biometrics,52(3,4), 591-611.

Slavik, B. (1982). Vestibular functiions in children with nonparalytic strabismus. Occupational Therapy Journal of Research, 2(4), 220-233.

Siegel, S. (1956). Nonparametric statistics for the behavior sciences. New York: McGraw Hill.

Singer, B. (1979). Distribution-free methods for nonparametric problems: a classified and selected bibliography. British Journal of Mathematical and Statistical Psychology, 32, 1-60.

Smith, J. E. K., (1976). Analysis of qualitative data. Review of Psychology, 27, 487-499.

Sprott, D. A. (1978). Robustness and nonparametric procedures are not the only or the safe alternative to normality. Canadian Journal of Psychology, 3, 180-185.

Statistical Analysis System User's Guide. (1982). Cary: North Carolina.

Statistical Package for the Social Sciences. (1983). New York: McGraw Hill.

Tikie, M. L., & Ssingh, M. (1981). Robust test for means when popuation variances are unequal. Communication in Statistics-Theoretical Methods. A10(20), 2057-2071.

Tukey, J. W. (1977). Exploratory data analysis. Reading, MA: Addison Wesley Co.

Tupper, D. E., & Rosenblood, L. K. (1984). Methodological commentary: methodological considerations in the use of attribute variables in medical research. Journal of Clinical Neurology, 6(4), 441-453.

Velleman, P., & Hoaglin, D. C. (1981). Application, basics and computing of exploratory data analysis. Boston: Duxbury Press.

Walberg, H. J., Strykowski, B. F., Rovai, E., & Hung, S. S. (1984). Exceptional performance, Review of Educational Research, 54(1), 87-112.

Welkowitz, J., Ewen, R. B., and Cohen, J. (1971). Introductory statistics for the behavioral sciences. New York: Academic Press.

West, W. L. (1981). The need, the response: Journal of Occupational Therapy Research. American Journal of Occupational Therapy, 35(1), 44.

Williamson, G. G. (1982). A heritage of activity: Development of theory. American Journal of Occupational Therapy, 36(11), 716-722.

Yerxa, E. J. (1984). Evaluation versus research: Outcomes or knowledge? American Journal of Occupational Therapy, 38(6), 407-408.

Yin, R. K. (1985). Case Study Methodology. Presentation to the Institute for Special Education Studies, Sponsored by the Division of Educational Services, U. S. Department of Education, November 6, Washington, D. C.

Yin, R. K. (1984). Case study research: Design and methods. Beverly Hills: Sage Publications

Appendix A

Frequency of Statistical Procedures in the

American Journal of Occupational Therapy 1980-1984

Table A-1

Frequencies By Years (1980-1984), _Totals and Percentage of Statistical Articles in the American Journal of Occupational Therapy_

	Frequency by Year						
Category	1980	1981	1982	1983	1984	Total	Percentage
Statistical Articles	25	22	21	24	22	114	43.18
Non Statis-tical Articles	35	30	36	28	21	150	56.82
Totals	60	52	57	52	43	264	100.00

Table A-2

Frequencies By Years (1980-1984), Totals and Percentages of Statistical Procedures in the American Journal of Occupational Therapy

	Frequency of Procedure by Year						
	1980	1981	1982	1983	1984	Total	Percentage
Parametric Procedures	41	34	23	32	42	172	77.13
Non-Parametric Procedures	4	9	8	8	10	39	17.49
Non Specified Tests	0	3	4	0	5	12	5.53
Totals	45	46	35	40	57	223	100.00

Note. Non specified tests refer to correlations or tests of difference which weren't specified by the author in the respective article.

Note. Articles were indexed according to types of statistical procedures employed. Thus, one article could have more than one statistical procedure and accounts for more statistical procedures than statistical articles.

Table A-3

Frequency of Parametric Statistical Procedures in the American Journal of Occupational Therapy by Year 1980-1984

	Frequency of Procedure by Year				
Procedure	1980	1981	1982	1983	1984
Pearson Correlation	8	7	4	8	5
Chi Square	6	5	2	3	6
Student t	0	1	1	0	0
Paired t	2	2	3	2	3
Independent t	10	7	1	7	2
One Way ANOVA	5	1	0	5	7
Factorial ANOVA	4	3	5	2	8
Planned Contrasts	0	0	0	0	2
Post Hoc Multiple Comparisons	2	2	3	1	5
ANCOVA	0	1	3	0	1
Regression	3	0	0	3	2
Discriminant Analysis	1	1	0	0	0
Canonical Correlation	0	0	0	1	0
Factor Analysis	0	2	0	0	1
MANOVA	0	1	0	0	0
Meta Analysis	0	0	1	0	0
Other	0	1	0	0	0
Totals	41	34	23	32	42

Table A-4

Frequencies Summed Across Years (1980-1984) and Percentages of Total Parametric Statistical Procedures in the American Journal of Occupational Therapy

Procedures	Total Frequency	Percentage
Pearson Correlation	32	18.61
Chi Square	22	12.79
Student t	2	1.16
Paired t	12	6.98
Independent t	27	15.70
One Way ANOVA	18	10.47
Factorial ANOVA	22	12.79
Planned Contrasts	2	1.16
Post-hoc Multiple Comparisons	13	7.56
ANCOVA	5	2.91
Regression	8	4.65
Discriminant	2	1.16
Canonical Correlation	1	.58
Factor Analysis	3	1.74
MANOVA	1	.58
Meta Analysis	1	.58
Other	1	.58
Totals	172	100.00

Table A-5

<u>Specification and Frequency of Nonparametric Statistical Procedures by Year in the American Journal of Occupational Therapy</u> 1980-1984

	Frequency of Procedures				
Procedure	1980	1981	1982	1983	1984
Sign Test	0	0	1	0	0
Spearman Rho	0	3	1	2	3
Kendall's Tau	0	1	2	0	2
Point Biserial	0	0	0	0	0
Cramer's V	0	2	0	0	0
Fisher's Exact Test	2	0	0	0	0
Mann Whitney U-Test	1	1	2	3	2
Wilcoxon Matched Pair Signed Rank Test	0	1	0	0	2
Kruskall-Wallis ANOVA	1	1	1	0	0
Randomized Permutation Test (ANOVA)	0	0	0	1	0
Split Middle Test of Trend	0	0	1	2	0
Totals	4	9	8	8	10

Table A-6

Frequencies Summed Across Years (1980-1984) *and Percentage of Total of Nonparametric Procedures in the American Journal of Occupational Therapy*

Procedures	Total Frequency	Percentage
Sign Test	1	2.56
Spearman's Rho	9	23.08
Kendall Tau	5	12.82
Point Biserial	0	0
Cramer's V	2	5.13
Fisher's Exact Test	2	5.13
Mann Whitney U-test	9	23.08
Wilcoxon Matched Pair Signed-Rank Test	3	7.69
Kruskall-Wallis ANOVA	4	10.26
Randomized Permutation Test (ANOVA)	1	2.56
Split Middle Test of Trend	3	7.69
Totals	39	100.00

Appendix B

Frequency of Statistical Procedures in the Occupational Therapy in Mental Health 1980-1984

Table B-1

Frequencies By Years (1980-1984), Totals and Percentage of Statistical Articles in Occupational Therapy in Mental Health

	Frequency by Year						
Category	1980	1981	1982	1983	1984	Total	Percentage
Statistical Articles	2	-	2	2	5	11	14.29
Non Statistical Articles	20	-	12	21	13	66	85.71
Totals	22	-	14	23	18	77	100.00

Note. Volume 1, Issue 4 was the only volume published in 1981 and was designated with a 1980/1981 publication date. Thus, it was included in 1980, which leaves no issues in 1981.

Table B-2

Frequencies By Years (1980-1984), Totals and Percentages of Statistical Procedures in Occupational Therapy in Mental Health

	Frequency of Procedure by Year						
	1980	1981	1982	1983	1984	Total	Percentage
Parametric Procedures	2	–	2	4	8	16	88.89
Non-Parametric Procedures	0	–	2	0	0	2	11.11
Non Specified Tests	0	–	0	0	0	0	0
Totals	2	–	4	4	8	18	100.00

Note. Non specified tests refer to correlations or tests of difference which weren't specified by the author in the respective article.

Note. Articles were indexed according to types of statistical procedure employed. Thus, one article could have more than one statistical procedure and accounts for more statistical procedures than statistical articles.

Table B-3

Frequency of Parametric Statistical Procedures in Occupational Therapy in Mental Health 1980-1984

	Frequency of Procedure by Year				
Procedure	1980	1981	1982	1983	1984
Pearson Correlation	0	-	0	0	3
Chi Square	0	-	0	0	1
Student t	0	-	0	0	0
Paired t	0	-	0	0	1
Independent t	0	-	1	0	0
One Way ANOVA	1	-	1	1	0
Factorial ANOVA	0	-	0	1	1
Planned Contrasts	0	-	0	0	0
Post Hoc Multiple Comparisons	0	-	0	1	1
ANCOVA	1	-	0	0	0
Regression	0	-	0	0	0
Discriminant Analysis	0	-	0	1	0
Canonical Correlation	0	-	0	0	0
Factor Analysis	0	-	0	0	0
MANOVA	0	-	0	0	0
Meta Analysis	0	-	0	0	0
Other	0	-	0	0	1
Totals	2	-	2	4	8

Table B-4

Frequencies Summed Across Years (1980-1984) and Percentages of Total Parametric Statistical Procedures Occupational Therapy in Mental Health

Procedures	Total Frequency	Percentage
Pearson Correlation	3	18.75
Chi Square	1	6.25
Student t	0	0
Paired t	1	6.25
Independent t	1	6.25
One Way ANOVA	3	18.75
Factorial ANOVA	2	12.5
Planned Contrasts	0	0
Post-hoc Multiple Comparisons	2	12.5
ANCOVA	1	6.25
Regression	0	0
Discriminant	1	6.25
Canonical Correlation	0	0
Factor Analysis	0	0
MANOVA	0	0
Meta Analysis	0	0
Other	1	6.25
Totals	16	100.00

Table B-5

Specification and Frequency of Nonparametric Statistical Procedures in Occupational Therapy in Mental Health by Year 1980-1984

	Frequency of Procedures				
Procedure	1980	1981	1982	1983	1984
Sign Test	0	-	0	0	0
Spearman Rho	0	-	1	0	0
Kendall's Tau	0	-	0	0	0
Point Biserial	0	-	0	0	0
Cramer's V	0	-	0	0	0
Fisher's Exact Test	0	-	0	0	0
Mann Whitney U-Test	0	-	0	0	0
Wilcoxon Matched Pair Signed Rank Test	0	-	0	0	0
Kruskall-Wallis ANOVA	0	-	1	0	0
Randomized Permutation Test (ANOVA)	0	-	0	0	0
Split Middle Test of Trend	0	-	0	0	0
Totals	0	-	2	0	0

Table B-6

Frequencies Summed Across Years (1980-1984) and Percentage of Total of Nonparametric Procedures in the Occupational Therapy in Mental Health

Procedures	Total Frequency	Percentage
Sign Test	0	0
Spearman's Rho	1	50
Kendall Tau	0	0
Point Biserial	0	0
Cramer's V	0	0
Fisher's Exact Test	0	0
Mann Whitney U-test	0	0
Wilcoxon Matched Pair Signed-Rank Test	0	0
Kruskall-Wallis ANOVA	1	50
Randomized Permutation Test (ANOVA)	0	0
Split Middle Test of Trend	0	0
Totals	2	100

Appendix C

Frequency of Statistical Procedures in the

Occupational Therapy Journal of Research 1980-1984

Table C-1

<u>Frequencies By Years</u> (1980-1984), <u>Totals and Percentage of Statistical Articles in the Occupational Therapy Journal of Research</u>

	Frequency by Year						
Category	1980	1981	1982	1983	1984	Total	Percentage
Statistical Articles	-	8	13	15	14	50	76.92
Non Statistical Articles	-	1	4	4	6	15	23.08
Totals	-	9	17	19	20	65	100.00

<u>Note</u>. <u>Occupational Therapy Journal of Research</u> began publication in 1981.

Table C-2

Frequencies By Years (1980-1984), Totals and Percentages of Statistical Procedures in the Occupational Therapy Journal of Research

	Frequency of Procedure by Year						
	1980	1981	1982	1983	1984	Total	Percentage
Parametric Procedures	-	24	17	23	32	96	90.56
Non-Parametric Procedures	-	1	3	5	1	10	9.43
Non Specified Tests	-	1	0	0	1	2	.94
Totals	-	25	20	28	34	108	100.00

Note. Non specified tests refer to correlations or tests of difference which weren't specified by the author in the respective article.

Note. Articles were indexed according to types of statistical procedure employed. Thus, one article could have more than one statistical procedure and accounts for more statistical procedures than statistical articles.

Table C-3

Frequency of Parametric Statistical Procedures in the Occupational Therapy Journal of Research 1981-1984

	Frequency of Procedure by Year				
Procedure	1980	1981	1982	1983	1984
Pearson Correlation	-	5	7	6	6
Chi Square	-	1	2	2	5
Student t	-	0	0	1	1
Paired t	-	1	0	2	4
Independent t	-	3	0	2	2
One Way ANOVA	-	4	2	2	2
Factorial ANOVA	-	1	1	2	5
Planned Contrasts	-	0	0	0	1
Post Hoc Multiple Comparisons	-	1	1	2	5
ANCOVA	-	2	0	0	0
Regression	-	1	2	2	0
Discriminant Analysis	-	0	1	0	0
Cannonical Correlation	-	0	0	0	0
Factor Analysis	-	1	0	0	0
MANOVA	-	0	0	0	0
Meta Analysis	-	0	1	0	0
Other	-	4	0	2	1
Totals	-	24	17	23	32

Table C-4

Frequencies Summed Across Years (1981-1984) and Percentages of Total Parametric Statistical Procedures the Occupational Therapy Journal of Research

Procedures	Total Frequency	Percentage
Pearson Correlation	24	25.00
Chi Square	10	10.42
Student t	2	2.08
Paired t	7	7.29
Independent t	7	7.29
One Way ANOVA	10	10.42
Factorial ANOVA	9	9.38
Planned Contrasts	1	1.04
Post-hoc Multiple Comparisons	9	9.38
ANCOVA	2	2.08
Regression	5	5.21
Discriminant	1	1.04
Canonical Correlation	0	0
Factor Analysis	1	1.04
MANOVA	0	0
Meta Analysis	1	1.04
Other	7	7.29
Totals	96	100.00

Table C-5

Specification and Frequency of Nonparametric Statistical Procedures in the Occupational Therapy Journal of Research by Year 1981-1984

	Frequency of Procedures				
Procedure	1980	1981	1982	1983	1984
Sign Test	-	0	0	0	0
Spearman Rho	-	0	0	1	0
Kendall's Tau	-	0	0	1	0
Point Biserial	-	0	0	0	0
Cramer's V	-	0	0	0	1
Fisher's Exact Test	-	0	0	3	0
Mann Whitney U-Test	-	0	2	0	0
Wilcoxon Matched Pair Signed Rank Test	-	0	0	0	0
Kruskall-Wallis ANOVA	-	0	1	0	0
Randomized Permutation Test (ANOVA)	-	0	0	0	0
Split Middle Test of Trend	-	1	0	0	0
Totals	-	1	3	5	1

Table C-6

Frequencies Summed Across Years (1981-1984) and Percentage of Total of Nonparametric Statistical Procedures in the Occupational Therapy Journal of Research

Procedures	Total Frequency	Percentage
Sign Test	0	0
Spearman's Rho	1	10
Kendall Tau	1	10
Point Biserial	0	0
Cramer's V	1	10
Fisher's Exact Test	3	30
Mann Whitney U-test	2	20
Wilcoxon Matched Pair Signed-Rank Test	0	0
Kruskall-Wallis ANOVA	1	10
Randomized Permutation Test (ANOVA)	0	0
Split Middle Test of Trend	1	10
Totals	10	100

Appendix D

<u>Frequency of Statistical Procedures in</u>

<u>Physical and Occupational Therapy in Geriatrics</u> 1980-1984

Table D-1

Frequencies By Years (1980-1984), Totals and Percentage of Statistical Articles in Physical and Occupational Therapy in Geriatrics

	Frequency by Year						
Category	1980	1981	1982	1983	1984	Total	Percentage
Statistical Articles	1	-	2	3	0	6	13.33
Non Statistical Articles	5	-	14	13	7	39	86.67
Totals	6	-	16	16	7	45	100.00

Note. Volume 1, Issue 4 was the only volume published in 1981 and was designated with a 1980/1981 publication date. Thus, it was included in 1980, which leaves no issues in 1981.

Table D-2

Frequencies By Years (1980-1984), *Totals and Percentages of Statistical Procedures in Physical and Occupational Therapy in Geriatrics*

	Frequency of Procedure by Year						
	1980	1981	1982	1983	1984	Total	Percentage
Parametric Procedures	1	-	4	7	1	13	92.86
Non-Parametric Procedures	0	-	0	1	0	1	7.14
Non Specified Tests	0	-	0	1	0	1	6.67
Totals	1	-	4	9	1	15	100.00

Note. Non specified tests refer to correlations or tests of difference which weren't specified by the author in the respective article.

Note. Articles were indexed according to types of statistical procedure employed. Thus, one article could have more than one statistical procedure and accounts for more statistical procedures than statistical articles.

Table D-3

Frequency of Parametric Statistical Procedures in Physical and Occupational Therapy in Geriatrics 1980-1984

Procedure	Frequency of Procedure by Year				
	1980	1981	1982	1983	1984
Pearson Correlation	0	-	0	2	0
Chi Square	1	-	2	1	0
Student t	0	-	0	0	0
Paired t	0	-	1	1	0
Independent t	0	-	0	1	0
One Way ANOVA	0	-	0	2	0
Factorial ANOVA	0	-	1	0	0
Planned Contrasts	0	-	0	0	0
Post-hoc Multiple Comparisons	0	-	0	0	0
ANCOVA	0	-	0	0	0
Regression	0	-	0	0	0
Discriminant Analysis	0	-	0	0	0
Canonical Correlation	0	-	0	0	0
Factor Analysis	0	-	0	0	0
MANOVA	0	-	0	0	0
Meta Analysis	0	-	0	0	0
Other	0	-	0	0	1
Totals	1	-	4	7	1

Table D-4

Frequencies Summed Across Years (1980-1984) and Percentages of Total Parametric Statistical Procedures in Physical and Occupational Therapy in Geriatrics

Procedures	Total Frequency	Percentage
Pearson Correlation	2	15.37
Chi Square	4	30.77
Student t	0	0
Paired t	2	15.39
Independent t	1	7.7
One Way ANOVA	2	15.39
Factorial ANOVA	1	7.7
Planned Contrasts	0	0
Post-hoc Multiple Comparisons	0	0
ANCOVA	0	0
Regression	0	0
Discriminant	0	0
Canonical Correlation	0	0
Factor Analysis	0	0
MANOVA	0	0
Meta Analysis	0	0
Other	1	7.7
Totals	13	100

Table D-5

Specification and Frequency of Nonparametric Statistical Procedures by Year in Physical and Occupational Therapy in Geriatrics 1980-1984

Procedure	Frequency of Procedures 1980	1981	1982	1983	1984
Sign Test	0	–	0	0	0
Spearman Rho	0	–	0	0	0
Kendall's Tau	0	–	0	0	0
Point Biserial	0	–	0	0	0
Cramer's V	0	–	0	0	0
Fisher's Exact Test	0	–	0	1	0
Mann Whitney U-Test	0	–	0	0	0
Wilcoxon Matched Pair Signed Rank Test	0	–	0	0	0
Kruskall-Wallis ANOVA	0	–	0	0	0
Randomized Permutation Test (ANOVA)	0	–	0	0	0
Split Middle Test of Trend	0	–	0	0	0
Totals	0	–	0	1	0

Table D-6

Frequencies Summed Across Years (1980-1984) and Percentage of Total of Nonparametric Procedures in Physical and Occupational Therapy in Geriatrics

Procedures	Total Frequency	Percentage
Sign Test	0	0
Spearman's Rho	0	0
Kendall Tau	0	0
Point Biserial	0	0
Cramer's V	0	0
Fisher's Exact Test	1	100
Mann Whitney U-test	0	0
Wilcoxon Matched Pair Signed-Rank Test	0	0
Kruskall-Wallis ANOVA	0	0
Randomized Permutation Test (ANOVA)	0	0
Split Middle Test of Trend	0	0
Totals	1	100

Appendix E

Frequency of Statistical Procedures in Physical and Occupational Therapy in Pediatrics 1980-1984

Table E-1

<u>Frequencies By Years</u> (1980-1984), <u>Totals and Percentage of Statistical Articles in Physical and Occupational Therapy in Pediatrics</u>

	Frequency by Year						
Category	1980	1981	1982	1983	1984	Total	Percentage
Statistical Articles	5	5	9	8	10	37	55.22
Non Statistical Articles	4	5	8	6	7	30	44.78
Totals	9	10	17	14	17	67	100.00

Table E-2

Frequencies By Years (1980-1984), Totals and Percentages of Statistical Procedures in Physical and Occupational Therapy in Pediatrics

	Frequency of Procedure by Year						
	1980	1981	1982	1983	1984	Total	Percentage
Parametric Procedures	5	5	13	12	14	49	81.67
Non-Parametric Procedures	0	3	2	3	2	10	16.67
Non Specified Tests	1	0	0	0	0	1	1.67
Totals	6	8	15	15	16	60	100.00

Note. Non specified tests refer to correlations or tests of difference which weren't specified by the author in the respective article.

Note. Articles were indexed according to types of statistical procedure employed. Thus, one article could have more than one statistical procedure and accounts for more statistical procedures than statistical articles.

Table E-3

Frequency of Parametric Statistical Procedures in Physical and Occupational Therapy in Pediatrics 1980-1984

	Frequency of Procedure by Year				
Procedure	1980	1981	1982	1983	1984
Pearson Correlation	0	1	3	6	3
Chi Square	1	0	2	0	1
Student t	0	0	0	0	0
Paired t	0	1	0	0	1
Independent t	2	1	0	3	4
One Way ANOVA	0	0	1	0	1
Factorial ANOVA	2	1	2	2	2
Planned Contrasts	0	0	0	0	0
Post-hoc Multiple Comparisons	0	0	3	0	1
ANCOVA	0	1	1	0	0
Regression	0	0	0	1	0
Discriminant Analysis	0	0	1	0	0
Canonical Correlation	0	0	0	0	0
Factor Analysis	0	0	0	0	0
MANOVA	0	0	0	0	0
Meta Analysis	0	0	0	0	0
Other	0	0	0	0	1
Totals	5	5	13	12	14

Table E-4

Frequencies Summed Across Years (1980-1984) and Percentages of Total Parametric Statistical Procedures Physical and Occupational Therapy in Pediatrics

Procedures	Total Frequency	Percentage
Pearson Correlation	13	26.53
Chi Square	4	8.16
Student t	0	0
Paired t	2	4.01
Independent t	10	20.41
One Way ANOVA	2	4.08
Factorial ANOVA	9	18.37
Planned Contrasts	0	0
Post-hoc Multiple Comparisons	4	8.16
ANCOVA	2	4.08
Regression	1	2.04
Discriminant	1	2.04
Canonical Correlation	0	0
Factor Analysis	0	0
MANOVA	0	0
Meta Analysis	0	0
Other	1	2.04
Totals	49	100.00

Table E-5

Specification and Frequency of Nonparametric Statistical Procedures by Year in Physical and Occupational Therapy in Pediatrics 1980-1984

	Frequency of Procedures				
Procedure	1980	1981	1982	1983	1984
Sign Test	0	0	0	0	0
Spearman Rho	0	0	0	0	0
Kendall's Tau	0	0	1	0	1
Point Biserial	0	1	0	0	0
Cramer's V	0	0	0	0	0
Fisher's Exact Test	0	0	0	1	0
Mann Whitney U-Test	0	1	0	2	0
Wilcoxon Matched Pair Signed Rank Test	0	1	0	0	1
Kruskall-Wallis ANOVA	0	0	1	0	0
Randomized Permutation Test (ANOVA)	0	0	0	0	0
Split Middle Test of Trend	0	0	0	0	0
Totals	0	3	2	3	2

Table E-6

Frequencies Summed Across Years (1980-1984) and Percentage of Total of Nonparametric Procedures in Physical and Occupational Therapy in Pediatrics

Procedures	Total Frequency	Percentage
Sign Test	0	0
Spearman's Rho	0	0
Kendall Tau	2	20.00
Point Biserial	1	10.00
Cramer's V	0	0
Fisher's Exact Test	0	0
Mann Whitney U-test	3	30.00
Wilcoxon Matched Pair Signed-Rank Test	2	20.00
Kruskall-Wallis ANOVA	1	10.00
Randomized Permutation Test (ANOVA)	0	0
Split Middle Test of Trend	0	0
Totals	10	100.00

Appendix F

<u>Total Frequency and Rank Ordering of Statistical</u>

<u>Procedures in the Occupational Therapy</u>

<u>Literature</u> 1980-1984

Table F-1

Frequency of Statistical Versus Nonstatistical Articles by Journal for All Years Combined 1980-1981

	Frequency by Year				
Frequency	AJOT	OTMH	OTJR	PTOTG	PTOTP
Statistical Articles	114	11	50	6	37
Non-statistical Articles	150	66	15	39	30
Totals	264	77	65	45	67

Table F-2

Frequency of Parametric, Nonparametric and Nonspecified Statistical Procedures by Journal for All Years Combined

1980-1984

Frequency	AJOT	OTMH	OTJR	PTOTG	PTOTP
Parametric Procedures	172	16	96	13	49
Non-Parametric Procedures	39	2	10	1	10
Non Specified Tests	12	0	2	1	1

Note. Non specified tests refer to correlations or tests of difference which weren't specified by the author in the respective article.

Note. Articles were indexed according to types of statistical procedure employed. Thus, one articles could have more than one statistical procedure and accounts for more statistical procedures than statistical articles.

Table F-3

Frequency of Parametric Statistical Procedure by Journal for All Years Combined 1980-1984

Procedure	AJOT	OTMH	OTJR	PTOTG	PTOTP
Pearson Correlation	32	3	24	2	13
Chi Square	22	1	10	4	4
Student t	2	0	2	0	0
Paired t	12	1	7	2	2
Independent t	27	1	7	1	10
One Way ANOVA	18	3	10	2	2
Factorial ANOVA	22	2	9	1	9
Planned Contrasts	2	0	1	0	0
Post-hoc Multiple Comparisons	13	2	9	0	4
ANCOVA	5	1	2	0	2
Regression	8	0	5	0	1
Discriminant Analysis	2	1	1	0	1
Canonical Correlation	1	0	0	0	0
Factor Analysis	3	0	1	0	0
MANOVA	1	0	0	0	0
Meta Analysis	1	0	1	0	0
Other	1	1	7	1	1
Totals	172	16	96	13	49

Table F-4

Frequencies of Nonparametric Procedures by Journal for All Years Combined 1980-1984

Procedure	AJOT	OTMH	OTJR	PTOTG	PTOTP
Sign Test	1	0	0	0	0
Spearman Rho	9	1	1	0	0
Kendall's Tau	5	0	1	0	2
Point Biserial	0	0	0	0	1
Cramer's V	2	0	1	0	0
Fisher's Exact Test	2	0	3	0	0
Mann Whitney U-Test	9	0	2	0	3
Wilcoxon Matched Pair Signed Rank Test	3	0	0	0	2
Kruskall-Wallis ANOVA	4	1	1	1	1
Randomized Permutation Test (ANOVA)	1	0	0	0	0
Split Middle Test of Trend	3	0	1	0	0
Totals	39	2	10	1	9

Appendix G

Specification of Boxplot Procedures

Specification of Boxplot Procedures

1. List data values.

2. Rank order data values.

3. Calculate mean.

4. Calculate median position using the formula (n + 1)/2.

5. Specify value of data point in median position.

6. Calculate quartile location using the formula (truncated median position + 1)/2.

7. Identify Q3 by calculating quartile location counting down from the largest value in the data set.

8. Identy Q1 by calculating quartile location counting up from the smallest value in the data set.

9. Calculate interquartile range by subtracting Q1 from Q3.

10. Calculate scale factor by multiplying the interquartile range times 1.5.

11. Calculate the lower fence by subtracting the scale from the value associated with Q1.

12. Calulate the upper fence by adding the scale factor to Q3.

13. Calculate the outer, lower fence by subtracting two times the scale factor from Q1.

14. Calculate the outer, upper fence by adding two times the scale factor to Q3.

15. Plot quartiles Q1 and Q3.

16. Plot the mean with a circle and the median with a square.

17. Plot adjacent values with an "x" and draw tails.

18. Plot mild outliers with a circle and connect.

19. Plot extreme outliers with a blackened circle and connect.

A more detailed description of the boxplot procedures can be found elsewhere (Royeen, in press-b).

Appendix H

Data for Paired t-test

Table H-1

Duration of Nystagmus for Each Condition by Subject

Case	No. Seconds Bright	No. Seconds Dim	Difference Scores
1	24	23	1
2	26	26	0
3	16	10	6
4	22	19	3
5	23	22	1
6	19	20	-1
7	16	20	-4
8	13	12	1
9	19	21	-2
10	19	20	-1
11	15	16	-1
12	16	17	-1
13	8	8	0
14	20	21	-1
15	16	17	-1
16	23	18	5
17	21	19	2
18	14	14	0

Apendix I

Data from Independent t-test

Table I-1

Duration of Nystagmus for Each Subject by Group

Group	Total Score in Seconds
1	53
1	41
1	45.5
1	61.5
1	40.5
2	21
2	14
2	15
2	17
2	22
2	26
2	21
2	27
2	25
2	25

Appendix J

Data for One Factor Analysis of Variance

Table J-1

Grade for Each Subject by Group Psychiatric Internship

Group	Psychiatric Internship Grade
1	4.0
1	4.0
1	4.0
1	4.0
1	4.0
1	4.0
2	4.0
2	3.7
3	3.8
3	4.0
3	4.0
4	3.7
4	4.0

Table J-2

Grade Point Average for Each Subject by Group

Group	Grade Point Average
1	3.406
1	2.650
1	3.265
1	3.755
1	3.103
1	3.390
2	3.487
2	3.256
3	3.053
3	3.343
3	3.440
4	3.418
4	3.675

Table J-3

Grade for Each Subject by Group in Physical Dysfunction Internship

Group	Physical Dysfunction Grade
1	4.0
1	4.0
1	4.0
1	4.0
1	4.0
1	4.0
2	4.0
2	3.9
3	4.0
3	3.6
3	4.0
4	3.8
4	3.5

Appendix K

Data from Pearson Product Moment Correlation

Table K-1

Data for Correlations

Table K-1

Data for Correlations

Subject	Group	Sex	Scores on Dichotic Listening Test
01	2	2	23298554304563671730785943427149
02	2	2	16358564353458491833856127405662
03	2	1	15256756374568611531775131466467
06	2	2	14256557374165552116624730486571
07	2	1	15368565364164561934886129516777
12	2	2	07479080175358860749938316555990
15	2	1	14388768274863732033886033527177
18	2	2	21247552354466611825725540497465
19	2	2	25278751395175702029825637497268
20	2	2	21288255464072412325805153387624
22	2	2	16317860424472531827755646477852
23	2	2	27238347405680831930825840588291
24	2	1	14378567174048682917774238225037
26	2	1	16226354343860541822675335376052
29	2	2	23258051314261622820804435346661
31	1	2	17328261454777541537876644467553
32	1	2	18338561443465382024735347457746
33	1	2	19237555515790751432776255569355
35	1	1	24217548505083502324785253569163
36	1	2	23278353424875601824705443517865
37	1	1	09418373465373671042877450599191
38	1	2	28187743445387702821825251558864
39	1	2	21257753445583761429726056478624
50	1	1	16328061465383671731806544588589
51	1	1	21176348422455332119674938255339
52	1	2	18328360383763492029825633406157
53	1	2	21278054373358462124755240447056
54	1	2	18236853334565641922685233486869
55	1	1	15368565324262611540926921536284
59	1	2	20288056395074682131875730477390

Table K-1

Data for Correlations

Subject	Group	Sex	Scores on Dichotic Listening Test
01	2	2	23288554304563671730785943427149
02	2	2	16358564353458491833856127405662
03	2	1	15256756374568611531775131466467
06	2	2	14256557374165552116624730486571
07	2	1	15368565364164561934886129516777
12	2	2	07479080175358860749938316555990
15	2	1	14388768274863732033886033527177
18	2	2	21247552354466611825725540497465
19	2	2	25278751395175702029825637497268
20	2	2	21288255464072412325805153387624
22	2	2	16317860424472531827755646477852
23	2	2	27238347405680831930825840588291
24	2	1	14378567174048682917774238225037
26	2	1	16226354343860541822675335376052
29	2	2	23258051314261622820804435346661
31	1	2	17328261454777541537876644467553
32	1	2	18338561443465382024735347457746
33	1	2	19237555515790751432776255569355
35	1	1	24217548505083502324785253569163
36	1	2	23278353424875601824705443517865
37	1	1	09418373465373671042877450599191
38	1	2	28187743445387702821825251558864
39	1	2	21257753445583761429726056478624
50	1	1	16328061465383671731806544588589
51	1	1	21176348422455332119674938255339
52	1	2	18328360383763492029825633406157
53	1	2	21278054373358462124755240447056
54	1	2	18236853334565641922685233486869
55	1	1	15368565324262611540926921536284
59	1	2	20288056395074682131875730477390

Apendix L

<u>Representation of Formulae for Dichotic Listening Test</u>

Table L-1

Formulae for Dichotic Listening Coded Alphabetically

Alphabetical Designation	Formulae
A1	RCV1
A2	RCV2
B1	PERCV1
B2	PERCV2
C1	POECV1
C2	POECV2
D1	LD1
D2	LD2
E1	RD1
E2	RD2
F1	PERD1
F2	PERD2
G1	POED1
G2	POED2
H1	RCV1/LCV1
H2	RCV2/LCV2
I1	RD1/LD1
I2	RD2/LD2
J1	(R-L)CV1/(R+L)CV1
J2	(R-L)CV2/(R+L)CV2
K1	(R-L)D1/(R+L)D1
K2	(R-L)D2/(R+L)D2
L1	RCV1-LCV1
L2	RCV2-LCV2
M1	RD1-LD1
M2	RD2-LD2

Note. Aphabetical designation with number 1 is the first test situation. Alphabetical designation 2 designates the retest situation.

Appendix M

<u>Correlations of Equilibrium Scores with Age by Group</u>

Table M-1

<u>Correlations of Angle Measures with Age for Normal Girls</u>

(n = 10)

Variable	Original Analysis	Replicated Analysis & p Value	Spearman Rho & p Value
TBRAA	.695	.69517 .0256 *	.69301 .0263 *
TBTA	.328	.32845 .3541	.18293 .6130
FBRA	.487	.48727 .1532	.32518 .3592
TBRT	.322	.32179 .3646	.43698 .2067
TBTT	.132	.12310 .7348	.16103 .6567
FBRT	.611	.61073 .0607	.55833 .0935
FBPT	-.086	-.08629 .8126	.04606 .8995

Table M-2

Correlations of Angle measures with Age for Normal Boys

(n = 14)

Variable	Original Analysis	Replicated Analysis & p Value	Spearman Rho & p Value
TBRAA	.548	.54753 .0427	.60620 .0216
TBTA	.384	.38363 .1757	.35320 .2154
FBRA	.392	.39178 .1659	.40421 .1517
TBRT	.299	.29892 .2992	.35980 .2064
TBTT	.267	.26676 .3566	.38722 .1714
FBRT	.088	.08839 .7638	.07259 .8052
FBPT	.635	.63473 .0147 *	.62390 .0171 *

Table M-3

Correlations of Angle Measures with Age for Learning Disabled Boys (n = 9)

Variable	Original Analysis	Replicated Analysis & p Value	Spearman Rho & p Value
TBRAA	.751	.75100 .0197 *	.71495 .0304 *
TBTA	.176	.17628	-.03361
FBRA	.470	.47011 .2016	.46862 .2032
TBRT	.450	.45002 .2242	.47633 .1949
TBTT	-.105	-.10544 .7872	-.41032 .2727
FBRT	.333	.33331 .3808	.37553 .3193
FBPT	.136	.13633 .7265	.11161 .7750

Appendix N

Data for Univariate Regression Analysis

Table N-1

Angle Scores and Age in Months for All Subject

Normal Girls		Normal Boys		Learning Disabled Boys	
TBRA	Age	FBPT	Age	TBRA	Age
40	76	10	96	35	93
50	74	8	93	48	82
57	117	4	59	64	114
37	55	9	58	48	92
62	85	9	58	48	92
60	126	5	59	22	80
52	100	8	114	42	82
35	83	10	126	37	79
58	91	10	110	33	75
60	137	10	98		
		6	53		
		5	64		
		10	64		
		7	75		

The 6 page vita has been removed from the scanned document