A Monte Carlo Study of the Robustness of Coefficient Alpha

by

Sharon G. Shultz
Faculty of Education

Thesis submitted to
the school of Graduate Studies and Research
in partial fulfilment of the requirements for the
MA degree in Education

University of Ottawa

ISBN 0-315-83840-X

ABSTRACT

The biasedness and efficiency of coefficient alpha were studied given various error score distributions, number of examinees, population reliability values, and number of subtests. A Monte Carlo methodology was used in which an additive true score matrix was constructed and an error score was added to each true score. The true scores were uniformly distributed while the error score distributions used were the normal, mixed normal, exponential, and the negative exponential. In addition, 1000 replications were used and the resulting sampling distributions were examined with respect to their skewness and kurtosis.

The results indicated that, when 50, 100, and 200 examinees and 5, 10, 20, and 30 subtests were used, coefficient alpha was an unbiased estimator of population reliability values of 0.4, 0.6, 0.8, and 0.9. The efficiency of the estimate increased as the number of examinees or the population reliability increased, regardless of the error score distribution used. Of the four error score distributions used the normal was the most efficient. In addition, the sampling distributions tended to become less skewed and more leptokurtic as the number of examinees and the population reliability increased. Surprisingly, an increase in the number of subtests did not appear to affect the biasedness and efficiency of coefficient alpha or the skewness and kurtosis of its sampling distribution. The results are related to the literature and suggestions for future study are given.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

Introduction

In many testing situations the use of parallel forms or test-retest reliability coefficients is impractical. Viable alternatives to these approaches are the internal consistency coefficients. One of the most commonly used internal consistency coefficients is coefficient alpha which is a function of sample variances. Since variances are influenced by nonnormality, it would stand to reason that coefficient alpha would be influenced by nonnormality. The purpose of this study is to investigate the robustness of coefficient alpha to nonnormality.

A review of the literature is presented in Chapter II. Prior to this review the derivation of coefficient alpha is provided. Following the review of the literature, the purpose of the study and the research questions are outlined. The methodology of the study is provided in Chapter III. Justification for the levels of the dependent and independent variables are given. Information concerning the Monte Carlo design and the data generation procedure is also given in Chapter III. The results are provided in Chapter IV. The final chapter, Chapter V, includes a discussion of the results, possible reasons for the results, limitations of the study, and suggestions for future study.

CHAPTER II

Review of the Literature

When educators require an estimate of reliability, parallel forms and test-retest approaches are often impractical. Coefficient alpha is often the reliability estimate of choice. Coefficient alpha is the mean of all possible split-half correlations and estimates the lower bound of the reliability. It can be used for both dichotomous and nondichotomous data and requires only one test administration. Only the nondichotomous case was considered in this study.

Derivation of Coefficient Alpha

Coefficient alpha can be derived in various ways, for example, derivations based on an ANOVA approach or a composite measures approach. For this study, the derivation based on composite measurements will be used. The derivation which follows is based on by Novick and Lewis (1967) and Crocker and Algina (1986).

The observed test score model in classical test theory is expressed as:

$$X = T + E \tag{1}$$

where X is the observed score, T is the true score, and E is the measurement error score. With this model it is assumed that the measurement error scores for an examinee are uncorrelated to that individual's true scores, the item error scores are uncorrelated, and the measurement error scores are expected to sum to zero over the population of examinees.

The reliability can be expressed as

$$\frac{\sigma_{T_C}^{2}}{\sigma_{C}^{2}} \qquad (2)$$

where $\sigma_{T_C}^{2}$ denotes the variance of the composite true score and

σ_{C}^{2} denotes the variance of the composite observed test score or the variance of the sum of **k** parallel subtest scores.

If all **k** parallel measures have equal true score variances and equal true score covariances and by definition, true score variance equals the sum of **k** X **k** matrix of true score variances and covariances, then:

$$\sigma_{T_C}^{2} = k\sigma^{2}(T_i) + k(k-1)\,\sigma(T_i, T_j) \qquad (3)$$

where $\sigma(T_i, T_j)$ denotes the covariance between T_i and T_j , and $\sigma^2(T_i)$ denotes the variance of T_i .

Based on the fact that $\sigma(T_i, T_j) = \sigma^2(T_i)$,

$$\sigma_{T_c}^2 = k\sigma(T_i, T_j) + k(k-1)\sigma(T_i, T_j) \tag{4}$$

or

$$\sigma_{T_c}^2 = k^2\sigma(T_i, T_j). \tag{5}$$

However, when the **k** subtests are not strictly parallel then

$$\sigma_{T_g}^2 \geq \sigma(T_i, T_g). \tag{6}$$

For nonparallel subtests at least one subtest (in this case subtest g) has a true score variance which is greater than or equal to its covariance with any other subtest. But for any two subtests that are not strictly parallel the sum of their true score variances is greater than or equal to twice their covariances,

$$\sigma_{T_i}^2 + \sigma_{T_j}^2 \geq 2\sigma(T_i, T_j). \tag{7}$$

In addition, with nonparallel subtests, the sum of **k** true score variances will be greater than or equal to the sum of **k(k-1)** covariances divided by **(k-1)**. This can be denoted as

$$\sum \sigma_{T_i}^2 \geq \frac{\sum_{i=1}^{k}\sum_{j=1}^{k}\sigma(T_i, T_j)}{k-1}, \ (i \neq j). \tag{8}$$

When the sum of the covariances is added to each side

of the equation then,

$$\sum \sigma_{T_i}^{2} + \sum_{i=1}^{k}\sum_{j=1}^{k} \sigma(T_i, T_j), \, (i \neq j) \geq \frac{\sum_{i=1}^{k}\sum_{j=1}^{k} \sigma(T_i, T_j)}{k-1} + \sum_{i=1}^{k}\sum_{j=1}^{k} \sigma(T_i, T_j), (i \neq j). \quad (9)$$

As seen in equation 3, the variance of the true score composite equals the sum of the variances of the true scores plus the sum of the covariances of the true scores. Therefore, equation (9) can be expressed as:

$$\sigma_{T_c}^{2} \geq \frac{k\sum_{i=1}^{k}\sum_{j=1}^{k} \sigma(T_i, T_j)}{k-1}, \, (i \neq j) \quad (10)$$

or

$$\sigma_{T_c}^{2} \geq \frac{k}{k-1}\sum_{i=1}^{k}\sum_{j=1}^{k} \sigma(T_i, T_j), \, (i, \neq j) \quad (11)$$

where $\sum_{i=1}^{k}\sum_{j=1}^{k} \sigma(T_i, T_j), i \neq j$ is the sum of **k(k-1)** covariances of non parallel subtests.

If both sides of equation (11) are divided by the variance of the composite then

$$\frac{\sigma_{T_c}^{2}}{\sigma_C^{2}} \geq \frac{k}{k-1}\left(\frac{\sum_{i=1}^{k}\sum_{j=1}^{k} \sigma(T_i, T_j)}{\sigma_C^{2}}\right), \, (i \neq j). \quad (12)$$

This is reliability since, by definition, reliability is the ratio of true score variance to observed score variance.

Since the true score covariance equals the observed score covariance,

$$\frac{\sigma_{T_C}^{2}}{\sigma_C^{2}} \geq \frac{k}{k-1}\left(\frac{\sum_{i=1}^{k}\sum_{j=1}^{k}\sigma(X_i, X_j)}{\sigma_C^{2}}\right), \ (i \neq j). \qquad (13)$$

In addition, it can be shown that the observed score variance equals the sum of the covariances and the item variances. Therefore,

$$\frac{\sigma_{T_C}^{2}}{\sigma_C^{2}} \geq \frac{k}{k-1}\left(\sigma_C^{2} - \sum \frac{\sigma_{X_i}^{2}}{\sigma_C^{2}}\right), \quad (i \neq j). \qquad (14)$$

Hence, coefficient alpha is the lower bound estimate of reliability.

Assumptions Underlying Coefficient Alpha

For the purposes of this study, the underlying assumptions of coefficient alpha can be categorized as assumptions underlying the derivation of alpha and assumptions underlying the estimation of alpha. A discussion of these categories follows.

Assumptions Underlying the Derivation.

Derivations of coefficient alpha, such as the derivation of Lord & Novick (1967), used a classical test theory approach to the derivation. This method does not assume a normal observed score distribution. However, an ANOVA approach to the derivation of coefficient alpha, such as the derivation presented by Feldt (1965), involves the assumptions of ANOVA which include normality. The assumptions underlying the derivation also include additivity (Novick & Lewis, 1967; Lord & Novick, 1968; Zimmerman, 1969; Zimmerman, Zumbo, & Lalonde, in press). Additivity means that the matrix of true scores must be additive in nature. Zimmerman, Zumbo, and Lalonde (in press) found that violation of the assumptions of uncorrelated errors between subtests and additivity resulted in a greater variability of the estimator. When additivity was violated, alpha underestimated the reliability. However, when the uncorrelated errors assumption was violated, alpha overestimated the reliability.

Assumptions Underlying the Estimation.

Although normality is not an assumption of the derivation of coefficient alpha presented above, it is an assumption of the estimation of coefficient alpha. The assumptions of estimation stem from least squares estimation. As seen in equation (15) coefficient alpha is a function of sample variances. Therefore, the assumptions of least squares estimation, which include normality, apply to coefficient alpha. This study will investigate how nonnormal and normal error score distributions affect the estimation of the population reliability using coefficient alpha.

Review of Studies

Commonly used approaches to most problems in classical test theory are based on statistics that are not robust or resistant to nonnormality. The terms "robust" or "resistant" are defined as insensitivity to changes in the underlying distribution (Huber, 1981; Mosteller & Tukey, 1977). For example, the variance is not robust or resistant to nonnormality since small deviations from normality can greatly affect its value (Lind & Zumbo, in press; Shoemaker & Hettmansperger, 1982). Consequently, Lind and Zumbo (in press) called for the investigation of the robustness or resistance of measures from classical test theory, such as coefficient alpha, to nonnormal data. A recent study by Wilcox (1992), developed a new measure of reliability which is a robust analogue of coefficient alpha. Wilcox (1992) derived this analogue based on the assumption that coefficient alpha is not robust to even slightly nonnormal observed score distributions. However, the robustness of coefficient alpha has not been studied in depth.

To the author's knowledge very little empirical or analytical research has been conducted to investigate the violation of the normality assumption underlying coefficient alpha. As indicated in the review by Feldt, Woodruff, and Salih (1987), a sampling theory of coefficient alpha would allow researchers to obtain an unbiased estimate of the population value, to establish confidence intervals for coefficient alpha, and to determine if coefficient alpha has a specific value for a given population.

The sampling distribution for coefficient alpha has not yet been determined. However, a transformation of coefficient alpha has been derived independently by Feldt (1965) and Kristof (1963). This transformation was based on an ANOVA derivation and has been proven to be distributed as an F distribution. Therefore, assumptions of least squares estimation, including normality, apply. In other words, if a normal distribution of observed scores is used, a transformation of coefficient alpha can be used to determine confidence intervals or to test hypotheses. The transformation of coefficient alpha is:

$$\frac{1 - \textit{population reliability}}{1 - \textit{coefficient } \alpha} .$$

Based on this transformation, Feldt et al. (1987) stated that coefficient alpha would tend to underestimate the population reliability and would be a biased estimator when the number of examinees is small (e.g. n = 50).

Feldt et al. (1987) presented a formula that would give an unbiased estimate of the population reliability:

$$[(N - 3)\ \alpha\ /\ (N - 1)] + 2\ /\ N - 1.$$

Bay (1973) also derived coefficient alpha using a mixed model ANOVA approach. He also derived the same transformation equation that was presented by Feldt (1965). Based on these derivations, Bay (1973) suggested that the sampling distribution of the reliability estimate would be robust against the violation of the normality assumption if (a) large numbers of examinees are used, (b) the reliability is close to zero, or (c) a large number of subtests is used and the true score or test score kurtosis is close to zero.

Bay (1973) also performed a Monte Carlo computer simulation. The simulation involved 30 examinees and eight subtests with 2000 replications. Six different true score distributions were used: normal, uniform, exponential, the sum of two independent uniform distributions, the sum of three independent uniform distributions and the sum of six independent uniform distributions. In addition, three different error score distributions were used, the normal, exponential, and uniform. The means, variances, and the mean squared errors of the alphas were obtained. The results of the computer simulation allowed Bay to conclude that a leptokurtic true score distribution could cause coefficient alpha to seriously underestimate the population reliability, and the effect of nonnormality of error score distributions is negligible when a large number of subtests is used.

A small component of Zimmerman, Zumbo, and Lalonde (in press) involved the estimation of coefficient alpha under a more general test score model than Bay. Forty examinees, ten subtests, and 2000 replications were used. The error score distributions used were the normal, uniform, exponential, and the mixed-normal. Reliability values of 0.65, 0.75, and 0.90 were used. Based on these parameters, Zimmerman et al. (in press) found that coefficient alpha was unbiased and its efficiency did not change over the distributions.

One difference between Bay (1973) and Zimmerman et al. (in press) should be noted. Bay (1973) varied the true score and the error score distributions while Zimmerman et al. (in press) only varied the error score distributions.

To date it has been shown that a transformation of coefficient alpha is distributed as a F distribution. This transformation is based on the assumption of normally distributed observed scores. Therefore, this transformation can only be used for hypothesis testing and to determine confidence intervals when the underlying observed score distribution is normal. In addition, the sampling distribution of coefficient alpha is unknown. If the sampling distribution of coefficient alpha were known, confidence intervals could be calculated from the data and hypotheses could be tested directly. This is similar to the commonly found discussion of computing confidence intervals of the mean based on the z-distribution.

Not only is the sampling distribution of coefficient alpha unknown, but also, the affects of nonnormality on the estimation of coefficient alpha have only been preliminarily tested by Bay (1973) and Zimmerman et al. (in press). Both Bay (1973) and Zimmerman et al. (in press) suggest that under certain conditions the estimation of reliability is robust. These conclusions are restricted, however, to the very limited parameters investigated. The effects of nonnormality with various numbers of examinees, numbers of subtests, and population reliabilities have not been examined.

Purpose of the Study

The purpose of this study was to examine the unbiasedness and efficiency of coefficient alpha under violation of the normality assumption underlying the estimation of coefficient alpha. In this way the robustness of coefficient alpha to nonnormal error score distributions was investigated.

More specifically, the following questions are posed. Are the biasedness and efficiency of coefficient alpha and the skewness and kurtosis of its sampling distribution affected by the:

(1) shape of the error score distribution?

(2) number of examinees?

(3) number of subtests?

(4) population reliability?

CHAPTER III

Methodology

A discussion of the methodology is presented in this chapter. More specifically, the independent variables, the Monte Carlo simulation methodology, the dependent variables, and the computer program used are discussed.

Independent Variables

The independent variables for this study were the number of examinees, the number of subtests, the population reliability, and the distribution of the error scores.

Number of Examinees and Number of Subtests

As stated in the review of the literature, two studies investigated the robustness of coefficient alpha. Bay (1973) used 30 examinees and eight subtests in his Monte Carlo study. Zimmerman et al. (in press) used 40 examinees and 10 subtests. The purpose of this study was to expand on the previous literature. The expansion involved increasing the number of independent variables to include the number of examinees, number of subtests, level of reliability, and error score distributions.

No written guidelines were found which suggested minimum values for the number of examinees and the number of subtests when coefficient alpha was being used to estimate the population reliability. Practitioners often state that a sizeable number of examinees (100 or more) is required when using coefficient alpha.

One way of ascertaining what levels of these variables should be investigated was to review current published articles which used coefficient alpha. In total 73 articles were reviewed. These articles often used multiple instruments and multiple sample sizes of examinees. A total of 122 samples of examinees and 210 samples of subtests were reviewed. The articles were published in 1991 and 1992 in Educational Psychology, British Journal of Educational Psychology, Psychological Assessment, Psychological Assessment: A Journal of Consulting and Clinical Psychology, Journal of Personality Assessment, Canadian Journal of Behavioural Science, and Personality and Social Psychology Bulletin. As seen in Tables 1 and 2, the majority of the samples were of less than 200 examinees and less than 30 subtests. Based on this information, the number of examinees used for this study were 25, 50, 100, and 200. The number of subtests used were 5, 10, 20, and 30. The spacing of the levels would permit

Table 1

Frequency of Number of Examinees for Recently Published Studies Using Coefficient Alpha

Number of Examinees	Frequency
1 to 25	3
26 to 50	11
51 to 75	24
76 to 100	16
101 to 200	38
201 to 300	9
301 to 400	9
401 to 500	2
501 to 600	5
601 to 700	1
701 to 800	1
801 to 900	1
OVER 901	2
TOTAL	122

Table 2

Frequency of Number of Subtests for Recently Published Studies Using Coefficient Alpha

Number of Subtests	Frequency
1 to 10	133
11 to 20	44
21 to 30	18
31 to 40	3
41 to 50	4
51 to 60	4
61 to 70	2
OVER 71	2
TOTAL	210

the detection of trends.

Reliability

The levels of the population reliability used were 0.4, 0.6, 0.8, and 0.9. These levels were chosen to determine the robustness of coefficient alpha at both high and low values of population reliability. The spacing of the levels was chosen so that trends across the levels could be detected.

Error Score Distributions

The levels of this independent variable, error score distribution, are similar to those in Zimmerman et al. (in press). Zimmerman et al.(in press) used normal, mixed normal, uniform, and exponential distributions. The present study used a negative exponential distribution instead of the uniform distribution. This change allowed this study to investigate the robustness of coefficient alpha for a highly negatively skewed distribution which was not tested in Zimmerman et al. (in press). Therefore, the four distributions used in this study were the normal, mixed normal, exponential, and negative exponential.

There were two reasons for including a normal distribution. First, empirical information about how normality affected coefficient alpha was limited to the situations tested in the Monte Carlo studies of Bay

(1973) and Zimmerman et al. (in press). Second, if coefficient alpha was shown to be affected by nonnormality, the results obtained by using the normal distribution could be used as a baseline against which the nonnormal distributions could be compared.

The mixed normal distribution was used since it is relatively normal in shape except it had a larger proportion of error scores in its tails. In addition, the mixed normal distribution has highly leptokurtic which would test Bay's (1973) conclusion that a kurtosis close to zero is required for coefficient alpha to be robust to nonnormality.

The exponential and negative exponential distributions were used because of their skewness. The exponential distribution is highly positively skewed while the negative exponential distribution is highly negatively skewed.

The result was a completely crossed four (number of examinees) by four (number of subtests) by four (alpha values) by four (error score distributions) design.

Monte Carlo Simulation

In the simulation used, the error scores of a given person over replications were distributed according to the various distributions. This implied that for any given replication, the error scores over examinees were distributed in the same way. The distributions which were used in this study were the following:

1. Normal Distribution Normal deviates were obtained by the method of Box and Muller (1958).
2. Exponential Distribution This highly positively skewed distribution was based on

$$X = -\log X_1 - 1.$$

3. Negative-Exponential Distribution This was a highly negatively skewed distribution based on

$$X = (\log X_1 - 1) + 2.$$

4. Mixed-Normal Distribution This distribution is also known as the contaminated normal. Sample variances were obtained from a normal distribution with a mean of zero and a variance of one, with a probability of 0.90, and from a normal distribution with a mean of zero and a variance of 100 with a probability of 0.10. The resulting distribution appeared bell-shaped but had a larger proportions of its area in the tails. Further details about this distribution are given in Zimmerman and Zumbo (1993). The normal deviates were generated by the Box-Muller (1958) method. The random number generator used in this study was

multiplicative congruential generator described by Lewis and Oraw (1989) and Lewis, Goodman, and Miller (1969).

This study utilized a computer program created by Zumbo and Zimmerman which involved the following steps:

Step 1: The population reliability, the number of subtests, and the number of examinees, as well as the error score distribution were input into the computer.

Step 2: A true score matrix was created so that the assumption of additivity was met. The true score matrix was created by constructing a matrix with test items on one axis and examinees on the other axis and adding the respective terms. In this way the true score distribution was uniform. Table 3 is an example of an additive true score matrix.

Table 3

Example of An Additive True Score Matrix

Examinee Number	Item Number			
	1	2	3	4
1	2	3	4	5
2	3	4	5	6
3	4	5	6	7
4	5	6	7	8

<u>Step 3</u>: The mean true score and its variance were calculated. Then using an error distribution with a mean of zero, independent error scores from the specified distribution were selected and added to the true score items. The variance of the error distribution was initially set to one and was then modified so that a specified reliability could be achieved.

<u>Step 4</u>: Coefficient alpha was calculated.

<u>Step 5</u>: The selection of an independent error score and the calculation of coefficient alpha were repeated to achieve 1000 replications.

<u>Step 6</u>: The mean, variance, mean squared error, skewness, and kurtosis of coefficient alpha over the 1000 replications were calculated. The 0.95 confidence interval of the mean was also calculated. The variance calculated is the variance from the sample mean of coefficient alpha while the MSE is the averaged squared deviation from the population value of alpha.

The methodology was an adaptation of the methodology used in Zimmerman et al. (in press). It is important to note that given the model in Zimmerman et al. (in press) the distribution of observed scores over replications has the same shape as the distribution of error scores over replications. Finally, in these simulations, the additivity and uncorrelated error assumptions were satisfied.

The methodology of this study is different than the methodology used by Bay (1973). Bay (1973) varied both the true score and the error score distribution. The present study only varied the error score distribution while the true score distribution remained uniformly distributed. It can be clearly seen by examining Table 3 where the total score across the four items is uniformly distributed for that sample.

Dependent Variables

The dependent variables for this study were biasedness, efficiency, skewness, and kurtosis of the sampling distribution of coefficient alpha over each set of 1000 replications.

Biasedness

Given that there is not a perfect estimator which always gives the correct answer, it is reasonable that an estimator should do so on the average. That is, it is desirable that the expected value or the average over a long series of replications of an estimator equals the parameter which it is supposed to estimate.

If this is the case the estimator is said to be unbiased. The calculation of the mean and its 0.95 confidence interval over each set of 1000 replications was used to determine if coefficient alpha was a biased or unbiased estimator of the population reliability. If the population reliability was within the 0.95 confidence interval, coefficient alpha was considered to be an unbiased estimator. However, if the population reliability was not within the 0.95 confidence interval then coefficient alpha was a biased estimator.

Efficiency

The efficiency of coefficient alpha was determined by the variance and mean squared error of the computed alphas over the 1000 replications. The variance indicates how much the estimation of coefficient alpha varied over the 1000 replications while the MSE indicates the accuracy of the estimation with respect to the population value. This study investigated how the independent variables affected the variance and MSE of coefficient alpha thereby affecting the efficiency of its estimation.

Skewness and Kurtosis

The use of 1000 replications resulted in a sampling distribution of coefficient alpha for each situation tested. The skewness and kurtosis of each sampling distribution was calculated. As a result, the trends in the skewness and kurtosis at the various levels of the independent variables were examined.

CHAPTER IV

Results

The results are presented in this chapter. The mean, 0.95 confidence interval, variance, mean squared error, skewness, and kurtosis for the computed alphas over the 1000 replications were calculated. The resulting data matrices from the 1000 replications are summarized in Appendix A (Tables 11 to 26).

In the text the results are presented for each of the dependent variables when there were 25 examinees, five subtests and the population reliability was 0.4.

Biasedness

Coefficient alpha was considered biased when the population reliability was not within the 0.95 confidence interval. The cases in which coefficient alpha was a biased estimator of the population reliability are presented in Table 4. Table 4 shows the 0.95 confidence interval for all subtest sizes and all distributions when the population reliability was 0.4 and the number of examinees were 25 and 50. When the number of examinees was 25, all distributions had cases which were biased. These cases differed from

Table 4

Confidence Intervals of Computed Coefficient Alphas By Error Score Distribution for 25 Subtests, 50 Examinees and Population Reliability of 0.4

	Exponential		Negative Exponential		Normal		Mixed Normal	
	N=25	N=50	N=25	N=50	N=25	N=50	N=25	N=50
k = 5	0.345, 0.394 *	0.363, 0.412	0.345, 0.394 *	0.364, 0.414	0.345, 0.395 *	0.358, 0.407	0.376, 0.426	0.380, 0.430
k = 10	0.336, 0.386 *	0.368, 0.418	0.350, 0.400	0.358, 0.408	0.343, 0.392 *	0.356, 0.410	0.367, 0.417	0.371, 0.420
k = 20	0.331, 0.381 *	0.356, 0.405	0.339, 0.389 *	0.352, 0.402	0.333, 0.383 *	0.359, 0.409	0.340, 0.389 *	0.366, 0.416
k = 30	0.354, 0.400	0.356, 0.405	0.354, 0.403	0.358, 0.408	0.346, 0.396 *	0.358, 0.408	0.344, 0.394 *	0.362, 0.411

* denotes that the population reliability of 0.40 did not fall within the 0.95 confidence interval.

distribution to distribution. When the error scores were normally distributed, coefficient alpha was a biased estimator for all levels of number of subtests. When the error scores were exponentially distributed, coefficient alpha was a biased estimator for all numbers of subtests except 30. In the same situation when a negatively exponential distribution of the error scores was used, a biased estimate occurred when 5 and 20 subtests were used. Lastly, when a mixed normal distribution of error scores was used, coefficient alpha was a biased estimator when 20 and 30 subtests were used. Of the four distributions used, the normal distribution performed the poorest.

When the number of examinees increased to 50 and the population reliability remained 0.4, coefficient alpha was unbiased for all distributions at all subtest sizes. In addition, when the population reliability values of 0.6, 0.8 and 0.9 were used, coefficient alpha was an unbiased estimator regardless of the number of subtests, the number of examinees or the distribution used. In other words the only times that coefficient alpha was a biased estimator occurred when the population reliability was 0.4 and there were 25 examinees. In all of these cases of biasedness, coefficient alpha underestimated the population reliability.

Efficiency

Efficiency of the estimation of coefficient alpha was measured by the variance and mean squared error. If the variance and mean squared error decreased as an independent variable increased, the efficiency of the estimation increased. If the variance and mean squared error decreased as one of the independent variables increased, the efficiency of the estimation decreased. Therefore, efficiency must be discussed for each of the independent variables. In addition, the way in which the efficiency changed as the population reliability and the number of examinees increased is discussed.

Of all of the distributions, the normal distribution tended to have variance and mean squared error values equal to or less than the other distributions (See Tables 5, 6, and 7). Therefore, the normal distribution has higher efficiency than the other distributions.

Throughout the data the normal error score distribution tended to be the most efficient and the mixed normal tended to be the least efficient. The remaining distributions tended to be between the normal and the mixed normal.

Table 5 shows the variance and mean squared error as the population reliability increased. The number of examinees was 25 and there were five subtests. For all distributions the variance and the mean squared error decreased as the population reliability increased. This result occurred at all levels of the number of examinees and the number of subtests. Therefore, the efficiency of coefficient alpha increased as the population reliability increased.

Table 5

Variance and Mean Squared Error By Population Reliability By Error Score Distribution for 25 Examinees and 5 Subtests

Distribution	Population Reliability	Variance	Mean Squared Error
Normal	0.4	0.042	0.043
	0.6	0.014	0.014
	0.8	0.002	0.002
	0.9	0.000	0.000
Mixed Normal	0.4	0.047	0.047
	0.6	0.026	0.026
	0.8	0.007	0.007
	0.9	0.002	0.002
Exponential	0.4	0.048	0.049
	0.6	0.018	0.018
	0.8	0.003	0.003
	0.9	0.001	0.001
Negative Exponential	0.4	0.045	0.046
	0.6	0.018	0.019
	0.8	0.003	0.003
	0.9	0.001	0.001

It is important to note, especially for the practitioner, that the confidence interval reported in Table 4 was based on 1000 replications and is only one way of determining the variance to expect. Another way would be to consider the variance for one particular combination of the independent variables. For example in Table 5 the variance for 25 examinees, 5 subtests, 0.4 population reliability, and a normal error score distribution was 0.042 and the standard deviation was 0.205. Given that the mean alpha in this case is 0.37 and assuming a normal sampling distribution of coefficient alpha, 68 percent of the expected alphas should fall within one standard deviation from the mean or between 0.17 and 0.57. This indicates a much larger variability than the confidence interval of the mean (Table 4). It is therefore suggested that practitioners consider the standard deviation especially when interpreting the results found at the lower levels of the independent variables.

The variance and mean squared error as the number of examinees increased are displayed in Table 6. The number of subtests was five while the population reliability was 0.4. For all distributions the variance and mean squared error decreased as the number

Table 6

Variance and Mean Squared Error By The Number of Examinees By Error Score Distribution for 5 Subtests and Population Reliability of 0.4

Distribution	No. of Examinees	Variance	Mean Squared Error
Normal	25	0.042	0.043
	50	0.018	0.018
	100	0.008	0.009
	200	0.004	0.004
Mixed Normal	25	0.047	0.047
	50	0.021	0.021
	100	0.011	0.011
	200	0.005	0.005
Exponential	25	0.048	0.049
	50	0.019	0.019
	100	0.008	0.008
	200	0.004	0.004
Negative Exponential	25	0.045	0.046
	50	0.020	0.020
	100	0.009	0.009
	200	0.004	0.004

of examinees increased. The same results were found regardless of the population reliability or the number of subtests used. Therefore, the efficiency of coefficient alpha increased as the number of examinees increased.

The variance and mean squared error as the number of subtests increased are shown in Table 7. The number

Table 7

Variance and Mean Squared Error By the Number of Subtests By Error Score Distribution for 25 Examinees and Population Reliability of 0.4

Distribution	No. of Subtests	Variance	Mean Squared Error
Normal	5	0.042	0.043
	10	0.039	0.040
	20	0.033	0.035
	30	0.031	0.032
Mixed Normal	5	0.047	0.047
	10	0.046	0.046
	20	0.043	0.044
	30	0.038	0.039
Exponential	5	0.048	0.049
	10	0.036	0.038
	20	0.039	0.041
	30	0.036	0.038
Negative Exponential	5	0.045	0.046
	10	0.036	0.037
	20	0.035	0.036
	30	0.034	0.034

of examinees was 25 and the population reliability was 0.4. For all distributions the variance and the mean squared error stayed the same or decreased slightly as the number of subtests increased. Similar results occurred for all population reliability values and for all levels of the number of examinees. Therefore, increasing the number of subtests had little effect on the efficiency of coefficient alpha.

The way in which the efficiency of the estimate changed as the population reliability and the number of examinees increased is shown in Figure 1. Figure 1 is composed of four graphs. Each graph plots the variance against the population reliability for each sample size at one level of the number of subtests for one of the four distributions. The way in which the variance decreases as the population reliability and the number of examinees increase can be seen. This pattern was similar at all number of subtests and for all distributions.

Skewness and Kurtosis

The skewness and kurtosis of the 1000 computed alphas as the population reliability increased are presented in Table 8. The number of examinees was 25 and there were five subtests. For the normal and exponential distributions, the skewness was less extreme as the population reliability increased. No pattern was seen for skewness of the negative exponential and the mixed normal distribution as the population reliability increased. For the normal and the exponential distributions the kurtosis decreased as the population reliability increased. There was no clear pattern for the kurtosis of the mixed normal and negative exponential distributions as the population reliability increased.

Figure 1

Variance as a Function of Reliability and Number of Subtests By Error Score Distribution

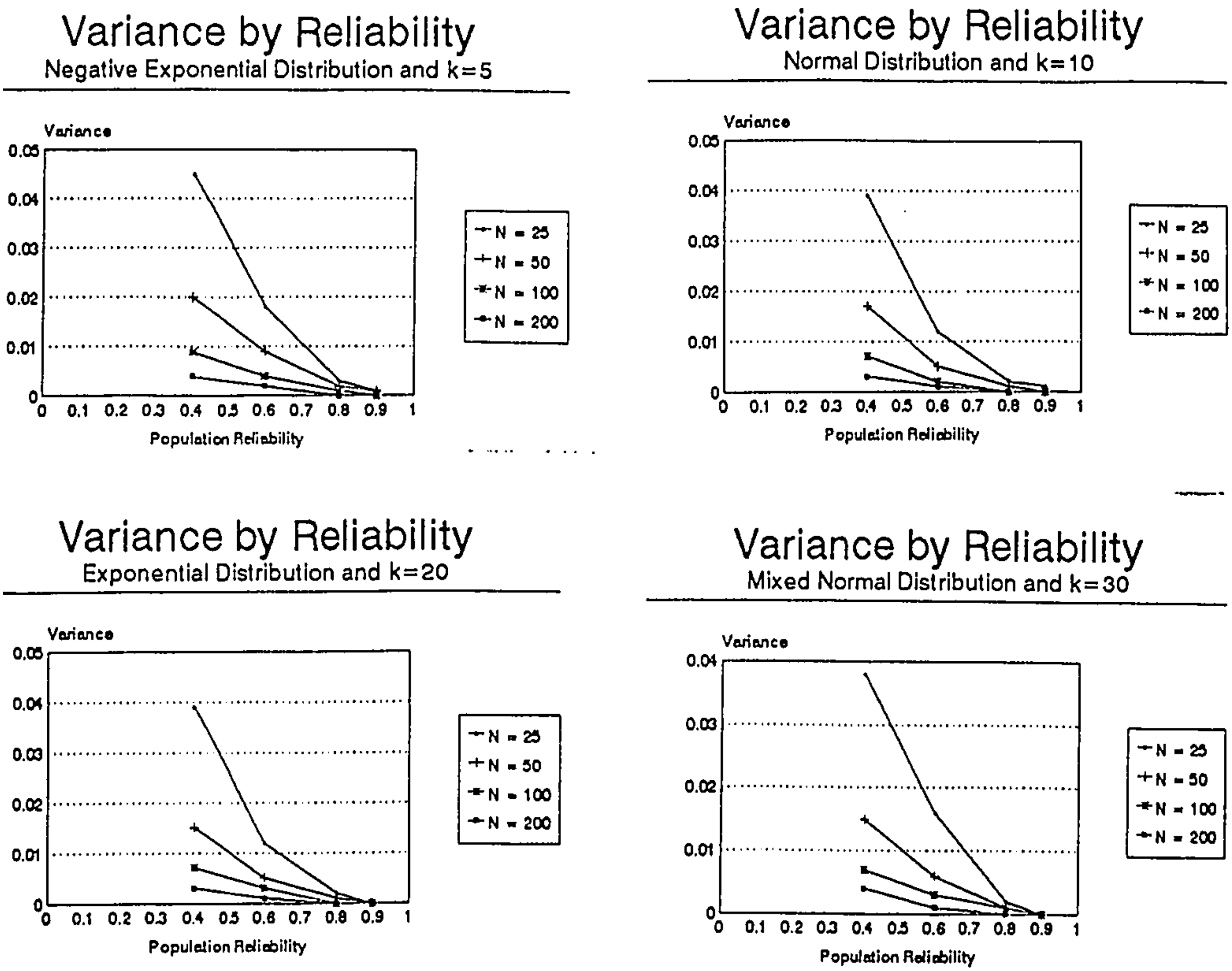

Table 8

Skewness and Kurtosis By Population Reliability By Error Score Distribution for 25 Examinees and 5 Subtests

Distribution	Population Reliability	Skewness	Kurtosis
Normal	0.4	-1.527	5.001
	0.6	-1.107	1.980
	0.8	-0.819	1.474
	0.9	-0.650	0.677
Mixed Normal	0.4	-0.930	1.331
	0.6	-1.166	2.405
	0.8	-0.891	0.989
	0.9	-0.920	1.919
Exponential	0.4	-1.339	2.799
	0.6	-1.181	2.094
	0.8	-1.073	2.063
	0.9	-0.988	1.568
Negative Exponential	0.4	-1.176	2.154
	0.6	-1.207	2.534
	0.8	-1.319	4.255
	0.9	-0.968	1.674

The skewness and kurtosis of the 1000 computed alphas as the number of examinees increased are provided in Table 9. There were five subtests and the population reliability was 0.4. For all distributions the skewness was less extreme as the number of examinees increased. For the normal and the negative exponential distributions the kurtosis decreased as the number of examinees increased. When the error scores

Table 9

Skewness and Kurtosis By the Number of Examinees By Error Score Distribution for 5 Subtests and Population Reliability of 0.4

Distribution	N	Skewness	Kurtosis
Normal	25	-1.527	5.001
	50	-0.907	1.127
	100	-0.648	0.764
	200	-0.374	0.390
Mixed Normal	25	-0.930	1.331
	50	-0.684	0.875
	100	-0.539	1.164
	200	-0.348	-0.063
Exponential	25	-1.339	2.799
	50	-0.660	0.609
	100	-0.334	-0.151
	200	-0.285	-0.090
Negative Exponential	25	-1.176	2.154
	50	-0.900	1.262
	100	-0.580	0.624
	200	-0.433	0.193

were exponentially distributed, the kurtosis tended to decrease until the 200 examinees were used, at which point there was a slight increase in the kurtosis. However, this increase in kurtosis brought the kurtosis closer to zero. For the mixed normal distribution, as the number of examinees increased no clear pattern for the kurtosis was seen.

In Table 9 the normal distribution tends to be more negatively skewed and more leptokurtic than the other distributions. This trend was not seen throughout the data.

Table 10 shows the skewness and kurtosis of the 1000 computed alphas for 25 examinees and a population reliability of 0.4 for all distributions and all number of subtests. No clear pattern was seen in the skewness or kurtosis as the number of subtests increased. Throughout the data there was an overall trend for skewness to increase and become less extreme and for kurtosis to decrease and

Table 10

Skewness and Kurtosis By Subtest Size By Error Score Distribution for 25 Examinees and Population Reliability of 0.4

Distribution	k	Skewness	Kurtosis
Normal	5	-1.527	5.001
	10	-1.471	4.554
	20	-1.194	2.250
	30	-1.244	2.664
Mixed Normal	5	-0.930	1.331
	10	-1.592	5.134
	20	-1.436	3.434
	30	-1.145	2.602
Exponential	5	-1.339	2.799
	10	-1.092	2.199
	20	-1.292	2.752
	30	-1.461	6.509
Negative Exponential	5	-1.176	2.154
	10	-1.130	2.050
	20	-1.070	1.600
	30	-1.184	2.090

become less leptokurtic as the number of examinees increased or the population reliability increased. Skewness and kurtosis tended to approach zero as the number of examinees increased. In addition, there was an overall trend for skewness to be negative while, to a lesser extent, kurtosis tended to be positive.

CHAPTER V

Discussion

This chapter includes a discussion of the results. The research questions are addressed and the significance of this study with respect to previous research is discussed. Since the levels of the independent variables used simulate what could occur in actual testing situations, the implications of the results for users of coefficient alpha will be discussed in the significance of the study section. In addition, the limitations of this study are discussed along with suggestions for future study.

Error Score Distributions

Four error score distributions, the normal, mixed normal, negative exponential, and exponential were used in this study. The mixed normal distribution was similar to the normal except that it had a higher proportion of error scores within its tails, thereby bilaterally lengthening the tails and making them more pronounced. The negative exponential distribution was negatively skewed while the exponential distribution was positively skewed. Some differences among these distributions did occur in relation to the dependent variables. Surprisingly, in the cases of biasedness, the normal error score distribution performed the poorest. Because of the lack of consistency, it is difficult to arrive at any conclusions for the three other distributions. For the exponential error score distribution, coefficient alpha appeared to become an unbiased estimator as the number of subtests increased. On the other hand, for the mixed normal error score distribution, coefficient alpha appeared to become a biased estimator as the number of subtests increased. For the negative exponential error score distribution no pattern was seen. For the biased cases why did coefficient alpha tend to be a biased estimator when the error scores were normally distributed? This may have occurred because the normal distribution was also the most efficient which would tend to produce a smaller confidence interval. However, the size of the confidence intervals in Table 4 did not vary considerably across the distributions.

Some additional speculation can be made as to why the results occurred. One possibility is that for smaller numbers of examinees, perhaps the margin of error was larger. However, this margin of error should have been relatively consistent across the different distributions.

With respect to efficiency, the differences among the distributions were more noticeable at the lower levels of the independent variables, number of examinees and population reliability. The variance and mean squared error approached zero for all of the distributions as the other independent variables increased. However, throughout the data, when the error scores were normally distributed, coefficient alpha was the most efficient while the mixed normal error score distribution tended to be the least efficient.

Although the normal distribution tended be more skewed and more leptokurtic than the other distributions, this trend only occurred when 25 examinees, five subtests and a population reliability of 0.4 were used. Therefore, this result applied to this situation only.

As the population reliability increased, skewness became less extreme when a normal, exponential, or negative exponential distribution was used. No pattern was seen for the mixed normal error score distributions. When the number of examinees increased, the skewness became less extreme regardless of the error score distribution. It is interesting to note that as the number of subtests increased no pattern for skewness could be seen for any of the error score distributions.

The sampling distribution tended to become less leptokurtic for the normal and exponential error score distributions as the population reliability increased. No patterns could be seen for the negative exponential and the mixed normal distributions. When the number of examinees increased, the sampling distribution tended to become less leptokurtic for all distributions except the mixed normal which had no discernable pattern. As was the case with skewness, no patterns in kurtosis were seen as the number of subtests increased.

Therefore, for the normal and exponential distributions, the sampling distribution became less skewed and less leptokurtic as the population reliability increased. The same pattern occurred for the normal, negative exponential, and the exponential distributions when the number of examinees increased. It is interesting to note that no pattern was seen for any of the distributions as the number of subtests increased. Overall, the mixed normal tended to have no discernable pattern as the independent variables increased.

Number of Examinees

In some instances when 25 examinees were used and the population reliability was 0.4, coefficient alpha was a biased estimator of the population reliability. However, when the number of examinees was increased to 50, coefficient alpha became an unbiased estimator.

Based on these results, users of coefficient alpha could consider coefficient alpha an unbiased estimator of the population reliability if 50 or more examinees are used. Interestingly, Feldt et al. (1987) suggested that coefficient alpha would be a biased estimator if 50 examinees were used. Further research to confirm and expand upon the results of this present study is required before firm conclusions can be made.

Throughout the data, the efficiency of the estimation tended to increase as the number of examinees increased regardless of the error score distribution. In addition, as the number of examinees increased, the sampling distribution tended to become less skewed and less leptokurtic for all error score distributions.

Population Reliability

The results as the population reliability increased are similar to the results as the number of examinees increased. When the population reliability was 0.4 and there were 25 examinees, coefficient alpha was a biased estimator of the population reliability. However, when the population reliability was increased to 0.6, coefficient alpha became an unbiased estimator of the population reliability.

Throughout the data, for all of the error score distributions, the efficiency of the estimation tended to increase as the population reliability increased. In addition, the sampling distribution tended to become less skewed and less leptokurtic for the normal and the negative exponential distributions as the population reliability increased.

Number of Subtests

In the cases of biasedness, no pattern could be seen as the number of subtests increased. In addition, increasing the number of subtests had little effect on the efficiency of coefficient alpha. Also, no pattern occurred in the skewness and kurtosis of the sampling distribution as the number of subtests increased.

These results were surprising. Would the same results occur if dichotomous data, which would yield relatively larger error scores than nondichotomous data, were used? Is the present belief that a large number of subtests is required for an accurate estimation of population reliability valid? Further research is required before these questions can be addressed.

Research Questions

The research questions which were the basis of this study can now be answered based on the results obtained. The results show that given 50, 100 or 200 examinees, 5, 10, 20, or 30 subtests, and population reliability values of 0.4, 0.6, 0.8, or 0.9, coefficient alpha was an unbiased estimator of the population reliability. Coefficient alpha remained robust to nonnormal error score distributions except in some cases when the population reliability was 0.4 and the number of examinees was 25. However, in all of these exceptions coefficient alpha underestimated the population reliability. The skewness and kurtosis of the sampling distribution were affected by some of the independent variables. For the normal and the exponential distributions there was a tendency for the sampling distribution to be less skewed and less leptokurtic as the population reliability increased. For the normal, exponential, and negative exponential distributions, the sampling distribution tended to become less skewed and less leptokurtic as the number of examinees increased. As the number of subtests increased, there was no clear pattern for the skewness or kurtosis. Of all of the distributions used the mixed normal was most likely to have no clear pattern in the skewness and kurtosis.

Therefore, the biasedness and efficiency of coefficient alpha and the skewness and kurtosis of its sampling distribution were effected by the shape of the error score distribution, the number of examinees and the population reliability. In addition, the biasedness and efficiency did not appear effected by the number of subtests.

Significance of the Study

The results of this study expand upon the work of Zimmerman et al. (in press) who found that for forty examinees, ten subtests, and population reliability values of 0.65, 0.75, and 0.90 with uniform, normal, exponential, and mixed normal error score distributions, coefficient alpha was an unbiased estimator and its efficiency did not change over the distributions. In comparison this study also found that coefficient alpha was an unbiased estimator for 50, 100, and 200 examinees with 5, 10, 20, and 30 subtests at population reliability values of 0.4, 0.6, 0.8, and 0.9 with underlying error score distributions that were normal, exponential, negative exponential, and mixed normal. However, when there were 25 examinees, coefficient alpha was an unbiased estimator except when the population reliability was 0.4. The efficiency of coefficient alpha as an estimator of the population reliability was found to improve as the number of examinees and population reliability increased.

In order to compare Bay's (1973) conclusions based on his Monte Carlo results to the results of the present study, the methodologies should be compared. In Bay's (1973) Monte Carlo study, he varied both the true score distributions and the error score distributions. The present study varied the error score distributions but the true score distributions were always uniformly distributed. Therefore, the true score distributions were platykurtic. Based on his Monte Carlo study, Bay (1973) concluded that a leptokurtic true score distribution could cause coefficient alpha to seriously underestimate the population reliability, and that the effect of a nonnormal error score distribution would be negligible if a large number of subtests were used. Since the true score distributions for the present study were always distributed uniformly, the above conclusion was not tested. However, the present study did test Bay's (1973) conclusion that the effect of a nonnormal error score distribution would be negligible if a large number of subtests were used. The present study found that the number of subtests did not tend to improve the unbiasedness or the efficiency of coefficient alpha. Given its methodology, the present study showed that a large number of subtests was not required for the robustness of coefficient alpha.

This study has contributed to the literature on the robustness of coefficient alpha. It has resulted in more questions than it has answered. Given the methodology used, this study has shown that coefficient alpha is robust to nonnormality. It has also shown that coefficient alpha is an unbiased estimator of the population reliability for 50 examinees or more and in some cases for as small as 25 examinees.

Although generalizations should not be based on only one study, some implications can be made for users of coefficient alpha. Based on the results of this study, coefficient alpha would be an unbiased and efficient estimator of the population reliability value if 50 or more examinees are used. The use of a test which has been shown to yield a high estimate of reliability would also improve the robustness of coefficient alpha to nonnormality. Given the levels of the independent variables used, coefficient alpha was robust to nonnormality, although the normal distribution was the most efficient. However, further research is suggested to verify these results and expand upon them. Specifically, future research should investigate the robustness of coefficient alpha given real data which is not normally distributed.

Another way in which this study has contributed to the literature is by suggesting future research based on the limitations of this study.

Limitation

There are many different ways of constructing true score matrices which meet the additivity assumption. The true score matrix used in this study was created in such a way that the true scores were distributed uniformly. Micceri (1989) found that uniform observed score distributions were rare in large-sample achievement and psychometric measures. Also in practical usage a uniform distribution of observed scores is not an expected distributional shape. Therefore, the use of uniform true score distributions in this study is likely unrealistic and restricts the generalizability of these results. Perhaps a better way of studying the robustness of coefficient alpha would be to vary the true score and the error score distributions.

In conclusion, this study has empirically examined the robustness of coefficient alpha. Based on the results and limitations of this study, suggestions for future study have been made.

REFERENCES

Bay, K.S. (1973). The effect of non-normality on the sampling distribution and standard error or reliability coefficient estimates under an analysis of variance model. British Journal of Mathematical and Statistical Psychology, 26, 45-57.

Box, G.E.P. & Muller, M. (1958). A note on the generation of random normal deviates. Annals of Mathematical Statistics, 29, 610-611.

Crocker, L. & Algina, J. (1986). Introduction to classical and modern test theory. Fort Worth: Holt, Rinehart and Winston, Inc.

Feldt, L.S. (1969). A test of the hypothesis that Cronbach's alpha or Kuder-Richardson coefficient twenty is the same for two tests. Psychometrika, 34(3), 363-373.

Feldt, L.S. (1980). A test of the hypothesis that Cronbach's alpha reliability coefficient is the same for two tests administered to the same sample. Psychometrika, 45(1), 99-105.

Feldt, L.S., Woodruff, D.J., & Salih, F.A. (1987). Statistical inference for coefficient alpha. Applied Psychological Measurement, 93-103.

Huber, P.J. (1981). Robust Statistics. New York: Wiley.

Kristof, W. (1963). The statistical theory of stepped-up reliability coefficients when a test has been divided into several equivalent parts. Psychometrika, 28, 221-238.

Lewis, P.A.W., Goodman, A.S., & Miller, J.M. (1969). A pseudorandom number generator for the system 360. IBM Systems Journal, 8, 136-146.

Lewis, P.A.W. & Oraw, E.J. (1989). Simulation methodology for statisticians, operations analysts, and engineers: Vol. 7. Pacific Grove, California: Wadsworth.

Lind, J.C. & Zumbo, B.D. (in press). The continuity principle in psychological research: An introduction to robust statistics. Canadian Psychology.

Lord, F. & Novick, M. (1968). Statistical theories of mental test scores. Reading, Mass.: Addison-Wesley.

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

Mosteller, F. & Tukey, J.W. (1977). Data analysis and regression. Reading, Mass.: Addison-Wesley.

Novick, M.R. & Lewis, C. (1967). Coefficient alpha and the reliability of composite measurements. Psychometrika, 32(1), 1-13.

Shoemaker, L.H. & Hettmansperger, T.P. (1982). Robust estimates and tests for the one- and two-sample problem. Biometrika, 69, 47-53.

Wilcox, R.R. (1992). Robust generalizations of classical test reliability and Cronbach's alpha. British Journal of Mathematical and Statistical Psychology, 45, 239-254.

Wilcox, R.R. (1990). Comparing variances and means when distributions have non-identical shapes. Communications in Statistics--Simulation and Computation, 19, 155-173.

Woodruff, D.J. & Feldt, L.S. (1986). Tests for equality of several alpha coefficients when their sample estimates are dependent. Psychometrika, 51(3), 393-413.

Zimmerman, D.W. (1969). An item sampling model for the reliability of composite test. Educational and Psychological Measurement, 29, 49-59.

Zimmerman, D.W. & Zumbo, B.D. (1993). Relative power of parametric and nonparametric statistical methods. In G. Keren & C. Lewis (Eds.), A Handbook for data analysis in the behavioural sciences (pp. 481-517). Hillsdale, N.J.: Lawrence Erlbaum.

Zimmerman, D.W., Zumbo, B.D., & Lalonde, C. (1993). Coefficient alpha as an estimate of test reliability under violation of two assumptions. Educational and Psychological Measurement, 53, 33-49.

Zumbo, B.D. (in press). The lurking assumptions in using generalizability theory to monitor individual's progress. In D. Laveault, B.D. Zumbo, M. Gessaroli, and M. Boss (Eds.), Modern theories of measurement: Problems and issues. Ottawa, Ontario: University of Ottawa Press.

APPENDIX A: Summary Data for all Combinations of Independent Variables

Table 11

Summary Data for Exponential Error Score Distribution and 25 Examinees

k	Stats	alpha			
		0.40	0.60	0.80	0.90
5	Mean	0.370	0.589	0.806	0.902
	Variance	0.048	0.018	0.003	0.001
	MSE	0.049	0.018	0.003	0.001
	Skewness	-1.339	-1.181	-1.073	-0.988
	Kurtosis	2.799	2.094	2.063	1.568
	95% CI	0.345, 0.394	0.551, 0.626	0.757, 0.856	0.847, 0.958
10	Mean	0.361	0.589	0.803	0.903
	Variance	0.036	0.014	0.002	0.000
	MSE	0.038	0.014	0.002	0.000
	Skewness	-1.092	-1.172	-0.801	-0.753
	Kurtosis	2.199	2.644	1.007	1.095
	95% CI	0.336, 0.386	0.552, 0.626	0.753, 0.852	0.847, 0.959
20	Mean	0.356	0.588	0.803	0.902
	Variance	0.039	0.012	0.002	0.000
	MSE	0.041	0.012	0.002	0.000
	Skewness	-1.292	-1.109	-0.854	-0.793
	Kurtosis	2.752	1.576	1.954	2.188
	95% CI	0.331, 0.381	0.551, 0.625	0.753, 0.853	0.846, 0.958
30	Mean	0.358	0.590	0.799	0.903
	Variance	0.036	0.003	0.001	0.000
	MSE	0.038	0.011	0.002	0.000
	Skewness	-1.461	-1.328	-0.820	-0.586
	Kurtosis	6.509	3.985	0.826	0.544
	95% CI	0.354, 0.400	0.551, 0.626	0.752, 0.851	0.847, 0.959

Table 12
Summary Data for Exponential Error Score Distribution and 50 Examinees

k	Stats	alpha			
		0.40	0.60	0.80	0.90
5	Mean	0.387	0.599	0.801	0.902
	Variance	0.019	0.007	0.002	0.000
	MSE	0.019	0.007	0.002	0.000
	Skewness	-0.660	-0.770	-0.694	-0.578
	Kurtosis	0.609	1.123	0.787	0.578
	95% CI	0.363, 0.412	0.562, 0.636	0.751, 0.851	0.846, 0.958
10	Mean	0.393	0.599	0.801	0.902
	Variance	0.016	0.005	0.001	0.000
	MSE	0.017	0.005	0.000	0.000
	Skewness	-0.906	-0.619	-0.562	-0.527
	Kurtosis	1.719	0.800	0.186	0.441
	95% CI	0.368, 0.418	0.562, 0.636	0.751, 0.850	0.846, 0.958
20	Mean	0.380	0.596	0.800	0.901
	Variance	0.015	0.005	0.001	0.000
	MSE	0.015	0.005	0.001	0.000
	Skewness	-0.725	-0.624	-0.467	-0.727
	Kurtosis	0.588	0.451	0.536	1.256
	95% CI	0.356, 0.405	0.559, 0.634	0.750, 0.849	0.845, 0.957
30	Mean	0.381	0.592	0.799	0.901
	Variance	0.015	0.005	0.001	0.000
	MSE	0.015	0.005	0.001	0.000
	Skewness	-0.775	-0.776	-0.654	-0.509
	Kurtosis	1.204	1.544	0.797	0.943
	95% CI	0.356, 0.405	0.555, 0.630	0.749, 0.848	0.845, 0.957

Table 13

Summary Data for Exponential Error Score Distribution and 100 Examinees

k	Stats	alpha			
		0.40	0.60	0.80	0.90
5	Mean	0.397	0.598	0.801	0.900
	Variance	0.008	0.004	0.001	0.000
	MSE	0.008	0.004	0.001	0.000
	Skewness	-0.334	-0.534	-0.457	-0.541
	Kurtosis	-0.151	0.568	0.494	1.163
	95% CI	0.372, 0.422	0.560, 0.635	0.751, 0.850	0.844, 0.956
10	Mean	0.392	0.597	0.800	0.901
	Variance	0.008	0.003	0.001	0.000
	MSE	0.008	0.003	0.001	0.000
	Skewness	-0.723	-0.604	-0.484	-0.393
	Kurtosis	0.858	0.688	0.541	0.111
	95% CI	0.367, 0.417	0.560, 0.634	0.750, 0.849	0.845, 0.957
20	Mean	0.391	0.599	0.801	0.900
	Variance	0.007	0.003	0.000	0.000
	MSE	0.007	0.003	0.000	0.000
	Skewness	-0.601	-0.492	-0.537	-0.249
	Kurtosis	0.801	0.289	0.361	0.137
	95% CI	0.366, 0.416	0.561, 0.636	0.751, 0.850	0.844, 0.956
30	Mean	0.391	0.597	0.801	0.900
	Variance	0.007	0.002	0.000	0.000
	MSE	0.007	0.002	0.000	0.000
	Skewness	-0.616	-0.634	-0.435	-0.256
	Kurtosis	0.755	0.638	0.408	0.151
	95% CI	0.366, 0.415	0.560, 0.634	0.751, 0.850	0.844, 0.956

Table 14

Summary Data for Exponential Error Score Distribution and 200 Examinees

k	Stats	alpha			
		0.40	0.60	0.80	0.90
5	Mean	0.394	0.599	0.801	0.901
	Variance	0.004	0.002	0.000	0.000
	MSE	0.004	0.002	0.000	0.000
	Skewness	-0.285	-0.152	-0.419	-0.329
	Kurtosis	-0.090	-0.033	0.114	-0.088
	95% CI	0.370, 0.419	0.562, 0.636	0.751, 0.851	0.845, 0.956
10	Mean	0.399	0.599	0.800	0.900
	Variance	0.004	0.001	0.000	0.000
	MSE	0.004	0.001	0.000	0.000
	Skewness	-0.319	-0.327	-0.290	-0.180
	Kurtosis	0.091	0.154	-0.297	-0.183
	95% CI	0.375, 0.424	0.562, 0.636	0.750, 0.850	0.844, 0.956
20	Mean	0.401	0.598	0.801	0.900
	Variance	0.003	0.001	0.00	0.000
	MSE	0.003	0.001	0.000	0.000
	Skewness	-0.263	-0.392	-0.233	-0.230
	Kurtosis	-0.108	0.303	0.015	0.148
	95% CI	0.376, 0.426	0.561, 0.635	0.751, 0.850	0.844, 0.956
30	Mean	0.397	0.598	0.800	0.900
	Variance	0.003	0.001	0.000	0.000
	MSE	0.003	0.001	0.000	0.000
	Skewness	-0.253	-0.358	-0.461	-0.356
	Kurtosis	0.121	-0.085	0.258	0.209
	95% CI	0.372, 0.422	0.562, 0.636	0.750, 0.849	0.844, 0.956

Table 15
Summary Data for Normal Error Score Distribution and 25 Examinees

k	Stats	alpha			
		0.40	0.60	0.80	0.90
5	Mean	0.370	0.585	0.798	0.901
	Variance	0.042	0.014	0.002	0.000
	MSE	0.043	0.014	0.002	0.000
	Skewness	-1.527	-1.107	-0.819	-0.650
	Kurtosis	5.001	1.980	1.474	0.677
	95% CI	0.345, 0.395	0.547, 0.622	0.748, 0.847	0.845, 0.957
10	Mean	0.368	0.584	0.800	0.903
	Variance	0.039	0.012	0.002	0.000
	MSE	0.040	0.012	0.002	0.000
	Skewness	-1.471	-0.975	-0.741	-0.936
	Kurtosis	4.554	1.514	0.629	2.048
	95% CI	0.343, 0.392	0.547, 0.621	0.750, 0.850	0.847, 0.958
20	Mean	0.358	0.591	0.800	0.902
	Variance	0.033	0.010	0.001	0.000
	MSE	0.035	0.010	0.000	0.000
	Skewness	-1.194	-0.937	-0.744	-0.464
	Kurtosis	2.250	1.619	0.736	0.239
	95% CI	0.333, 0.383	0.554, 0.628	0.751, 0.850	0.846, 0.958
30	Mean	0.371	0.585	0.800	0.902
	Variance	0.031	0.011	0.001	0.000
	MSE	0.032	0.011	0.001	0.000
	Skewness	-1.244	-1.060	-0.664	-0.770
	Kurtosis	2.664	2.392	0.847	1.614
	95% CI	0.346, 0.396	0.548, 0.622	0.751, 0.850	0.846, 0.958

Table 16

Summary Data for Normal Error Score Distribution and 50 Examinees

k	Stats	alpha			
		0.40	0.60	0.80	0.90
5	Mean	0.383	0.590	0.800	0.901
	Variance	0.018	0.006	0.001	0.000
	MSE	0.018	0.006	0.001	0.000
	Skewness	-0.907	-0.681	-0.423	-0.338
	Kurtosis	1.127	0.656	0.291	0.292
	95% CI	0.358, 0.407	0.553, 0.627	0.750, 0.849	0.845, 0.957
10	Mean	0.381	0.595	0.802	0.902
	Variance	0.017	0.005	0.001	0.000
	MSE	0.017	0.005	0.000	0.000
	Skewness	-0.978	-0.891	-0.378	-0.422
	Kurtosis	1.666	1.775	0.343	0.290
	95% CI	0.356, 0.410	0.557, 0.632	0.752, 0.851	0.852, 0.951
20	Mean	0.384	0.594	0.800	0.901
	Variance	0.014	0.005	0.001	0.000
	MSE	0.014	0.005	0.000	0.000
	Skewness	-0.753	-0.702	-0.552	-0.190
	Kurtosis	0.753	1.053	0.567	0.115
	95% CI	0.359, 0.409	0.557, 0.631	0.751, 0.850	0.845, 0.957
30	Mean	0.383	0.594	0.801	0.901
	Variance	0.013	0.005	0.001	0.000
	MSE	0.014	0.005	0.000	0.000
	Skewness	-0.694	-0.746	-0.706	-0.350
	Kurtosis	0.760	1.004	0.696	-0.017
	95% CI	0.358, 0.408	0.557, 0.631	0.751, 0.850	0.845, 0.957

Table 17

Summary Data for Normal Distribution and 100 Examinees

k	Stats	alpha			
		0.40	0.60	0.80	0.90
5	Mean	0.388	0.597	0.801	0.901
	Variance	0.008	0.003	0.000	0.000
	MSE	0.009	0.003	0.000	0.000
	Skewness	-0.648	-0.392	-0.475	-0.311
	Kurtosis	0.764	0.154	0.749	0.224
	95% CI	0.363, 0.413	0.560, 0.635	0.751, 0.850	0.845, 0.957
10	Mean	0.391	0.598	0.800	0.900
	Variance	0.007	0.002	0.000	0.000
	MSE	0.007	0.002	0.000	0.000
	Skewness	-0.504	-0.427	-0.615	-0.301
	Kurtosis	0.063	0.308	1.027	-0.062
	95% CI	0.366, 0.416	0.560, 0.635	0.750, 0.849	0.844, 0.956
20	Mean	0.392	0.599	0.800	0.901
	Variance	0.007	0.002	0.000	0.000
	MSE	0.007	0.002	0.000	0.000
	Skewness	-0.515	-0.391	-0.399	-0.280
	Kurtosis	0.726	0.106	0.282	0.012
	95% CI	0.367, 0.417	0.562, 0.636	0.751, 0.850	0.845, 0.956
30	Mean	0.394	0.598	0.801	0.900
	Variance	0.006	0.002	0.000	0.000
	MSE	0.006	0.002	0.000	0.000
	Skewness	-0.687	-0.459	-0.425	-0.274
	Kurtosis	1.250	0.403	0.686	0.148
	95% CI	0.369, 0.419	0.560, 0.635	0.751, 0.851	0.844, 0.956

Table 18

Summary Data for Normal Error Score Distribution and 200 Examinees

k	Stats	alpha			
		0.40	0.60	0.80	0.90
5	Mean	0.396	0.598	0.801	0.900
	Variance	0.004	0.001	0.000	0.000
	MSE	0.004	0.001	0.000	0.000
	Skewness	-0.374	-0.287	-0.256	0.062
	Kurtosis	0.390	0.043	0.061	0.183
	95% CI	0.371, 0.421	0.561, 0.635	0.751, 0.850	0.844, 0.956
10	Mean	0.393	0.597	0.800	0.900
	Variance	0.003	0.001	0.000	0.000
	MSE	0.003	0.001	0.000	0.000
	Skewness	-0.269	-0.414	-0.300	-0.211
	Kurtosis	0.465	0.219	0.188	-0.017
	95% CI	0.368, 0.417	0.560, 0.634	0.750, 0.850	0.844, 0.956
20	Mean	0.396	0.598	0.800	0.900
	Variance	0.003	0.001	0.000	0.000
	MSE	0.003	0.001	0.000	0.000
	Skewness	-0.322	-0.414	-0.306	0.043
	Kurtosis	0.153	0.336	0.097	-0.114
	95% CI	0.371, 0.421	0.561, 0.635	0.750, 0.849	0.845, 0.956
30	Mean	0.394	0.599	0.800	0.900
	Variance	0.004	0.001	0.000	0.000
	MSE	0.004	0.001	0.000	0.000
	Skewness	-0.339	-0.293	-0.141	-0.165
	Kurtosis	0.035	-0.070	-0.197	0.419
	95% CI	0.369, 0.419	0.562, 0.636	0.751, 0.850	0.844, 0.956

Table 19

Summary Data for Negative Exponential Error Score Distribution and 25 Examinees

k	Stats	alpha			
		0.40	0.60	0.80	0.90
5	Mean	0.370	0.585	0.800	0.903
	Variance	0.045	0.018	0.003	0.001
	MSE	0.046	0.019	0.003	0.001
	Skewness	-1.176	-1.207	-1.319	-0.968
	Kurtosis	2.154	2.534	4.255	1.674
	95% CI	0.345, 0.394	0.547, 0.622	0.750, 0.850	0.847, 0.959
10	Mean	0.375	0.596	0.805	0.902
	Variance	0.036	0.013	0.002	0.000
	MSE	0.037	0.013	0.002	0.000
	Skewness	-1.130	-0.969	-0.769	-0.893
	Kurtosis	2.050	1.492	0.940	1.613
	95% CI	0.350, 0.400	0.559, 0.634	0.755, 0.854	0.846, 0.958
20	Mean	0.364	0.591	0.799	0.903
	Variance	0.035	0.012	0.002	0.000
	MSE	0.036	0.012	0.002	0.000
	Skewness	-1.070	-1.030	-0.841	-0.763
	Kurtosis	1.600	1.470	0.909	0.896
	95% CI	0.339, 0.389	0.554, 0.628	0.750, 0.849	0.847, 0.959
30	Mean	0.379	0.589	0.801	0.902
	Variance	0.034	0.011	0.002	0.000
	MSE	0.034	0.011	0.002	0.000
	Skewness	-1.184	-1.082	-0.909	-0.492
	Kurtosis	2.090	1.975	1.761	0.317
	95% CI	0.354, 0.403	0.551, 0.626	0.752, 0.851	0.847, 0.958

Table 20
Summary Data for Negative Exponential Error Score Distribution and 50 Examinees

k	Stats	alpha			
		0.40	0.60	0.80	0.90
5	Mean	0.389	0.599	0.801	0.900
	Variance	0.020	0.009	0.002	0.000
	MSE	0.020	0.009	0.002	0.000
	Skewness	-0.900	-1.012	-0.818	-0.510
	Kurtosis	1.262	2.426	1.152	0.391
	95% CI	0.364, 0.414	0.562, 0.637	0.751, 0.851	0.844, 0.956
10	Mean	0.383	0.596	0.800	0.901
	Variance	0.018	0.006	0.001	0.000
	MSE	0.018	0.006	0.001	0.000
	Skewness	-0.945	-0.880	-0.777	-0.567
	Kurtosis	1.573	1.677	1.439	0.464
	95% CI	0.358, 0.408	0.558, 0.633	0.750, 0.850	0.845, 0.957
20	Mean	0.377	0.597	0.800	0.901
	Variance	0.018	0.005	0.001	0.000
	MSE	0.018	0.005	0.001	0.000
	Skewness	-1.058	-0.754	0.583	-0.498
	Kurtosis	2.360	1.579	0.723	0.532
	95% CI	0.352, 0.402	0.560, 0.634	0.750, 0.850	0.845, 0.957
30	Mean	0.383	0.595	0.800	0.902
	Variance	0.015	0.005	0.001	0.000
	MSE	0.015	0.005	0.001	0.000
	Skewness	-0.802	-0.647	-0.639	-0.395
	Kurtosis	1.146	0.812	1.015	0.247
	95% CI	0.358, 0.408	0.558, 0.633	0.750, 0.849	0.846, 0.958

Table 21
Summary Data for Negative Exponential Error Score Distribution and 100 Examinees

k	Stats	alpha			
		0.40	0.60	0.80	0.90
5	Mean	0.393	0.597	0.800	0.900
	Variance	0.009	0.004	0.001	0.000
	MSE	0.009	0.004	0.001	0.000
	Skewness	-0.580	-0.689	-0.558	-0.494
	Kurtosis	0.624	0.851	0.426	0.319
	95% CI	0.3681, 0.4178	0.560, 0.634	0.750, 0.850	0.844, 0.956
10	Mean	0.392	0.599	0.801	0.901
	Variance	0.008	0.003	0.001	0.000
	MSE	0.008	0.003	0.001	0.000
	Skewness	-0.495	-0.766	-0.498	-0.381
	Kurtosis	0.336	2.177	0.364	0.286
	95% CI	0.367, 0.417	0.562, 0.637	0.751, 0.851	0.845, 0.956
20	Mean	0.395	0.599	0.801	0.900
	Variance	0.007	0.003	0.000	0.000
	MSE	0.007	0.003	0.000	0.000
	Skewness	-0.731	-0.572	-0.306	-0.347
	Kurtosis	1.731	0.853	0.250	0.127
	95% CI	0.370, 0.420	0.561, 0.636	0.751, 0.850	0.844, 0.956
30	Mean	0.396	0.598	0.800	0.900
	Variance	0.007	0.002	0.000	0.000
	MSE	0.007	0.002	0.000	0.000
	Skewness	-0.528	-0.428	-0.430	-0.265
	Kurtosis	0.681	0.353	0.168	0.088
	95% CI	0.371, 0.421	0.560, 0.635	0.751, 0.850	0.844, 0.956

Table 22

Summary Data for Negative Exponential Error Score Distribution and 200 Examinees

k	Stats	alpha			
		0.40	0.60	0.80	0.90
5	Mean	0.398	0.600	0.800	0.900
	Variance	0.004	0.002	0.000	0.000
	MSE	0.004	0.002	0.000	0.000
	Skewness	-0.433	-0.349	-0.418	-0.264
	Kurtosis	0.193	0.075	0.686	0.213
	95% CI	0.373, 0.423	0.563, 0.637	0.751, 0.850	0.845, 0.956
10	Mean	0.394	0.600	0.800	0.900
	Variance	0.004	0.001	0.001	0.000
	MSE	0.004	0.001	0.000	0.000
	Skewness	-0.414	-0.501	-0.487	-0.151
	Kurtosis	0.365	0.779	0.484	-0.137
	95% CI	0.370, 0.419	0.563, 0.637	0.751, 0.850	0.844, 0.956
20	Mean	0.395	0.599	0.800	0.900
	Variance	0.004	0.001	0.000	0.00
	MSE	0.004	0.001	0.0[illegible]	0.000
	Skewness	-0.418	-0.409	-0.237	-0.062
	Kurtosis	0.169	0.272	0.185	-0.075
	95% CI	0.370, 0.420	0.561, 0.636	0.751, 0.850	0.844, 0.956
30	Mean	0.397	0.598	0.800	0.900
	Variance	0.003	0.001	0.000	0.000
	MSE	0.003	0.001	0.000	0.000
	Skewness	-0.538	-0.554	-0.339	-0.216
	Kurtosis	0.598	1.131	0.283	-0.096
	95% CI	0.372, 0.421	0.561, 0.635	0.750, 0.850	0.844, 0.956

Table 23

Summary Data for Mixed Normal Error Score Distribution and 25 Examinees

k	Stats	alpha			
		0.40	0.60	0.80	0.90
5	Mean	0.401	0.604	0.807	0.903
	Variance	0.047	0.026	0.007	0.002
	MSE	0.047	0.026	0.007	0.002
	Skewness	-0.930	-1.166	-0.891	-0.920
	Kurtosis	1.331	2.405	0.989	1.919
	95% CI	0.376, 0.426	0.567, 0.641	0.758, 0.857	0.847, 0.959
10	Mean	0.392	0.594	0.802	0.903
	Variance	0.046	0.017	0.004	0.001
	MSE	0.046	0.017	0.004	0.001
	Skewness	-1.592	-0.945	-0.960	-0.838
	Kurtosis	5.134	1.382	1.299	1.296
	95% CI	0.367, 0.417	0.557, 0.631	0.752, 0.852	0.847, 0.959
20	Mean	0.364	0.593	0.803	0.902
	Variance	0.043	0.015	0.003	0.001
	MSE	0.044	0.015	0.003	0.001
	Skewness	-1.436	-1.256	-0.936	-0.651
	Kurtosis	3.434	2.817	2.259	0.579
	95% CI	0.340, 0.389	0.556, 0.631	0.753, 0.853	0.846, 0.958
30	Mean	0.369	0.585	0.801	0.903
	Variance	0.038	0.016	0.002	0.000
	MSE	0.039	0.016	0.002	0.000
	Skewness	-1.145	-1.390	-0.934	-0.633
	Kurtosis	2.602	3.071	1.755	0.561
	95% CI	0.344, 0.394	0.548, 0.623	0.751, 0.851	0.847, 0.959

Table 24
Summary Data for Mixed Normal Error Score Distribution and 50 Examinees

k	Stats	alpha			
		0.40	0.60	0.80	0.90
5	Mean	0.405	0.601	0.803	0.901
	Variance	0.021	0.011	0.004	0.001
	MSE	0.021	0.011	0.004	0.001
	Skewness	-0.684	-0.630	-0.592	-0.611
	Kurtosis	0.875	0.848	0.668	0.827
	95% CI	0.380, 0.430	0.576, 0.626	0.753, 0.852	0.846, 0.957
10	Mean	0.396	0.600	0.802	0.901
	Variance	0.020	0.008	0.002	0.001
	MSE	0.020	0.008	0.002	0.001
	Skewness	-1.035	-0.738	-0.710	-0.645
	Kurtosis	2.060	1.182	1.343	1.110
	95% CI	0.371, 0.420	0.563, 0.638	0.752, 0.851	0.845, 0.957
20	Mean	0.391	0.596	0.801	0.901
	Variance	0.014	0.007	0.001	0.000
	MSE	0.014	0.007	0.001	0.000
	Skewness	-0.808	-0.952	-0.718	-0.416
	Kurtosis	1.251	1.415	0.722	0.166
	95% CI	0.366, 0.416	0.559, 0.633	0.751, 0.850	0.845, 0.957
30	Mean	0.387	0.600	0.804	0.901
	Variance	0.015	0.006	0.001	0.000
	MSE	0.015	0.006	0.001	0.000
	Skewness	-0.742	-0.909	-0.592	-0.410
	Kurtosis	0.900	1.494	0.713	0.414
	95% CI	0.362, 0.411	0.563, 0.637	0.754, 0.853	0.845, 0.957

Table 25
Summary Data for Mixed Normal Error Score Distribution and 100 Examinees

k	Stats	alpha			
		0.40	0.60	0.80	0.90
5	Mean	0.400	0.601	0.802	0.901
	Variance	0.011	0.006	0.002	0.000
	MSE	0.011	0.006	0.002	0.000
	Skewness	-0.539	-0.386	-0.444	-0.394
	Kurtosis	1.164	0.144	0.397	0.326
	95% CI	0.375, 0.425	0.563, 0.639	0.752, 0.851	0.845, 0.957
10	Mean	0.399	0.600	0.801	0.900
	Variance	0.009	0.004	0.001	0.000
	MSE	0.009	0.004	0.001	0.000
	Skewness	-0.647	-0.627	-0.608	-0.317
	Kurtosis	1.452	1.275	1.440	0.394
	95% CI	0.374, 0.423	0.563, 0.638	0.751, 0.850	0.844, 0.956
20	Mean	0.392	0.598	0.799	0.901
	Variance	0.008	0.003	0.001	0.000
	MSE	0.008	0.003	0.001	0.000
	Skewness	-0.554	-0.480	-0.[illegible]11	-0.397
	Kurtosis	0.594	0.104	-0.062	0.298
	95% CI	0.367, 0.417	0.560, 0.635	0.750, 0.849	0.845, 0.957
30	Mean	0.391	0.598	0.801	0.901
	Variance	0.007	0.003	0.001	0.000
	MSE	0.008	0.003	0.001	0.000
	Skewness	-0.487	-0.605	-0.355	-0.253
	Kurtosis	0.309	0.585	0.379	-0.017
	95% CI	0.366, 0.416	0.561, 0.636	0.751, 0.851	0.845, 0.957

Table 26

Summary Data for Mixed Normal Error Score Distribution and 200 Examinees

k	Stats	alpha			
		0.40	0.60	0.80	0.90
5	Mean	0.399	0.602	0.800	0.901
	Variance	0.005	0.003	0.001	0.000
	MSE	0.005	0.003	0.001	0.000
	Skewness	-0.348	-0.393	-0.171	-0.262
	Kurtosis	-0.063	-0.024	-0.200	-0.227
	95% CI	0.374, 0.423	0.565, 0.639	0.751, 0.850	0.845, 0.957
10	Mean	0.396	0.600	0.802	0.901
	Variance	0.005	0.002	0.000	0.000
	MSE	0.005	0.002	0.001	0.000
	Skewness	-0.443	-0.430	-0.218	-0.298
	Kurtosis	0.270	0.210	0.206	0.402
	95% CI	0.371, 0.421	0.563, 0.637	0.752, 0.851	0.845, 0.957
20	Mean	0.397	0.600	0.799	0.900
	Variance	0.004	0.001	0.000	0.000
	MSE	0.004	0.001	0.000	0.000
	Skewness	-0.330	-0.195	-0.414	-0.300
	Kurtosis	0.034	0.025	0.182	0.241
	95% CI	0.372, 0.421	0.562, 0.637	0.750, 0.849	0.845, 0.956
30	Mean	0.399	0.600	0.801	0.901
	Variance	0.004	0.001	0.000	0.000
	MSE	0.004	0.001	0.000	0.000
	Skewness	-0.305	-0.467	-0.384	-0.139
	Kurtosis	-0.018	0.451	0.394	-0.014
	95% CI	0.375, 0.424	0.563, 0.637	0.751, 0.851	0.845, 0.956