DOCUMENT RESUME

ED 471 346 TM 034 637

AUTHOR	Leech, Nancy L.; Onwuegbuzie, Anthony J.
TITLE	A Call for Greater Use of Nonparametric Statistics.
PUB DATE	2002-11-00
NOTE	25p.; Paper presented at the Annual Meeting of the Mid-South Educational Research Association (Chattanooga, TN, November 6-8, 2002).
PUB TYPE	Reports - Descriptive (141) -- Speeches/Meeting Papers (150)
EDRS PRICE	EDRS Price MF01/PC02 Plus Postage.
DESCRIPTORS	Classification; Computer Software; Educational Research; *Effect Size; *Nonparametric Statistics; *Statistical Inference

ABSTRACT

This paper advocates the use of nonparametric statistics. First, the consequence of using parametric inferential techniques under nonnormality is described. Second, the advantages of using nonparametric techniques are presented. The third purpose is to demonstrate empirically how infrequently nonparametric statistics appear in studies, even those published in the most reputable journals. Fourth, a typology of nonparametric statistics is presented for all univariate general linear model analyses. The nonparametric statistics that are available in the most commonly used statistical software are delineated, and finally, nonparametric effect size indices are outlined. (Contains 1 table and 52 references.) (Author/SLD)

Running head: NONPARAMETRIC STATISTICS

A Call for Greater Use of Nonparametric Statistics

Nancy L. Leech

University of Colorado at Denver

Anthony J. Onwuegbuzie

Howard University

Correspondence should be addressed to Nancy L. Leech, University of Colorado at Denver, P. O. Box 173364, Campus box 106, Denver, CO 80217-3364, or E-Mail: nancy_leech@ceo.cudenver.edu

Paper presented at the annual meeting of the Mid-South Educational Research Association, Chattanooga, TN, November 7, 2002.

Abstract

This paper advocates use of nonparametric statistics. First, the consequence of using parametric inferential techniques under non-normality is described. Next, the advantages of using nonparametric techniques are presented. The third purpose is to demonstrate empirically how infrequently nonparametric statistics appear in studies, even those published in the most reputable journals. Fourth, a typology of nonparametric statistics is presented for all univariate GLM analyses. Fifth, the nonparametric statistics that are available in the most commonly used statistical software are delineated. Finally, nonparametric effect size indices are outlined.

A Call for Greater Use of Nonparametric Statistics

Whether to choose a parametric or nonparametric statistic can be one of the most difficult steps in analyzing data. Many researchers struggle with this step, or just ignore this step, by proceeding on to use the more common parametric statistic. The process of checking assumptions in order to justify use of the parametric statistic, and being certain that the data fit the assumptions, is paramount and should be undertaken regularly (Kerlinger, 1964, 1973; Nunnally, 1975; Tukey, 1977).

If the data violate the assumptions that justify use of the desired parametric statistic, then transformation of the data could be used that more adequately fits the assumption (Kirk, 1982). Indeed, a member of the family of Box-Cox transformations could be used (Box & Cox, 1964). For example, if the score distribution has moderate positive skew, then the square root transformation might be most appropriate; for severe positive skew a logarithm transformation might be useful; for moderate negative skew, reflecting the variable (i.e., subtracting each score from the largest score plus one) and then taking the square root might suffice; for severe negative skew, reflecting the variable and then taking the logarithm might be effective; finally, for *J*-shaped distributions, the inverse transformation might be the most adequate (Box & Cox, 1964; Bradley, 1982; Tabachnick & Fidell, 1996).

Whatever, transformation is used, it is essential that the same assumptions are checked on the transformed data. Providing an appropriate transformation is selected, transforming the data can be an extremely useful method for dealing with outliers, as well as for deviations from the assumptions of normality, linearity, and homoscedasticity (i.e., variability of scores for one continuous variable being approximately the same at all values of another continuous variable). However, the transformations can be problematic for two major reasons. First and foremost, it is not unusual that several transformations often must be attempted before the most appropriate one is found. This can be extremely time-consuming and frustrating for researchers working under tight deadlines for their research. Second, even when a suitable transformation is found, the subsequent inferential analysis is often more difficult to interpret than would have been the case if the variable had been analyzed in its original form. Thus, although some textbook authors (e.g., Tabachnick & Fidell, 1996) strongly recommend transforming data when assumptions are violated, very few researchers use this technique. In fact, the use of transformations typically is not considered important by instructors of graduate-level statistics and research methodology courses, nor is this topic even covered in these classes taught at the introductory level (Mundfrom, Shaw, Thomas, Young, & Moore, 1998). Such a lack of coverage likely leads to a lack of awareness of the potential usefulness of data transformations.

Because of the lack of awareness of data transformations coupled with the problems described above when they are used, few researchers transform their data. Consistent with this assertion, Keselman et al. (1998), who examined articles published in 17 prominent educational and behavioral science research journals in the 1994 or 1995, reported that data transformations were used in only 7.59% of articles involving between-subjects univariate designs (n = 79). Instead of using data transformations, some researchers decide to utilize the other option available for addressing assumption violations, namely, the nonparametric statistic (Gliner & Morgan, 2000; Hinkle, Wiersma, & Jurs, 1998; Newton & Rudestam, 1999). As stated by Hollander and Wolfe (1973, p.

1), nonparametric statistics represent a class of statistical methods that have specific "desirable properties that hold under relatively mild assumptions regarding the underlying populations(s) from which the data are obtained." Hotelling and Pabst (1936) are credited for developing this field (Savage, 1953).

Since the publication of the first textbook devoted exclusively to nonparametric procedures approximately one-half a century ago (Kendall, 1948), there has been a proliferation of textbooks dedicated to this topic. Yet, use of nonparametric statistics is extremely scant among researchers (Elmore & Woehlke, 1996; Jenkins, Fuqua, & Froehle, 1984). There are many possible reasons for this lack of use (Anderson, 1961; Blair, 1985). One reason stems from the fact that many researchers had graduate-level instruction in statistics that was taught in a rote manner. Another reason is that some researchers do not remember or do not know how to check their data for possible assumption violations (Sawilowsky, 1990). A third explanation for the lack of use of nonparametric statistics might arise from the fact that many graduate-level programs minimize students' exposure to statistical content and methodology. This started during the 1970's when nonparametric statistics were given a secondary role to parametric statistics in many textbooks (Sawilowsky, 1990; Winn & Johnson, 1978). Indeed, Aiken, West, Sechrest, and Reno (1990) reported that statistical and methodological curricula had advanced little in the previous 20 years. Thus, it is likely that many current researchers did not have much exposure to nonparametric statistics in their graduate courses.

A fourth reason for the scarcity of use of nonparametric statistics is that many researchers believe parametric statistics are extremely robust to violations to data assumptions (Boneau, 1960; Box, 1954; Glass, Peckman, & Sanders, 1972; Lindquist, 1953). Kerlinger (1964, 1973) and Nunnally (1975) discussed the lack of use of nonparametric statistics as stemming from the assumption that nonparametric tests are less powerful than parametric statistics. Finally, the dearth of use of nonparametric methods might have arisen from a failure, fear, or even refusal to recognize that analytical techniques that were once popular no longer reflect best practices, and, moreover, may now be deemed inappropriate, misleading, invalid, or obsolete.

Thus, this paper advocates use of nonparametric statistics. First, the role of statistical assumptions is described. Then, the consequence of using parametric inferential techniques under non-normality is presented. Next, the advantages of utilizing nonparametric techniques are presented. The third purpose is to demonstrate empirically how infrequently nonparametric statistics appear in studies, even those published in the most reputable journals. Fourth, a typology of nonparametric statistics is presented for all univariate GLM analyses. Fifth, the nonparametric statistics that are available in the most commonly used statistical software are delineated. Finally, nonparametric effect size indices are outlined.

Nonparametric Statistics and the Role of Statistical Assumptions

Most data in social science research fail to meet the assumptions for parametric statistics (Micceri, 1989). For these cases, if the data are not transformed, then nonparametric techniques should be utilized. To know when to use nonparametric statistics, a basic understanding of the role of statistical assumptions is necessary. Statistical assumptions can be thought of as "rules" or "guidelines" for a given statistic. Before a statistic is to be used, the assumptions for the statistic need to be checked to see if they have been met. All univariate parametric analyses, including analyses of bivariate relationships, are subsumed by the general linear model (GLM), and are therefore bounded by its assumptions. An important assumption that prevails for all univariate GLM analyses is that the dependent variable is normally distributed.

The more the normality assumption is violated, the less justified it is to rely on parametric statistics to conduct null hypothesis significance tests. However, many parametric statistics are assumed to be "robust" against reasonable violations of assumptions (Boneau, 1960, 1962; Box, 1953; Hinkle et al., 1998; Gardner, 1975; Minium, 1978; Newton & Rudestam, 1999). A statistical procedure or test is considered to be robust with respect to the particular underlying assumption, if it is reasonably insensitive to slight departures from the assumption (Hollander & Wolfe, 1973). If a parametric statistic is robust, then it can still be used when the assumption is not adequately met. Yet, Bradley (1978) and Singer (1979) contend that parametric statistics are not truly robust. Moreover, Bradley (1982) demonstrated that statistical inference becomes increasingly less robust as distributions depart from normality. Further, Tabachnick and Fidell (1996, p. 70) noted that "even when the statistics are used purely descriptively, normality, linearity, and homoscedasticity of variables enhance the analysis." In fact, according to some methodologists (e.g., Bradley, 1978; Singer, 1979), assumption violations are only tolerated by the overwhelming majority of researchers so that the parametric statistics can be used.

Disturbingly, the majority of studies in the social and behavioral sciences do not utilize random samples (Shaver & Norton, 1980a, 1980b), even though "inferential statistics is based on the assumption of random sampling from populations" (Glass &

Hopkins, 1984, p. 177). In fact, randomness, in the form of random error, is the basis for sampling distributions against which observed findings are compared (Carver, 1993). Further, use of nonrandom samples increases the chances that scores will be non-normal. Another factor that contributes to violations of normality is that many data sets are generated from small samples. These problems render it likely that the underlying samples yield scores from dependent measures that depart from normality. Thus, it is not surprising that the majority of data in the social and behavioral sciences are not normally distributed (Micceri, 1989).

When the dependent variable deviates from normality, the parametric GLM analysis should not be used. Bradley (1968) defined a nonparametric statistic as being a "distribution-free test ...which makes no assumptions about the precise form of the sampled population" (p. 15). Alternatively stated, nonparametric methods are termed distribution-free because they can be employed for variables whose joint distribution represents any specified distribution, including the bivariate normal, or whose joint distribution is not known and therefore is unspecified (Gibbons, 1993). Therefore, when the assumption of normality is not met, a nonparametric statistic is the more appropriate choice.

An important question to be asked is how much should scores deviate from normality before nonparametric statistics become essential. With regard to univariate inferential statistical techniques, Onwuegbuzie and Daniel (2002) have provided objective but simple criteria for determining whether scores deviate from normality. Specifically, these methodologists stated the following:

Additionally, for adequate sample sizes, a formal test of statistical significance can be conducted by utilizing the fact that the ratio of the skewness and kurtosis coefficients to their respective standard errors (i.e., standardized skewness and standardized kurtosis coefficients) are themselves normally distributed. Most other statistical packages print as options skewness and kurtosis coefficients but not their standard errors. However, these standard errors can be approximated manually (the standard error for skewness is approximately equal to the square root of $6/n$, and the standard error for kurtosis is approximately equal to the square toot of $24/n$, where n is the sample size). For both small and large sample sizes, rather than conducting a test of statistical significance, criteria can be used for assessing whether the standardized skewness and/or kurtosis coefficients are unacceptably large. One rule of thumb that we offer is that (a) standardized skewness and kurtosis coefficients which lie within ± 2 suggest no serious departures from normality, (b) coefficients outside this range but within the ± 3 boundary signify slight departures from normality, and (c) standardized coefficients outside the ± 3 range indicate important departures from normality. Using such a rule provides an objective method of assessing normality that is based on effect sizes (i.e., standardized coefficients). (pp. 75)

Consequences of not Meeting the Assumption

Problems arise when a parametric statistic is used with data that are not normally distributed. Labovitz (1967) points out that "a word of caution is necessary…it frequently turns out that the violation of one assumption does not appreciably alter the test, [although] the violation of two or more assumptions frequently does have a marked effect" (p.158). Thus, when the assumptions are not met, using a parametric statistic likely will generate invalid results (Field, 2000; Hinkle, Wiersma, & Jurs, 1998; Newton & Rudestam, 1999).

The use of a parametric statistic when the assumption of normality is grossly violated can have serious consequences (Siegel, 1956). In fact, large skewness and kurtosis coefficients affect Type I and Type II error rates. For instance, a non-normal kurtosis coefficient typically produces an underestimate of the variance of a variable, which, in turn, increases the Type I error rate (Tabachnick & Fidell, 1996). Although the parametric *t*-test is typically robust with regard to Type I error under the conditions of large and equal samples sizes, this test is not powerful for when data are characterized by skewed distributions. In fact, under skewed conditions, the Wilcoxon Rank Sum test, a nonparametric counterpart of the *t*-test is three to four times more powerful (Blair & Higgins, 1980; Bridge & Sawilowsky, 1999; Nanna & Sawilowsky , 1998)—a finding of which researchers appear to be unaware.

Advantages of Using Nonparametric Techniques

There are many advantages of using nonparametric techniques. Siegel (1956) outlined six main advantages. The first advantage is that for most nonparametric statistics, the "accuracy of the probability statement does not depend on the shape of the population" (p. 32). Further, the size of the sample is not as important, because small sample sizes will not cause the results to be misleading to the extent that small samples unduly affect parametric tests. The third advantage is that nonparametric statistics can be used when observations come from several different populations.

Next, nonparametric statistics can be used with data that are ordinal, or ranked, as well as with interval- and ratio-scaled data. Nonparametric statistics can be used with nominal data as well. Finally, for many researchers, nonparametric statistics can be easily learned and applied, at least at the univariate level. Most statistical computer software packages, such as the Statistical Package for the Social Sciences (SPSS; SPSS Inc., 2001) and the Statistical Analysis System (SAS Institute Inc., 2002), include nonparametric statistics.

McSeeney and Katz (1978) summarized the reasons for using nonparametric statistics. These include (a) nonparametric statistics have fewer assumptions, (b) nonparametric statistics can be used with rank-ordered data, (c) nonparametric statistics can be used with small samples, (d) data do not need to be normally distributed, and (e) outliers can be present.

Hollander and Wolfe (1973) provided six reasons for using nonparametric statistics. Specifically, they contended that nonparametric methods (a) require few assumptions about the underlying population from which the data are collected; (b) do not necessitate the assumption of normality; (c) are often easier to apply than are their parametric counterparts; (d) are typically easy to understand; (e) are appropriate when parametric methods cannot be employed; and (f) are only slightly less efficient than parametric methods under normality, while being more efficient under non-normality.

Further, when approximate normality is met, nonparametric tests are still relatively efficient--the asymptotic relative efficiency of nonparametric tests with respect to parametric tests can be as high as 95.5% (Gibbons, 1993; Hollander & Wolfe, 1973). Consequently, in many cases, researchers have relatively little to lose by using nonparametric tests if the distribution is normal. If the distribution is not normal, tests based on nonparametric tests likely are more efficient than are their parametric counterparts. It is thus surprising that researchers do not utilize nonparametric tests more than they do.

Use of Nonparametric Statistics in Published Journal Articles

Many graduate-level statistics and research methodology courses in the past have not included extensive information regarding nonparametric statistics (Aiken et al., 1990; Sawilowsky, 1990; Winn & Johnson, 1978). In fact, Mundfrom et al. (1998) found that the chi-square statistic was the only nonparametric statistic presented in introductory-level statistics and research methodology classes. Further, of the inferential statistics cited, the statistics and research methodology instructors indicated that the chi-square test was the fourth least most covered technique and was considered the fourth least important topic (Mundfrom, et al., 1998). Thus, nonparametric statistics infrequently appear in published articles, including those in the most reputable journals (Elmore & Woehlke, 1996; Jenkins et al., 1984). Moreover, many researchers do not report whether assumptions were checked, or whether the data fit the assumptions. For example, Keselman et al. (1998) reported that less than one-fifth of articles (i.e., 19.7%) "indicated some concern for distributional assumption violations" (p. 356). Similarly, Onwuegbuzie (in press) found that only 11.1% of researchers discussed the extent to which analysis of variance, analysis of covariance, multivariate analysis of variance, or multivariate analysis of covariance were violated.

To better understand this phenomenon, Royeen (1986) identified five published studies that used parametric statistics. For each study, the data were checked for whether it met the assumptions for the parametric statistic used. In three of the five studies, the assumptions were not met. Next, the appropriate nonparametric statistic was computed on the data. For each of the three studies that did not meet the assumptions, there were large differences in the results yielded by the nonparametric statistic when compared with the published results from the parametric statistic. Thus, this examination demonstrates that if the assumptions are not met, the results can be very misleading. Furthermore, this examination exemplifies the problem that many studies have: if the assumptions have not been checked and they have not been met for the parametric statistics utilized, then the results are invalid. This is important to note when reading published studies that do not include information about whether or not the assumptions have been checked.

A Typology of Nonparametric Statistics

A myriad of nonparametric statistics exists for conducting distribution-free tests. The vast majority of these tests are readily available from the major statistical software (e.g., SPSS, SAS). A selection of some of the most common tests is provided in Table 1.

Insert Table 1 about here

Nonparametric Effect Size Indices

As stipulated by the current edition of the American Psychological Association (APA) Publication Manual (2001):

> When reporting inferential statistics (e.g., *t* tests, *F* tests, and chi-square), include information about the obtained magnitude or value of the test statistic, the degrees of freedom, the probability of obtaining a value as extreme as or more extreme than the one obtained, and the direction of the effect.... Neither of the two types of probability value directly reflects the magnitude of an effect or the strength of a relationship. For the reader to fully understand the importance of your findings, it is almost always necessary to include some index of effect size or strength of relationship in your Results section...The general principle to be followed, however, is to provide the reader not only with information about statistical significance but also with enough information to assess the magnitude of the observed effect or relationship. (pp. 22-26)

Reporting effect sizes is no less important for statistically significant nonparametric findings than it is for statistically significant parametric results. However, when the few researchers who use nonparametric methods observe a statistically significant *p*-value, typically either they do not provide effect sizes, or they compute parametric-based effect sizes. First and foremost, statistically significant nonparametric statistics *always* should be followed up by some measure of effect size. However, it should be noted that just as parametric tests are adversely affected by departures from GLM assumptions, so too are parametric effect sizes (e.g., d, ω^2, ϵ^2). For example, as noted by Onwuegbuzie and Levin (2002), parametric effect sizes are affected by non-normality and heterogeneity of scores. Thus, whatever assumptions were violated that led to the use of nonparametric methods also would distort the parametric effect size. In fact, Hogarty and Kromrey (2001), using Monte Carlo methods, demonstrated that the most frequently used effect-size estimates (e.g., *d*) are extremely sensitive to departures from normality and homogeneity. Even trimmed effect-size measures (Hedges & Olkin, 1985; Yuen, 1974) exhibit extreme bias when the sample is small.

Therefore, researchers should consider following up statistically significant nonparametric *p*-values with nonparametric effect sizes. Nonparametric effect sizes include Cramer's *V*, the phi coefficient, and the odds ratio. These effect sizes indices, which are appropriate for chi-square analyses, are readily available on SPSS and SAS. Other nonparametric effect size estimates include (a) γ_1 (Kraemer & Andrews, 1982), which is based on the degree of overlap between samples; (b) the Common Language (CL) effect-size statistic (McGraw & Wong, 1992), which indicates the relative frequency with which a score sampled from one distribution is greater than a score sampled from a second distribution; (c) Vargha and Delaney's (2000) *A*, which is a measure of stochastic superiority that is appropriate for ordinally scaled distributions; (d) Cliff's (1993) *d*, appropriate for comparing two groups, which assesses the equivalence of probabilities of scores in each group being larger than scores in the other group (i.e., dominance); and (e) Wilcox and Muska's (1999) *W*, a nonparametric analogue of ω^2, which estimates the degree of certainty with which an observation can be linked to one population rather than the other. Of these five measures, Cliff's *d* and Vargha and Delaney's *A* appear to be the most robust to violations of normality and heterogeneity of variance (Hogarty & Kromrey, 2001). Unfortunately, none of these five nonparametric measures are computed by the major statistical software programs. Thus, software development companies can play an important role here in motivating researchers to follow up statistically significant nonparametric statistics with nonparametric effect sizes.

Conclusions

For the last 50 years, nonparametric techniques have been underutilized, despite the fact that statistical software routinely allows the computation of an array of nonparametric statistics, and despite the fact that parametric techniques are extremely sensitive to extreme violations to GLM assumptions. Unfortunately, many researchers are not being made adequately aware that nonparametric statistics provide viable alternatives to their parametric counterparts. Clearly, instructors, journal editors, and statistical software developers can play vital roles in promoting the nonparametric movement. In any case, much more work is needed to promote the use of distribution-free statistics. As such, we hope that the present call for the use of nonparametric techniques represents one small step in the right direction.

References

Aiken, L. S., West, S. G., Sechrest, L., & Reno, R. R. (1990). Graduate training in statistics, methodology, and measurement in psychology: A survey of Ph.D. programs in North America. *American Psychologist, 45*(6), 721-734.

American Psychological Association. (2001). *Publication manual of the American Psychological Association* (5th ed.). Washington, DC: Author.

Anderson, N. H. (1961). Scales and statistics: Parametric and nonparametric. *Psychological Bulletin, 58*, 305-316.

Blair, R. C. (1985, April). *Some comments on the statistical treatment of rank data*. Paper presented at the annual meeting of the American Educational Research Association and the National Council on Measurement in Education, Chicago, IL.

Blair, R. C., & Higgins, J. J. (1980). A comparison of power of Wlicoxon's rank-sum statistic to that of Student's *t* statistic under various nonnormal distributions. *Journal of Educational Statistics, 5*, 309-335.

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

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

Box, G. E. P. (1953). Non-normality and tests on variances. *Biometrika, 40*, 318-335.

Box, G. E. P. (1954). Some theories on quadratic forms applied to the study of analysis of variance problems: Effect of unequality of variance in the one-way classification. *Annals of Mathematical Statistics, 25*, 290-302.

Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society, 26*(Series B), 211-243.

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

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

Bradley, J. V. (1982). Robustness? The insidious L-shaped distribution. *Bulletin of the Psychonomic Society, 20*(2), 85-88.

Bridge, P. K., & Sawilowsky, S. S. (1999). Increasing physician's awareness of the impact of statistical tests on research outcomes: Investigating the comparative power of the Wilcoxon Rank-Sum test and independent samples *t* test to violations from normality. *Journal of Clinical Epidemiology, 52*, 229-235.

Carver, R. P. (1993). The case against statistical significance testing, revisited. *Journal of Experimental Education, 61, 287-292.*

Cliff, N. (1993). Dominance statistics: Ordinal analyses to answer ordinal questions. *Psychological Bulletin, 114*, 494-509.

Elmore, P. B., & Woehlke, P. L. (1996, April). Research methods employed in "American Educational Research Journal," "Educational Researcher," and "Review of Educational Research" from 1978 to 1995. Paper presented at the Annual Meeting of the American Educational Research Association, New York.

Field, A. (2000). *Discovering statistics; Using SPSS for Windows*. London: Sage.

Gardner, P. L. (1975). Scales and statistics. *Review of Educational Research, 45,* 43-57.

Gibbons, J. D. (1993). *Nonparametric measures of association* (Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07B091). Newbury Park, CA: Sage.

Glass, G.V., & Hopkins, K. D. (1984). *Statistical methods in education and psychology*. (2nd ed.) Englewood Cliffs, NJ: Prentice-Hall.

Glass,G., Peckman, 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*, 237-288.

Gliner, J. A., & Morgan, G. A. (2000). *Research methods in applied settings: An integrated approach to design and analysis*. Mahwah, NJ: Erlbaum.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.

Hinkle, D. E., Wiersma, W., & Jurs, S. G. (1998). *Applied statistics for the behavioral science*. Boston, MA: Houghton Mifflin.

Hogarty, K. Y., & Kromrey, J. D. (2001, April). We've *been reporting some effect sizes: Can you guess what they mean?* Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.

Hollander, M., & Wolfe, D. A. (1973). Nonparametric statistical methods. New York: John Wiley & Sons.

Hotelling, H., & Pabst, M. R. (1936). Rank correlation and tests of significance involving no assumptions of normality. *Annals for the Mathematical Statistician, 7*, 29-43.

Jenkins, S. J., Fuqua, D. R., & Froehle, T. C. (1984). A critical examination of the use of nonparametric statistics in the "Journal of Counseling Psychology." *Perceptual and Motor Skills, 59*, 31-35.

Kendall, M. G. (1948). *Rank correlation methods*. London, Griffin.

Kerlinger, F. N. (1964). *Foundations of behavioral research*. New York: Holt, Rinehart, and Winston.

Kerlinger, F. N. (1973). *Foundations of behavioral research* (2nd ed.) New York: Holt, Rinehart, and Winston.

Kraemer, H. C., & Andrews, G. A. (1982). A nonparametric technique for meta analysis effect size calculation. *Psychological Bulletin, 91*, 404-412.

Labovitz, S. (1967). Some observations on measurement and statistics. *Social Forces, 46*, 151-160.

Lindquist, E. F. (1953). *Design and analysis of experiment in psychology and education*. New York: Houghton Mifflin.

McGraw, K. O., & Wong, S. P. (1992). A common language effect size statistic. *Psychological Bulletin, 111*, 361-365.

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

Micceri, T. (1989). The unicorn, the normal curve, and other improbable creatures. *Psychological Bulletin, 105*(1), 156-166.

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

Mundfrom, D. J., Shaw, D. G., Thomas, A., Young, S., & Moore, A. D. (1998, April). *Introductory graduate research courses: An examination of the knowledge base*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Nanna, M. J., & Sawilowsky , S. S. (1998). Analysis of Likert scale data in disability and medical rehabilitation evaluation. *Psychometric Methods, 3*, 55-67.

Newton, R. R., & Rudestam, K. E. (1999). *Your statistical consultant: Answers to your data analysis questions*. Thousand Oaks, CA: Sage.

Nunnally, J. (1975). *Introduction to statistics for psychology and education*. New York: McGraw-Hill.

Onwuegbuzie, A. J. (in press). Common analytical and interpretational errors in educational research: an analysis of the 1998 volume of the British Journal of Educational Psychology. *Research for Educational Reform*.

Onwuegbuzie, A. J., & Daniel, L. G. (2002). Uses and misuses of the correlation coefficient. *Research in the Schools*, *9*, 73-90.

Royeen, C. B. (1986, April). A comparison of parametric versus nonparametric statistics. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA.

SAS Institute Inc. (2002). *SAS/STAT User's Guide* (Version 8.2) [Computer software]. Cary, NC: SAS Institute Inc.

Savage, I. R. (1953). Bibliography of nonparametric statistics and related topics.

Journal of the American Statistical Association, 48, 844-906.

Sawilowsky, S. S. (1990). Nonparametric tests of interaction in experimental design. *Review of Educational Research, 60*, 91-126.

Shaver, J. P., & Norton, R. S. (1980a). Populations, samples, randomness, and replication in two social studies journals. *Theory and Research in Social Education, 8*(2), 1-20.

Shaver, J. P., & Norton, R. S. (1980b). Randomness and replication in ten years of the American Educational Research Journal. *Educational Researcher, 9*(1), 9-15.

Siegel, S. (1956). *Nonparametric statistics for the behavioral 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.

SPSS Inc. (2001). *SPSS 11.0 for Windows.* [Computer software]. Chicago, IL: SPSS Inc.

Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics* (3rd ed.). New York: HarperCollins College Publishers.

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

Vargha, A., & Delaney, H. D. (2000). A critique and improvement of the CL Common Language effect size statistics of McGraw and Wong. *Journal of Educational and Behavioral Statistics, 25*, 101-132.

Wilcox, R. R., & Muska, J. (1999). Measuring effect size: A nonparametric analogue of ω^2. *British Journal of Mathematical and Statistical Psychology, 52*, 93-110.

Winn, P. R., & Johnson, R. H. (1978). *Business statistics*. New York: Macmillan.

Yuen, K. K. (1974). The two-sample trimmed t for unequal population variances. *Biometrika, 61*, 165-170.

Table 1: Typology of Nonparametric Statistics

Method	Test
Measures of Association:	
Spearman's Rank Correlation Coefficient	consistency
Kendall's Rank Correlation Coefficient	consistency
Chi-square Test of Independence	concordance/discordance
Tau	consistency
Theil test	slope of regression line
Cochran Test	consistency
Fisher's exact test	relationships
Single Population Tests:	
Binomial	proportions
Kolmogorov-Smirnov Goodness-of-Fit Test	goodness of fit test for continuous data
Sign Test	paired replicates
Wilcoxon Signed Rank Test	symmetry and equality of location
Gupta test	symmetry
Hodges-Lehman One-sample Estimator	median
Comparison of Two Populations:	
Chi-square Test of Homogeneity	differences in proportions
Wilcoxon (Mann-Whitney) Test	differences in location and spread
Kolmogorov-Smirnov Two-Sample Test	differences between population distributions
Rosenbaum's Test	differences in location
Tukey's Test	differences in spread
Hodges-Lehman Two-sample Estimator	difference in medians
Savage Test	differences in spread when medians equal
Ansari-Bradley Test	differences in dispersion
Moses Confidence Interval	differences in location
Comparison of Several Populations:	
Kruskal-Wallis Test	symmetry and equality of location
Friedman's Test	symmetry and location (two-way data)
Terpstra-Jonckheere Test	medians equal vs. changing median
Page's Test	ordered alternatives
The Match Test for Ordered Alternatives	medians equal vs. medians ordered
Miller's jackknife Test	unknown squared ratio of scale differs from 1
Hollander Test	X and Y variables are interchangeable

TM034637

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

ERIC
Educational Resources Information Center

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: A CALL FOR GREATER USE OF NONPARAMETRIC STATISTICS	
Author(s): NANCY L. LEECH + ANTHONY J. ONWUEGBUZIE	
Corporate Source:	Publication Date: 2002

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2A documents	The sample sticker shown below will be affixed to all Level 2B documents
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 1	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2A	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2B
Level 1 ✓	Level 2A ☐	Level 2B ☐
Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.	Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only	Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, please

Signature: AJO	Printed Name/Position/Title: ANTHONY J. ONWUEGBUZIE PhD	
Organization/Address: SCHOOL OF EDUCATION HOWARD UNIVERSITY	Telephone: 202-249-2416	FAX: 202-249-2416
	E-Mail Address: TONYONWUEGBUZIE@aol.com	Date: 11/6/02

2441 4TH STREET NW
Washington, DC 20059

(Over)

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

ERIC CLEARINGHOUSE ON ASSESSMENT AND EVALUATION
UNIVERSITY OF MARYLAND
1129 SHRIVER LAB
COLLEGE PARK, MD 20742-5701
ATTN: ACQUISITIONS

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
4483-A Forbes Boulevard
Lanham, Maryland 20706

Telephone: 301-552-4200
Toll Free: 800-799-3742
FAX: 301-552-4700
e-mail: info@ericfac.piccard.csc.com
WWW: http://ericfacility.org