

2004

An Analysis of the Effects That Several Design Variables Have on the Accuracy and Precision of Glass' Delta and Hedges' g

Joshua N. Dryer

Embry-Riddle Aeronautical University - Daytona Beach

Follow this and additional works at: <https://commons.erau.edu/db-theses>



Part of the [Ergonomics Commons](#)

Scholarly Commons Citation

Dryer, Joshua N., "An Analysis of the Effects That Several Design Variables Have on the Accuracy and Precision of Glass' Delta and Hedges' g" (2004). *Theses - Daytona Beach*. 48.

<https://commons.erau.edu/db-theses/48>

This thesis is brought to you for free and open access by Embry-Riddle Aeronautical University – Daytona Beach at ERAU Scholarly Commons. It has been accepted for inclusion in the Theses - Daytona Beach collection by an authorized administrator of ERAU Scholarly Commons. For more information, please contact commons@erau.edu.

An analysis of the effects that several design variables have on the accuracy
and precision of Glass' delta
and Hedges' g.

Joshua N. Dryer
Embry-Riddle Aeronautical University

A thesis proposal submitted to the Department of Graduate Studies of Human Factors and
Systems at Embry-Riddle Aeronautical University for the fulfillment of the requirements
for the degree of

Master of Human Factors and Systems

Daytona Beach, Florida
2004

UMI Number: EP32061

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI[®]

UMI Microform EP32061
Copyright 2011 by ProQuest LLC
All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Copyright 2004 Joshua N. Dyer
All Rights Reserved

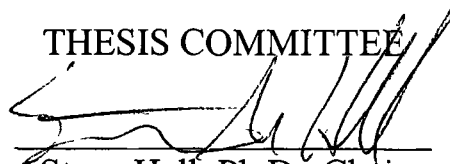
AN ANALYSIS OF THE EFFECTS THAT SEVERAL DESIGN
VARIABLES HAVE ON THE ACCURACY AND PRECISION OF
GLASS' DELTA AND HEDGES' G

by


Joshua N. Dryer

This thesis was prepared under the direction of the candidate's thesis committee chair, Steve Hall, Ph. D., Department of Human Factors & Systems, and has been approved by the members of the thesis committee. It was submitted to the Department of Human Factors & Systems and has been accepted in partial fulfillment of the requirements for the degree of Master of Science in Human Factors & Systems.


THESIS COMMITTEE



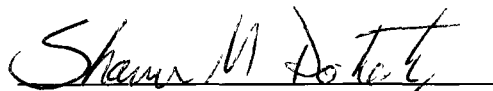
Steve Hall, Ph.D., Chair



Linda Trocine, Ph.D., Member



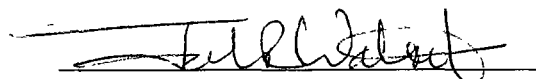
Roger Kayser, Member



MS HFS Program Coordinator



Department Chair, Department of Human Factors and Systems



Associate Dean of Student Academics

ABSTRACT

Effect size is the standardized effect that some treatment has on a sample of a population. In particular, Hedges' g and Glass delta are mean difference effect size estimators that are used to compute the effect sizes found in an experimental situation. A confidence interval is an interval placed around a point estimate that indicates the precision with which the point estimate can be made. This paper provides an explanation of the concept of effect size estimation and confidence interval calculation, the different methods that can be used to calculate effect sizes and confidence intervals, and applies these methods in a Monte Carlo simulation. It was found that under most conditions the method of effect size and confidence interval calculation was not relevant.

ACKNOWLEDGMENTS

I must acknowledge first and foremost Jesus Christ, my Lord and Savior for without Him, we would be lost. I also want to thank my loving wife Darlene for her faith in me and my work. To my mentor, Steve Hall, who truly cares about his students and taught me everything I know about statistics. Thanks

Steve, for all of the work that you put into me and this project, I was so blessed to have been able to work with such a fine professor. Finally, thanks to my mom, dad, and family for the support that you have given me through my years of study.

TABLE OF CONTENTS

	PAGE
INTRODUCTION.....	1
EFFECT SIZE.....	3
THE BIG THREE.....	6
COHEN'S <i>d</i>	7
HEDGES' <i>g</i>	8
GLASS' delta.....	9
EFFECT SIZE SELECTION.....	10
CONFIDENCE INTERVALS.....	12
METHODS.....	18
RESULTS.....	24
DISCUSSION.....	29
REFERENCES.....	35
APPENDIX A.....	39
APPENDIX B.....	40
APPENDIX C.....	46
APPENDIX D.....	65

INTRODUCTION

The field of social science is still in its infant stages when being compared to such well established fields as chemistry and physics. One can expect many mistakes as well as the field makes changes leading to improvements. This is particularly true when referring to statistical analysis and how it is applied to social science research.

Anyone educated in the field of the social sciences is familiar with how experiments and observations are conducted. Both of these scenarios usually end with some sort of statement referring to statistical significance. “The differences between drug A and drug B are statistically significant” is a good example of a conclusion to typical experimental research. Recently, statistical significance testing has been the subject of debate between the statisticians and professionals within the social sciences. Many say that it should not be used at all anymore while others say it should be used in conjunction with other, more meaningful analyses (Cohen, 1988).

In response to this current debate the American Psychological Association (APA) Task Force on Statistical Inference has provided some guidance stating, “Always provide some effect-size estimate when reporting a p value” (Wilkinson & APA Task Force on Statistical Inference, 1999, p. 599). It is now the position of the APA that a significance test alone is not adequate enough analysis and many of the top scientific journals now require effect size estimation for publication. The fifth edition of the APA (2001) *Publication Manual* stressed the importance of effect size reporting:

“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. You can estimate the magnitude of the effect or the strength of the relationship with a number of common effect size estimates...The general principle to be followed...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. 25-26)”

Effect size estimation is not only important to provide the reader with enough information to assess the situation correctly but also for future studies. Cooper and Hedges (1994) defined a study and effect size best: “if we define a study to mean a set of observations taken on a subject sample on one or more occasions, there are three possible dimensions of variation in study results. There may be different results for different measures, for different sub-samples of people, and for different times of measurement. All of these variations have implications from conceptualizing and coding effect sizes” (Cooper & Hedges, 1994, p. 112). They illustrated the importance for effect size measurements and how they can help to standardize the results obtained in a study to allow the reader to generalize the results as well as compare the results across studies.

Comparing the results across studies is called *meta-analysis*. Rosenthal (1991) stated that, “When we ask whether two studies are telling the same story, what we usually mean is whether the results (in terms of the estimated effect size) are reasonably consistent with each other or whether they are significantly heterogeneous” (Rosenthal, 1991, p. 63). Meta-analysis is becoming the most popular statistical technique for analyzing data in the social sciences. The main goal of any experiment or observation is to analyze a sample of a population and in turn infer these results onto the population. Meta analysis takes statistical inference one step further by aggregating the results of individual studies to better estimate the relationship among constructs at the population level. The resulting inference based on many separate values is more precise and meaningful than any one of the experiments could have produced on their own.

It seems apparent then that reporting effect sizes should be a common practice in every experimental study; however, this is not the case and effect size reporting is still not part of the template of data interpretation. Researchers may be reluctant to report effect sizes because of the numerous effect sizes that can be computed and the complexity surrounding the proper computation of effect sizes. Which effect size should be used? Which effect size is the most precise (confidence intervals placed around effect size estimation)? Is this precision based on design characteristics? The current study will address these questions. First, an overview of what an effect size is and problems with effect size estimation will be discussed. Second, each of the major effect size indices will be described. Finally, the appropriate times to use each of the effect size estimators will be discussed.

Effect Size

Cohen (1988) perhaps stated the definition of an effect size the best: “it can now be readily made clear that when the null hypothesis is false, it is false to some specific degree, i.e., *the effect size (ES) is some specific nonzero value in the population.* The larger the value, the greater the *degree* to which the phenomenon under study is manifested” (Cohen, 1988, p. 10). So, in other words the effect size is the degree to which the experimental group is affected by the treatment, only presented in a standardized form. Cohen later stated how the effect size is related to the null hypothesis: “Thus, whether measured in one unit or another, whether expressed as a difference between two population parameters or the departure of a population parameter from a constant or in any other suitable way, the effect size can itself be treated as a parameter which takes the value zero when the null hypothesis is true and *some other specific*

nonzero value when the null hypothesis is false, and in this way the effect size serves as an index of degree of departure from the null hypothesis” (Cohen, 1988, p. 10). Effect size is so useful, as Cohen stated, due to the fact that no matter what way the experimental design was composed, the effect size can standardize the results and make them comparable across different studies. Finally, Cohen stated why an effect size is so important to statistical analysis: “we need a ‘pure’ number, one free of our original measurement unit, with which to index that can be alternately called the degree of departure from the null hypothesis of the alternate hypothesis, or the effect size we wish to detect. This is accomplished by standardizing the raw effect size as expressed in the measurement unit of the dependent variable by dividing it by the (common) standard deviation of the measures in their respective populations, the latter also in the original measurement unit” (Cohen, 1988, p. 20).

Effect size estimation, however, is not a perfect instrument of statistical analysis. This type of estimation can be affected by error just like any other statistical inference. Restriction of range or unreliability in measurement instruments can lead to attenuation of the magnitude of the effect size estimate. Error variance, or any aspect of the experiment that is not controlled by the experimenter, may lead to an incorrect estimate of effect size. The effect size based on data collected during an experiment is called the observed effect size because it is only an estimation of the true effect size present in the population of interest. The observed effect size is different from the true effect size due to the experimenter’s inability to gather the entire population or any type of experimental bias, as discussed above (Cooper & Hedges, 1994; Hunter & Schmidt, 1990). This is why the term effect size estimator is used; the true effect size of a population is unknown.

An experiment just takes a sample from the population and the effect size found from that sample is the best estimate that the experimenter can make about the true effect size of the population.

In addition to sampling error, other factors may bias an effect size estimate. Hunter and Schmidt (1990) listed 10 other factors (11 factors in all), but the most common are measurement unreliability and restriction of range. Theoretically, effect size estimates can be corrected for some of Hunter and Schmidt's 11 artifacts, but these corrections are often estimates in and of themselves (see Hall & Brannick, 2002).

Rosenthal (1991) commented on the Glass, McGaw, and Smith (1981) corrections and he basically stated that such tests as repeated measures, analysis of covariance, and blocked variable designs do have a tendency to produce larger effect sizes as well as larger amounts of significant test statistics (Rosenthal, 1991). Hunter and Schmidt (1990) also proposed a set of corrections in which they adjust for unreliability for the studies' variables, dichotomization of variables, restriction of range, poor construct validity, and unequal sample sizes. The work of Hunter and Schmidt is important because it reminds us that there are many different factors in an experiment that can lower the obtained effect size estimate and these factors can be corrected to obtain the most accurate effect size estimate (Rosenthal, 1991). Rosenthal made a good point questioning how important an effect size is if it does not represent the true effect size in a real experimental situation. The adjusted effect size will probably give a closer estimation of the real population level effect size in a perfect experimental environment. Those environments, however, do not exist, nor will they ever exist. The corrections do seem to control for the errors mentioned above, but do they just change the surface variables and leave all the affected

variables within the study unchanged? Do the corrections change the integrity of the study or the interpretation of the results? What scientists and readers really want to know is, “what is really happening in the real world where there are not perfect sample pools, unequal sample pools, and imperfect construct validity?”

The Big Three

Effect size estimation can be done in several ways. Specifically, an effect size can be computed by using the relationship between the variables or differences between the group means. The relationship between the variables is formed with such formulas as the Pearson Product-Moment Correlation Coefficient (r), odds ratios, risk rates, and risk differences (Law, Schmidt, & Hunter, 1994). Group mean differences, which are the focus of this study, are used to form an effect size estimator. There are three main mean differences effect size indices: Cohen’s d , Glass’s delta, Hedges’ g . These effect size indices are all very similar; with minor differences in the denominator used by each.

Two of the most common mean difference effect size estimates are Glass’s delta and Hedges’ g (equations 2 and 3, respectively). Both of these estimates use s in the denominator and are hence appropriate to use when working with sample data. Because of the use of pooled s , Hedges’ g assumes homogeneity of variance. When this assumption is violated, as is often the case, Glass’ delta is a more appropriate estimate to use. Delta uses the control group standard deviation in the denominator, addressing the violation of the homogeneity of variance assumption at the expense of precise estimation of the standard deviation.

Cohen's d

Cohen's d is basically a theoretical definition of effect size and uses pooled σ as the standardizing unit. In practice, σ is usually unknown and is estimated using s . Nonetheless, Cohen's d provides the basic formulation for mean difference effect size estimates (equation 1).

$$\frac{\bar{X}_{large} - \bar{X}_{small}}{\sigma_{pooled}} \quad [1]$$

As stated above Cohen's d standardizes the study results with σ pooled. Two assumptions must be met to use Cohen's d : equal sample sizes and that the standard deviations of both groups are similar (homogeneity of variance).

Cohen proposed the use of sigma as opposed to the actual standard deviation of the experimental population. Since this number is not attainable, the experimenter's best estimate is most appropriate. Cohen knew that this figure could not be attained but wrote the formula in broad population oriented terms. A Cohen's d obtained from the population parameters would not be an effect size estimator, it would be the effect size of the population, but since it is impossible to obtain these numbers the best estimate will do. In a simulation study, Rosenthal (1991) indicated that sample groups with equal sample sizes produced very precise effect size estimates, however as sample sizes varied Cohen's d tended to underestimate the effect size (Rosenthal, 1991).

Cohen's d has also been found to have a small amount of positive bias. Sawilowsky and Yoon (2001) ran a simulation study of Cohen's d and found a small, positive bias in d values. Specifically, average d estimates were .17 even though the true modeled effect size was 0. This is concerning because an effect size of .20 is often

labeled as a “small effect.” Roberts and Henson (2002) criticized this finding because Sawilowsky and Yoon did not allow computed d values in the simulation to be negative, thus artificially inflating the resulting average d value.

Roberts and Henson (2002) replicated the Sawilowsky and Yoon simulation, but allowed the computed d values to be negative. They found a much smaller positive bias in d , and also found that the existing Ezekiel correction procedure tended to overcorrect the d values for this bias. While the small positive bias in d estimates is expected, there are few situations in which correction for this bias is actually recommended (Hunter & Schmidt, 1990).

Hedges g

While Cohen’s d provides a theoretical and conceptual definition of standardized effect size, Hedge’s g represents the practical implementation of Cohen’s definition,

$$\frac{\bar{X}_{large} - \bar{X}_{small}}{S_{pooled}} \quad [2].$$

As stated above the only difference between Hedges g and Cohen’s d is that Hedges g pools the sample standard deviation (S) while Cohen’s d pools the population standard deviation (σ). The only difference between these is that the sample standard deviation is divided by $n-1$ participants and the population standard deviation is divided by n participants. Indeed, Hedge’s g is essentially the practical application of Cohen’s d . While Hedge’s g addresses the practical issue of estimating the standard deviation, it is still subject to the same assumptions of Cohen’s d , namely equal sample sizes and equality of variances.

Rosenthal (1991) supported the use of the pooled sample standard deviation: “The pooled S —that is, the one computed from both groups—tends to provide a better

estimate in the long run of the population standard deviation. However, when the S 's based on the two different conditions differ greatly from each other, choosing the control group S as the standardizing quantity is a very reasonable alternative" (Rosenthal, 1991, p. 16). It seems as if Rosenthal is stating that the pooled S is the best standard error figure to use if the sample group fulfills both assumptions. If the standard deviations are not homogeneous than the control group standard deviation would be sufficient.

Glass' Delta

Glass' delta addresses the issue of unequal group variances by using the control group standard deviation as the denominator (equation 3).

$$\frac{\bar{X}_{large} - \bar{X}_{small}}{S_{control}} \quad [3]$$

The use of the control group standard deviation is partially justified by the notion that the control group variance represents the amount of variability in scores expected in the general untreated population. It is not uncommon for the implantation of some treatment to either increase or even decrease score variability (Grissom & Kim, 2001). The former will occur when some sort of treatment by subject interaction occurs, meaning that the treatment does not work for everyone. The later occurs when the treatment acts as a mechanism to equate performance or some other outcome across skill levels or some other innate characteristic. In some cases, second order sampling error may explain differences between treatment and control group variances, but this is most likely to occur with small sample sizes. Indeed, with less than 20 degrees of freedom, point estimates of σ can be in error by hundreds of percent (Hoaglin, Mosteller, & Tukey, 1991).

Grissom & Kim argued that sometimes it is necessary to use only the control group standard deviation but that caution should be exercised because the control group and experimental groups may indeed have the same standard deviations in which case both the control and experimental group's standard deviations should be used (Grissom & Kim, 2001). The authors cautioned the use of using only the control group standard deviation when the experimental group standard deviation may be equivalent. This is important to know because if both groups' standard deviations could be used as a pooled standard deviation, the pooled standard deviation would provide a better estimate of the true population standard deviation. In other words omitting the experimental group's standard deviation when it should not be omitted would weaken the ability to generalize the study's results to the population. The decision to use the pooled estimate versus the control group only estimate should be based on educated insight. If there is no reason to believe that the treatment should cause a change in score variability, the pooled estimate should be used even if the observed sample-based variance estimates appear different.

Glass' delta should only be used in specific situations, as mentioned above, and in those situations it may be the only effect size estimator that can consistently report an accurate effect size. However, caution should be taken to ensure that it is not used in situations like those previously mentioned for fear that it will weaken the study or give an inaccurate estimate of the population effect size.

Effect Size Selection

So how will a confidence interval help a researcher decide which effect size estimator to use in an experiments analysis? Rosenthal (1991) supported the use of confidence intervals for both meta-analysis as well as blocked meta-analysis because it

would better facilitate transference of the statistic to the population level (Rosenthal, 1991). A confidence interval gives the reader an idea of how precise that estimate really is, specifically with what level of certainty the author can place around that estimate. Rosenthal supported the use of confidence intervals for not only meta-analysis, but also for individual studies in order to enable the scientist to evaluate the results on a population level. The argument for the incorporation of confidence intervals into t tests for example is also applicable to confidence intervals and effect sizes. A confidence interval allows the researcher to apply his or her results to the population as well as indicate how precise the point estimate is.

Picking the appropriate effect size estimator and forming a confidence interval is also important due to selective publishing. Selective publishing is an unfortunate side effect of publishing only those results that reach a certain level of significance. Most statisticians know that almost any effect can be found significant with a large enough sample size. The reason for using effect size estimators and confidence intervals is to select those studies not by their level of significance but by their effect size and confidence interval. Cooper and Hodges (1994) also caution the use of effect sizes and confidence intervals. They noted that using effect sizes can cause studies with extreme effect sizes to be published and thus encourage selective publication. Also, studies with smaller sample sizes have the capability of producing much larger effect size estimates than their larger sample size counterparts; an occurrence such as this is termed an induced association (Cooper & Hodges, 1994). Using effect sizes with confidence intervals will help to combat this problem; however it will not solve it. There is no solution to the problem that Cooper and Hodges spoke of; however the results of this study will help the

social sciences, specifically fields such as Human Factors to be able to produce more consistent publications. This means that similar experiments will produce similar results and those results will be analyzed in a regulated fashion and in turn earn more respect from other more developed fields of study. A side effect of more consistent publications and a greater respect is industries' need for human factors engineering and other related fields still in their infancy.

Confidence Intervals

Confidence intervals were first constructed by Neyman in the 1930's (Cumming & Finch, 2001). A confidence interval (CI) is a range of numbers placed around a point estimate in order to state with a certain level of confidence (C) that that interval contains the population parameter being estimated. If all of the possible samples of equal sample size were taken from the population approximately 95% of those intervals would contain the population parameter and 5% would not (Hinkle, Wiersma, & Jurs, 1998). Now, seven decades later, the importance of confidence interval computation is once again being realized in the social sciences. Confidence interval formulation is the future of statistical analysis particularly within null hypothesis and effect size estimation testing in the social sciences (Wilkenson et al., 1999).

The formulation of confidence intervals for effect size estimates is not an easy task and there is much speculation as to which procedure is best. Proper CI's for effect size estimates are not computed using a central distribution. They must be computed differently using a non-central distribution because the alternative hypothesis distribution is usually centered on some non-zero value. For example, group mean comparisons are often examined using the t -test in which the observed value of t serves as the center of the

expected distribution of t . This distribution is not normally distributed, thus the upper and lower bounds, given some level of confidence, are not equidistant from the center. Non-central distributions occur when a normally distributed variable (effect size) has a mean not equal to zero and is divided by a standard deviation estimate. The standard deviation is distributed approximately as χ^2 , which is not a normal distribution. Another important aspect of non-central distributions is that they change shape for every different value of the hypothesized parameter (Smithson, 2001). Smithson also stated that a CI for an effect size estimation is really constructed on the basis of a sample statistic and that it is not constructed around a sample statistic. There are no common tables like those of central test statistics for non-central distributions, so iteration must be used to find the best estimate from the sample pool because the values are not exact and maximum and minimum values (upper and lower bounds for the CI) are being estimated. Feingold (1995) found that this difference occurs due to the difference in variability and central tendency between both groups being compared. Feingold also reported that it is necessary to compute each tail of the CI for an effect size one at a time (Feingold, 1995). The amount that the distribution is skewed depends on the non-centrality parameter (Δ), which is the distance that the mean of the normal distribution is displaced from zero (Cumming & Finch, 2001). This non-centrality parameter is estimated using the observed value of t (Smithson, 2001). Cumming and Finch illustrated that there are several characteristics of non-central t distributions that must be remembered when forming a CI for an effect size: 1) When $\Delta = 0$ there is a central t distribution which is symmetric and centered at zero. 2) Non-central t is centered on Δ , which can be positive or negative. 3) For any given Δ , as df increases non-central t approaches central t in

shape. 4) This change, however, is very gradual. For example, when $\Delta = 2$ and $df = 60$, non-central t is just visibly skewed. 5) Larger Δ values make the approach to symmetric shape slower. 6) When n is at least 30 it can be assumed that the normal distribution will be a good approximation (n being all group sample sizes assuming equal sample sizes). 7) The curve is more skewed for smaller df and decreases as df increases. 8) If Δ is positive, the skew will be positive (Cumming & Finch, 2001). These characteristics are very important in the formation of CI's specifically in determining which type of distribution to assume (central or non-central t) as well as what one can expect from different degrees of Δ .

The calculation of a CI for a non-central t distribution is somewhat different than that of a central distribution. A central t distribution is affected by df , while a non-central t distribution is affected by df and the non-centrality parameter. In order for a distribution to be central, the population group means in question must be equivalent. Cumming and Finch illustrated this location shift by rewriting the t statistic formula as

$$\frac{(\bar{X} - \mu) + (\mu - \mu_0)}{S / \sqrt{n}} \sim t_{n-1, \Delta} \quad [4].$$

The location shift represents a measure of some kind of effect, which is estimated by an effect size estimator. In Null Hypothesis Significant Testing, it is assumed that the null hypothesis is true and so $\mu = \mu_0$, thus it is assumed that there is no effect. If there is an effect and a CI is to be placed around that effect then the non-centrality parameter (Δ) must first be estimated using

$$\Delta = \frac{(\mu - \mu_0)}{\frac{\sigma}{\sqrt{n}}} \quad [5].$$

Next, the lower and upper bounds must be estimated at one time (equations 6 and 7, respectively).

$$\Pr \left(t_{n-1, \Delta L} \geq \frac{\bar{X} - \mu_0}{S / \sqrt{n}} \right) = .025 \quad [6]$$

$$\Pr \left(t_{n-1, \Delta U} \geq \frac{\bar{X} - \mu_0}{S / \sqrt{n}} \right) = .975 \quad [7]$$

So, $t_{n-1, \Delta L}$ is the non-central t distribution that gives the estimated t value with a probability of .025 in its lower tail and $t_{n-1, \Delta U}$ is the non-central t distribution that gives the estimated t value with a probability of .975 in its upper tail. Since t estimates the non-centrality parameter, delta, the upper and lower bounds of delta are essentially the upper and lower bounds of the t distribution assuming H_a is true. These t values can be converted into Hedge's g values using equation 8.

$$g = \frac{2t}{\sqrt{N}} \quad [8]$$

The estimation of the upper and lower bounds of a non-central t distribution is made practical by the use of iteration methods performed using some sort of software program. Smithson (2001) has created syntax code for SPSS that estimates the CI for an effect size estimate and Cumming & Finch (2001) cited their use of ESCI graphical software, which runs under the Microsoft EXCEL program.

In conclusion, the process of forming a CI around an effect size estimate can be broken down into steps. An estimate of an effect size can be expressed in terms of a t distribution, which is centered on observed t . The observed t value serves as an estimate of the non-centrality parameter. The upper and lower bounds of the CI around t , given

some level of confidence, is estimated using an iterative process. These upper and lower values can then be transformed into an effect size estimate such as g .

An effect size estimate can be calculated for any group mean difference regardless of the whether that difference is statistically significant or not. However, it is imperative that a CI be placed around that effect size estimate so that the user can determine whether or not the effect size is of interest to him or her. The use of CIs allows the researcher to justify the presentation of an effect size estimate even if the corresponding NHST is not significant. If a CI is not presented, the reader will be unable to adequately evaluate the meaningfulness of the reported effect size in light of the precision with which it was estimated. That is why it is so important for the social sciences to blend the already in place method of NHST with effect size estimation and the formation of CI's around all point estimates.

The Present Study

Smithson (2001) and Cumming and Finch (2001) made an impressive argument about the need to create CIs around effect size estimates and furthermore, that non-central CIs should be constructed. Unfortunately, the construction of non-central intervals adds a layer of complexity to interval construction that may discourage some researchers from constructing CIs at all. Additionally, the complexity of effect size estimation is further complicated by the need to correct effect size estimates for inherent bias. It is no wonder that many researchers have simply chosen to not report confidence intervals around standardized estimates of effect sizes, given the number of choice points involved in the process. Therefore, the goal of the present study is to evaluate the accuracy and precision

of the Hedge's g and Glass' delta estimates using both central and non-central interval construction procedures as well as applying correction for bias.

A variety of factors will impact the accuracy and precision of each effect size estimate and they will also determine the extent to which non-central interval construction and bias correction will impact the final results. Therefore, the accuracy and precision of the various methods of effect size estimation will be examined across factorial combinations of raw effect size, variance ratio (i.e. unequal variance), and sample size.

Based on the past literature, these factors should affect the precision and accuracy of each of the effect size estimators in the following ways:

- 1) Hedge's g will be the most accurate and precise estimator if the sample sizes are equal and homogeneity of variance is present.
- 2) Glass' delta will be the most accurate and precise estimator if the standard deviations of the experimental group and control differ greatly (heterogeneity of variance) or if the experimental group is affected by uncontrolled variables making it different from the control group.
- 3) For those samples with $n < 30$, a non-central t distribution will better estimate the proper confidence interval around the effect size estimate than will a central t distribution, regardless of the effect size estimator that is used.
- 4) As the observed group mean difference changes so does the distribution in that the higher the observed group mean difference, the higher the non-centrality parameter and the slower the distribution approaches a symmetrical shape.

METHODS

A Monte Carlo simulation was programmed using MINITAB statistical software (MINITAB Inc., 1997). Samples of data were generated using pre-defined population parameters of the variance ratio, sample size, and observed group mean difference. In total, 72 unique combinations were simulated separately.

Monte Carlo Simulation

The simulation program began by defining several population parameters to generate virtual populations from which individual sample data would be drawn. These sample data were then used to compute effect size estimates and confidence intervals using the Glass' delta and Hedges' g approaches.

As discussed in the introduction, a variety of conditions were simulated in order to determine how the various effect size estimation procedures would perform under a variety of likely scenarios.

Mean Difference

The nominal mean difference was manipulated in order to generate a wide range of g scores. These values were selected so that both small and large effect sizes were simulated. The values of .25, .5, .75, 1, 2, and 3 were used. In the simulation, this was achieved by setting the parameter of μ_C to 1 and then adjusting the parameter, μ_T , to achieve the desired mean difference value, where the sub-scripts C and T indicate control and treatment groups, respectively.

Variance Ratio

Equality of variance across the treatment and control groups is preferable as it allows for the pooling of standard deviation estimates, thus enhancing the precision with

which sigma can be estimated. In many situations, such equality does not exist. Three ratios of treatment to control group variances were used: .25, 1, and 4. In the Monte Carlo simulation, this was achieved by keeping a constant control group standard deviation ($\sigma = 1$) and manipulating the treatment group standard deviation ($\sigma = .5, 1, 2$ respectively). The general rule of thumb is that a variance ratio of 3 or greater indicates heterogeneity (Keppel, 1991). The variance ratios selected created bias ($vr = .25$), homogeneity between the groups ($vr = 1$), and heterogeneity between the groups ($vr = 4$). In a practical sense, heterogeneity usually develops when the treatment being administered impacts the participants at different rates. In some cases, the treatment may actually reduce score variability. In any event, heterogeneity of variance is not uncommon and poses specific problems for the computation of standardized effect size estimates.

Sample Size

The sample sizes of the two groups were manipulated at four levels ($n = 5, 10, 30,$ and 50). The sample sizes of the treatment and control groups were always equivalent ($n_C = n_T$). The levels of group sample sizes were chosen for three reasons. First, group sample sizes of 5 and 10 were chosen because they represent group sample sizes of the majority of experimental studies. Second, sample sizes of 30 were chosen because Cumming and Finch (2001) found that group sample sizes of 30 or more produced a non-central distribution that closely resembled a central distribution. In other words, the sample distribution was basically no longer non-central but central in form. Lastly, a group sample size of 50 was used to go above and beyond Cumming and Finch's findings to observe how much closer the non-central distribution would change towards a central distribution.

Process

The process began by defining population parameters for two separate populations (control and treatment populations). For example, one combination of parameters included:

Control population: $\mu_C = 1, \sigma_C = 1, n_C = 5$

Treatment population: $\mu_T = 2, \sigma_T = 1, n_T = 5$

Next, 'n' numbers were randomly generated from the two population distributions and these two groups of 'n' numbers were used to calculate the sample statistics (mean and standard deviation):

Sample number	Control	Treatment
1	1	2
2	1.5	2.1
3	1.1	2.3
4	1.2	1.9
5	.9	1.8
Sample Mean	1.14	2.02
Standard Dev.	.230	.192

After that, computations were performed using the sample statistics which rendered a Hedges' *g* effect size estimate and a Glass' *delta* effect size estimate. Each of these point estimates were stored in an array. This process was repeated 1,000 times, each time the Hedges' *g* and Glass' *delta* point estimates were stored in the array.

The result was an array 2 x 1,000 in size. The mean was computed for each vector, representing the point estimate of *g* and *delta*, separately. The values in each of the two vectors were then rank-ordered and the 25th and 975th values were identified. These values represented the lower and upper tails of the distribution of observed effect

size values, thus identifying the bounds of the 95% confidence interval for g and delta. The various population parameters in Minitab were adjusted to create the next experimental condition and the sampling process was repeated for that experimental condition. This continued until all 72 unique experimental conditions were completed. The Minitab source code to produce one of the 72 experimental conditions is presented in Appendix A.

Estimation Procedures

The Monte Carlo simulation procedure was used to generate empirical effect size values to represent “truth” given some set of parameters. In a real study, the researcher usually has access to only one set of sample-based statistics from which effect size estimates and confidence intervals can be computed. The parameters defined in Minitab were used as “sample-based” information and were used to generate estimates of Hedge’s g and Glass’ delta. The equations provided by Hunter and Schmidt (1990, pgs. 267 – 290) were used throughout this process. The point estimates of g and delta were computed using the equations 9 through 11.

$$g = \frac{\bar{x}_1 - \bar{x}_2}{S_{Pooled}} \quad [9]$$

$$delta = g\sqrt{[(1 + v^2) / 2]} \quad [10]$$

$$v = \frac{s_t}{s_c} \quad [11]$$

The standard error of g was then computed and was used to create the central confidence interval for g (see equations 12 and 13). The corresponding lower and upper bound Glass’ delta values were computed by applying equation 10 to the lower and upper bound values of g .

$$s_{\bar{g}} = \sqrt{\left[\frac{(N-1)}{(N-3)} \right] \cdot \left[\left(\frac{4}{N} \right) \cdot \left(\frac{1+g^2}{8} \right) \right]} \quad [12]$$

$$CI_g = g \pm t_{(05, N-2, 2)} \cdot s_{\bar{g}} \quad [13]$$

Several authors have suggested the use of correction values for the various standardized effect size estimates (Hedges & Olkin, 1985; Hunter & Schmidt, 1990). The impact of such corrections has been the source of some debate, but most would agree that correction processes have little impact unless n is very small. Given that this simulation included small sample size situations, the correction equation described by Hunter and Schmidt (1990) was applied to the Hedge's g computations to produce adjusted point estimates and confidence intervals. The correction factor, a , is applied to the point estimate, g , as well as to $s_{\bar{g}}$ (see equations 14 to 16).

$$a = 1 + \frac{.75}{N-3} \quad [14]$$

$$g' = g / a \quad [15]$$

$$s_{\bar{g}}' = s_{\bar{g}} / a \quad [16]$$

Thus far, mathematically derived central confidence intervals and adjusted central confidence intervals for both g and delta have been computed. The computation of the non-central distributions and adjusted non-central distributions for g and delta were more complex and required a computer program to perform the calculations. Fortunately, Smithson (2003) provided detailed instructions on how to perform these calculations as well as a link to SPSS syntax code that would do the computations. The basic premise of the program is to find the t value for a given circumstance and use that t value as an estimate of the non-centrality parameter. This information was then used in an iterative

fashion to find the lower and upper ends of the non-central t distribution that corresponds with the given t value. The upper and lower t values were then converted into g values. The g values were converted into delta values using equation 10 and a complete set of adjusted values for both g and delta were computed using equations 6 through 8. The complete syntax used to generate the non-central and adjusted non-central distributions of g and delta is presented in Appendix B. The result of the procedures thus far can be summarized as follows:

1. Empirically derived confidence intervals for g and delta based on the Monte Carlo simulation.
2. Mathematically derived central confidence intervals for g and delta.
3. Mathematically derived adjusted central confidence intervals for g and delta.
4. Non-central confidence intervals for g and delta based on Smithson's (2003) iteration method.
5. Adjusted non-central confidence intervals for g and delta.

Empirical vs. Estimate Comparisons

The primary purpose of this study was to determine which estimation procedure would provide the best estimate of the true distribution of effect size values under a variety of circumstances. Therefore, the empirically derived means and upper and lower bound estimates were used as the standard against which all other estimates were compared. The adjusted and non-adjusted mathematical and iteration estimates were compared to their empirical counterparts by subtracting the empirical value from the estimated value. The difference between the two numbers represents how far off the

estimate is relative to the likely population value. This process was done for all of the various estimates for the lower and upper bound and point estimate values.

RESULTS

The results of the Monte Carlo simulation and appropriate calculations are presented in Figures 1 through 18 (Appendix C) and in Tables 1 through 6 (see Appendix D). The 18 graphs are pictorial representations of the overlap between the empirical confidence intervals (Hedges' g and Glass' δ) generated by the Monte Carlo simulation and the calculated confidence intervals (central g , non-central g , adjusted central g , adjusted non-central g , central δ , non-central δ , adjusted central δ , and adjusted non-central δ). The empirical and computed interval data were used to find the differences between the empirical outcomes and the various estimation procedures. Tables 7 through 12 present these difference values across the 72 possible combinations of variance ratio, mean difference, and sample size (labeled appropriately). The difference values were computed by subtracting the computed LB, point estimate, or UB from the empirically derived data set from the corresponding value as estimated by both Hedges' g and Glass' δ procedures. Corrected values are also presented in this table.

It should also be noted that when analyzing the results of the manipulation of the mean difference variable it was imperative that only the data in the groups with variance ratio = 1 were used to insure that there was no variance ratio by mean difference interaction.

Hedges g

Central g

As can be seen in Figures 9 and 6 and Tables 9 and 12, central *g* point and confidence interval estimates were closest to the empirical values when the VR = 1, MD = .5, and $n = 50$ and deviated the most when the VR = .25, MD = 3, and $n = 5$. Note: this was computed by adding the absolute value of the differences between each of the upper and lower bounds of the confidence intervals to the difference between the computed effect size and the empirical effect size.

Variance Ratio. When the variance ratio was not equal to 1, the central *g* confidence interval became too wide and the point estimate tended to be too low.

Mean Difference. The central *g* confidence interval did not seem to be affected by the change in mean difference; however the point estimate tended to become too low as the mean difference increased.

Sample Size. When the sample size was less than $n = 50$, the central *g* confidence interval was too wide or too narrow depending on the variable combination and the point estimate tended to be too low.

Non-central g

As can be seen in Figures 8 and 6 and Tables 8 and 12, non-central *g* point and confidence interval estimates were closest to the empirical values when the VR = 1, MD = .25, and $n = 50$ and deviated the most when the VR = .25, MD = 3, and $n = 5$.

Variance Ratio. When the variance ratio was not equal to 1, the non-central *g* confidence interval became too narrow and the point estimate tended to be too low.

Mean Difference. The non-central g confidence interval did not seem to be affected by the change in mean difference until mean difference was 2 or greater, at which time it became too narrow. The point estimate tended to become too low as the mean difference increased.

Sample Size. When the sample size was less than $n = 50$, the non-central g confidence interval became too narrow and the point estimate tended to be too low.

Adjusted Central g

As can be seen in Figures 6 and 10 and Tables 10 and 12, adjusted central g point and confidence interval estimates were closest to the empirical values when the VR = 1, MD = .75, and $n = 50$ and deviated the most when the VR = .25, MD = 3, and $n = 5$.

Variance Ratio. When the variance ratio was not equal to 1, the adjusted central g confidence interval became too narrow and the point estimate tended to be too low.

Mean Difference. The adjusted central g confidence interval did not seem to be affected by the change in mean difference until mean difference was 2 or greater, at which time it became too narrow. The point estimate tended to become too low as the mean difference increased.

Sample Size. When the sample size was less than $n = 50$, the adjusted central g confidence interval became too narrow and the point estimate tended to be too low.

Adjusted Non-central g

As can be seen in Figures 8 and 10 and Tables 8 and 12, adjusted non-central g point and confidence interval estimates were closest to the empirical values when the VR = 1, MD = .25, and $n = 50$ and deviated the most when the VR = .25, MD = 3, and $n = 5$.

Variance Ratio. When the variance ratio was not equal to 1, the adjusted non-central g confidence interval became too narrow and if the variance ratio was less than 1 the point estimate tended to be too low, but if it was more than 1, the point estimate remained basically unchanged..

Mean Difference. The adjusted non-central g confidence interval did not seem to be affected by the change in mean difference until mean difference was 2 or greater, at which time it became narrower. The point estimate tended to become too low as the mean difference increased.

Sample Size. When the sample size was less than $n = 50$, the adjusted non-central g confidence interval became too narrow and the point estimate tended to be too low.

Glass' delta

Central delta

As can be seen in Figures 8 and 12 and Tables 8 and 12, central delta point and confidence interval estimates were closest to the empirical values when the VR = .1, MD = .25, and $n = 30$ and deviated the most when the VR = 1, MD = 3, and $n = 5$.

Variance Ratio. The central delta confidence interval and point estimate seemed to be unaffected by the change in variance ratio, with most values being too narrow and low.

Mean Difference. The central delta confidence interval was only slightly affected by the increase in mean difference until mean difference was 2 or greater, at which time it became narrower. The point estimate tended to become too low as the mean difference increased.

Sample Size. When the sample size was less than $n = 50$, the central delta confidence interval became too narrow and the point estimate tended to be too low.

Non-central delta

As can be seen in Figures 2, 3, 8, and 12 and Tables 8, 9, and 12, non-central delta point and confidence interval estimates were closest to the empirical values when: (VR = .1, MD = .25, n = 50), (VR = .25, MD = .25, n = 50), (VR = .25, MD = .5, n = 50) and deviated the most when the VR = 1, MD = 3, and n = 5.

Variance Ratio. When the variance ratio was not equal to .25, the non-central delta confidence interval became too narrow and point estimate seemed to be unaffected by the change in variance ratio, with most values being too low.

Mean Difference. The non-central delta confidence interval was only slightly affected by the increase in mean difference until mean difference was 2 or greater, at which time it became narrower. The point estimate tended to become too low as the mean difference increased.

Sample Size. When the sample size was less than $n = 50$, the non-central delta confidence interval became too narrow and the point estimate tended to be too low.

Adjusted central delta

As can be seen in Figures 3 and 12 and Tables 9 and 12, adjusted central delta point and confidence interval estimates were closest to the empirical values when VR = .25, MD = .5 and n = 50 and deviated the most when the VR = 1, MD = 3, and n = 5.

Variance Ratio. When the variance ratio was not equal to .25, the adjusted central delta confidence interval became too narrow and point estimate seemed to be unaffected by the change in variance ratio, with most values being too low.

Mean Difference. The adjusted central delta confidence interval was only slightly affected by the increase in mean difference until mean difference was 2 or greater, at

which time it became narrower. The point estimate tended to become too low as the mean difference increased.

Sample Size. When the sample size was less than $n = 50$, the adjusted central delta confidence interval became too narrow and the point estimate tended to be too low.

Adjusted non-central delta

As can be seen in Figures 2, 8, and 12 and Tables 8 and 12, adjusted non-central delta point and confidence interval estimates were closest to the empirical values when: (VR = .1, MD = .25, $n = 50$) and (VR = .25, MD = .25, $n = 50$) and deviated the most when the VR = 1, MD = 3, and $n = 5$.

Variance Ratio. When the variance ratio was not equal to .25, the adjusted non-central delta confidence interval became too narrow and point estimate seemed to be unaffected by the change in variance ratio, with most values being too low.

Mean Difference. The adjusted non-central delta confidence interval was only slightly affected by the increase in mean difference until mean difference was 2 or greater, at which time it became narrower. The point estimate tended to become too low as the mean difference increased.

Sample Size. When the sample size was less than $n = 50$, the adjusted non-central delta confidence interval became too narrow and the point estimate tended to be too low.

DISCUSSION

The accuracy and precision of two different effect size estimation procedures were compared against empirical values derived using Monte Carlo simulation. Both central and non-central intervals were created in order to assess when the creation of the more complex non-central interval is warranted. Furthermore, the effect size estimates

and confidence intervals were computed using both non-adjusted and adjusted techniques. The goal of the current research was to identify the method of estimation and confidence interval construction that produces maximal overlap with the empirically derived intervals under a variety of circumstances (i.e. different sample sizes, unequal variances, and various mean differences).

Table 13 illustrates the average differences between the calculated point estimates and confidence intervals with the empirical point estimates and confidence intervals. For instance, at the variance ratio = 1 level, the average central d differences include the average of the differences between central d , its confidence interval, and the empirical values for each at all of the mean difference levels (1, .25, .5, .75, 2, and 3) and all of the sample size levels (5, 10, 30, and 50). Table 13 basically indicates which effect size estimator and confidence interval formation method is best for which situation.

It is quite clear that effect size estimation is part of the future of experimental testing (Wilkinson & APA Task Force on Statistical Inference, 1999). This simulation study has also highlighted the importance of confidence interval calculation for effect size estimates. Effect size estimates and confidence intervals around those estimates are good indicators as to the degree of effect (in a common language) as well as the precision of that estimate. The main reason for performing this simulation was to determine if there was a superior effect size estimation and confidence interval calculation method. The variables selected for the Monte Carlo simulation helped to indicate which methods were best for each situation.

In light of the literature review, it is evident that Hedges' g and Glass' delta cannot be compared as equal effect size estimators (Rosenthal, 1991 & 1994). Each has

specific assumptions that should be met in order to be an accurate estimator. Each of these should be used according to those assumptions.

Central versus Non-central distribution

When central and non-central distributions for the computation of confidence intervals are compared, the average differences in the confidence intervals formed around the point estimates are on the order of a few hundredths (see Table 13). Notice that the central distribution calculations seem to predict the upper bound of the confidence interval better and the non-central distribution seems to predict the lower bound of the confidence interval better. There is less difference between the two compared (central vs. non-central) within Glass' delta than Hedges' g. Depending on how accurate a scientist would like to be (more or less than the hundredths place), it seems as if using a central or a non-central distribution to compute a confidence interval around an effect size point estimate does not really matter.

Unadjusted versus Adjusted Point Estimates

When the unadjusted and adjusted point estimates are compared to the empirical point estimate values, the unadjusted point estimates are more accurate (see Table 13). This higher level of accuracy extends across all of the simulated conditions. The lower bound confidence interval estimates across the non-adjusted and adjusted processes were almost equivalent (within .03 in all cases) and the unadjusted calculations produced closer estimates of the upper bounds of the confidence intervals in all cases. These data indicate that the use of unadjusted point estimates and CI construction are preferable to the adjusted counterparts.

Conclusion

The proper estimation of effect sizes and their corresponding confidence intervals is somewhat confusing. Many researchers may debate on which estimate to use (Hedge's g or Glass' δ), whether or not bias correction should be applied, and whether or not a non-central confidence interval should be constructed. Indeed, the perceived need to calculate a non-central confidence interval may persuade some researchers to forgo confidence interval construction altogether. While there are certainly theoretical reasons to use bias correction and non-central distribution assumptions, the actual benefits of implementing these procedures may be outweighed by the cost of additional complexity. Indeed, in some situations, the use of bias correction or non-central interval construction may actually reduce the accuracy of the resulting estimates.

The findings of this study can be interpreted in two general ways. First, the reader can use the data in Table 13 to identify the most accurate and precise procedure to follow given a specific circumstance. Second, general conclusions about the overall appropriateness of each estimation procedure, presented below and also in Table 14, can be used as a guide for selecting the best approach for effect size estimation.

The results suggest that the use of Glass' δ will produce more accurate point estimates and confidence intervals when the variance ratio is less than one and estimates similar to Hedges' g with variance ratios of 1 and 4. In practice though, the choice to use δ or g is usually made based on theoretical grounds.

Theoretically, a non-central distribution should be used when creating confidence intervals around point estimates of effect size. The data, however, illustrate that using a central distribution would produce very similar results in most cases. This is not to say

that the use of non-central distributions for confidence interval construction will be counter productive, only that the added complexities involved in creating such distributions may not be warranted by the small differences in outcomes. Those who are comfortable with the construction of such intervals may wish to continue their use.

There seems to be a good deal of consensus on when bias correction is practically useful (Hunter & Schmidt, 1990) and the data bear this out. Specifically, bias correction is most effective when small sample sizes are used. In most other situations, the use of bias correction actually reduces the accuracy of point estimates and confidence intervals.

Available Method	Conclusion
Glass' delta vs. Hedges' g	Glass' delta will produce more accurate point estimates and confidence intervals when the variance ratio is less than one and estimates similar to Hedges' g with variance ratios of 1 and 4.
Central vs. Non-central distribution assumptions	A central distribution produces very similar results as a non-central distribution in most cases which can make using a non-central distribution less attractive (factoring in the added calculations required to form the CI).
Adjusted vs. un-adjusted effect size estimations	Bias correction is most effective when small sample sizes are used. In most other situations, the use of bias correction actually reduces the accuracy of point estimates and confidence intervals.

Table 14
The basic conclusions derived from the study about effect size estimators, distribution assumptions, and bias adjustments.

Future Research

This study focused on effect size estimates and confidence intervals around those estimates for mean difference calculations. It will be interesting to see how the social sciences will use the current research that the field of statistics has recently presented and if it will indeed increase the utility of publications within the field. It would be very informative to re-analyze published study results and compute effect size estimates and corresponding confidence intervals using the various methods described in this paper in order to assess the extent to which methodological choices impact final results. Such analyses may indeed support the notion that methodological choices in the estimation of effect sizes have little impact on the outcome.

REFERENCES

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

Bird, K. D. (2002). Confidence Intervals For Effect Sizes in Analysis of Variance. *Educational and Psychological Measurement*, 62 (2), 197-226.

Brandstatter, E. (1999). Confidence intervals as an alternative to significance testing. *Methods of Psychological Research Online*. 4 (2). <http://www.ipn.uni-kiel.de/mpr/>

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Cooper, H., Hedges, L. V., (1994). *The Handbook of Research Synthesis*. New York, New York: Russell Sage Foundation.

Cumming, G., Finch, S. (2001). A primer on the understanding, use and calculation of confidence intervals that are based on central and non-central distributions. *Educational and Psychological Measurement*. 61 (4), 532-574.

Feingold, A. (1995). The additive effects of differences in central tendency and variability are important in comparisons between groups. *American Psychologist*. 50 (1), 5-13.

Finch, S., Cumming, G., Thomason, N. (2001). Reporting of statistical inference in the *Journal of Applied Psychology*: little evidence of reform. *Educational and Psychological Measurement*, 61 (2), 181-210.

Glass, G. V., McGaw, B., Smith, M. L. (1981). *Meta-analysis in Social Research*. Beverly Hills, California: Sage Publications.

- Grissom, R. J., Kim, J. J. (2001). Review of Assumptions and Problems in the Appropriate Conceptualization of Effect Size. *Psychological Methods*, 6, 135-146.
- Hall, S. M., Brannick, M. T. (2002). Comparison of Two Random-Effects Methods of Meta-Analysis. *Journal of Applied Psychology*, 87 (2), 377-389.
- 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 sciences* (4th ed.). Boston, Mass.: Houghton Mifflin Company.
- Hittner, J. B., May, K., Silver, N., C. (2003). A Monte Carlo evaluation of tests for comparing dependent correlations. *The Journal of General Psychology*. 130 (2), 149-168.
- Hoaglin, D.C., Mosteller, F. and Tukey ,J.W. (Eds). (1991). *Fundamentals of exploratory analysis of variance*. New York: John Wiley & Sons.
- Huberty, C. J. (2002). A History of Effect Size Indices. *Educational and Psychological Measurement*, 62 (2), 227-240.
- Hunter, J. E., and Schmidt, F. L. (1990). *Methods of meta-analysis: correcting error and bias in research findings*. Sage, California.
- Hyde, J. S. (2001). Reporting Effect Sizes: The Roles of Editors, Textbook Authors, and Publication Manuals. *Educational and Psychological Measurement*, 61 (2), 225-228.
- Keppel, G. (1991). *Design and analysis: A researcher's handbook* (3rd ed.). Upper Saddle River, New Jersey: Prentice-Hall, Inc.

- Kirk, R. E. (2001). Promoting good statistical practices: some suggestions. *Educational and Psychological Measurement*, 61 (2), 213-218.
- Kromrey, J., D., Hogarty, K., Y. (1998). Analysis options for testing group differences on ordered categorical variables: An empirical investigation of type I error control and statistical power. *Multiple Linear Regression Viewpoints*. 25, 70-82.
- Law, K. S., Schmidt, F. L., & Hunter, J. E. (1994b). A test of two refinements in procedures in meta-analysis. *Journal of Applied Psychology*, 79, 978-986.
- Riopelle, A., J. (2000). Are effect sizes and confidence levels problems for or solutions to the null hypothesis tests? *The Journal of General Psychology*. 127 (2), 198-216.
- Roberts, J. K., Henson, R. K. (2002). Correction for bias in estimating effect sizes. *Educational and Psychological Measurement*, 62, 241-253.
- Rosenthal, R. (1994). Parametric measures of effect size. In H. Cooper & L.V. Hedges (Eds.), *The handbook of research synthesis* (pp. 231-244. New York: Russell Sage Foundation.
- Rosenthal, R. (1991). *Meta-analytic procedures for social research* (revised). Newbury Park : Sage.
- Rosenthal, R. & Rosnow, R. L. (1991). *Essentials of behavioral research: Methods and data analysis* (2nd ed.). New York: McGraw Hill
- Rosnow, R. L., & Rosenthal, R. (1996). Computing contrasts, effect sizes, and counternulls on other people's published data: General procedures for research consumers. *Psychological Methods*, 1, 331-340

Smithson, M. (2001). Correct confidence intervals for various regression effect sizes and parameters: The importance of non-central distributions in computing intervals. *Educational and Psychological Measurement*, 61 (4), 605-632.

Thompson, B. (2002). What future quantitative social science research could look like: Confidence intervals for effect sizes. *Educational Researcher*, 25-32.

Thompson, B. (2001). Editor's not on the "Colloquium on effect sizes: The roles of editors, textbook authors, and the publication manual. *Educational and Psychological Measurement*, 61 (2), 211-212.

Wilkinson, L., & American Psychological Association Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanation. *American Psychologist*, 54, 594-604.

Wilson, S.A., Becker, L. A., Tinker, R. H. (1995). Eye movement desensitization and reprocessing (EMDR) treatment for psychologically traumatized individuals. *Journal of Consulting and Clinical Psychology*, 63, 928-937.

Vacha-Haase, T. (2001). Statistical significance should not be considered one of life's guarantees: effect sizes are needed. *Educational and Psychological Measurement*, 61 (2), 219-224.

APPENDIX A

```
Name c1 = 'control'
Random 5 'control';
  Normal 0.0 1.0.
Name c2 = 'treat'
Random 5 'treat';
  Normal 1.0 1.0.
Name C3 = 'effectsize'
Let 'effectsize' = Mean(treat)-Mean(control)
Name C5 = 'controlstd'
Let 'controlstd' = STDEV(control)
Name C6 = 'treatstd'
Let 'treatstd' = STDEV(treat)
Name C7 = 'stdevpool'
Let 'stdevpool' = (STDEV(control) + STDEV(treat)) / 2
Name C8 = 'cohen'
Name C4 = 'hedges'
Name C9 = 'glass'
Let c4(1) = c3 / c7
let c9(1) = c3/c5
let c13 = count (control)
let c14 = count (treat)
Name C10 = 'stdc'
let 'stdc' =sqrt(rsum((c1(1)-mean(control))** 2,(c1(2)-mean(control))** 2,(c1(3)-mean(control))** 2,(c1(4)-mean(control))** 2,(c1(5)-mean(control))** 2)/c13)
name c11 = 'stdc2'
let 'stdc2' = sqrt(rsum((c2(1)-mean(treat))** 2,(c2(2)-mean(treat))** 2,(c2(3)-mean(treat))** 2,(c2(4)-mean(treat))** 2,(c2(5)-mean(treat))** 2)/c14)
name c12 = 'pstdcn'
let 'pstdcn' = rsum(c10,c11) / 2
let c8(1) = c3/c12
let k2 = count (c8)
add k2 1 k2
execute "program1.mtj" 20
```

APPENDIX B

COMMENT BEGIN BY SUPPLYING THE COMPUTER WITH THE MEAN DIFFERENCE
BETWEEN TWO GROUP (MD), THE SAMPLE SIZE IN EACH GROUP (n).
COMMENT AND THE ESTIMATED SD FOR EACH GROUP; THIS PROGRAM ASSUMES
EQUAL GROUP SAMPLE SIZES.
COMMENT NOTE THAT THE VARIANCE RATIO (vr) IS BASED ON SDt / SDc AND IS
EXPRESSED IN VARIANCE (not sd) UNITS.
COMMENT THUS A VR=4 INDICATES THAT THE TREATMENT GROUP VARIANCE WAS
4X'S THAT OF THE CONTROL GROUP VARIANCE.
COMMENT IN sd TERMS, SD FOR TREATMENT IS DOUBLE THAT OF SD CONTROL.

```
COMPUTE tval = MD/sqrt(((sigmac**2+sigmat**2)/2)*(2/n)) .  
COMPUTE df = (2*n)-2.  
COMPUTE conf = .95.  
EXECUTE .
```

```
COMMENT THIS SCRIPT COMPUTES CONFIDENCE INTERVALS FOR THE  
NONCENTRALITY PARAMETER  
COMMENT FOR THE NONCENTRAL T DISTRIBUTION.  
COMMENT IT USES THE SPSS NONCENTRAL T CALCULATOR AND LAUBSCHER'S (1960)  
NORMAL APPROXIMATION  
COMMENT TO THE NONCENTRAL F WITH 1 DF FOR THE SPECIAL CASE WHERE F = T^2,  
COMMENT WITH A DECISION RULE FOR CHOOSING BETWEEN THEM. THE REASON FOR  
THIS IS THAT THE  
COMMENT NONCENTRAL T ALGORITHM FAILS FOR LARGE SAMPLE SIZE OR EFFECT  
SIZE.  
COMMENT THE FIRST PART USES THE NONCENTRAL T CALCULATOR IN SPSS.  
COMMENT THIS COMPUTES THE LOWER LIMIT ON THE T STATISTIC.  
COMPUTE #LC3 = TVAL .  
COMPUTE LC2 = TVAL/2 .  
COMPUTE #LC1 = -TVAL .  
COMPUTE #CUMF1 = NCDF.T(TVAL,DF,#LC1) .  
COMPUTE #ULIM = 1-(1-CONF)/2 .  
LOOP IF (#CUMF1 LT #ULIM) .  
+ COMPUTE LC2 = #LC1 .  
+ COMPUTE #LC1 = #LC1 TVAL .  
+ COMPUTE #CUMF1 = NCDF.T(TVAL,DF,#LC1) .  
END LOOP .  
COMPUTE #CUMF2 = NCDF.T(TVAL,DF,LC2) .  
COMPUTE #DIFF = 1 .  
LOOP IF (#DIFF GT .00005)  
+ DO IF (#CUMF2 LT #ULIM) .  
+ COMPUTE #LC3 = LC2 .  
+ COMPUTE LC2 = (LC2 + #LC1)/2 .  
+ COMPUTE #CUMF2 = NCDF.T(TVAL,DF,LC2) .  
+ ELSE .  
+ COMPUTE #LC1 = LC2 .  
+ COMPUTE LC2 = (LC2 + #LC3)/2 .  
+ COMPUTE #CUMF2 = NCDF.T(TVAL,DF,LC2) .  
+ END IF .  
+ COMPUTE #DIFF = ABS(#CUMF2 - #ULIM) .  
END LOOP .  
COMPUTE UCDF = NCDF.T(TVAL,DF,LC2) .  
EXECUTE .
```



```

COMMENT
COMMENT      THIS COMPUTES THE UPPER LIMIT ON THE T STATISTIC.
COMPUTE #UC3 = 2*TVAL .
COMPUTE UC2 = 1.5*TVAL .
COMPUTE #UC1 = TVAL .
COMPUTE #LLIM = (1-CONF)/2 .
COMPUTE #CUMF3 = NCDF.T(TVAL,DF,#UC3)
LOOP IF (#CUMF3 GT #LLIM) .
+   COMPUTE UC2 = #UC3
+   COMPUTE #UC3 = #UC3 + TVAL .
+   COMPUTE #CUMF3 = NCDF.T(TVAL,DF,#UC3) .
END LOOP
COMPUTE #CUMF2 = NCDF.T(TVAL,DF,UC2) .
COMPUTE #DIFF = 1 .
LOOP IF (#DIFF GT .00001) .
+   DO IF (#CUMF2 LT #LLIM) .
+       COMPUTE #UC3 = UC2 .
+       COMPUTE UC2 = (UC2 + #UC1)/2 .
+       COMPUTE #CUMF2 = NCDF.T(TVAL,DF,UC2) .
+   ELSE
+       COMPUTE #UC1 = UC2 .
+       COMPUTE UC2 = (UC2 + #UC3)/2 .
+       COMPUTE #CUMF2 = NCDF.T(TVAL,DF,UC2) .
+   END IF .
+   COMPUTE #DIFF = ABS(#CUMF2 - #LLIM) .
END LOOP .
COMPUTE LCDF = NCDF.T(TVAL,DF,UC2) .
COMMENT
COMMENT      THIS NEXT STATEMENT COMPUTES THE POWER IN RELATION TO THE T
VALUE.
COMPUTE POWER = 1 - NCDF.T(IDF.T(1-(1-CONF)/2,DF),DF,TVAL) .
EXECUTE .
COMMENT
COMMENT      THE SECOND PART USES LAUBSCHER'S SQUARE-ROOT APPROXIMATION.
COMMENT      THIS COMPUTES THE LOWER LIMIT ON THE F NONCENTRALITY
PARAMETER.
COMPUTE #LLC3 = TVAL**2 .
COMPUTE #LLC2 = TVAL**2/2 .
COMPUTE #LLC1 = .001
COMPUTE #ULIM = 1-(1-CONF)/2 .
COMPUTE #CUMF1 = 1-CDFNORM((Sqrt(2*(1+#LLC1)-((1+2*#LLC1)/(1+#LLC1)))-Sqrt((2*DF-
1)*TVAL**2*1/DF))/
  Sqrt(1*TVAL**2/DF+((1+2*#LLC1)/(1+#LLC1))))
LOOP IF (#CUMF1 LT #ULIM) .
+   COMPUTE #LLC2 = #LLC1 .
+   COMPUTE #LLC1 = #LLC1/4 .
+   COMPUTE #CUMF1 = 1-CDFNORM((Sqrt(2*(1+#LLC1)-((1+2*#LLC1)/(1+#LLC1)))-
Sqrt((2*DF-1)*TVAL**2*1/DF))/
  Sqrt(1*TVAL**2/DF+((1+2*#LLC1)/(1+#LLC1)))) .
END LOOP .
COMPUTE #CUMF3 = 1-CDFNORM((Sqrt(2*(1+#LLC3)-((1+2*#LLC3)/(1+#LLC3)))-Sqrt((2*DF-
1)*TVAL**2*1/DF))/
  Sqrt(1*TVAL**2/DF+((1+2*#LLC3)/(1+#LLC3)))) .
LOOP IF (#CUMF3 GT #ULIM) .
+   COMPUTE #LLC2 = #LLC3 .
+   COMPUTE #LLC3 = #LLC3 + TVAL**2 .

```

```

+      COMPUTE #CUMF3 = 1-CDFNORM((Sqrt(2*(1+#LLC3)-((1+2*#LLC3)/(1+#LLC3)))-
Sqrt((2*DF-1)*TVAL**2*(1/DF))/
  Sqrt(1*TVAL**2/DF+((1+2*#LLC3)/(1+#LLC3)))) .
END LOOP .
COMPUTE #CUMF2 = 1-CDFNORM((Sqrt(2*(1+#LLC2)-((1+2*#LLC2)/(1+#LLC2)))-Sqrt((2*DF-
1)*TVAL**2*(1/DF))/
  Sqrt(1*TVAL**2/DF+((1+2*#LLC2)/(1+#LLC2)))) .
COMPUTE #DIFF = 1 .
LOOP IF (#DIFF GT .00005) .
+      DO IF (#CUMF2 LT #ULIM) .
+          COMPUTE #LLC3 = #LLC2 .
+          COMPUTE #LLC2 = (#LLC2 + #LLC1)/2 .
+          COMPUTE #CUMF2 = 1-CDFNORM((Sqrt(2*(1+#LLC2)-
((1+2*#LLC2)/(1+#LLC2)))-Sqrt((2*DF-1)*TVAL**2*(1/DF))/
  Sqrt(1*TVAL**2/DF+((1+2*#LLC2)/(1+#LLC2)))) .
+      ELSE .
+          COMPUTE #LLC1 = #LLC2 .
+          COMPUTE #LLC2 = (#LLC2 + #LLC3)/2 .
+          COMPUTE #CUMF2 = 1-CDFNORM((Sqrt(2*(1+#LLC2)-
((1+2*#LLC2)/(1+#LLC2)))-Sqrt((2*DF-1)*TVAL**2*(1/DF))/
  Sqrt(1*TVAL**2/DF+((1+2*#LLC2)/(1+#LLC2)))) .
+      END IF .
+      COMPUTE #DIFF = ABS(#CUMF2 - #ULIM) .
END LOOP .
COMPUTE #UUCDF = 1-CDFNORM((Sqrt(2*(1+#LLC2)-((1+2*#LLC2)/(1+#LLC2)))-Sqrt((2*DF-
1)*TVAL**2*(1/DF))/
  Sqrt(1*TVAL**2/DF+((1+2*#LLC2)/(1+#LLC2))))
COMMENT
COMMENT      THIS COMPUTES THE UPPER LIMIT ON THE T NONCENTRALITY
PARAMETER.
COMPUTE #UUC3 = 3*TVAL**2 .
COMPUTE #UUC2 = 2*TVAL**2 .
COMPUTE #UUC1 = TVAL**2 .
COMPUTE #LLIM = (1-CONF)/2 .
COMPUTE #CUMF1 = 1-CDFNORM((Sqrt(2*(1+#UUC1)-((1+2*#UUC1)/(1+#UUC1)))-Sqrt((2*DF-
1)*TVAL**2*(1/DF))/
  Sqrt(1*TVAL**2/DF+((1+2*#UUC1)/(1+#UUC1)))) .
LOOP IF (#CUMF1 LT #LLIM) .
+      COMPUTE #UUC2 = #UUC1 .
+      COMPUTE #UUC1 = #UUC1/4 .
+      COMPUTE #CUMF1 = 1-CDFNORM((Sqrt(2*(1+#UUC1)-((1+2*#UUC1)/(1+#UUC1)))-
Sqrt((2*DF-1)*TVAL**2*(1/DF))/
  Sqrt(1*TVAL**2/DF+((1+2*#UUC1)/(1+#UUC1)))) .
END LOOP .
COMPUTE #CUMF3 = 1-CDFNORM((Sqrt(2*(1+#UUC3)-((1+2*#UUC3)/(1+#UUC3)))-Sqrt((2*DF-
1)*TVAL**2*(1/DF))/
  Sqrt(1*TVAL**2/DF+((1+2*#UUC3)/(1+#UUC3)))) .
LOOP IF (#CUMF3 GT #LLIM) .
+      COMPUTE #UUC2 = #UUC3
+      COMPUTE #UUC3 = #UUC3 + TVAL**2 .
+      COMPUTE #CUMF3 = 1-CDFNORM((Sqrt(2*(1+#UUC3)-((1+2*#UUC3)/(1+#UUC3)))-
Sqrt((2*DF-1)*TVAL**2*(1/DF))/
  Sqrt(1*TVAL**2/DF+((1+2*#UUC3)/(1+#UUC3)))) .
END LOOP .
COMPUTE #CUMF2 = 1-CDFNORM((Sqrt(2*(1+#UUC2)-((1+2*#UUC2)/(1+#UUC2)))-Sqrt((2*DF-
1)*TVAL**2*(1/DF))/

```

```

    Sqrt(1*TVAL**2/DF+((1+2*#UUC2)/(1+#UUC2)))) .
COMPUTE #DIFF = 1
LOOP IF (#DIFF GT .00001) .
+   DO IF (#CUMF2 LT #LLIM) .
+       COMPUTE #UUC3 = #UUC2 .
+       COMPUTE #UUC2 = (#UUC2 + #UUC1)/2 .
+       COMPUTE #CUMF2 = 1-CDFNORM((Sqrt(2*(1+#UUC2)-
((1+2*#UUC2)/(1+#UUC2)))-Sqrt((2*DF-1)*TVAL**2*1/DF))/
    Sqrt(1*TVAL**2/DF+((1+2*#UUC2)/(1+#UUC2)))) .
+   ELSE .
+       COMPUTE #UUC1 = #UUC2 .
+       COMPUTE #UUC2 = (#UUC2 + #UUC3)/2 .
+       COMPUTE #CUMF2 = 1-CDFNORM((Sqrt(2*(1+#UUC2)-
((1+2*#UUC2)/(1+#UUC2)))-Sqrt((2*DF-1)*TVAL**2*1/DF))/
    Sqrt(1*TVAL**2/DF+((1+2*#UUC2)/(1+#UUC2)))) .
+   END IF .
+   COMPUTE #DIFF = ABS(#CUMF2 - #LLIM) .
END LOOP .
COMPUTE #LLCDF = 1-CDFNORM((Sqrt(2*(1+#UUC2)-((1+2*#UUC2)/(1+#UUC2)))-Sqrt((2*DF-
1)*TVAL**2*1/DF))/
    Sqrt(1*TVAL**2/DF+((1+2*#UUC2)/(1+#UUC2)))) .
COMMENT
COMMENT     THIS NEXT STATEMENT COMPUTES THE POWER IN RELATION TO THE F
VALUE.
COMPUTE #PPOWER = CDFNORM((Sqrt(2*(1+TVAL**2*(1/DF)*(1+DF+1))-
((1+2*TVAL**2*(1/DF)*(1+DF+1))/(1+TVAL**2*(1/DF)*(1+DF+1)))-Sqrt((2*DF-1)*IDF.F(1-(1-
CONF)/2,1,DF)*1/DF))/
    Sqrt(1*IDF.F(1-(1-
CONF)/2,1,DF)/DF+((1+2*TVAL**2*(1/DF)*(1+DF+1))/(1+TVAL**2*(1/DF)*(1+DF+1)))) .
COMMENT
COMMENT     NOW CHOOSE THE METHOD TO USE FOR THE FINAL ESTIMATES.
COMMENT     THIS DECISION IS BASED ON THE OBSERVATION THAT SPSS NONCENTRAL
T
COMMENT     PRODUCES UPPER BOUND ESTIMATES THAT FALL BELOW ACCURATE
VALUES UNTIL
COMMENT     ARE SIMPLY EQUAL TO THE TVAL THAT IS GIVEN AS THE STARTING
POINT.
COMMENT     THE NORMAL APPROXIMATION, ON THE OTHER HAND, TENDS TO
UNDERESTIMATE
COMMENT     THE TRUE NONCENTRALITY PARAMETER FOR SMALL N AND/OR SMALL
EFFECTS.
COMMENT     NONCENTRAL T ALSO PRODUCES A LOWER BOUND THAT HITS A CEILING
OF APPROXIMATELY
COMMENT     17.5. I HAVE ALLOWED A SMALL MARGIN BELOW THAT IN CASE THE
ROUTINE DECLINES
COMMENT     IN ACCURACY IN THE NEIGHBORHOOD OF THIS CEILING.
DO IF ((SQRT(#UUC2) > UC2 OR SQRT(#LLC2) > 16.5) AND LC2 > 0) .
+   COMPUTE     LC2 = SQRT(#LLC2) .
+   COMPUTE     UCDF = #UUCDF .
END IF
DO IF (SQRT(#UUC2) > UC2 OR (SQRT(#LLC2) > 16.5 AND LC2 > 0)) .
+   COMPUTE     POWER = #PPOWER .
END IF .
DO IF (SQRT(#UUC2) > UC2 OR UC2 <= TVAL)
+   COMPUTE     UC2 = SQRT(#UUC2)
+   COMPUTE     LCDF = #LLCDF .

```

END IF .
EXECUTE .

COMMENT THESE COMPUTATIONS ARE BASED ON THE HUNTER & SCHMIDT EQUATIONS AND ARE FUNCTIONALLY EQUIVALENT TO THE COMPUTATIONS OFFERED BY SMITHSON; THE PROCESS IS ONE OF CONVERTING T VALUES INTO D VALUES; SEE P 272 OF H&S.

COMMENT BEGIN BY COMPUTING THE CENTRAL VALUES FOR D (NOTE THAT 'D' IS EQUIVALENT TO 'G' IN THIS PROGRAM).

COMMENT HERE ARE THE UNADJUSTED D VALUES.

COMPUTE #totn = 2*n.

COMMENT THIS COMPUTATION COMPUTES THE SE OF HEDGE'S G AND SUBSEQUENTLY BUILDS A CENTRAL 95% CI AROUND THAT VALUE.

COMMENT THIS EQUATION COMES FROM HUNTER & SCHMIDT (1991) (P 281) AND IS BASED ON THE "SMALL SAMPLE" EQUATION.

COMMENT NOTE THAT HUNTER & SCHMIDT CALL HEDGE'S G -- 'd'.

COMPUTE D = 2*tval/SQRT(#totn).

COMPUTE sehedg = SQRT(((#totn-1)/(#totn-3)) * ((4/#totn)* (1+(D**2/8)))).

COMMENT COMPUTE UB AND LB VALUES BASED ON TCRT AND SE VALUES.

COMMENT COMPUTE THE TCRT VALUE BASED ON DF AND CONFIDENCE

LEVEL.

COMPUTE tcrit = ABS(IDF.T(((1-conf)/2),(df))) .

COMPUTE CD = D .

COMPUTE CDUB = D + tcrit*sehedg.

COMPUTE CDLB = D - tcrit*sehedg.

COMMENT NOW COMPUTE THE NON-CENTRAL D VALUES BY CONVERTING THE LB AND UB VALUES OF T INTO D UNITS.

COMPUTE NCD = D .

COMPUTE NCDLB = 2*LC2/SQRT(#totn) .

COMPUTE NCDUB = 2*UC2/SQRT(#totn) .

COMMENT HUNTER & SCHMIDT ENCOURAGE THE USE OF CORRECTED d VALUES BASED ON THE CORRECTION FACTOR a.

COMMENT THIS ROUTINE WILL ALSO COMPUTE CORRECTED VALUES OF HEDGE'S G AND SE ESTIMATES.

COMMENT COMPUTE ADJUSTMENT FACTOR 'a' AS DESCRIBE BY H&S.

COMPUTE #a = 1 + .75 / (#totn-3).

COMMENT ADJUST D AND SE VALUES USING CORRECTION FACTOR.

COMPUTE #ASE = sehedg / #a.

COMPUTE #AD = D / #a.

COMMENT COMPUTE THE ADJUSTED CENTRAL UPPER AND LOWER BOUNDS OF D.

COMPUTE ACD = #AD.

COMPUTE ACDLB = ACD - tcrit*#ASE.

COMPUTE ACDUB = ACD + tcrit*#ASE.

COMMENT THE FOLLOWING "TEST" PROVES THAT ADJUSTING CLB AND CUB BY DIVIDING BY 'a' IS EQUIVALENT TO BUILDING A NEW.

COMMENT INTERVAL BASED ON ACD AND ASE; RESULTS HAVE BEEN REMOVED FROM OUTPUT.

COMPUTE #TEST = CDLB/#A.

COMPUTE #TEST2 = ACDLB-#TEST.

COMMENT COMPUTE THE ADJUSTED NON-CENTRAL UPPER AND LOWER BOUNDS OF D.

COMPUTE ANCD = NCD/#a.

COMPUTE ANCDLB = NCDLB/#a.

COMPUTE ANCDUB = NCDUB/#a.

COMMENT THIS NEXT SECTION WILL COMPUTE GLASS' DELTA BY CONVERTING THE VARIOUS D VALUES INTO DELTA VALUES.

COMMENT THIS CONVERSION EQUATION IS PROVIDED IN H&S (P 277) AND IS BASED ON THE VARIANCE RATIO.

COMMENT WE WILL COMPUTE A CONVERSION VARIABLE WHICH WILL BE MULTIPLIED BY THE D VALUE.

COMPUTE #v = sigmat/sigmac.

COMPUTE #cor = SQRT((1/#v**2)/2).

COMMENT COMPUTE NON-ADJUSTED CENTRAL AND NON-CENTRAL GLASS' DELTA VALUES.

COMPUTE CG = CD/#COR.

COMPUTE CGLB = CDLB/#COR.

COMPUTE CGUB = CDUB/#COR.

COMPUTE NCG = NCD/#COR.

COMPUTE NCGLB = NCDLB/#COR.

COMPUTE NCGUB = NCDUB/#COR.

COMMENT COMPUTE ADJUSTED CENTRAL AND NON-CENTRAL GLASS' DELTA VALUES.

COMPUTE ACG = ACD/#COR.

COMPUTE ACGLB = ACDLB/#COR.

COMPUTE ACGUB = ACDUB/#COR.

COMPUTE ANCG = ANCD/#COR.

COMPUTE ANCGLB = ANCDLB/#COR.

COMPUTE ANCGUB = ANCDUB/#COR.

EXECUTE.

APPENDIX C

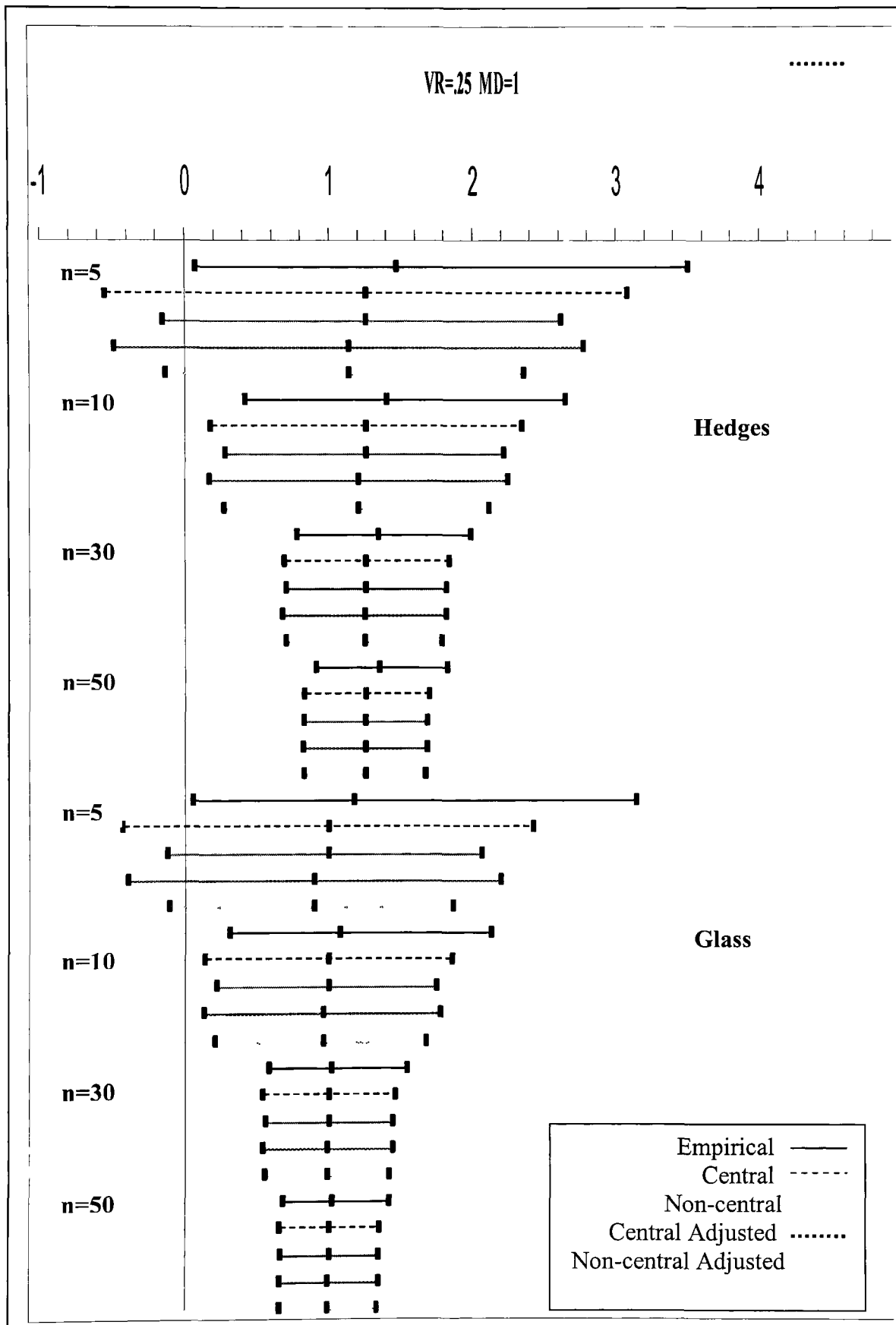


Figure 1. Confidence intervals for Hedges g and Glass' δ

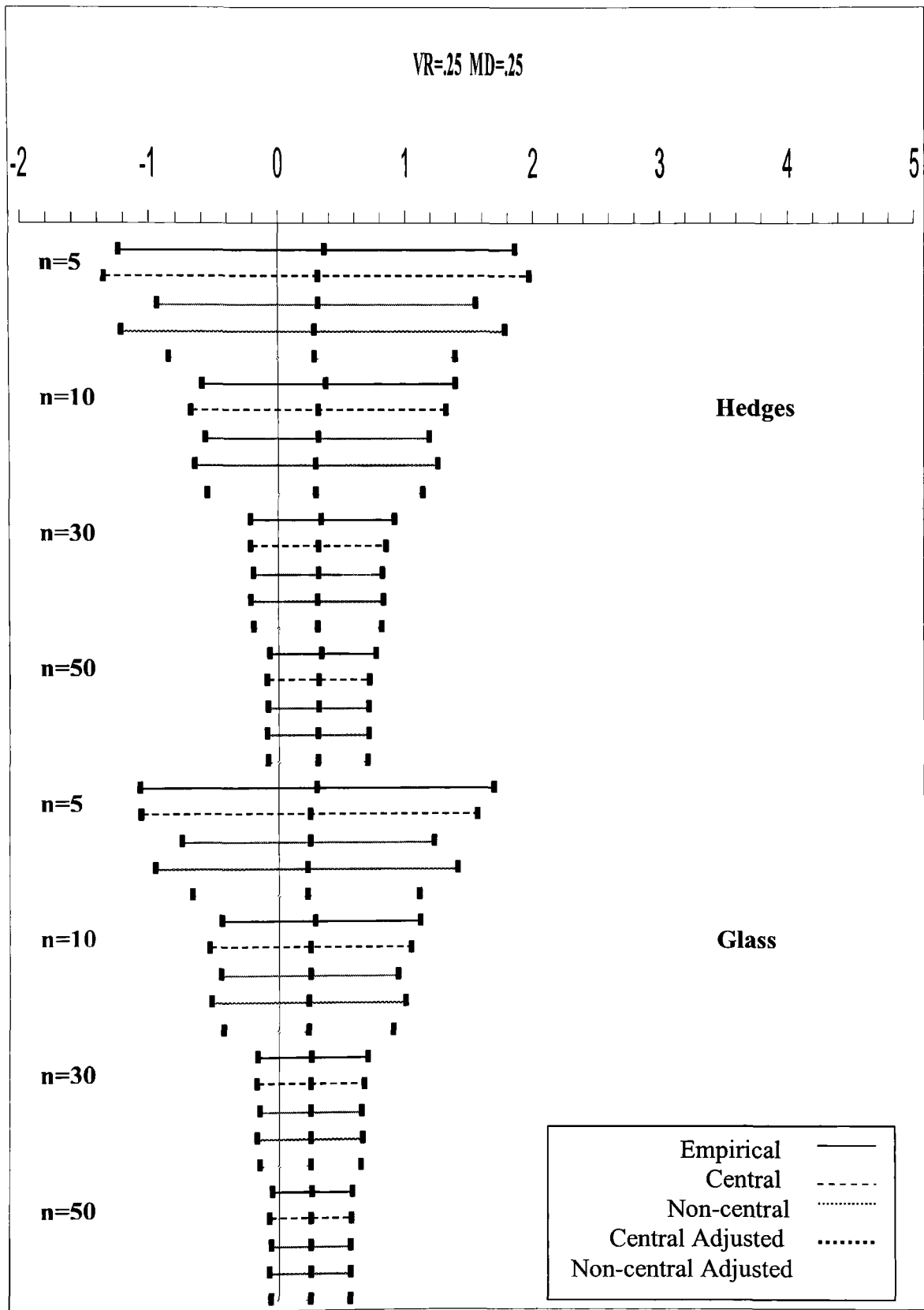


Figure 2. Confidence intervals for Hedges g and Glass' delta

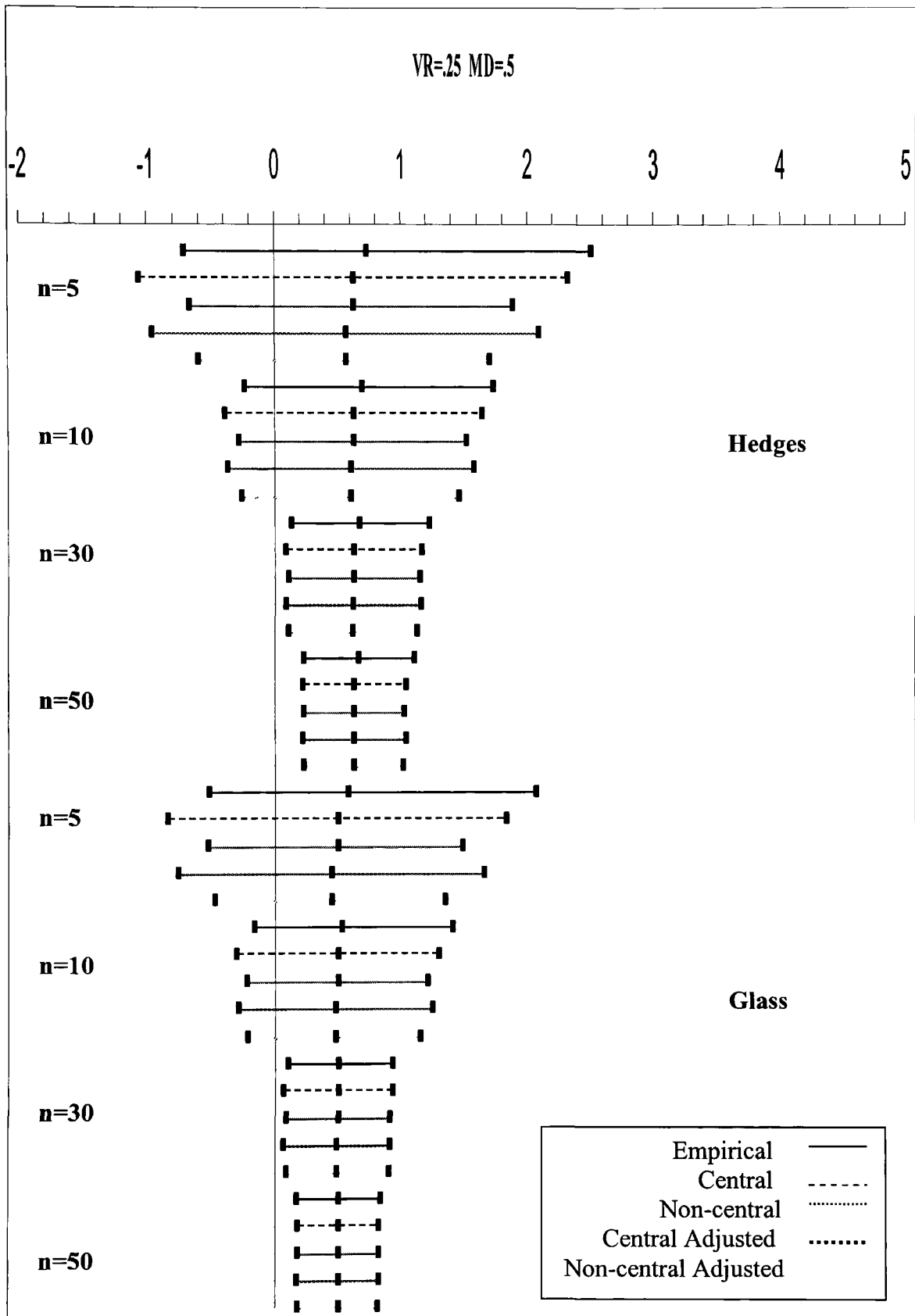


Figure 3. Confidence intervals for Hedges g and Glass' delta

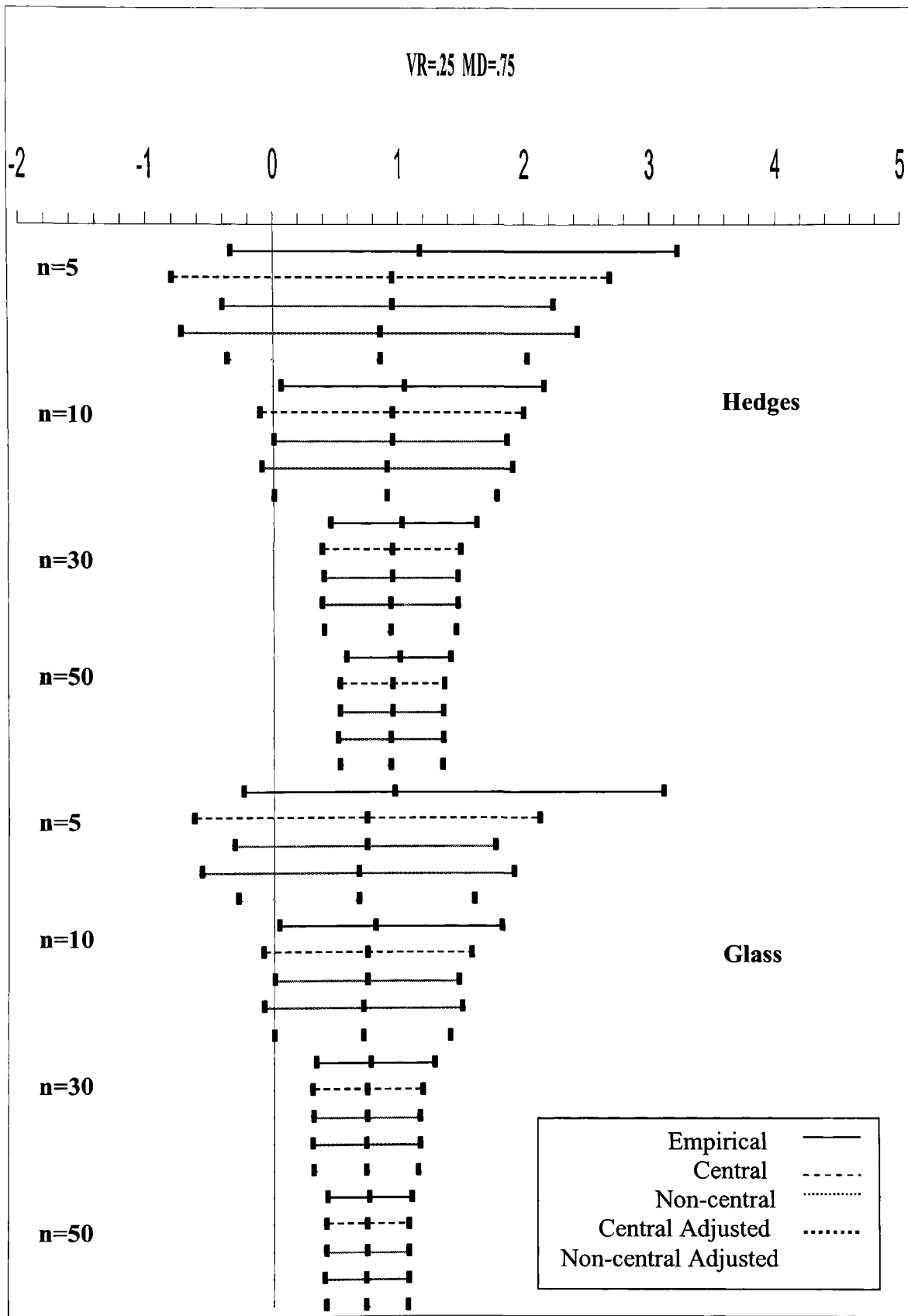


Figure 4. Confidence intervals for Hedges g and Glass' delta

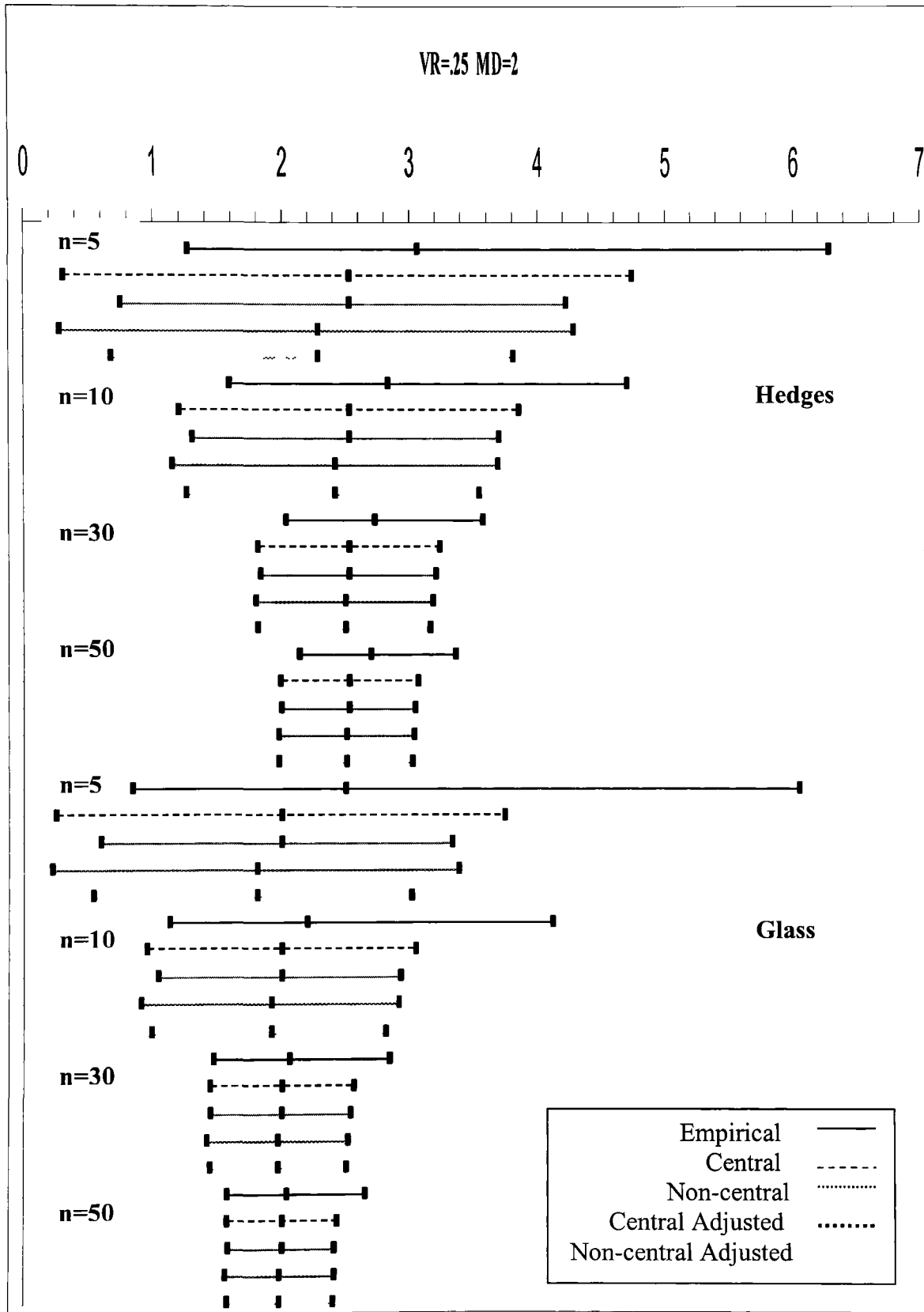


Figure 5. Confidence intervals for Hedges g and Glass' δ

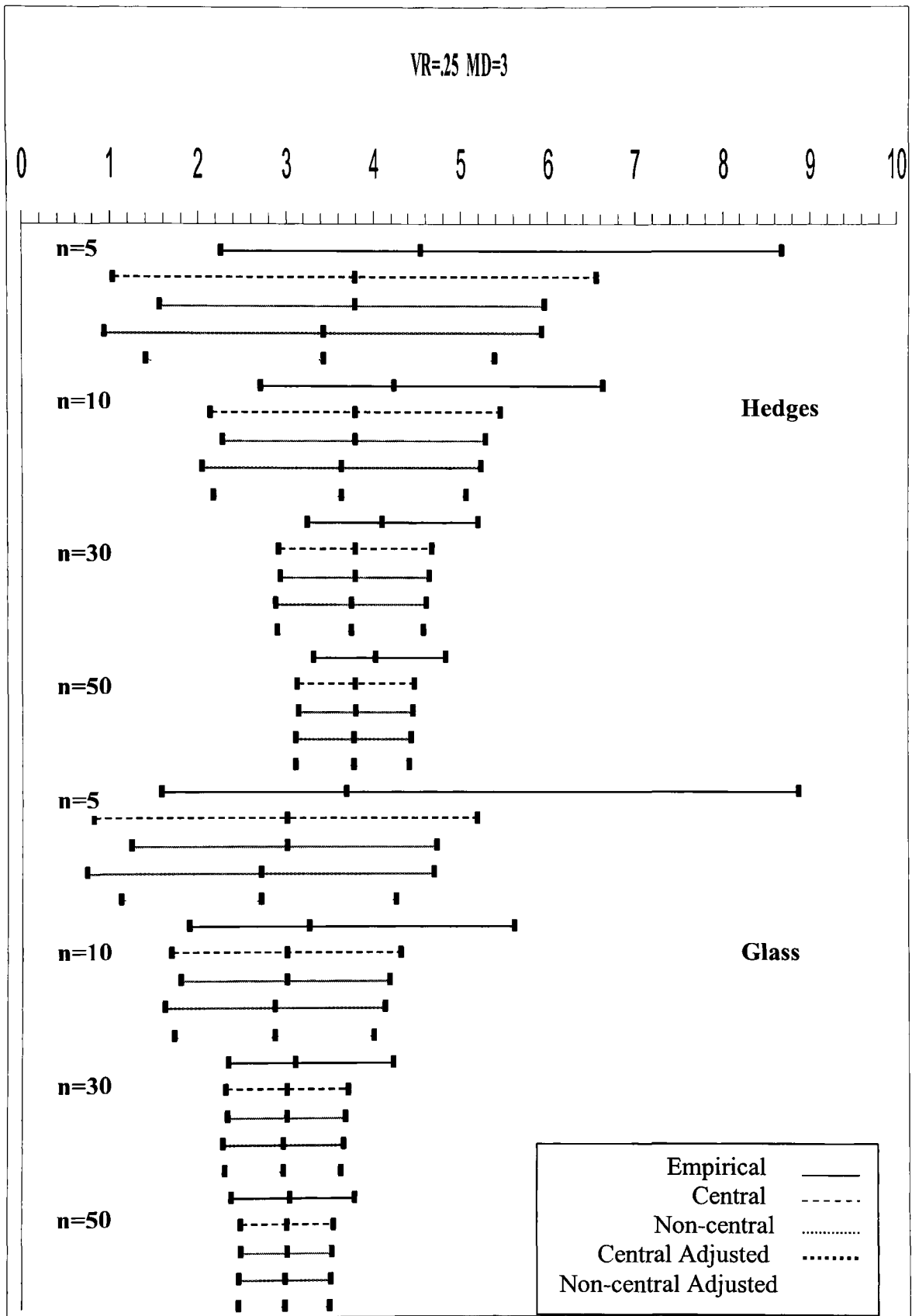


Figure 6. Confidence intervals for Hedges g and Glass' δ

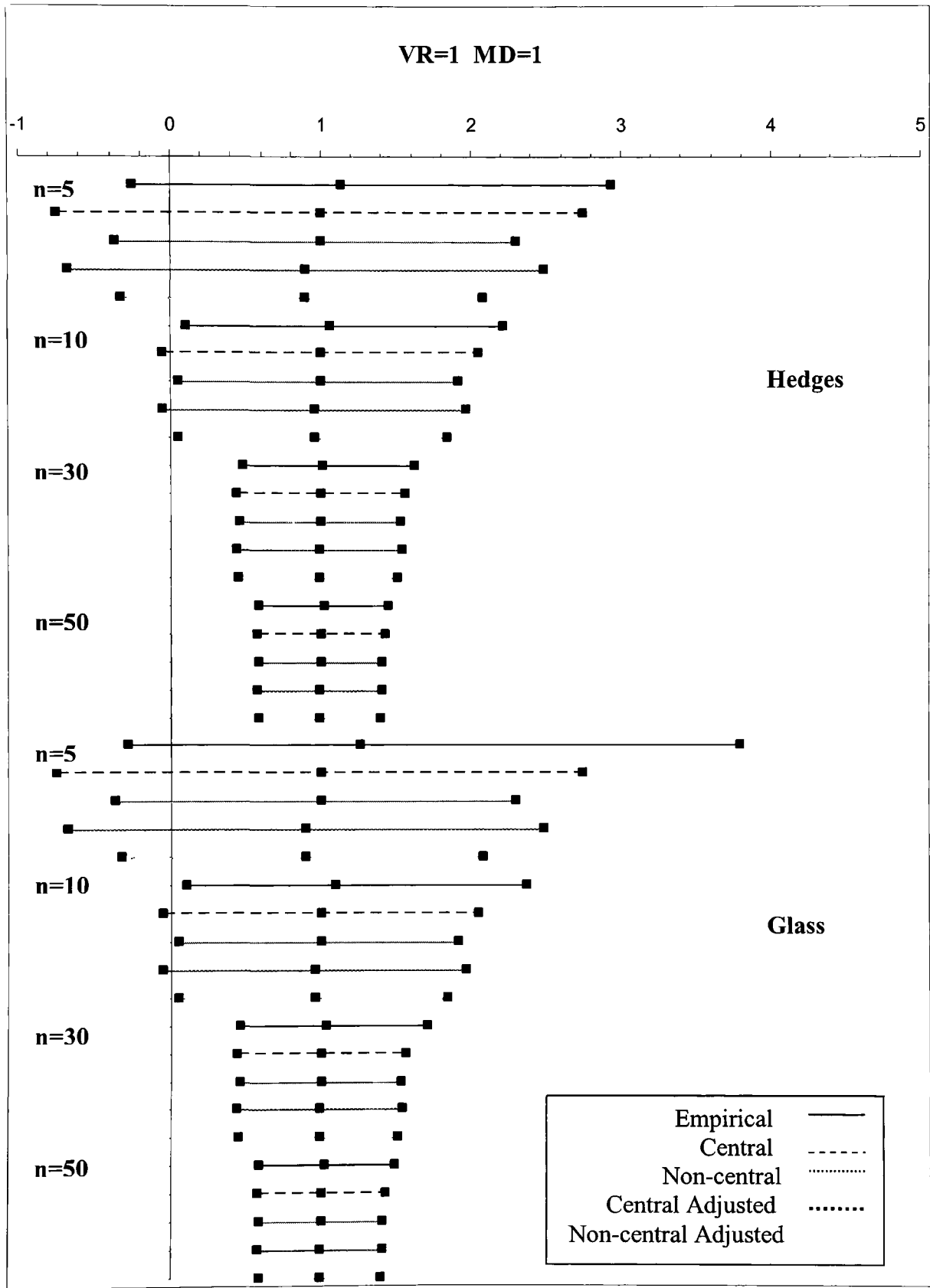


Figure 7. Confidence intervals for Hedges g and Glass' delta

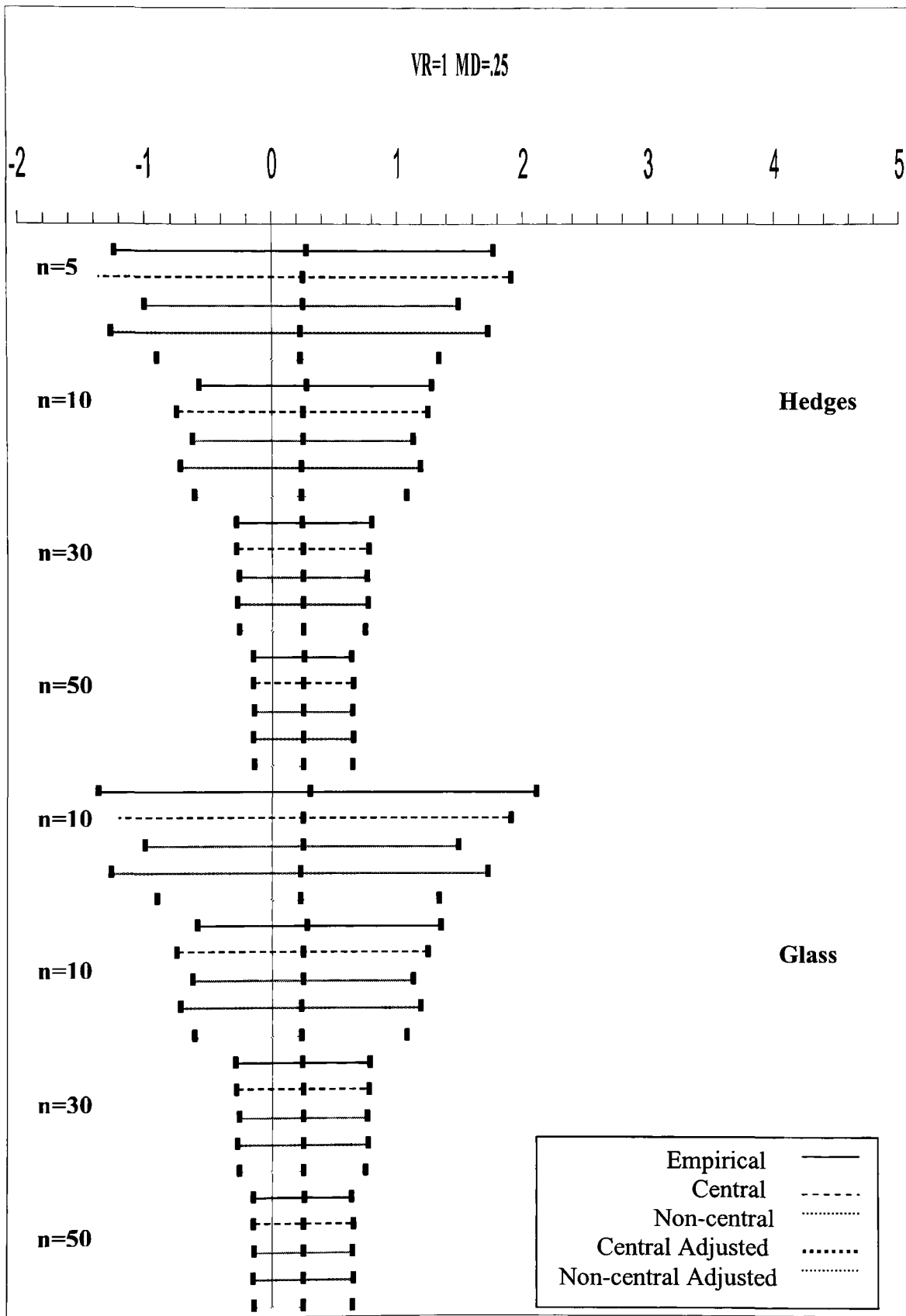


Figure 8. Confidence intervals for Hedges g and Glass' delta

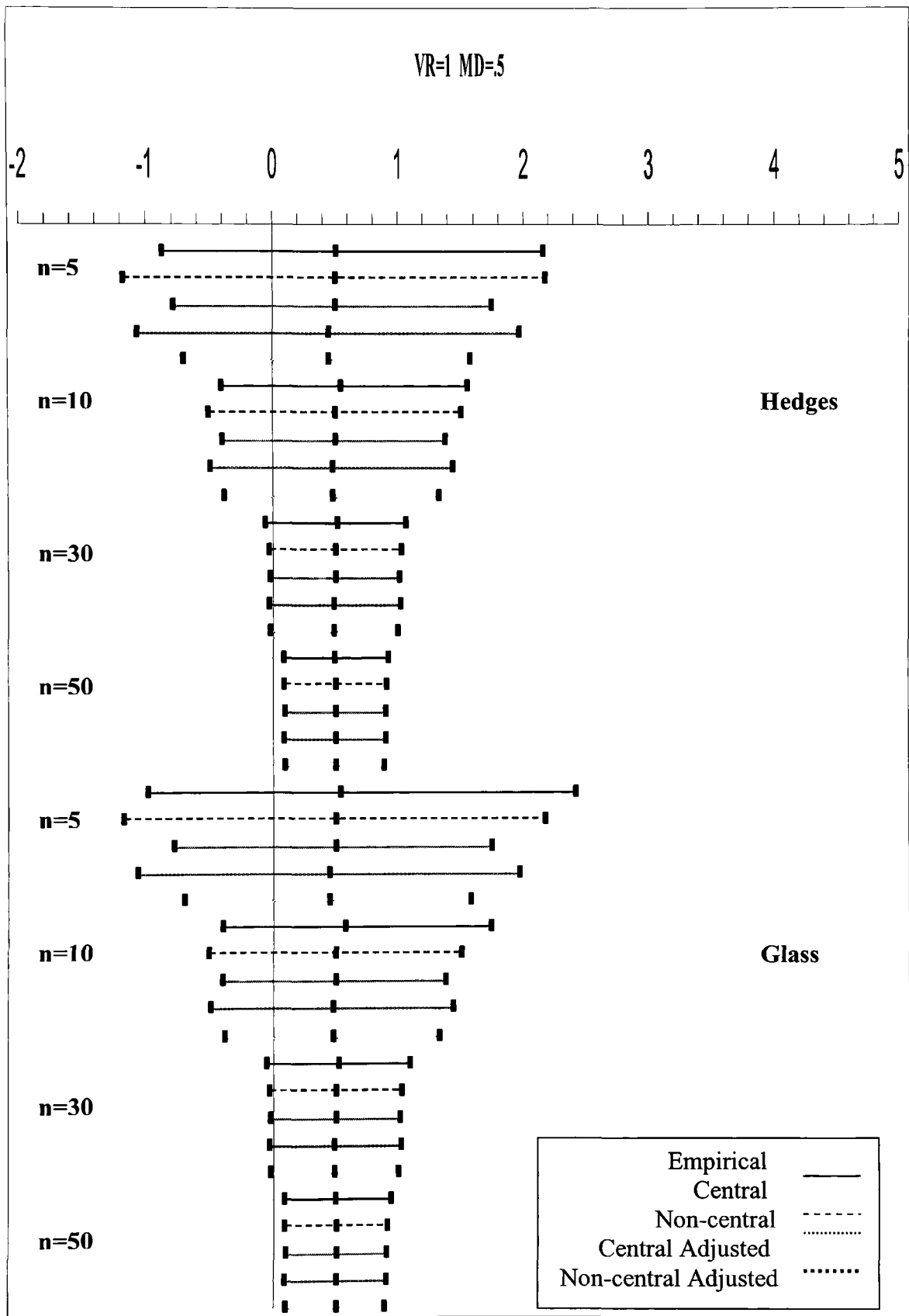


Figure 9. Confidence intervals for Hedges g and Glass' delta

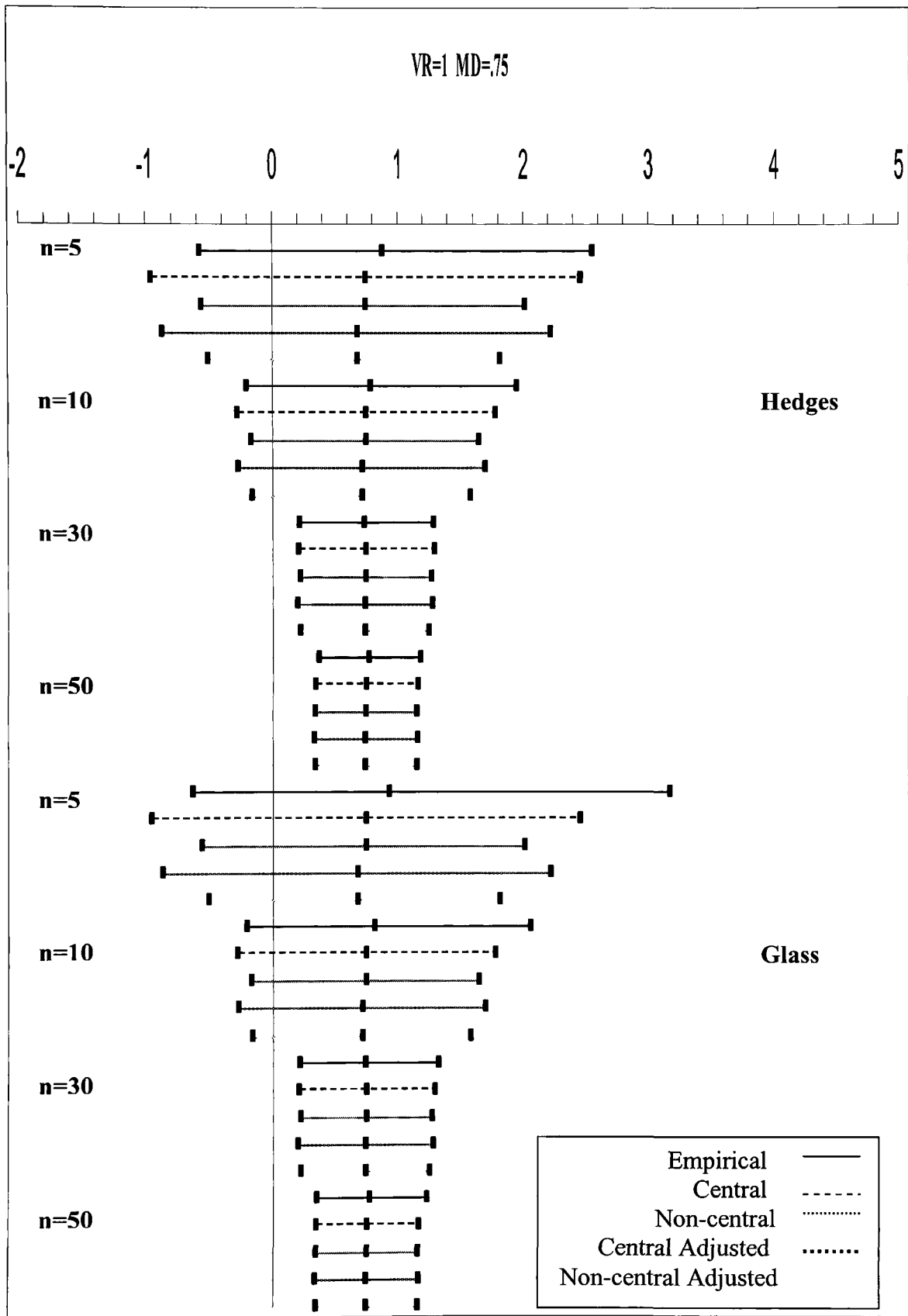


Figure 10. Confidence intervals for Hedges g and Glass' δ

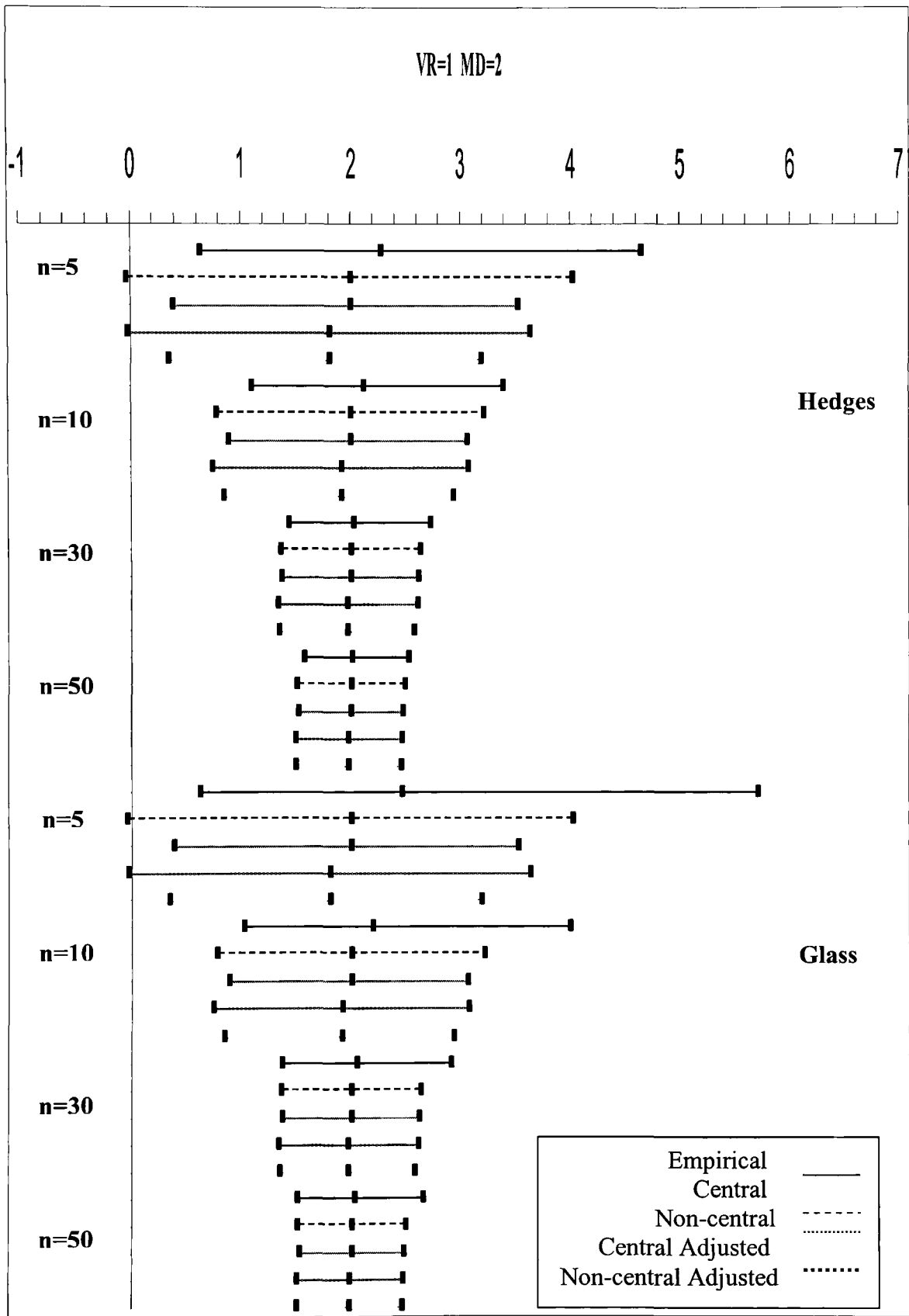


Figure 11. Confidence intervals for Hedges g and Glass' delta

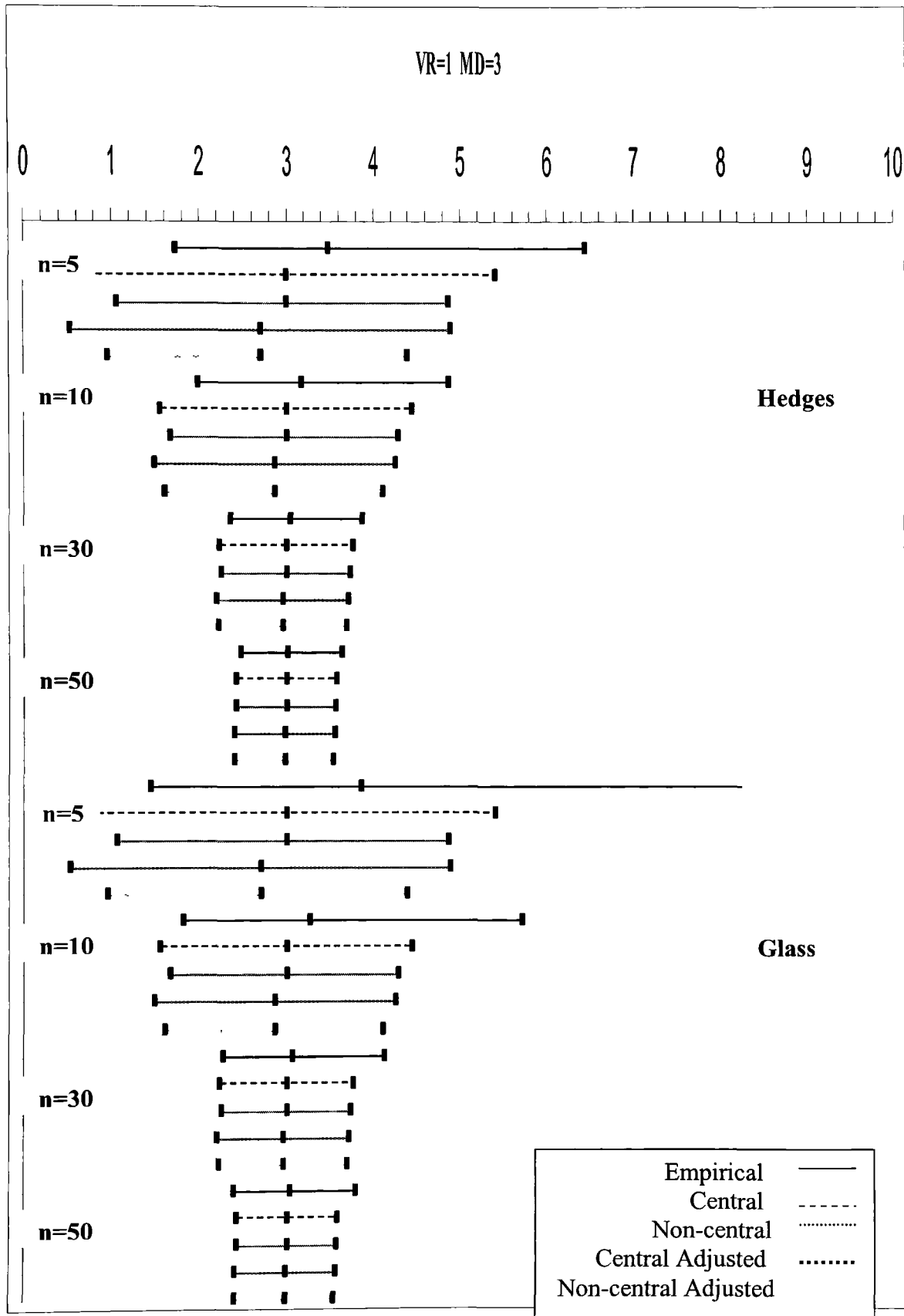


Figure 12. Confidence intervals for Hedges g and Glass' δ

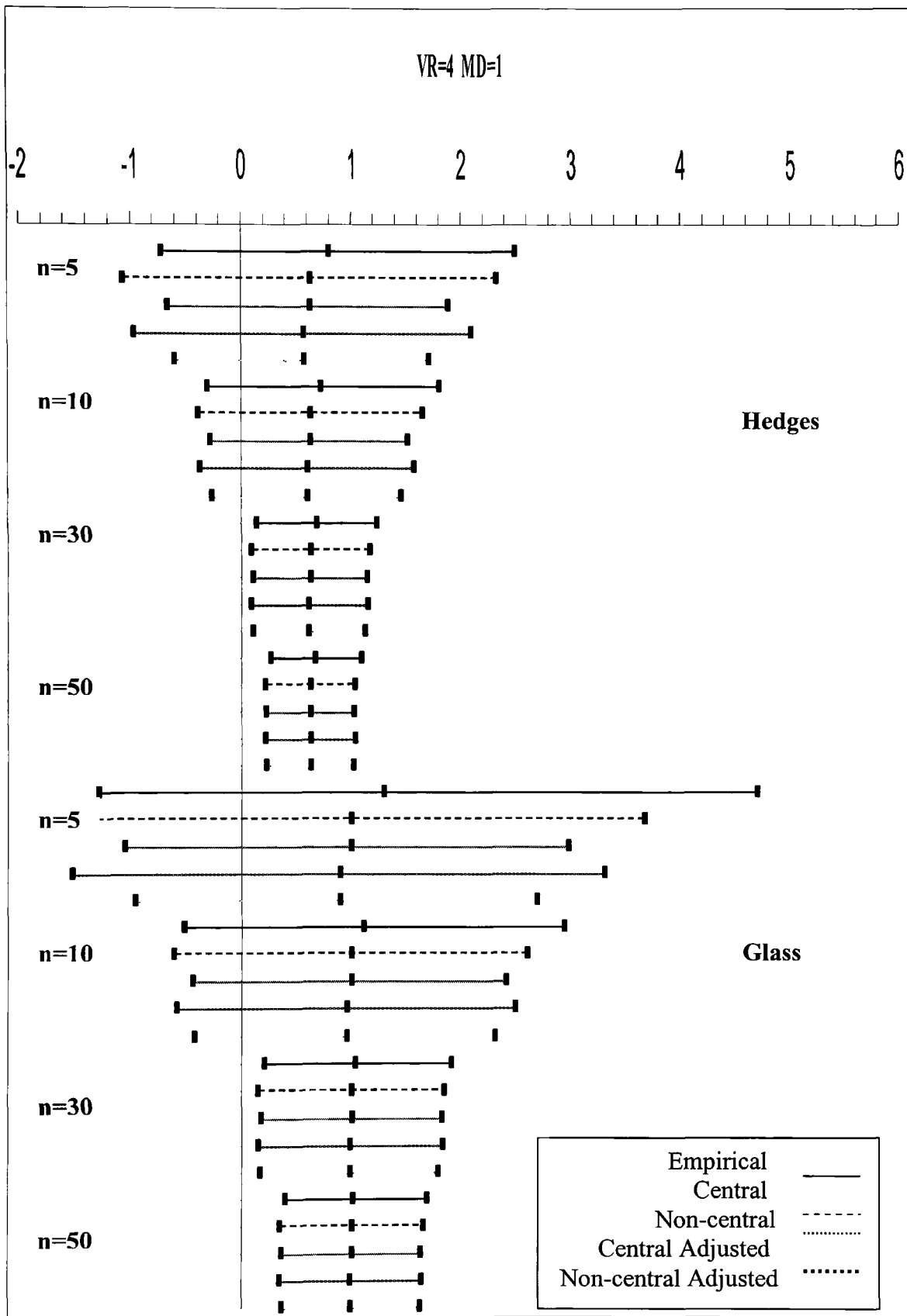


Figure 13. Confidence intervals for Hedges g and Glass' δ

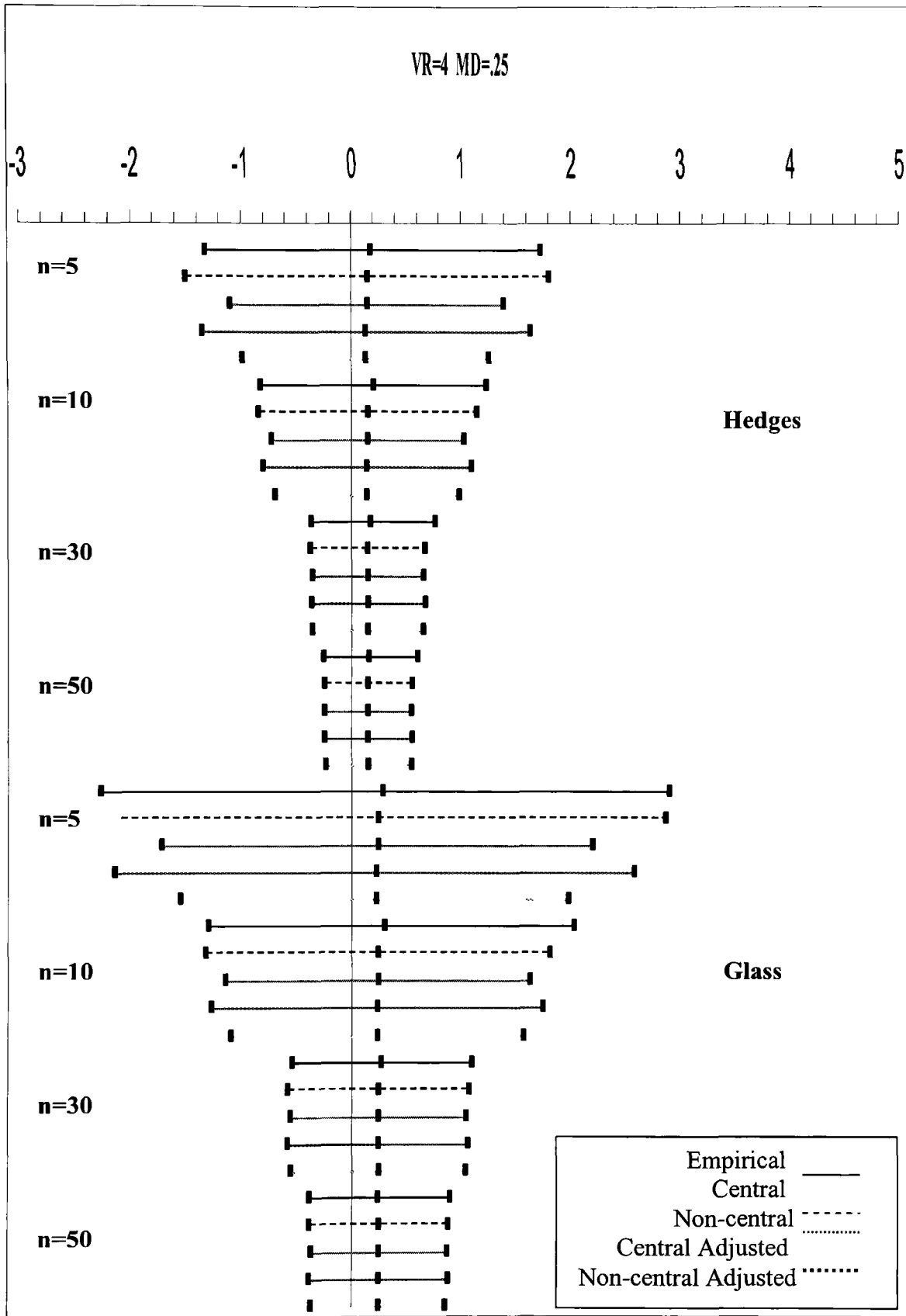


Figure 14. Confidence intervals for Hedges g and Glass' δ

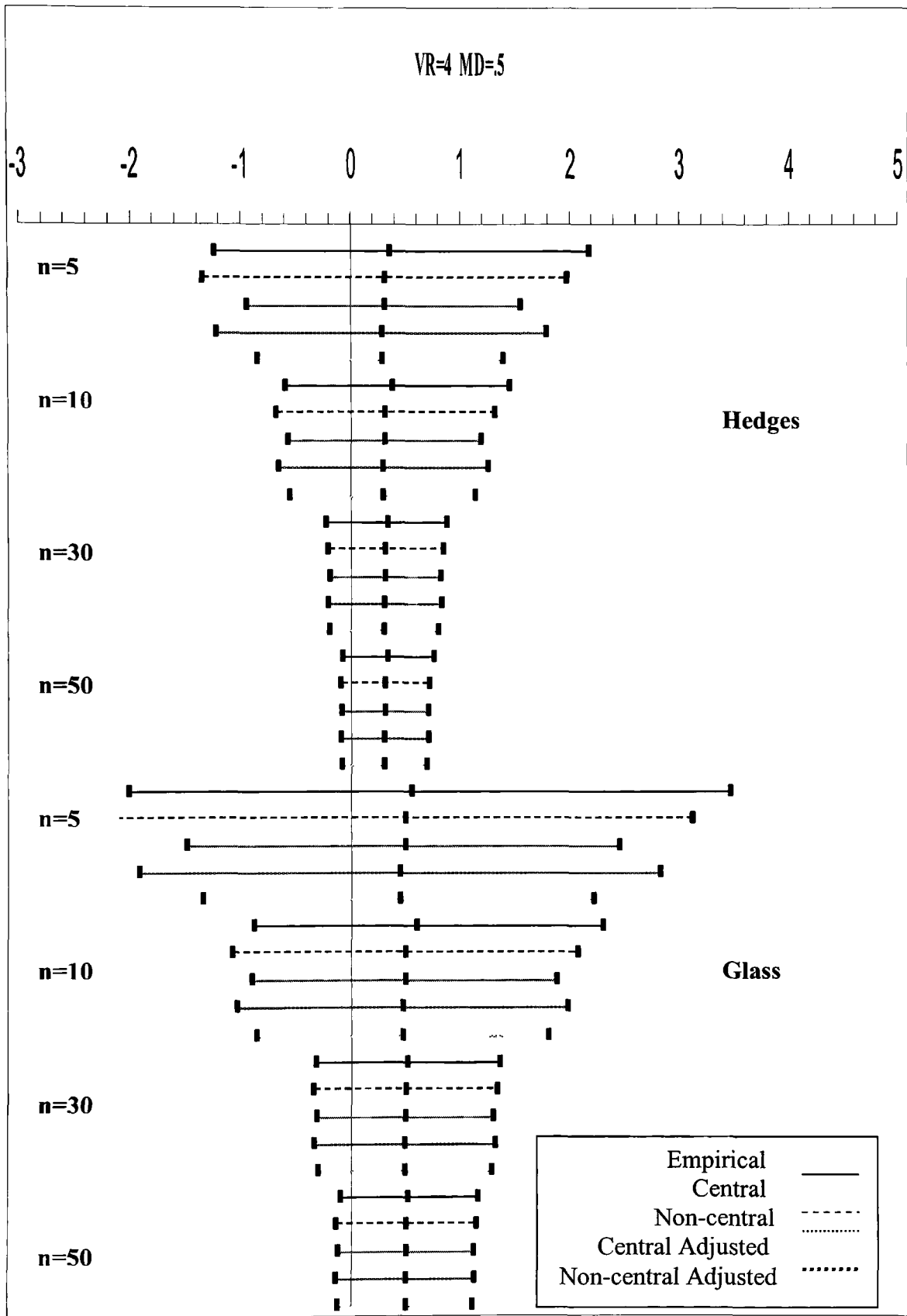


Figure 15. Confidence intervals for Hedges g and Glass' delta

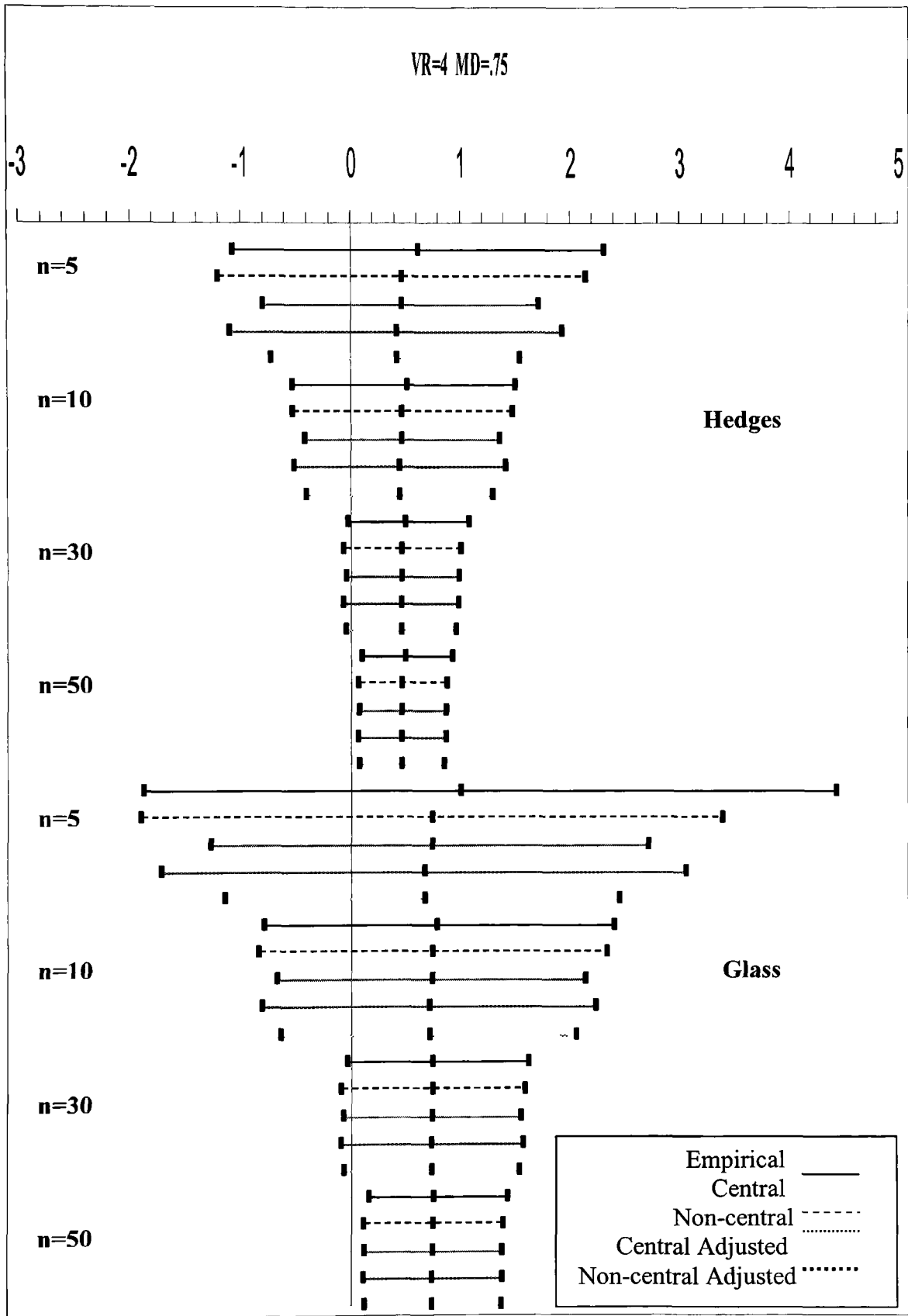


Figure 16. Confidence intervals for Hedges g and Glass' δ

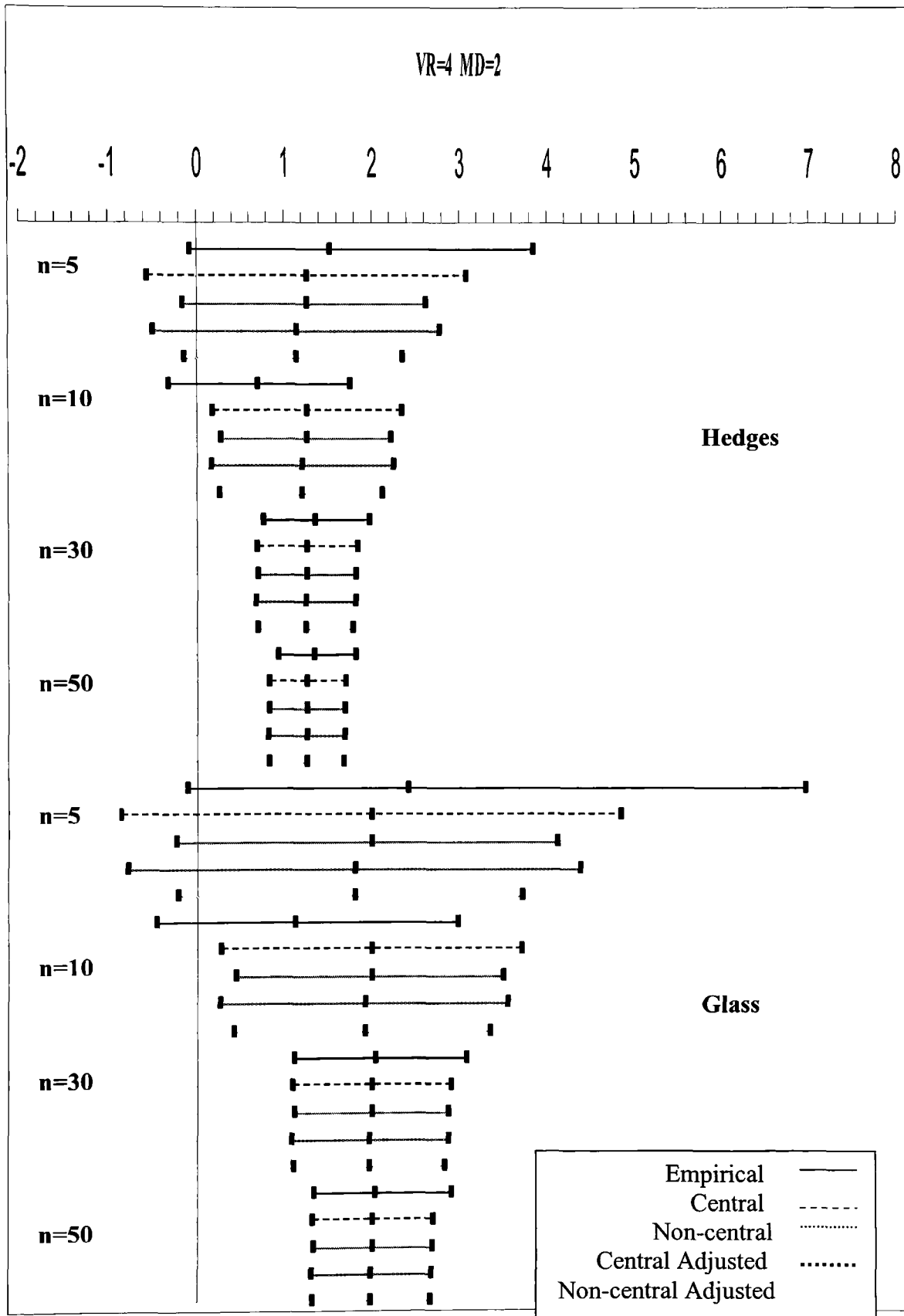


Figure 17. Confidence intervals for Hedges g and Glass' delta

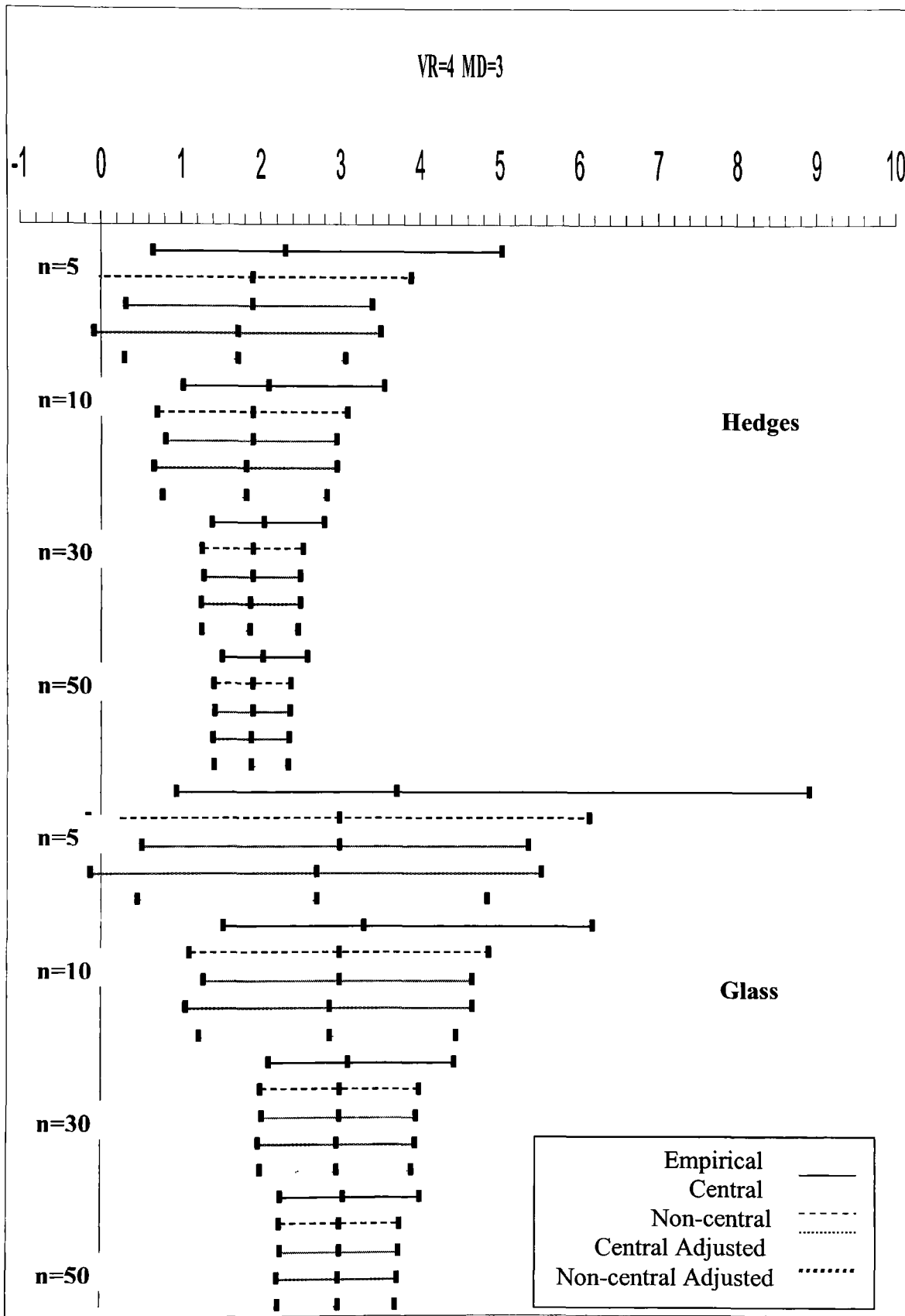


Figure 18. Confidence intervals for Hedges g and Glass' δ

APPENDIX D

Variance	Mean	n	Empirical g			Central g			Noncentral g			Adjusted Central g			Adjusted Noncentral g		
			Ratio	Difference	LB	C	UB	LB	C	UB	LB	C	UB	LB	C	UB	LB
1.00	1.00	5.00	-0.24515	1.13204	2.93809	-0.75	1.00	2.75	-0.36	1.00	2.30	-0.68	.90	2.49	-0.32	.90	2.08
1.00	1.00	10.00	0.10777	1.06393	2.22077	-0.05	1.00	2.05	.05	1.00	1.92	-0.05	.96	1.97	.05	.96	1.84
1.00	1.00	30.00	0.47852	1.01531	1.62251	.44	1.00	1.56	.46	1.00	1.53	.44	.99	1.54	.45	.99	1.51
1.00	1.00	50.00	0.58043	1.01615	1.44926	.57	1.00	1.43	.58	1.00	1.41	.57	.99	1.41	.58	.99	1.40
.25	1.00	5.00	0.07434	1.47629	3.50854	-0.55	1.26	3.08	-0.15	1.26	2.62	-0.49	1.14	2.78	-0.13	1.14	2.36
.25	1.00	10.00	0.4167	1.40417	2.65248	.18	1.26	2.35	.28	1.26	2.22	.17	1.21	2.25	.27	1.21	2.12
.25	1.00	30.00	0.78067	1.35035	1.99662	.69	1.26	1.84	.70	1.26	1.82	.68	1.25	1.82	.70	1.25	1.79
.25	1.00	50.00	0.91114	1.3549	1.83006	.83	1.26	1.70	.83	1.26	1.69	.82	1.26	1.69	.83	1.26	1.68
4.00	1.00	5.00	-0.71795	0.800352	2.49865	-1.06	.63	2.33	-0.66	.63	1.89	-0.96	.57	2.10	-0.60	.57	1.71
4.00	1.00	10.00	-0.3037	0.723015	1.806	-0.39	.63	1.65	-0.28	.63	1.52	-0.37	.61	1.58	-0.26	.61	1.46
4.00	1.00	30.00	0.13994	0.689446	1.23757	.09	.63	1.17	.11	.63	1.15	.09	.62	1.16	.11	.62	1.13
4.00	1.00	50.00	0.26776	0.671195	1.09616	.22	.63	1.04	.23	.63	1.03	.22	.63	1.04	.23	.63	1.02

Variance	Mean	n	Empirical delta			Central delta			Noncentral delta			Adj. Central delta			Adj. Noncentral delta		
			Ratio	Difference	LB	C	UB	LB	C	UB	LB	C	UB	LB	C	UB	LB
1.00	1.00	5.00	-0.282	1.25524	3.805	-0.75	1.00	2.75	-0.36	1.00	2.30	-0.68	.90	2.49	-0.32	.90	2.08
1.00	1.00	10.00	0.10727	1.09585	2.37542	-0.05	1.00	2.05	.05	1.00	1.92	-0.05	.96	1.97	.05	.96	1.84
1.00	1.00	30.00	0.45399	1.02737	1.71025	.44	1.00	1.56	.46	1.00	1.53	.44	.99	1.54	.45	.99	1.51
1.00	1.00	50.00	0.5885	1.02419	1.49385	.57	1.00	1.43	.58	1.00	1.41	.57	.99	1.41	.58	.99	1.40
.25	1.00	5.00	0.05521	1.17973	3.15457	-0.43	1.00	2.43	-0.12	1.00	2.07	-0.39	.90	2.20	-0.11	.90	1.87
.25	1.00	10.00	0.31109	1.07336	2.13269	.14	1.00	1.86	.22	1.00	1.75	.13	.96	1.78	.21	.96	1.68
.25	1.00	30.00	0.58321	1.02036	1.54678	.54	1.00	1.46	.56	1.00	1.44	.54	.99	1.44	.55	.99	1.42
.25	1.00	50.00	0.67467	1.0211	1.41632	.65	1.00	1.35	.66	1.00	1.34	.65	.99	1.34	.65	.99	1.33
4.00	1.00	5.00	-1.2793	1.2969	4.7067	-1.68	1.00	3.68	-1.05	1.00	2.99	-1.52	.90	3.32	-0.95	.90	2.70
4.00	1.00	10.00	-0.51506	1.1133	2.94809	-0.61	1.00	2.61	-0.44	1.00	2.41	-0.58	.96	2.50	-0.42	.96	2.31
4.00	1.00	30.00	0.21338	1.03817	1.91863	.15	1.00	1.85	.18	1.00	1.82	.15	.99	1.83	.17	.99	1.79
4.00	1.00	50.00	0.39594	1.01277	1.68442	.35	1.00	1.65	.36	1.00	1.63	.35	.99	1.64	.36	.99	1.62

Table 1
Empirically defined and computed confidence intervals

Variance Ratio	Mean Difference	n	Empirical g			Central g			Noncentral g			Adjusted Central g			Adjusted Noncentral g		
			LB	C	UB	LB	C	UB	LB	C	UB	LB	C	UB	LB	C	UB
1.00	.25	5.00	-1.24302	0.2852	1.7762	-1.41	.25	1.91	-1.00	.25	1.49	-1.27	.23	1.73	-.91	.23	1.34
1.00	.25	10.00	-0.57176	0.2789	1.27816	-.75	.25	1.25	-.63	.25	1.13	-.72	.24	1.19	-.61	.24	1.08
1.00	.25	30.00	-0.27664	0.24657	0.79653	-.28	.25	.78	-.26	.25	.76	-.27	.25	.77	-.26	.25	.75
1.00	.25	50.00	-0.14598	0.25698	0.63336	-.15	.25	.65	-.14	.25	.64	-.15	.25	.65	-.14	.25	.64
.25	.25	5.00	-1.22851	0.36829	1.87157	-1.35	.32	1.98	-.94	.32	1.56	-1.22	.29	1.79	-.85	.29	1.40
.25	.25	10.00	-0.58831	0.37808	1.39325	-.68	.32	1.32	-.57	.32	1.19	-.65	.30	1.26	-.55	.30	1.14
.25	.25	30.00	-0.21317	0.34073	0.91491	-.21	.32	.85	-.19	.32	.82	-.21	.31	.83	-.19	.31	.81
.25	.25	50.00	-0.06892	0.34168	0.77221	-.09	.32	.72	-.08	.32	.71	-.09	.31	.71	-.08	.31	.70
4.00	.25	5.00	-1.31737	0.18358	1.73336	-1.50	.16	1.81	-1.09	.16	1.40	-1.35	.14	1.64	-.98	.14	1.26
4.00	.25	10.00	-0.82067	0.2035	1.23688	-.84	.16	1.15	-.72	.16	1.03	-.80	.15	1.10	-.69	.15	.99
4.00	.25	30.00	-0.35869	0.18285	0.76966	-.37	.16	.68	-.35	.16	.66	-.36	.16	.68	-.35	.16	.66
4.00	.25	50.00	-0.24549	0.16037	0.6079	-.24	.16	.56	-.24	.16	.55	-.24	.16	.56	-.23	.16	.55

Variance Ratio	Mean Difference	n	Empirical delta			Central delta			Noncentral delta			Adj. Central delta			Adj. Noncentral delta		
			LB	C	UB	LB	C	UB	LB	C	UB	LB	C	UB	LB	C	UB
1.00	.25	5.00	-1.37512	0.30561	2.11119	-1.41	.25	1.91	-1.00	.25	1.49	-1.27	.23	1.73	-.91	.23	1.34
1.00	.25	10.00	-0.59115	0.28392	1.35371	-.75	.25	1.25	-.63	.25	1.13	-.72	.24	1.19	-.61	.24	1.08
1.00	.25	30.00	-0.28672	0.24711	0.78397	-.28	.25	.78	-.26	.25	.76	-.27	.25	.77	-.26	.25	.75
1.00	.25	50.00	-0.14823	0.25859	0.6328	-.15	.25	.65	-.14	.25	.64	-.15	.25	.65	-.14	.25	.64
.25	.25	5.00	-1.07432	0.30534	1.70064	-1.07	.25	1.57	-.75	.25	1.23	-.96	.23	1.41	-.67	.23	1.11
.25	.25	10.00	-0.44195	0.29254	1.11758	-.54	.25	1.04	-.45	.25	.94	-.52	.24	1.00	-.43	.24	.90
.25	.25	30.00	-0.15873	0.25843	0.70076	-.17	.25	.67	-.15	.25	.65	-.17	.25	.66	-.15	.25	.64
.25	.25	50.00	-0.05053	0.25718	0.57451	-.07	.25	.57	-.06	.25	.56	-.07	.25	.56	-.06	.25	.56
4.00	.25	5.00	-2.26945	0.29476	2.90346	-2.37	.25	2.87	-1.72	.25	2.21	-2.14	.23	2.59	-1.55	.23	1.99
4.00	.25	10.00	-1.29718	0.30993	2.03464	-1.32	.25	1.82	-1.14	.25	1.63	-1.27	.24	1.75	-1.09	.24	1.57
4.00	.25	30.00	-0.53913	0.27639	1.1006	-.58	.25	1.08	-.55	.25	1.05	-.58	.25	1.07	-.55	.25	1.04
4.00	.25	50.00	-0.38731	0.24057	0.89561	-.38	.25	.88	-.37	.25	.87	-.38	.25	.88	-.37	.25	.86

Table 2
Empirically defined and computed confidence intervals

Variance Ratio	Mean Difference	n	Empirical g			Central g			Noncentral g			Adjusted Central g			Adjusted Noncentral g		
			LB	C	UB	LB	C	UB	LB	C	UB	LB	C	UB	LB	C	UB
1.00	.50	5.00	-0.87168	0.51416	2.16577	-1.18	.50	2.18	-.78	.50	1.75	-1.07	.45	1.97	-.70	.45	1.58
1.00	.50	10.00	-0.4016	0.54793	1.55896	-.51	.50	1.51	-.40	.50	1.38	-.49	.48	1.44	-.38	.48	1.33
1.00	.50	30.00	-0.05616	0.52102	1.06486	-.03	.50	1.03	-.02	.50	1.01	-.03	.49	1.02	-.02	.49	1.00
1.00	.50	50.00	0.09072	0.49504	0.92652	.09	.50	.91	.10	.50	.90	.09	.50	.90	.10	.50	.89
.25	.50	5.00	-0.70748	0.73219	2.51247	-1.06	.63	2.33	-.66	.63	1.89	-.96	.57	2.10	-.60	.57	1.71
.25	.50	10.00	-0.23194	0.69427	1.7369	-.39	.63	1.65	-.28	.63	1.52	-.37	.61	1.58	-.26	.61	1.46
.25	.50	30.00	0.13046	0.66985	1.22647	.09	.63	1.17	.11	.63	1.15	.09	.62	1.16	.11	.62	1.13
.25	.50	50.00	0.22853	0.66543	1.10652	.22	.63	1.04	.23	.63	1.03	.22	.63	1.04	.23	.63	1.02
4.00	.50	5.00	-1.23442	0.3559	2.17922	-1.35	.32	1.98	-.94	.32	1.56	-1.22	.29	1.79	-.85	.29	1.40
4.00	.50	10.00	-0.59159	0.3849	1.45503	-.68	.32	1.32	-.57	.32	1.19	-.65	.30	1.26	-.55	.30	1.14
4.00	.50	30.00	-0.22206	0.34485	0.88388	-.21	.32	.85	-.19	.32	.82	-.21	.31	.83	-.19	.31	.81
4.00	.50	50.00	-0.07241	0.34386	0.76181	-.09	.32	.72	-.08	.32	.71	-.09	.31	.71	-.08	.31	.70

Variance Ratio	Mean Difference	n	Empirical delta			Central delta			Noncentral delta			Adj. Central delta			Adj. Noncentral delta		
			LB	C	UB	LB	C	UB	LB	C	UB	LB	C	UB	LB	C	UB
1.00	.50	5.00	-0.98711	0.54355	2.42096	-1.18	.50	2.18	-.78	.50	1.75	-1.07	.45	1.97	-.70	.45	1.58
1.00	.50	10.00	-0.39451	0.57498	1.74513	-.51	.50	1.51	-.40	.50	1.38	-.49	.48	1.44	-.38	.48	1.33
1.00	.50	30.00	-0.05024	0.52464	1.09605	-.03	.50	1.03	-.02	.50	1.01	-.03	.49	1.02	-.02	.49	1.00
1.00	.50	50.00	0.09108	0.49851	0.93693	.09	.50	.91	.10	.50	.90	.09	.50	.90	.10	.50	.89
.25	.50	5.00	-0.51821	0.58368	2.07577	-.84	.50	1.84	-.52	.50	1.49	-.76	.45	1.66	-.47	.45	1.35
.25	.50	10.00	-0.16289	0.53421	1.41052	-.30	.50	1.30	-.22	.50	1.21	-.29	.48	1.25	-.21	.48	1.15
.25	.50	30.00	0.10387	0.50628	0.9298	.07	.50	.93	.09	.50	.91	.07	.49	.91	.09	.49	.90
.25	.50	50.00	0.16908	0.50184	0.83742	.18	.50	.82	.18	.50	.82	.17	.50	.82	.18	.50	.81
4.00	.50	5.00	-2.0108	0.56052	3.4752	-2.13	.50	3.13	-1.49	.50	2.46	-1.92	.45	2.83	-1.35	.45	2.22
4.00	.50	10.00	-0.8838	0.60177	2.31115	-1.08	.50	2.08	-.90	.50	1.89	-1.03	.48	1.99	-.86	.48	1.81
4.00	.50	30.00	-0.31432	0.52072	1.36601	-.34	.50	1.34	-.31	.50	1.30	-.33	.49	1.32	-.30	.49	1.29
4.00	.50	50.00	-0.09446	0.51562	1.16188	-.14	.50	1.14	-.12	.50	1.12	-.14	.50	1.13	-.12	.50	1.11

Table 3
Empirically defined and computed confidence intervals

Variance Ratio	Mean Difference	n	Empirical g			Central g			Noncentral g			Adjusted Central g			Adjusted Noncentral g		
			LB	C	UB	LB	C	UB	LB	C	UB	LB	C	UB	LB	C	UB
1.00	.75	5.00	-0.57617	0.87878	2.55364	-.96	.75	2.46	-.56	.75	2.02	-.87	.68	2.22	-.51	.68	1.82
1.00	.75	10.00	-0.20234	0.78674	1.94768	-.28	.75	1.78	-.17	.75	1.65	-.27	.72	1.70	-.16	.72	1.58
1	0.75	30.00	0.21451	0.73211	1.2848	0.21	0.75	1.29	0.22	0.75	1.27	0.2	0.7	1.28	0.22	0.74	1.25
1	0.75	50.00	0.37155	0.7686	1.17995	0.34	0.75	1.16	0.34	0.75	1.15	0.33	0.7	1.16	0.34	0.74	1.15
.25	.75	5.00	-0.33173	1.17576	3.23129	-.80	.95	2.69	-.40	.95	2.24	-.72	.86	2.43	-.36	.86	2.03
.25	.75	10.00	0.06588	1.05124	2.1668	-.10	.95	2.00	.01	.95	1.87	-.09	.91	1.91	.01	.91	1.79
.25	.75	30.00	0.45621	1.02936	1.6245	.39	.95	1.50	.41	.95	1.48	.39	.94	1.48	.41	.94	1.46
.25	.75	50.00	0.58342	1.00906	1.41928	.53	.95	1.37	.53	.95	1.36	.52	.94	1.36	.53	.94	1.35
4.00	.75	5.00	-1.06964	0.61673	2.31485	-1.20	.47	2.15	-.80	.47	1.72	-1.09	.43	1.94	-.72	.43	1.55
4.00	.75	10.00	-0.52701	0.51815	1.50289	-.53	.47	1.48	-.42	.47	1.36	-.51	.45	1.42	-.40	.45	1.30
4.00	.75	30.00	-0.01916	0.49812	1.08615	-.06	.47	1.01	-.04	.47	.99	-.06	.47	.99	-.04	.47	.97
4.00	.75	50.00	0.10752	0.50059	0.93143	.07	.47	.88	.08	.47	.87	.07	.47	.87	.08	.47	.86

Variance Ratio	Mean Difference	n	Empirical delta			Central delta			Noncentral delta			Adj. Central delta			Adj. Noncentral delta		
			LB	C	UB	LB	C	UB	LB	C	UB	LB	C	UB	LB	C	UB
1.00	.75	5.00	-0.63216	0.92857	3.17703	-.96	.75	2.46	-.56	.75	2.02	-.87	.68	2.22	-.51	.68	1.82
1.00	.75	10.00	-0.2056	0.81522	2.05784	-.28	.75	1.78	-.17	.75	1.65	-.27	.72	1.70	-.16	.72	1.58
1	0.75	30.00	0.21684	0.74145	1.32558	.21	.75	1.29	.22	.75	1.27	.20	.74	1.28	.22	.74	1.25
1	0.75	50.00	0.34629	0.76973	1.22992	0.34	0.75	1.16	0.34	0.75	1.15	0.33	0.7	1.16	0.34	0.74	1.15
.25	.75	5.00	-0.2383	0.96751	3.1165	-0.63	0.75	2.13	-0.31	0.75	1.77	-0.57	0.7	1.92	-0.28	0.68	1.6
.25	.75	10.00	0.04506	0.81348	1.82674	-.08	.75	1.58	.01	.75	1.48	-.07	.72	1.51	.01	.72	1.41
.25	.75	30.00	0.34206	0.77792	1.28923	.31	.75	1.19	.32	.75	1.17	.31	.74	1.17	.32	.74	1.15
.25	.75	50.00	0.42565	0.75932	1.09831	.42	.75	1.08	.42	.75	1.08	.41	.74	1.08	.42	.74	1.07
4.00	.75	5.00	-1.8774	1.00922	4.4427	-1.90	.75	3.40	-1.27	.75	2.72	-1.72	.68	3.07	-1.14	.68	2.46
4.00	.75	10.00	-0.7884	0.7877	2.40979	-.84	.75	2.34	-.67	.75	2.15	-.81	.72	2.24	-.64	.72	2.06
4.00	.75	30.00	-0.03017	0.74854	1.62196	-.09	.75	1.59	-.06	.75	1.56	-.09	.74	1.57	-.06	.74	1.54
4.00	.75	50.00	0.16089	0.75424	1.42836	.11	.75	1.39	.12	.75	1.38	.11	.74	1.38	.12	.74	1.37

Table 4
Empirically defined and computed confidence intervals

Variance	Mean		Empirical g			Central g			Noncentral g			Adjusted Central g			Adjusted Noncentral g		
	Ratio	Difference	n	LB	C	UB	LB	C	UB	LB	C	UB	LB	C	UB	LB	C
1.00	2.00	5.00	0.63784	2.27984	4.65358	-.03	2.00	4.03	.39	2.00	3.53	-.02	1.81	3.64	.35	1.81	3.19
1.00	2.00	10.00	1.09753	2.12263	3.39973	.78	2.00	3.22	.89	2.00	3.07	.75	1.92	3.08	.85	1.92	2.94
1.00	2.00	30.00	1.43562	2.03271	2.72583	1.36	2.00	2.64	1.37	2.00	2.62	1.34	1.97	2.61	1.35	1.97	2.58
1.00	2.00	50.00	1.57499	2.0109	2.53157	1.51	2.00	2.49	1.52	2.00	2.48	1.50	1.98	2.47	1.50	1.98	2.46
.25	2.00	5.00	1.2665	3.06529	6.2933	.31	2.53	4.75	.75	2.53	4.23	.28	2.29	4.29	.68	2.29	3.82
.25	2.00	10.00	1.59733	2.83681	4.70884	1.20	2.53	3.86	1.31	2.53	3.71	1.15	2.42	3.70	1.26	2.42	3.55
.25	2.00	30.00	2.03483	2.73099	3.58187	1.82	2.53	3.24	1.84	2.53	3.21	1.80	2.50	3.19	1.82	2.50	3.17
.25	2.00	50.00	2.14398	2.7051	3.36808	1.99	2.53	3.07	2.00	2.53	3.05	1.98	2.51	3.04	1.98	2.51	3.03
4.00	2.00	5.00	-0.0671	1.52331	3.8562	-.55	1.26	3.08	-.15	1.26	2.62	-.49	1.14	2.78	-.13	1.14	2.36
4	2	10.00	-0.31173	0.70329	1.75883	0.18	1.26	2.35	0.28	1.26	2.22	0.17	1.2	2.25	0.27	1.21	2.12
4	2	30.00	0.76455	1.35763	1.97211	0.69	1.26	1.84	0.7	1.26	1.82	0.68	1.3	1.82	0.7	1.25	1.79
4.00	2.00	50.00	0.93135	1.34555	1.81909	.83	1.26	1.70	.83	1.26	1.69	.82	1.26	1.69	.83	1.26	1.68

Variance	Mean		Empirical delta			Central delta			Noncentral delta			Adj. Central delta			Adj. Noncentral delta		
	Ratio	Difference	n	LB	C	UB	LB	C	UB	LB	C	UB	LB	C	UB	LB	C
1.00	2.00	5.00	0.6303	2.46833	5.7131	-.03	2.00	4.03	.39	2.00	3.53	-.02	1.81	3.64	.35	1.81	3.19
1.00	2.00	10.00	1.02922	2.19893	4.00068	.78	2.00	3.22	.89	2.00	3.07	.75	1.92	3.08	.85	1.92	2.94
1.00	2.00	30.00	1.37234	2.05074	2.91323	1.36	2.00	2.64	1.37	2.00	2.62	1.34	1.97	2.61	1.35	1.97	2.58
1.00	2.00	50.00	1.5062	2.02795	2.6521	1.51	2.00	2.49	1.52	2.00	2.48	1.50	1.98	2.47	1.50	1.98	2.46
.25	2.00	5.00	0.8394	2.50071	6.0727	.25	2.00	3.75	.60	2.00	3.34	.22	1.81	3.39	.54	1.81	3.02
.25	2.00	10.00	1.12841	2.19998	4.12561	.95	2.00	3.05	1.04	2.00	2.93	.91	1.92	2.92	.99	1.92	2.81
.25	2.00	30.00	1.47255	2.06359	2.84542	1.44	2.00	2.56	1.45	2.00	2.54	1.42	1.97	2.52	1.44	1.97	2.50
.25	2.00	50.00	1.57402	2.03917	2.64919	1.57	2.00	2.43	1.58	2.00	2.41	1.56	1.98	2.41	1.57	1.98	2.40
4.00	2.00	5.00	-0.1046	2.42102	6.9813	-.86	2.00	4.86	-.23	2.00	4.13	-.78	1.81	4.39	-.21	1.81	3.73
4	2	10.00	-0.45066	1.12503	2.98568	.28	2.00	3.72	.45	2.00	3.51	.27	1.92	3.56	.43	1.92	3.36
4	2	30.00	1.11472	2.04446	3.08141	1.09	2	2.91	1.11	2	2.87	1.08	2	2.87	1.1	1.97	2.83
4.00	2.00	50.00	1.32928	2.02925	2.90768	1.31	2	2.69	1.32	2	2.68	1.3	2	2.67	1.31	1.98	2.66

Table 5
Empirically defined and computed confidence intervals

Variance Ratio	Mean Difference	n	Empirical g			Central g			Noncentral g			Adjusted Central g			Adjusted Noncentral g		
			LB	C	UB	LB	C	UB	LB	C	UB	LB	C	UB	LB	C	UB
1.00	3.00	5.00	1.73508	3.47762	6.44085	.59	3.00	5.41	1.06	3.00	4.87	.53	2.71	4.89	.96	2.71	4.39
1.00	3.00	10.00	1.98297	3.17297	4.86922	1.55	3.00	4.45	1.67	3.00	4.29	1.49	2.87	4.26	1.60	2.87	4.11
1.00	3.00	30.00	2.35503	3.04338	3.87521	2.23	3.00	3.77	2.25	3.00	3.74	2.20	2.96	3.72	2.22	2.96	3.69
1.00	3.00	50.00	2.4768	3.01564	3.63671	2.42	3.00	3.58	2.42	3.00	3.57	2.40	2.98	3.56	2.40	2.98	3.54
.25	3.00	5.00	2.257	4.54309	8.6787	1.03	3.79	6.56	1.56	3.79	5.97	.93	3.43	5.93	1.40	3.43	5.39
.25	3.00	10.00	2.70838	4.23794	6.63431	2.13	3.79	5.46	2.27	3.79	5.29	2.04	3.63	5.23	2.17	3.63	5.06
.25	3.00	30.00	3.23493	4.0931	5.20379	2.91	3.79	4.67	2.93	3.79	4.64	2.88	3.75	4.61	2.90	3.75	4.58
.25	3.00	50.00	3.31682	4.02654	4.82805	3.12	3.79	4.47	3.13	3.79	4.45	3.10	3.77	4.43	3.10	3.77	4.41
4.00	3.00	5.00	0.65634	2.30808	5.03593	-.09	1.90	3.89	.32	1.90	3.40	-.08	1.71	3.51	.29	1.71	3.07
4.00	3.00	10.00	1.02753	2.09906	3.5594	.70	1.90	3.09	.81	1.90	2.95	.67	1.82	2.96	.77	1.82	2.83
4.00	3.00	30.00	1.38569	2.04244	2.79473	1.26	1.90	2.53	1.28	1.90	2.50	1.25	1.87	2.50	1.26	1.87	2.47
4.00	3.00	50.00	1.51892	2.03113	2.58387	1.41	1.90	2.38	1.42	1.90	2.37	1.40	1.88	2.36	1.41	1.88	2.35

Variance Ratio	Mean Difference	n	Empirical delta			Central delta			Noncentral delta			Adj. Central delta			Adj. Noncentral delta		
			LB	C	UB	LB	C	UB	LB	C	UB	LB	C	UB	LB	C	UB
1.00	3.00	5.00	1.4473	3.86014	9.3989	.59	3.00	5.41	1.06	3.00	4.87	.53	2.71	4.89	.96	2.71	4.39
1.00	3.00	10.00	1.8162	3.26676	5.71954	1.55	3.00	4.45	1.67	3.00	4.29	1.49	2.87	4.26	1.60	2.87	4.11
1.00	3.00	30.00	2.26969	3.06563	4.1255	2.23	3.00	3.77	2.25	3.00	3.74	2.20	2.96	3.72	2.22	2.96	3.69
1.00	3.00	50.00	2.39103	3.03125	3.79352	2.42	3.00	3.58	2.42	3.00	3.57	2.40	2.98	3.56	2.40	2.98	3.54
.25	3.00	5.00	1.5675	3.68675	8.8847	.81	3.00	5.19	1.23	3.00	4.72	.73	2.71	4.69	1.11	2.71	4.26
.25	3.00	10.00	1.89262	3.26429	5.61781	1.69	3.00	4.31	1.79	3.00	4.18	1.61	2.87	4.13	1.72	2.87	4.00
.25	3.00	30.00	2.33535	3.10128	4.2246	2.30	3.00	3.70	2.32	3.00	3.67	2.27	2.96	3.65	2.29	2.96	3.62
.25	3.00	50.00	2.36831	3.03625	3.77545	2.47	3.00	3.53	2.47	3.00	3.51	2.45	2.98	3.50	2.45	2.98	3.49
4.00	3.00	5.00	0.9471	3.72345	8.9366	-.15	3.00	6.15	.51	3.00	5.38	-.13	2.71	5.55	.46	2.71	4.86
4.00	3.00	10.00	1.5304	3.30482	6.1929	1.11	3.00	4.89	1.28	3.00	4.67	1.06	2.87	4.68	1.23	2.87	4.47
4.00	3.00	30.00	2.10373	3.09825	4.44102	2.00	3.00	4.00	2.02	3.00	3.96	1.97	2.96	3.95	2.00	2.96	3.91
4.00	3.00	50.00	2.24731	3.04628	4.012	2.24	3.00	3.76	2.25	3.00	3.74	2.22	2.98	3.73	2.23	2.98	3.71

Table 6
Empirically defined and computed confidence intervals

Variance Ratio	Mean Difference	n	Empirical - Central d			Empirical - Noncentral d			Empirical - Adj. Central d			Empirical - Adj. Noncentral d		
			LB	C	UB	LB	C	UB	LB	C	UB	LB	C	UB
1.00	1.00	5.00	0.50	0.13	0.19	0.11	0.13	0.64	0.43	0.23	0.45	0.07	0.23	0.86
1.00	1.00	10.00	0.16	0.06	0.17	0.06	0.06	0.30	0.16	0.10	0.25	0.06	0.10	0.38
1.00	1.00	30.00	0.04	0.02	0.06	0.02	0.02	0.09	0.04	0.03	0.08	0.03	0.03	0.11
1.00	1.00	50.00	0.01	0.02	0.02	0.00	0.02	0.04	0.01	0.03	0.04	0.00	0.03	0.05
.25	1.00	5.00	0.62	0.22	0.43	0.22	0.22	0.89	0.56	0.34	0.73	0.20	0.34	1.15
.25	1.00	10.00	0.24	0.14	0.30	0.14	0.14	0.43	0.25	0.19	0.40	0.15	0.19	0.53
.25	1.00	30.00	0.09	0.09	0.16	0.08	0.09	0.18	0.10	0.10	0.18	0.08	0.10	0.21
.25	1.00	50.00	0.08	0.09	0.13	0.08	0.09	0.14	0.09	0.09	0.14	0.08	0.09	0.15
4.00	1.00	5.00	0.34	0.17	0.17	-0.06	0.17	0.61	0.24	0.23	0.40	-0.12	0.23	0.79
4.00	1.00	10.00	0.09	0.09	0.16	-0.02	0.09	0.29	0.07	0.11	0.23	-0.04	0.11	0.35
4.00	1.00	30.00	0.05	0.06	0.07	0.03	0.06	0.09	0.05	0.07	0.08	0.03	0.07	0.11
4.00	1.00	50.00	0.05	0.04	0.06	0.04	0.04	0.07	0.05	0.04	0.06	0.04	0.04	0.08

Variance Ratio	Mean Difference	n	Empirical - Central g			Empirical - Noncentral g			Empirical - Adj. Central g			Empirical - Adj. Noncentral g		
			LB	C	UB	LB	C	UB	LB	C	UB	LB	C	UB
1.00	1.00	5.00	0.47	0.26	1.06	0.08	0.26	1.51	0.40	0.36	1.32	0.04	0.36	1.73
1.00	1.00	10.00	0.16	0.10	0.33	0.06	0.10	0.46	0.16	0.14	0.41	0.06	0.14	0.54
1.00	1.00	30.00	0.01	0.03	0.15	-0.01	0.03	0.18	0.01	0.04	0.17	0.00	0.04	0.20
1.00	1.00	50.00	0.02	0.02	0.06	0.01	0.02	0.08	0.02	0.03	0.08	0.01	0.03	0.09
.25	1.00	5.00	0.49	0.18	0.72	0.18	0.18	1.08	0.45	0.28	0.95	0.17	0.28	1.28
.25	1.00	10.00	0.17	0.07	0.27	0.09	0.07	0.38	0.18	0.11	0.35	0.10	0.11	0.45
.25	1.00	30.00	0.04	0.02	0.09	0.02	0.02	0.11	0.04	0.03	0.11	0.03	0.03	0.13
.25	1.00	50.00	0.02	0.02	0.07	0.01	0.02	0.08	0.02	0.03	0.08	0.02	0.03	0.09
4.00	1.00	5.00	0.40	0.30	1.03	-0.23	0.30	1.72	0.24	0.40	1.39	-0.33	0.40	2.01
4.00	1.00	10.00	0.09	0.11	0.34	-0.08	0.11	0.54	0.06	0.15	0.45	-0.10	0.15	0.64
4.00	1.00	30.00	0.06	0.04	0.07	0.03	0.04	0.10	0.06	0.05	0.09	0.04	0.05	0.13
4.00	1.00	50.00	0.05	0.01	0.03	0.04	0.01	0.05	0.05	0.02	0.04	0.04	0.02	0.06

Table 7
Differences between the calculated values and the empirical values.

Variance Ratio	Mean Difference	n	Empirical - Central d			Empirical - Noncentral d			Empirical - Adj. Central d			Empirical - Adj. Noncentral d		
			LB	C	UB	LB	C	UB	LB	C	UB	LB	C	UB
1.00	.25	5.00	0.17	0.04	-0.13	-0.24	0.04	0.29	0.03	0.06	0.05	-0.33	0.06	0.44
1.00	.25	10.00	0.18	0.03	0.03	0.06	0.03	0.15	0.15	0.04	0.09	0.04	0.04	0.20
1.00	.25	30.00	0.00	0.00	0.02	-0.02	0.00	0.04	-0.01	0.00	0.03	-0.02	0.00	0.05
1.00	.25	50.00	0.00	0.01	-0.02	-0.01	0.01	-0.01	0.00	0.01	-0.02	-0.01	0.01	-0.01
.25	.25	5.00	0.12	0.05	-0.11	-0.29	0.05	0.31	-0.01	0.08	0.08	-0.38	0.08	0.47
.25	.25	10.00	0.09	0.06	0.07	-0.02	0.06	0.20	0.06	0.08	0.13	-0.04	0.08	0.25
.25	.25	30.00	0.00	0.02	0.06	-0.02	0.02	0.09	0.00	0.03	0.08	-0.02	0.03	0.10
.25	.25	50.00	0.02	0.02	0.05	0.01	0.02	0.06	0.02	0.03	0.06	0.01	0.03	0.07
4.00	.25	5.00	0.18	0.02	-0.08	-0.23	0.02	0.33	0.03	0.04	0.09	-0.34	0.04	0.47
4.00	.25	10.00	0.02	0.04	0.09	-0.10	0.04	0.21	-0.02	0.05	0.14	-0.13	0.05	0.25
4.00	.25	30.00	0.01	0.02	0.09	-0.01	0.02	0.11	0.00	0.02	0.09	-0.01	0.02	0.11
4.00	.25	50.00	-0.01	0.00	0.05	-0.01	0.00	0.06	-0.01	0.00	0.05	-0.02	0.00	0.06

Variance Ratio	Mean Difference	n	Empirical - Central g			Empirical - Noncentral g			Empirical - Adj. Central g			Empirical - Adj. Noncentral g		
			LB	C	UB	LB	C	UB	LB	C	UB	LB	C	UB
1.00	.25	5.00	0.03	0.06	0.20	-0.38	0.06	0.62	-0.11	0.08	0.38	-0.47	0.08	0.77
1.00	.25	10.00	0.16	0.03	0.10	0.04	0.03	0.22	0.13	0.04	0.16	0.02	0.04	0.27
1.00	.25	30.00	-0.01	0.00	0.00	-0.03	0.00	0.02	-0.02	0.00	0.01	-0.03	0.00	0.03
1.00	.25	50.00	0.00	0.01	-0.02	-0.01	0.01	-0.01	0.00	0.01	-0.02	-0.01	0.01	-0.01
.25	.25	5.00	0.00	0.06	0.13	-0.32	0.06	0.47	-0.11	0.08	0.29	-0.40	0.08	0.59
.25	.25	10.00	0.10	0.04	0.08	0.01	0.04	0.18	0.08	0.05	0.12	-0.01	0.05	0.22
.25	.25	30.00	0.01	0.01	0.03	-0.01	0.01	0.05	0.01	0.01	0.04	-0.01	0.01	0.06
.25	.25	50.00	0.02	0.01	0.00	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.01	0.01
4.00	.25	5.00	0.10	0.04	0.03	-0.55	0.04	0.69	-0.13	0.06	0.31	-0.72	0.06	0.91
4.00	.25	10.00	0.02	0.06	0.21	-0.16	0.06	0.40	-0.03	0.07	0.28	-0.21	0.07	0.46
4.00	.25	30.00	0.04	0.03	0.02	0.01	0.03	0.05	0.04	0.03	0.03	0.01	0.03	0.06
4.00	.25	50.00	-0.01	-0.01	0.02	-0.02	-0.01	0.03	-0.01	-0.01	0.02	-0.02	-0.01	0.04

Table 8

Differences between the calculated values and the empirical values.

Variance Ratio	Mean Difference	n	Empirical - Central d			Empirical - Noncentral d			Empirical - Adj. Central d			Empirical - Adj. Noncentral d		
			LB	C	UB	LB	C	UB	LB	C	UB	LB	C	UB
1.00	.50	5.00	0.31	0.01	-0.01	-0.09	0.01	0.42	0.20	0.06	0.20	-0.17	0.06	0.59
1.00	.50	10.00	0.11	0.05	0.05	0.00	0.05	0.18	0.09	0.07	0.12	-0.02	0.07	0.23
1.00	.50	30.00	-0.03	0.02	0.03	-0.04	0.02	0.05	-0.03	0.03	0.04	-0.04	0.03	0.06
1.00	.50	50.00	0.00	0.00	0.02	-0.01	0.00	0.03	0.00	0.00	0.03	-0.01	0.00	0.04
.25	.50	5.00	0.35	0.10	0.18	-0.05	0.10	0.62	0.25	0.16	0.41	-0.11	0.16	0.80
.25	.50	10.00	0.16	0.06	0.09	0.05	0.06	0.22	0.14	0.08	0.16	0.03	0.08	0.28
.25	.50	30.00	0.04	0.04	0.06	0.02	0.04	0.08	0.04	0.05	0.07	0.02	0.05	0.10
.25	.50	50.00	0.01	0.04	0.07	0.00	0.04	0.08	0.01	0.04	0.07	0.00	0.04	0.09
4.00	.50	5.00	0.12	0.04	0.20	-0.29	0.04	0.62	-0.01	0.07	0.39	-0.38	0.07	0.78
4.00	.50	10.00	0.09	0.06	0.14	-0.02	0.06	0.27	0.06	0.08	0.20	-0.04	0.08	0.32
4.00	.50	30.00	-0.01	0.02	0.03	-0.03	0.02	0.06	-0.01	0.03	0.05	-0.03	0.03	0.07
4.00	.50	50.00	0.02	0.02	0.04	0.01	0.02	0.05	0.02	0.03	0.05	0.01	0.03	0.06

Variance Ratio	Mean Difference	n	Empirical - Central g			Empirical - Noncentral g			Empirical - Adj. Central g			Empirical - Adj. Noncentral g		
			LB	C	UB	LB	C	UB	LB	C	UB	LB	C	UB
1.00	.50	5.00	0.19	0.04	0.24	-0.21	0.04	0.67	0.08	0.09	0.45	-0.29	0.09	0.84
1.00	.50	10.00	0.12	0.07	0.24	0.01	0.07	0.37	0.10	0.09	0.31	-0.01	0.09	0.42
1.00	.50	30.00	-0.02	0.02	0.07	-0.03	0.02	0.09	-0.02	0.03	0.08	-0.03	0.03	0.10
1.00	.50	50.00	0.00	0.00	0.03	-0.01	0.00	0.04	0.00	0.00	0.04	-0.01	0.00	0.05
.25	.50	5.00	0.32	0.08	0.24	0.00	0.08	0.59	0.24	0.13	0.42	-0.05	0.13	0.73
.25	.50	10.00	0.14	0.03	0.11	0.06	0.03	0.20	0.13	0.05	0.16	0.05	0.05	0.26
.25	.50	30.00	0.03	0.01	0.00	0.01	0.01	0.02	0.03	0.02	0.02	0.01	0.02	0.03
.25	.50	50.00	-0.01	0.00	0.02	-0.01	0.00	0.02	0.00	0.00	0.02	-0.01	0.00	0.03
4.00	.50	5.00	0.12	0.06	0.35	-0.52	0.06	1.02	-0.09	0.11	0.65	-0.66	0.11	1.26
4.00	.50	10.00	0.20	0.10	0.23	0.02	0.10	0.42	0.15	0.12	0.32	-0.02	0.12	0.50
4.00	.50	30.00	0.03	0.02	0.03	0.00	0.02	0.07	0.02	0.03	0.05	-0.01	0.03	0.08
4.00	.50	50.00	0.05	0.02	0.02	0.03	0.02	0.04	0.05	0.02	0.03	0.03	0.02	0.05

Table 9

Differences between the calculated values and the empirical values.

Variance Ratio	Mean Difference	n	Empirical - Central d			Empirical - Noncentral d			Empirical - Adj. Central d			Empirical - Adj. Noncentral d		
			LB	C	UB	LB	C	UB	LB	C	UB	LB	C	UB
1.00	.75	5.00	0.38	0.13	0.09	-0.02	0.13	0.53	0.29	0.20	0.33	-0.07	0.20	0.73
1.00	.75	10.00	0.08	0.04	0.17	-0.03	0.04	0.30	0.07	0.07	0.25	-0.04	0.07	0.37
1	0.75	30.00	0.00	-0.02	-0.01	-0.01	-0.02	0.01	0.01	-0.01	0.00	-0.01	-0.01	0.03
1	0.75	50.00	0.03	0.02	0.02	0.03	0.02	0.03	0.04	0.03	0.02	0.03	0.03	0.03
.25	.75	5.00	0.47	0.23	0.54	0.07	0.23	0.99	0.39	0.32	0.80	0.03	0.32	1.20
.25	.75	10.00	0.17	0.10	0.17	0.06	0.10	0.30	0.16	0.14	0.26	0.06	0.14	0.38
.25	.75	30.00	0.07	0.08	0.12	0.05	0.08	0.14	0.07	0.09	0.14	0.05	0.09	0.16
.25	.75	50.00	0.05	0.06	0.05	0.05	0.06	0.06	0.06	0.07	0.06	0.05	0.07	0.07
4.00	.75	5.00	0.13	0.15	0.16	-0.27	0.15	0.59	0.02	0.19	0.37	-0.35	0.19	0.76
4.00	.75	10.00	0.00	0.05	0.02	-0.11	0.05	0.14	-0.02	0.07	0.08	-0.13	0.07	0.20
4.00	.75	30.00	0.04	0.03	0.08	0.02	0.03	0.10	0.04	0.03	0.10	0.02	0.03	0.12
4.00	.75	50.00	0.04	0.03	0.05	0.03	0.03	0.06	0.04	0.03	0.06	0.03	0.03	0.07

Variance Ratio	Mean Difference	n	Empirical - Central g			Empirical - Noncentral g			Empirical - Adj. Central g			Empirical - Adj. Noncentral g		
			LB	C	UB	LB	C	UB	LB	C	UB	LB	C	UB
1.00	.75	5.00	0.33	0.18	0.72	-0.07	0.18	1.16	0.24	0.25	0.96	-0.12	0.25	1.36
1.00	.75	10.00	0.07	0.07	0.28	-0.04	0.07	0.41	0.06	0.10	0.36	-0.05	0.10	0.48
1	0.75	30.00	0.01	-0.01	0.04	0.00	-0.01	0.06	0.02	0.00	0.05	0.00	0.00	0.08
1	0.75	50.00	0.01	0.02	0.07	0.01	0.02	0.08	0.02	0.03	0.07	0.01	0.03	0.08
.25	.75	5.00	0.39	0.22	0.99	0.07	0.22	1.35	0.33	0.29	1.20	0.04	0.29	1.52
.25	.75	10.00	0.13	0.06	0.25	0.04	0.06	0.35	0.12	0.09	0.32	0.04	0.09	0.42
.25	.75	30.00	0.03	0.03	0.10	0.02	0.03	0.12	0.03	0.04	0.12	0.02	0.04	0.14
.25	.75	50.00	0.01	0.01	0.02	0.01	0.01	0.02	0.02	0.02	0.02	0.01	0.02	0.03
4.00	.75	5.00	0.02	0.26	1.04	-0.61	0.26	1.72	-0.16	0.33	1.37	-0.74	0.33	1.98
4.00	.75	10.00	0.05	0.04	0.07	-0.12	0.04	0.26	0.02	0.07	0.17	-0.15	0.07	0.35
4.00	.75	30.00	0.06	0.00	0.03	0.03	0.00	0.06	0.06	0.01	0.05	0.03	0.01	0.08
4.00	.75	50.00	0.05	0.00	0.04	0.04	0.00	0.05	0.05	0.01	0.05	0.04	0.01	0.06

Table 10
Differences between the calculated values and the empirical values.

Variance Ratio	Mean Difference	n	Empirical - Central d			Empirical - Noncentral d			Empirical - Adj. Central d			Empirical - Adj. Noncentral d		
			LB	C	UB	LB	C	UB	LB	C	UB	LB	C	UB
1.00	2.00	5.00	0.67	0.28	0.62	0.25	0.28	1.12	0.66	0.47	1.01	0.29	0.47	1.46
1.00	2.00	10.00	0.32	0.12	0.18	0.21	0.12	0.33	0.35	0.20	0.32	0.25	0.20	0.46
1.00	2.00	30.00	0.08	0.03	0.09	0.07	0.03	0.11	0.10	0.06	0.12	0.09	0.06	0.15
1.00	2.00	50.00	0.06	0.01	0.04	0.05	0.01	0.05	0.07	0.03	0.06	0.07	0.03	0.07
.25	2.00	5.00	0.96	0.54	1.54	0.52	0.54	2.06	0.99	0.78	2.00	0.59	0.78	2.47
.25	2.00	10.00	0.40	0.31	0.85	0.29	0.31	1.00	0.45	0.42	1.01	0.34	0.42	1.16
.25	2.00	30.00	0.21	0.20	0.34	0.19	0.20	0.37	0.23	0.23	0.39	0.21	0.23	0.41
.25	2.00	50.00	0.15	0.18	0.30	0.14	0.18	0.32	0.16	0.20	0.33	0.16	0.20	0.34
4.00	2.00	5.00	0.48	0.26	0.78	0.08	0.26	1.24	0.42	0.38	1.08	0.06	0.38	1.50
4	2	10.00	-0.49	-0.56	-0.59	-0.59	-0.56	-0.46	-0.48	-0.51	-0.49	-0.58	-0.51	-0.36
4	2	30.00	0.07	0.10	0.13	0.06	0.10	0.15	0.08	0.11	0.15	0.06	0.11	0.18
4.00	2.00	50.00	0.10	0.09	0.12	0.10	0.09	0.13	0.11	0.09	0.13	0.10	0.09	0.14

Variance Ratio	Mean Difference	n	Empirical - Central g			Empirical - Noncentral g			Empirical - Adj. Central g			Empirical - Adj. Noncentral g		
			LB	C	UB	LB	C	UB	LB	C	UB	LB	C	UB
1.00	2.00	5.00	0.66	0.47	1.68	0.24	0.47	2.18	0.65	0.66	2.07	0.28	0.66	2.52
1.00	2.00	10.00	0.25	0.20	0.78	0.14	0.20	0.93	0.28	0.28	0.92	0.18	0.28	1.06
1.00	2.00	30.00	0.01	0.05	0.27	0.00	0.05	0.29	0.03	0.08	0.30	0.02	0.08	0.33
1.00	2.00	50.00	0.00	0.03	0.16	-0.01	0.03	0.17	0.01	0.05	0.18	0.01	0.05	0.19
.25	2.00	5.00	0.59	0.50	2.32	0.24	0.50	2.73	0.62	0.69	2.68	0.30	0.69	3.05
.25	2.00	10.00	0.18	0.20	1.08	0.09	0.20	1.20	0.22	0.28	1.21	0.14	0.28	1.32
.25	2.00	30.00	0.03	0.06	0.29	0.02	0.06	0.31	0.05	0.09	0.33	0.03	0.09	0.35
.25	2.00	50.00	0.00	0.04	0.22	-0.01	0.04	0.24	0.01	0.06	0.24	0.00	0.06	0.25
4.00	2.00	5.00	0.76	0.42	2.12	0.13	0.42	2.85	0.68	0.61	2.59	0.11	0.61	3.25
4	2	10.00	-0.73	-0.87	-0.73	-0.90	-0.87	-0.52	-0.72	-0.79	-0.57	-0.88	-0.79	-0.37
4	2	30.00	0.02	0.04	0.17	0.00	0.04	0.21	0.03	0.07	0.21	0.01	0.07	0.25
4.00	2.00	50.00	0.02	0.03	0.22	0.01	0.03	0.23	0.03	0.05	0.24	0.02	0.05	0.25

Table 11

Differences between the calculated values and the empirical values.

Variance Ratio	Mean Difference	n	Empirical - Central d			Empirical - Noncentral d			Empirical - Adj. Central d			Empirical - Adj. Noncentral d		
			LB	C	UB	LB	C	UB	LB	C	UB	LB	C	UB
1.00	3.00	5.00	1.15	0.48	1.03	0.68	0.48	1.57	1.21	0.77	1.55	0.78	0.77	2.05
1.00	3.00	10.00	0.43	0.17	0.42	0.31	0.17	0.58	0.49	0.30	0.61	0.38	0.30	0.76
1.00	3.00	30.00	0.13	0.04	0.11	0.11	0.04	0.14	0.16	0.08	0.16	0.14	0.08	0.19
1.00	3.00	50.00	0.06	0.02	0.06	0.06	0.02	0.07	0.08	0.04	0.08	0.08	0.04	0.10
.25	3.00	5.00	1.23	0.75	2.12	0.70	0.75	2.71	1.33	1.11	2.75	0.86	1.11	3.29
.25	3.00	10.00	0.58	0.45	1.17	0.44	0.45	1.34	0.67	0.61	1.40	0.54	0.61	1.57
.25	3.00	30.00	0.32	0.30	0.53	0.30	0.30	0.56	0.35	0.34	0.59	0.33	0.34	0.62
.25	3.00	50.00	0.20	0.24	0.36	0.19	0.24	0.38	0.22	0.26	0.40	0.22	0.26	0.42
4.00	3.00	5.00	0.75	0.41	1.15	0.34	0.41	1.64	0.74	0.60	1.53	0.37	0.60	1.97
4.00	3.00	10.00	0.33	0.20	0.47	0.22	0.20	0.61	0.36	0.28	0.60	0.26	0.28	0.73
4.00	3.00	30.00	0.13	0.14	0.26	0.11	0.14	0.29	0.14	0.17	0.29	0.13	0.17	0.32
4.00	3.00	50.00	0.11	0.13	0.20	0.10	0.13	0.21	0.12	0.15	0.22	0.11	0.15	0.23

Variance Ratio	Mean Difference	n	Empirical - Central g			Empirical - Noncentral g			Empirical - Adj. Central g			Empirical - Adj. Noncentral g		
			LB	C	UB	LB	C	UB	LB	C	UB	LB	C	UB
1.00	3.00	5.00	0.86	0.86	3.99	0.39	0.86	4.53	0.92	1.15	4.51	0.49	1.15	5.01
1.00	3.00	10.00	0.27	0.27	1.27	0.15	0.27	1.43	0.33	0.40	1.46	0.22	0.40	1.61
1.00	3.00	30.00	0.04	0.07	0.36	0.02	0.07	0.39	0.07	0.11	0.41	0.05	0.11	0.44
1.00	3.00	50.00	-0.03	0.03	0.21	-0.03	0.03	0.22	-0.01	0.05	0.23	-0.01	0.05	0.25
.25	3.00	5.00	0.76	0.69	3.69	0.34	0.69	4.16	0.84	0.98	4.19	0.46	0.98	4.62
.25	3.00	10.00	0.20	0.26	1.31	0.10	0.26	1.44	0.28	0.39	1.49	0.17	0.39	1.62
.25	3.00	30.00	0.04	0.10	0.52	0.02	0.10	0.55	0.07	0.14	0.57	0.05	0.14	0.60
.25	3.00	50.00	-0.10	0.04	0.25	-0.10	0.04	0.27	-0.08	0.06	0.28	-0.08	0.06	0.29
4.00	3.00	5.00	1.10	0.72	2.79	0.44	0.72	3.56	1.08	1.01	3.39	0.49	1.01	4.08
4.00	3.00	10.00	0.42	0.30	1.30	0.25	0.30	1.52	0.47	0.43	1.51	0.30	0.43	1.72
4.00	3.00	30.00	0.10	0.10	0.44	0.08	0.10	0.48	0.13	0.14	0.49	0.10	0.14	0.53
4.00	3.00	50.00	0.01	0.05	0.25	0.00	0.05	0.27	0.03	0.07	0.28	0.02	0.07	0.30

Table 12
Differences between the calculated values and the empirical values.

Variance Ratio	Central g			Noncentral g			Adj. Central g			Adj. Noncentral g		
	LB	C	UB	LB	C	UB	LB	C	UB	LB	C	UB
1	0.20	0.07	0.13	0.06	0.07	0.29	0.19	0.12	0.24	0.07	0.12	0.39
0.25	0.28	0.18	0.40	0.13	0.18	0.56	0.27	0.24	0.53	0.14	0.24	0.68
4.00	0.11	0.07	0.16	-0.03	0.07	0.31	0.08	0.10	0.25	-0.04	0.10	0.39

Variance Ratio	Central delta			Noncentral delta			Adj. Central delta			Adj. Noncentral delta		
	LB	C	UB	LB	C	UB	LB	C	UB	LB	C	UB
1	0.15	0.12	0.51	0.01	0.12	0.67	0.14	0.17	0.62	0.01	0.17	0.77
0.25	0.15	0.11	0.53	0.04	0.11	0.66	0.15	0.16	0.63	0.05	0.16	0.75
4.00	0.15	0.11	0.53	-0.09	0.08	0.66	0.09	0.13	0.56	-0.11	0.13	0.78

Table 13

The average differences between the lower bounds, upper bounds, and point estimates for each level of variance ratio for each effect size estimator