Multivariate Analysis of Voluntary Health and Welfare Organization Financial Performance Measures

Rebecca A. Goza
Anderson University
University of Central Oklahoma

Lee Tyner
University of Central Oklahoma

Gregory Kaufinger
Anderson University

This investigation seeks to ascertain if the three factor model used by Ritchie and Kolodinsky (2003) for hospital foundations applies similarly for a different nonprofit sector of Voluntary Health and Welfare Organizations (VHWO's) regarding financial performance categories derived from the Internal Revenue Service Form 990, or if a four factor model is more appropriate. 301 data sets obtained from GuideStar were evaluated using principal component analysis, and ultimately with confirmatory factor analysis using IBM AMOS Version 23.0 software. The results demonstrated that the higher order CFA three factor Ritchie and Kolodinsky (2003) model did fit to evaluate VHWOs.

INTRODUCTION

Are nonprofit organizations (NPOs) effective? In most stakeholders' minds, NPOs provide important services throughout the United States and on a global basis. However, some NPO stakeholders and even loyal donors question, and rightly debate the degree to which such organizations can be considered effective (Herman and Renz, 1999; Jackson and Holland, 1998; Murray and Tassie, 1994; Kanter and Summers, 1987). Additionally, even though NPO stakeholders and donors are extremely interested in seeing their affiliated organizations perform well, there is still uncertainty and disagreement regarding the appropriateness and relative value of current NPO financial performance measurement mechanisms within the individual NPO segments of interest (i.e. churches, hospitals, foundations, etc.).

As such, consistent nonprofit performance evaluation criteria remain elusive to both researchers and practitioners (Forbes, 1998; Tuckman and Chang, 1998; Herman and Renz, 1999; Stone, Bigelow, and Crittenden, 1999; Rojas, 2000; Hoefer, 2000). The elusiveness of consistent nonprofit performance evaluation criteria, combined with the lack of strongly articulated performance criteria position statements by both the watchdog organizations and policy standards boards has created an avenue for accounting and performance evaluation ambiguities that should not continue in this era of reputation risk and data availability.

Herman and Renz have spent a large portion of their research careers investigating this topic. In particular, their research concludes that there is indeed a general lack of convergence of financial and performance criteria, and this lack of convergence contributes to NPOs using a wide range of financial measures to satisfy a variety of stakeholder interests (Herman and Renz, 1998). They also conclude (1998) that NPOs currently have little impetus for refining or testing financial performance measures. Ritchie and Kolodinsky (2003, p.368) similarly conclude that the "general lack of empirical testing of financial measures has adversely affected researchers' confidence in any single set of measures, owing to the myriad of measures in use today." This lack of critical analysis regarding NPO financial performance creates difficulties for researchers, practitioners, and watchdog organizations. Ritchie and Kolodinsky conclude (2003) that researchers will continue to have difficulty forming conclusions regarding the specific NPO activities and characteristics that lead to higher or lower performance. This is likely attributable to the aforementioned lack of consistent measurements. As previously mentioned, the difficulty extends beyond the research community, into practice as well. In other words, it is not logical to require nonprofits to spend resources and donor dollars strengthening or transparently reporting NPO performance ratio results unless it is clearly stipulated which NPO performance ratios mean something unique, or have a significant relationship to the individual NPO under review.

Because of the research community's inability to effectively determine which measures are critical, the practitioner community increasingly finds it impossible to effectively assess performance subjectively, particularly when attempting to identify tested measures that enable the comparison of one organization with that of similar organizations. As a consequence, the result over the years has been the evolution and usage of a myriad of possible performance measure calculations, which leave NPOs with little alternative other than to portray a picture of performance that is considered acceptable to the greatest stakeholder community, regardless of the individual performance measure utilized.

PURPOSE OF THE STUDY

This investigation seeks to discover if the three factor model used by Ritchie and Kolodinsky (2003) for hospital foundations applies similarly for Voluntary Health and Welfare Organizations (VHWO's) regarding financial performance categories or, if a four factor model is more appropriate. Once the model fit can be tested, the appropriate measurements can be used to provide focus and accountability while additionally ensuring that the VHWO sector's accounting policies are refined for consistency in both application and measurement.

Therefore, the theoretical basis for this paper contains elements of both economic consequences theory and positive accounting to the extent that the initial requirement of this investigation is to confirm if a model developed for hospitals applies similarly for assessing VHWO financial performance. This study seeks to specifically investigate if the six measurement ratios studied for hospital foundations by Ritchie and Kolodinsky (2003) for the three constructs of fiscal performance, fundraising efficiency, and public support, fit similarly from a model perspective, for the Voluntary Health and Welfare Organization nonprofit sector. The study also seeks to extend the research by investigating a fourth factor of interest, that of impact.

RESEARCH APPROACH/METHODOLOGY

This study is planned as a prespecified design as it is a continuation of the research approach used by Ritchie and Kolodinsky (2003) but uses a different nonprofit industry segment for scope purposes. Additionally, this investigation differs from the work published by Ritchie and Kolodinsky (2003) by incorporating Brown's (2006, p.1) statement that "confirmatory factor analysis (CFA) is a type of structural equation modeling (SEM) that deals specifically with measurement models, that is, the relationships between observed measures or indicators and latent variables or factors." Ritchie and Kolodinsky (2003) used exploratory factor and principal component analysis in their study. The use of confirmatory factor analysis for the current research criteria is the option that makes sense because this investigation is focused on the relationship strength of specific financial measures to resulting VHWO operational performance. It also enforces a strict structure in this investigation and follows the suggestion of Watts and Zimmerman (1986) to use strong statistical methods to address validity concerns. When addressing validity, Peter stated that "valid measurement is the *sine qua non* of science.

In the same light, the concept of 'validity' refers to the degree to which instruments accurately measure the constructs they are intended to measure" (p. 6). Therefore, the use of confirmatory factor analysis, a type of SEM will be incorporated for the purpose of strengthening the validity criteria of the research study. The use of SEM works because as stated by Schumacker & Lomax (2012, p.2) "the goal of SEM analysis (in the form of CFA) is to determine the extent to which the theoretical model is supported by the sample data." Additionally, a confirmatory factor analysis has not been used in similar research associated with VHWO operational performance measurement. Since this investigation will explore the relationship between observed variables and their underlying latent constructs for two hypothesized models, the use of confirmatory factor analysis extends research in the topic of interest in an arena not previously explored.

The primary goal of this investigation is to describe how well the VHWO constructs measure and convey the health of specific VHWO organizations. This investigation's results will be helpful because charity watchdog organizations and other stakeholders are both interested in, and actively assessing NPO performance. Therefore, from a theoretical basis, economic consequences theory is the primary theory for this investigation and the main reason for this determination is that accounting standards, including (FASB ASC) 958-205-45-6 and FASB 117 have economic consequences. If an accounting standard has no economic consequences, it can be argued that the standard is not needed.

RESEARCH QUESTIONS

Given the brief introduction, the following research questions are addressed in this investigation:

> *Research Question One: What are the primary constructs underlying VHWO organizational financial performance? The answer to this research question informs two additional research questions.*

Murphy et al. (1996) identified nine distinct nonprofit performance constructs and Venkatraman and Ramanujam (1986) proposed in their research findings that growth and profitability were distinct nonprofit performance constructs. Ritchie and Kolodinsky (2003) identified three constructs for the hospital nonprofit sector. Therefore, a preliminary groundwork has been laid as a foundation upon which to extend research for this and two additional questions for VHWO's.

> *Research Question Two: Which measurements of the primary constructs represent VHWO organizational financial performance?*

As the introduction has established and the literature section further demonstrates, organizational performance is multi-dimensional for nonprofit organizations. Therefore, the second research question seeks to identify appropriate measures for each construct. In order to address this research question, the relative and incremental informational content of each measure is studied. Finally, the constructs and measurement indicators are operationalized into a model of nonprofit organizational performance to address the final research question.

> *Research Question Three: Is a four factor model of VHWO organizational performance more effective than the three factor model offered by Ritchie and Kolodinsky (2003) in distinguishing between high and low performing VHWO's?*

The four factor model will be utilized to address the third research question of whether or not the model can be used in the future to distinguish between high and low performing VHWO's. In summary, this investigation attempts to test a measurement model that accurately describes VHWO organizational performance with the purpose of examining the nature of the performance itself as a multidimensional construct.

OVERVIEW OF THE VARIABLES

Nonprofit performance is often discussed in the literature under two key faces of mission and fiscal performance, and derived from the dual objectives of mission accomplishment and financial sustainability. Donors and other stakeholders outside nonprofit organizations typically focus more on mission performance and stress that nonprofits should make every effort to achieve the expected mission impact. However, nonprofits must also maintain financial sustainability to provide services. Additionally, better fiscal performance can be used to demonstrate both the growth of the organization and its management capacity. Common variables and nonprofit performance categories are outlined in Table 1.

TABLE 1
VHWO IRS FORM 990 FACTORS

Fiscal Performance	
1	Total revenue divided by total expenses (Siciliano, 1996, 1997) (Part I line 12 /Part I line 18)
2	Total contributions (gifts, grants, and other contributions) divided by total expenses (Part VIII line 1h / Part I line 18)
Fundraising Efficiency	
3	Fundraising expenses divided by Total revenue (Part I line 16b / Part I line 12)
4	Fundraising expenses divided by Direct public support (Greenlee, 1998)
Public Support	
5	Direct public support divided by total assets (Part I line 8 / Part I line 20)
6	Total contributions (gifts, grants, and other contributions) divided by total assets (Part VIII line 1h/ Part I line 20)

Irs.gov

Data Source

The sample for all three research questions includes VHWO IRS Form 990 data in electronic form purchased specifically for this investigation's research purposes from GuideStar.org. GuideStar digitizes over 300 points of data from the IRS reporting forms. Any of the 300 points of data may be used by academic researchers in record selection criteria. For the purpose of identifying a robust data sample, this investigation's customized data request from GuideStar was specific in that it isolated any 501(c)3 or 501(c)4 that also completed Part IX of the IRS Form 990. Part IX is the Statement of Functional Expenses, currently required only for VHWO's.

GuideStar identified a total potential population of NPOs that were likely VHWO's based on completion of IRS Form 990 Part IX. In order to further refine the initial data into a sample of similar VHWO's, an additional filter was performed further reducing the data file to any National Taxonomy of Exempt Entities (NTEE) codes starting with "G" or "H" to identify those with research related missions.

The NTEE system is used by the IRS and other organizations such as GuideStar to classify nonprofit organizations. NTEE codes beginning with a “G” or “H” are used in order to select a sample of similar NPOs for multivariate analysis purposes. A further filter was performed on the data to isolate any NTEE “G” or “H” where support revenue is 50% or more of gross receipts. The resulting population size actually purchased from GuideStar includes 1,532 NPOs.

Data Collection

The methodology used in this study could potentially be used as a model to evaluate financial measures for both nonprofit researchers and practitioners. A summary of the initial proposed performance ratios and preliminary construct categories, as well as six calculated measurement ratios is included in Table 1. These factors were derived through analysis of the revised Internal Revenue Service IRS Form 990. These three constructs and six measurements are the same as those used by Ritchie and Kolodinsky (2003) but have been revised for changes to the IRS Form 990 after Ritchie and Kolodinsky’s 2003 publication.

In addition to the basic dataset fields included in Table 1, a custom research request for this study included Part I Line 12 (Total Revenue), Part I Line 9 (Program Service Revenue), Part I Line 20 (Total Assets), Part I Line 18 (Total Expenses), Part I Line 22 (Net position), Part I Line 16b (Total Fundraising Expenses), Part VIII Line 1a (Direct Public Support), Part VIII Line 1h (Total contributions), Part III Line 4e (total program service expenses), Part VI Line 13, Part VI Line 14, Part VI Line 5, Part VI Line 12a, Part VI Line 12b, Par VI Line 12c, Part XII Line 1, Part XII Line 2a, Part XII Line 2b. The customized research request was provided to GuideStar and GuideStar ultimately provided the final dataset used for testing purposes.

This investigation’s population data set was designed in two steps. First, accounting data were collected from the IRS Form 990s; and then financial ratios were computed. The purchased population of 1,532 NPOs was further reduced to eliminate multiple incorporated chapters of similar organizations (i.e. Multiple Sclerosis in different states versus a headquarter location). A reduced sample of 387 homogenous VHWO’s was ultimately identified. This represents the reduced population of VHWO’s without conducting transformations for normality.

DATA ANALYSIS METHOD(S) USED

This current investigation uses principal component analysis and confirmatory factor analysis as the primary data analysis methods. The goal of the research is to determine if the same factors identified by Ritchie and Kolodinsky (2003) for hospitals correlate similarly for VHWO’s. Therefore, an inferential principal component analysis is used to explore and confirm cross-validate factor structure between hospitals and VHWO’s. This initial inferential principal component analysis will help to conclude if the earlier reported factor structure identified by Ritchie and Kolodinsky (2003) may be generated across a different population of VHWO’s.

Exploratory factor analysis and principal component analysis was utilized for the previous research in the area of nonprofit financial performance measurement by Ritchie and Kolodinsky (2003), whose prior research was used as a foundation for this investigation. Brown (2006) suggests that confirmatory factor analysis (CFA) is a type of structural equation modeling that deals specifically with measurement models. CFA is used to study the relationships between a set of observed variables and a set of continuous latent variables. In addition to information regarding the accounting measurements in Table 1, the results of the Ritchie and Kolodinsky (2003) research are shown below in Table 2. Table 2 results are carried forward from the Ritchie and Kolodinsky (2003) study and serve to outline the financial measures and constructs for hospitals identified through exploratory factor analysis and the principal component analysis conducted by the researchers in 2003.

The IRS Form 990 data purchased from GuideStar is the source used for the ratios calculated in this investigation. The aforementioned ratios derived initially from Table 2 are used as well in the current investigation to identify relationships. Principal component analysis is used initially in this current investigation. Ultimately, confirmatory factor analysis is also used to clearly distinguish this investigation from the research previously provided by Ritchie and Kolodinsky (2003). Additionally, a limitation of simple path analysis is the use of a single measure of each construct represented in a model. An alternative to this investigation is the use of multiple measures of each construct as outlined by Kline (2005), which tends to reduce the effect of measurement error in any individual measurement on the accuracy of the results. This approach aligns well with the theoretical basis which underpins this investigation. According to Kline (2005), if a standard CFA model with a single factor has at least three indicators, the model is identified. Kline (p.172) concludes "if a standard model with two or more factors has at least two indicators per factor, the model is identified." The path diagram used the initial phase of this investigation with principal component analysis (research questions one and two and the first three hypotheses (not included)) is shown in Figure 1 below. A nested confirmatory factor model approach will ultimately be incorporated with the purpose of investigating hypothesis four (not included) and whether a three-factor or a four-factor model is superior (research question Three).

TABLE 2
FACTOR ANALYSIS: FINANCIAL MEASURES AND IRS FORM 990 LINE ITEM LABELS

	Mean	*SD*	*Fundraising Efficiency*	*Public Support*	*Fiscal Performance*
Direct public support divided by fundraising expenses (Greenlee, 1998) (line 1A ÷ line 44D)	84	312	.99	.06	.08
Total revenue divided by fundraising expenses (line 12 ÷ line 44D)	121	400	.99	.01	.05
Total contributions divided by total revenue (Siciliano, 1996; Greenlee, 1998) (line 1D ÷ line 12)	.65	.18	.10	.86	.22
Direct public support divided by total assets (line 1A ÷ line 59B)	.16	.11	−.02	.91	−.05
Total revenue divided by total expenses (Siciliano, 1996, 1997) (line 12 ÷ line 17)	2.54	2.89	.06	.10	.99
Total contributions divided by total expenses (line 1D ÷ line 17)	1.8	2.9	.07	.20	.98

Note: Principal component analysis (promax rotation). Rotation converged in four iterations.

Ritchie and Kolodinsky (2003, p.374)

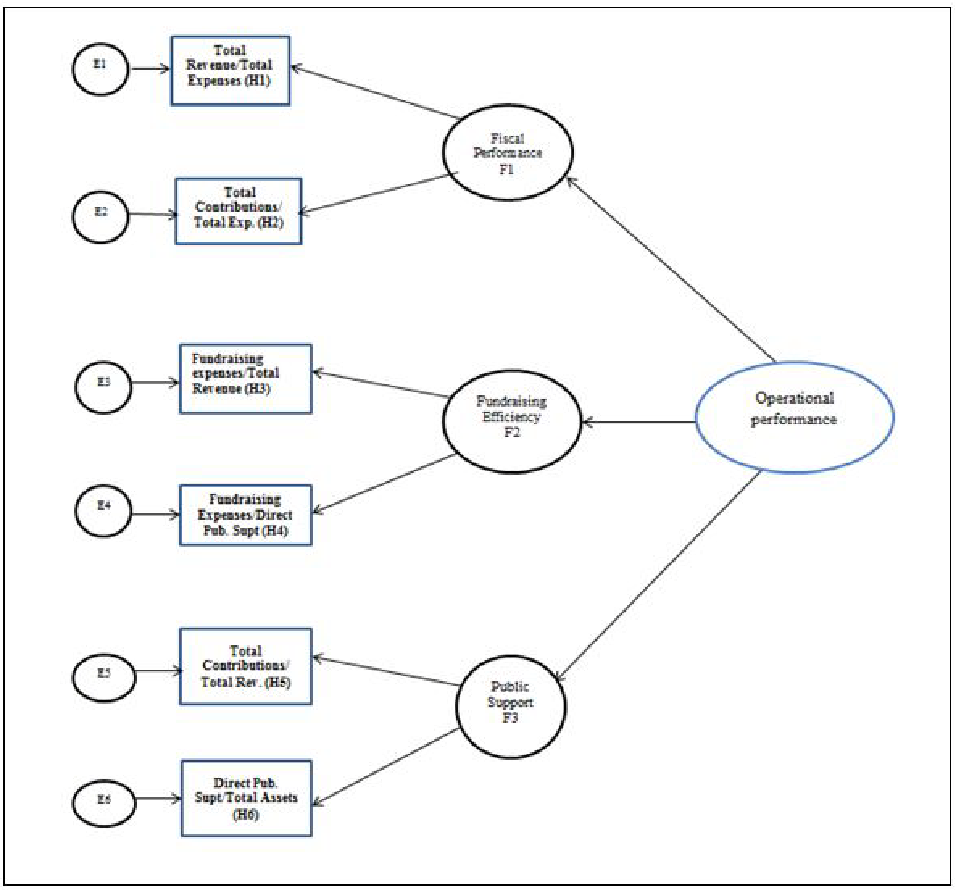
As stated by Flynn and Hodgkinson (2001), the nonprofit sector has relied largely on anecdotal evidence and general good will to publicize success and tax-exempt status. Based on the extensive review of the accounting literature, there appears to be insufficient guidance available for nonprofit management when attempting to assess the roles, functions and contributions of individual nonprofits beyond that generated at the institutional level. Flynn and Hodgkinson (2001, p.4) further point out that in the "increasingly competitive world in which nonprofits operate, there are new demands for impact analysis."

The unique difference between research questions 2 and 3 is the fact that in the third research question, a fourth construct of "impact" is introduced for further consideration in the VHWO NPO sector. Also, for additional informational background, two statistical tools of IBM SPSS Version 23 and Amos Version 23 are used for this investigation. IBM SPSS is used for principal component analysis and tests of normality and IBM Amos 23.0 is used for confirmatory factor analysis.

Second-Order CFA Model

In the Figure 1 path diagram, used for research questions one and two, there are three primary factors described below that operate as latent independent variables. Each of these three factors can be considered one level or one unidirectional arrow away from the observed variables "i.e., measurements." The current model also argues for a higher level factor of organizational performance that is accountable for the lower first-order factors. An important aspect of the second-order model included in this study is that the second-order factor of operational performance does not have its own set of measured indicators. Instead, it is linked indirectly to those indicators measuring the first-order factors. A second important aspect of the second-order model is that the three first-order latent independent variables appear to function as both dependent as well as independent variables. Since the first-order factors operate as dependent variables their variances and covariances are no longer estimable. The graphical depiction of the second-order CFA model in Figure 2 and is particularly important for the model fit analysis included in this study. A third aspect of the current model is the presence of single headed arrows leading from the second-order factor of operational performance to each of the three first-order factors.

FIGURE 1
PATH DIAGRAM – VHWO PERFORMANCE MEASURES – THREE FACTORS



These regression paths represent the second-order factor loadings; one second-order loading is fixed to unity for scaling purposes as depicted in Figure 2. It should be noted that the prediction of each of the first-order factors from the second-order factor is presumed not to be without error. Therefore, a residual disturbance term is associated with each of the three first-order factors and indicated in the model in Figure 2 as variables FPE, FEE, and PSE. In this instance, since the variances of the of the first-order factors are of interest, the regression paths for the residual disturbance terms FPE, FEE, and PSE are fixed to 1.0.

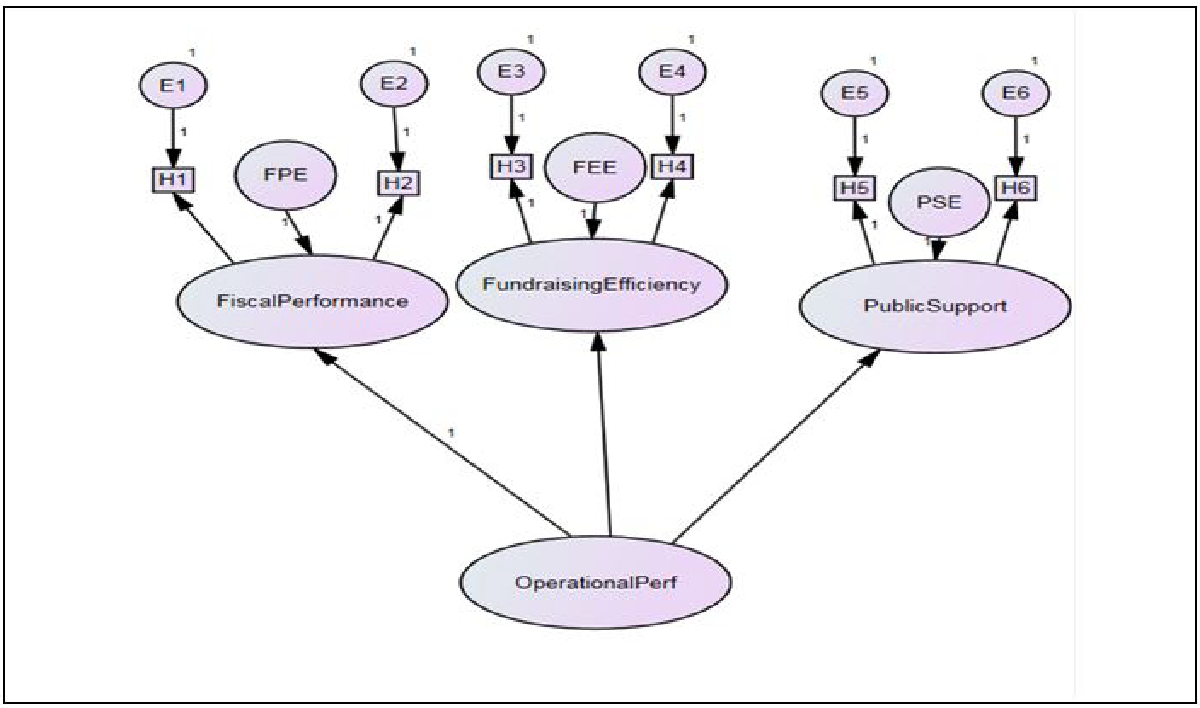
RESEARCH FINDINGS

The ultimate goal of this research is to test a model for assessing VHWO financial performance using ratios. Sori et al. (2006, p.71) pointed out that "financial ratios have long been used in various study areas in accounting and finance using either univariate or multivariate methodologies." This investigation includes an initial sample of 387 VHWO calculated performance ratio measurements to perform principle component factor analysis; and ultimately structured equation modeling, in the form of confirmatory factor analysis.

Outliers and Missing Data

Outliers were detected using Z-Score analysis. Any calculated Z score over 3.0 or less than -3.0 was considered to be an outlier. Outliers were removed from this study using listwise deletion prior to conducting further statistical analysis. A review of the outliers was conducted to determine if any were particularly appropriate for the current investigation. None of the excluded outliers were deemed salient and therefore the remaining VHWO's make up the final 301 sample size (n=301) analyzed further, including the measurement variables as follows: Total Revenue/Total Expense (H1 ratio), Total Contributions/Total Expense (H2 ratio), Fundraising Expenses/Total Revenue (H3 ratio), Fundraising Expenses/Direct Public Support (H4 ratio), Total Contributions/Total Revenue (H5 ratio), and Direct Public Support/Total Assets (H6 ratio).

FIGURE 2
THREE FACTOR – SIX MEASURE CFA MODEL USING AMOS



Assumption of Normality/Test of Sphericity

The descriptive statistics provide information regarding skewness and kurtosis. Skewness quantifies how symmetrical the distribution is around the sample mean. A symmetrical distribution has a skewness of zero. An asymmetrical distribution with a long tail to the right (higher values) has a positive skew. An asymmetrical distribution with a long tail to the left (lower values) has a negative skew. There is no standard threshold, however, Ruppert (2004) stated that if the skewness is greater than 1.0 (or less than -1.0), it is far from symmetrical.

As stated by Ruppert (2004) kurtosis is a measure of skewness that indicates whether or not the data distribution matches the Gaussian distribution (a kurtosis of 0). Normally distributed data have skewness of 0 and kurtosis of 0. Ruppert (2004) further clarifies that kurtosis indicates the heaviness of the tail. Ruppert (2004) indicates that normally distributed data has a kurtosis of 0. According to Ruppert (2004) heavier tails have positive kurtosis, and tails lighter than the normal distribution have a negative kurtosis. The results as determined for the current investigation (not included) using the standardized error, are indicative of data that is not normal.

According to Sori et al. (2006, p.72), the assumption of normality is important for the interpretation of the tests of significance, and if the data does not satisfy this assumption, the results obtained may be biased." However, with large enough sample sizes (>30 or 40), the violation of normality assumption should not cause major problems (Pallant, J., 2007).

Additionally, the Central Limit Theorem says that given random and independent samples of n observations, the distribution of sample means approaches normality as the size of n increases. Reliance on the Central Limit Theorem is the primary basis for proceeding with the current research, understanding that a limitation of the study will be the consideration of the possibility of bias in the estimates of standard errors of loadings provided with CFA.

Deakin (1976), Ezzamel et al. (1987) and So (1987) specifically considered the assumption of normality related to financial measures such as those used in this investigation, and conclude that most financial ratios tend to be skewed and abnormally distributed. Sori et al. (2006, p. 81) also states that "after necessary procedures have been taken such as outlier trimming and data transformations, many of the accounting bivariate variables tend to continue to depart from normality assumptions." Their conclusion strongly suggests that other factors exist that influence the normality of ratio variables. This conclusion was similarly suggested by Ezzamel (1987), with the proportionality assumption in ratio analysis factoring heavily for financial ratio research studies. The practical implications associated with the statistical properties of financial ratios and in particular with the issue of ratio proportionality were summarized by Foster (1986) in an overview of ratio analysis:

TABLE 3
SPSS TEST OF NORMALITY

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
H1	.208	301	.000	.767	301	.000
H2	.092	301	.000	.934	301	.000
H3	.159	301	.000	.773	301	.000
H4	.235	301	.000	.555	301	.000
H5	.187	301	.000	.843	301	.000
H6	.132	301	.000	.897	301	.000

a. Lilliefors Significance Correction

Tests of Normality after base-10 log transformation

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	Df	Sig.	Statistic	Df	Sig.
H1log10	.150	301	.000	.892	301	.000
H2Log10	.244	301	.000	.760	301	.000
H3Log10	.106	301	.000	.952	301	.000
H4Log10	.078	301	.000	.965	301	.000
H5Log10	.208	301	.000	.808	301	.000
H6Log10	.091	301	.000	.920	301	.000

a. Lilliefors Significance Correction

An important assumption underlying the use of ratios as a control for size differences is strict proportionality between the numerator and the denominator. This strict proportionality is assumed both in comparisons of ratios across firms at a point in time and in comparisons of the ratios of firms over time. (p. 96) Thus, under the assumption of ratio proportionality, inferences may be drawn directly from financial ratios but test of normality often indicate skewness. Fieldsend, et al. (1987) argue for the use of an alternative approach to the statistical modeling of accounting bivariates, claiming that basic accounting variables used in ratio analysis are bounded at zero and, for large samples, the evidence provides that their distribution may be Pareto-like, or lognormal (Ijiri and Simon, 1977). Q-Q plots and histograms generated using IBM SPSS are used to pictorially assess normality. In all six measures for this investigation, the Q-Q plots revealed the hump was too narrow and the data are not plausibly normal. This initial assessment on the current data aligns with the assumptions of ratio proportionality discussion above. Although the Q-Q Plot provides a visual basis for checking normality, it is still important to assess the degree of departure from normality using Kolmogorov-Smirnov and Shapiro-Wilk tests as depicted in Table 3 before the null hypotheses can be rejected.

The results indicate that the data is not normal due to the fact that the significance is <.05, as initially anticipated based on the previous discussion of ratio normality issues. Efforts to transform the data using a base-10 log transformation in IBM SPSS were attempted and resulted in no substantial improvement in normality. Notwithstanding the understanding of the data and tests of normality included in Table 3, the impact of the Central Limit Theorem remains. In probability theory, the Central Limit Theorem states that, given certain conditions, the arithmetic mean of a sufficiently large number of iterations of independent random variables, each with a well-defined expected value and well-defined variance, will be approximately normally distributed, regardless of the underlying distribution (Rice, J, 1995). Therefore, using this rationality coupled with the proportionality assumption of ratio analysis, the current sample is acceptable to use to proceed with the empirical testing.

Finally, Bartlett's Test of Sphericity and the Kaiser-Meyer-Olkin "KMO" tests were conducted. The Bartlett's Test of Sphericity tests the null hypothesis that the correlation matrix is an identity matrix. An identity matrix is a matrix where all the diagonal elements are 1 and all off diagonal elements are 0, implying that all of the variables are uncorrelated. If the significance value for this test is less than the alpha level, the null hypothesis is rejected that the population matrix is an identity matrix. The significance value for the current investigation provides a conclusion to reject the null hypothesis and conclude that there are correlations in the data set that are appropriate for factor analysis. Usually, the value of KMO more than .5 is considered sufficient for sampling. The results indicate the investigation will proceed.

VHWO Principle Component Analysis – Research Questions One and Two

Having determined that the VHWO sample data file is appropriate for use when relying on the ratio of proportionality assumption and the Central Limit Theorem, a correlation matrix was analyzed and included in Table 4. The component correlation matrix outlines correlations between the original variables, which are specified using the /variables subcommand in IBM SPSS. If any of the correlations are above 0.90, it indicates the need to remove one of the variables from the analysis, as the two variables seem to be measuring the same thing. This concern is not realized as shown in Table 4.

Based on the results in Table 4, the analysis can proceed. A principle component analysis was conducted without rotation using IBM SPSS. The unrotated results are included in Table 5 as a baseline for the investigation. Of particular interest in Table 5 is that for the first two research questions, 85.161 percent of the variance can be explained by the three factors investigated. Additionally, unrotated communalities are included in Table 5.

Communalities as stated by (Ramsey, 2006) represent the proportion of each variable's variance that can be explained by the principle components (i.e. underlying latent variables). Ramsey (2006) further states the values in the extraction column indicate the proportion of each variable's variance that can be explained by the principle components. Variables with high values are well represented in the common factor space according to Ramsey (2006).

TABLE 4
CORRELATION MATRIX – 3 FACTOR MODEL

Component Correlation Matrix

Component	1	2	3
1	1.000	.338	.029
2	.338	1.000	-.070
3	.029	-.070	1.000

Extraction Method: Principal Component Analysis.
Rotation Method: Promax with Kaiser Normalization.

TABLE 5
PRINCIPLE COMPONENT ANALYSIS – UNROTATED THREE FACTOR SIX MEASURE

Communalities

	Initial	Extraction
H1	1.000	.941
H2	1.000	.970
H3	1.000	.866
H4	1.000	.802
H5	1.000	.898
H6	1.000	.633

Extraction Method: Principal Component Analysis.

Component Matrix[a]

	Component		
	Public Support	Fiscal Performance	Fundraising Efficiency
Total Revenue/Total Expense (H1 ratio)	.366	-.668	.601
Total Contributions/Total Expense (H2 ratio)	.932	-.154	.279
Fundraising Expenses/Total Revenue (H3 ratio)	.262	.789	.418
Fundraising Expenses/Direct Public Support (H4 ratio)	-.566	.422	.551
Total Contributions/Total Revenue (H5 ratio)	.913	.245	-.059
Direct Public Support/Total Assets (H6 ratio)	.651	.302	-.343

Extraction Method: Principal Component Analysis.
a. 3 components extracted.

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.650	44.159	44.159	2.650	44.159	44.159
2	1.422	23.706	67.866	1.422	23.706	67.866
3	1.038	17.296	85.161	1.038	17.296	85.161
4	.583	9.710	94.871			
5	.296	4.927	99.798			
6	.012	.202	100.000			

Extraction Method: Principal Component Analysis.

In the unrotated results in Table 5, the variable variances are high. Next principle component analysis was conducted using a promax rotation. Promax is an "oblique" solution, meaning factors are correlated with one another. Even though a review of literature indicates that varimax rotation is the most common choice for rotation, it is an orthogonal method. According to Costello and Osborne (2005) orthogonal rotations produce factors that are uncorrelated. Costello and Osborne further suggest (2005) that in the social sciences a researcher should expect some correlation among factors, since behavior is rarely partitioned into units that function independently of one another.

Of important note, Costello and Osborne (2005) indicate that rotation cannot improve basic aspect of the analysis, such as the amount of variance extracted from the items. The goal of rotation in PCA is provide a simple interpretation of data structure. Bryant and Yarnold (1995, p. 132) define rotation as "a procedure in which the eigenvectors (factors) are rotated in an attempt to achieve simple structure." As indicated, the PCA results while using promax rotation are included in Table 6. Other rotation methods were attempted and did not result in improvement in variance explanation. The principle component analysis data reduction technique using a promax rotation results in a conclusion that three factors in the Table 5 rotation explain 85.161 percent of the variance.

Of additional note, initial eigenvalues for all three components were over >1.00 which provide support for the explanatory power of all three factors. According to Girden (2001) those components with eigenvalues less than 1.00 are not considered to be stable. They account for less variability than does a single variable and are not retained in PCA analysis. As further explained by Girden (2001) when a factor has an eigenvalue less than 1.00, in a sense it has less than one variable in it. The results indicate that all three initial factors are stable using the promax rotation.

A review of the results included in Table 5 indicates that the measurement Fundraising Expenses/Total Revenue (H3 ratio) and Fundraising Expenses/Direct Public Support (H4 ratio) load nicely into the fundraising efficiency factor. Similarly, the measurement variable Total Contributions/Total Revenue (H5 ratio) and measurement variable Direct Public Support/Total Assets (H6 ratio) load with the public support performance factor. Consistent loadings are noted using varimax rotation.

Also, the measurement variable of Total Revenue divided by Total Expenses (H1 ratio) and the measurement variable of Total Contributions divided by Total Expenses (H2 ratio) loads onto the fiscal performance factor. Interestingly, the measurement Total Contributions divided by Total Expense (H2 ratio) is one measurement that cross loads with the public support factor for VHWO's. As it is reasonable that a financial measurement using contributions would load into both a fiscal performance and public support factor, CFA will be performed.

VHWO Principle Component Analysis – Research Question Three

Since the results of the first two research questions revealed that the six measurements align similarly for VHWO's as they did for the hospital foundations studied by Ritchie and Kolodinsky (2003), further analysis was conducted given the nature of the specific sector under review to determine if a fourth factor is also important for VHWO's (research question Three). In an attempt to conduct further research, a path model (not included) was developed introducing a forth construct more specific to Voluntary Health and Welfare Organizations called impact.

Assumption of Normality

Since the first two research questions leads to the conclusion that most tend to be skewed on non-normally distributed, additional test of normality will not be conducted for the final research question. This is further justified when considering the previously confirmed assumption of ratio proportionality, wherein inferences may be drawn directly from financial ratios but tests of normality often indicate skewness.

TABLE 6
PRINCIPLE COMPONENT ANALYSIS - UNROTATED FOUR FACTOR EIGHT MEASURE

Communalities

	Initial	Extraction
H1	1.000	.936
H2	1.000	.952
H3	1.000	.911
H4	1.000	.791
H5	1.000	.890
H6	1.000	.599
H7	1.000	.801
H8	1.000	.182

Extraction Method: Principal Component Analysis.

Component Matrix[a]

	Component		
	Public Support	Fundraising Efficiency	Fiscal Performance
Total Revenue/Total Expense (H1 ratio)	.382	-.303	.836
Total Contributions/Total Expense (H2 ratio)	.919	-.130	.301
Fundraising Expenses/Total Revenue (H3 ratio)	.280	.911	-.048
Expenses/Direct Public Support (H4 ratio)	-.456	.745	.169
Total Contributions/Total Revenue (H5 ratio)	.905	.069	-.256
Direct Public Support/Total Assets (H6 ratio)	.597	-.021	-.491
Program Service Expense/Total Expense (H7 ratio)	-.284	-.794	-.300
Total Staff/Total Volunteer (H8 ratio)	-.370	-.119	.176

Extraction Method: Principal Component Analysis.
a. 3 components extracted.

Total Variance explained

	Initial Eigenvalues			Extraction Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.672	33.396	33.396	2.672	33.396	33.396
2	2.143	26.791	60.186	2.143	26.791	60.186
3	1.247	15.591	75.778	1.247	15.591	75.778
4	.949	11.861	87.638			
5	.485	6.066	93.704			
6	.358	4.473	98.178			
7	.133	1.668	99.846			
8	.012	.154	100.000			

Extraction Method: Principal Component Analysis.

Outliers and Missing Data

Similar to the procedures were utilized for the first two research questions, whereby outliers were detected using Z-Score analysis for the two measurements and one factor included in Hypothesis Four. Any calculated Z score over 3.0 or less than -3.0 was considered to be an outlier. The remaining VHWO's make up the 215 sample size (n=215), with a cutoff criterion of .40, analyzed further the newly introduced factor referenced as "impact".

Two additional measures and one additional construct for impact were considered in a separate principal component analysis to address the third research question and fourth hypothesis. The unrotated results for a four measure, eight measure model is included as a Table 6.

The results of the unrotated analysis represented in Table 6 revealed two considerations that preclude continued viability of further pursuit of a four factor model in the current investigation. First, unrotated communalities, which represent the proportion of each variable's variance that can be explained by the principle components (Ramsey, 2006) reveal that one of the two impact measurement variables (Total Staff/Total Volunteer (H8 ratio) measurement) returned a low value. As a reminder, Ramsey (2006) suggests that variables with high values are well represented in the common factor space. Since only two initial measurements were considered in this investigation for impact, the removal of one measurement would lead to the conclusion that PCA analysis could not continue without the minimum two measures per factor standard.

This concern is supported through analysis of eigenvalues > 1.0. As represented in Table 6, only three factors returned eigenvalues > 1.00 indicating that only three initial factors are stable and explain 75.778 percent of the variance. Since Costello and Osborne (2005) concluded that rotation cannot improve basic aspects of the analysis, such as the amount of variance, further rotation attempts will not produce results that will allow this current research to continue for hypothesis four and the third research question. Promax and varimax rotations were still conducted for research transparency purposes and the results did not contradict the conclusions reached through the unrotated results.

Confirmatory Factor Analysis (Structured Equation Modeling)

Since PCA results were successful for research questions 1 and 2, the associated confirmatory factor analysis model application for the current investigation represented in Figure 2 without depiction of covariances. The goal of CFA analysis in this current investigation is to further validate the model developed by Ritchie and Kolodinsky (2003) to confirm if the model would also fit for VHWOs.

As the initial step in determining whether the second-order model is identified, the number of parameters to be estimated is counted as follows: 6 first-order regression coefficients, 10 covariances and 3 second-order regression coefficients making a total of l9 distinct parameters to be estimated. Since there are 21 pieces of information in the variance/covariance matrix, this model is identified with 2 degrees of freedom. It was noted through use of IBM Amos that the Amos software automatically attempts to improve model fit and the Amos efforts defined 3 degrees of freedom.

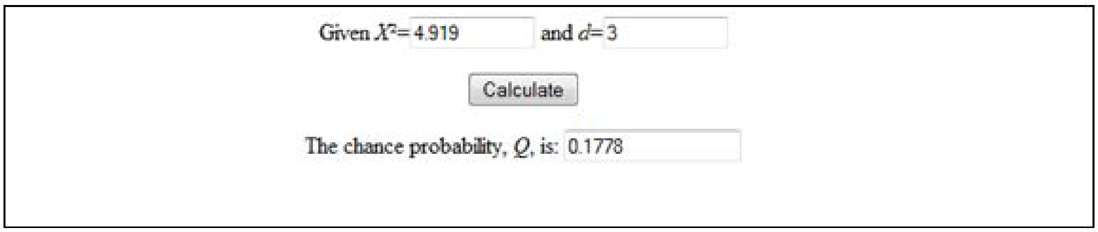
According to Kline (2005), one preferred option involved with analyzing indicators with nonnormal distributions is to use a corrected normal theory method. As Kline indicates (2005, p. 195) "this means to analyze the original data with a normal theory method such as Maximum Likelihood (ML), but use robust standard errors and corrected test statistics." This approach is utilized in this investigation using IBM Amos because attempts at meaningful data transformations revealed no significant changes in the results and were ineffective.

Having derived model estimates in IBM Amos the current investigation requires an evaluation of model fit. Aside from the Chi-square goodness-of-fit test, there are numerous ancillary indices of global fit identified in the literature. According to Kline (2005, p.134) when reporting results of SEM analysis, a "minimal set of fit indexes that should be reported include 1) model Chi-square, 2) the Steiger-Lind root mean square error of approximation (RMSEA; Steiger, 1990), the 3) Bentler comparative fit index (CFI; Bentler, 1990)."

Following the suggestions of Kline (2005), Figure 2 model evaluations were examined first using the Chi-square statistic and accompanying significance tests. The IBM Amos 23.0 analysis P-Value and Chi-square results for 3 constructs and 6 measure model are included in Figure 3. The Chi-square score identified with the CFA model is 4.919. To further clarify, CFA was performed using a normal theory method of maximum likelihood using the aforementioned assumption of normality for ratios and large sample sizes.

One of the most common Chi-square calculations is determining, given the measured Chi-square (X^2) value for an experiment with a degree of freedom d, the probability of the result being due to chance. If the Chi-square is not significant, the model is regarded as acceptable, and results are thought to be likely due to chance. The results included in Figure 3 denote that the P-Value of .1778 is not significant at $p < 0.05$ and suggests, at least initially the model fit is acceptable.

FIGURE 3
P-VALUE FROM CHI-SQUARE CALCULATOR FOR THREE FACTOR MODEL



The next model fit category analyzed was the root mean square error of approximation (RMSEA). According to Stieger (1990), RMSEA less than .05 is ideal. Browne & Cudeck (1993) have concluded that a RMSEA less than .08 is also acceptable for model fit determination. As indicated in Figure 4, the RMSEA for the current investigation was .046.

FIGURE 4
RMSEA FOR THREE FACTOR MODEL

RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	.046	.000	.117	.442
Independence model	.563	.539	.588	.000

According to Kline (2005), the comparative fit index compares the model of interest with some alternative, such as the null or independence model. Kline (2005) indicates that the CFI is also known as the Bentler Comparative Fit Index. Specifically, the CFI compares the fit of a target model to the fit of an independent model--a model in which the variables are assumed to be uncorrelated. In this context, fit refers to the difference between the observed and predicted covariance matrices, as represented by the Chi-square index. As further stated by Kline (2005) CFI represents the ratio between the discrepancies of the target model to the discrepancy of the independence model. Roughly, the CFI thus represents the extent to which the model of interest is better than is the independence model. According to Kline (2005) values that approach 1 indicate acceptable fit and Byrne (2004) states that CFI ideally should be > 0.93.

FIGURE 5
COMPARATIVE FIT INDEX FOR THREE FACTOR MODEL

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	.997	.983	.999	.993	.999
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

The CFI result produced in IBM Amos for the current investigation, as indicated in Figure 5 is 0.999 and this result continues to support model fit. Of additional note, while Chi-square is sometimes criticized in the literature as being sensitive to sample size, CFI is not found to have the same indication (Fan, Thompson, and Wang, 1999).

As a final test of model fit, the Goodness of Fit Index (GFI) was also assessed. According to Byrne (2004), a model is regarded as acceptable if GFI > 0.90. As represented in Figure 6 below, the GFI for the current investigation returned a value in IBM Amos of 0.995.

FIGURE 6
GOODNESS OF FIT INDEX FOR THREE FACTOR MODEL

RMR, GFI

Model	RMR	GFI	AGFI	PGFI
Default model	.003	.995	.963	.142
Saturated model	.000	1.000		
Independence model	.051	.569	.396	.406

SUMMARY OF RESULTS AND IMPLICATIONS

The empirical testing conducted during this investigation concludes that 1) the model studied by Ritchie and Kolodinsky (2003) applies similarly to VHWO's as it did in 2003 to the hospital nonprofit sector, and 2) a model which includes a fourth factor of interest called "impact" does not provide significant results for improved measurement of VHWO financial performance without conducting additional model revision.

Finally, expanded testing of the Ritchie and Kolodinsky (2003) model, conducted using CFA in this current investigation, did result in a conclusion that model fit could be empirically tested and validated. Using the methodology and data population outlined previously reveals that the three domain – six measure model used in the Ritchie and Kolodinsky (2003) study for hospital foundations does also measure VHWO financial performance. This conclusion is supported specifically by the PCA Promax loadings, and is supported further by confirmation of model fit as indicated by CFA. The prior study utilized exploratory factor analysis of multiple financial measurements to determine an appropriate model for use in hospitals.

This investigation did not recreate the exploratory factor analysis of the multiple financial measurements but rather starts with the conclusions reached by Ritchie and Kolodinsky (2003) in the hospital sector as a basis for both a similar analysis in a different nonprofit sector of VHWO's. Additionally, this investigation differs from the prior study in the use of confirmatory factor analysis to confirm model fit and to address validity concerns. Finally, this investigation is unique in the exploration of a fourth domain of interest called "impact".

REFERENCES

AICPA (2002). *Consideration of Fraud in a Financial Statement Audit*. Statement on Auditing Standards No. 99. New York: AICPA.

Bentler, P.M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin,* 107, 238-246.

Better Business Bureau (BBB). (2013). *Implementation guide to BBB wise giving alliance standards for charity accountability.* Retrieved December 22, 2013 from http://www.bbb.org/us/Charity-Evaluation/

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), Testing structural equation models (pp. 136-162). Newsbury Park, CA: Sage.

Brown, T. (2006). *Confirmatory factor analysis for applied research*. New York, NY: Guilford Press.

Bryant, F. B., & Yarnold, P. R. (1995). Principal-components analysis and confirmatory factor analysis. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding multivariate statistics* (pp. 99-136). Washington, DC: American Psychological Association.

Byrne, B. M. (2004). Structural Equation Modeling with EQS and EQS/Windows, Newbury Park, Sage.

Deakin, E.B. (1976). Distributions of financial accounting ratios: Some empirical evidence. *The Accounting Review,* 90-96. DOI: 10.1111/j.1468-5957.1987.tb00106.x

Ezzamel, M., Mar-Molinero, C., & Beech, A. (1987). On the distributional properties of financial ratios. *Journal of Business Finance & Accounting*, 14(4), 463-481. doi:10.1111/j.1468-5957.1987.tb00107.x

Fan, X., B. Thompson, and L. Wang (1999). Effects of sample size, estimation method, and model specification on structural equation modeling fit indexes. Structural Equation Modeling, 6, 56-83.

Fieldsend, S., Longford, N., McLeay, S. (1987). Industry effects and the proportionality assumption in ratio analysis: A variance component analysis. *Journal of Business Finance & Accounting*, 14(4), 497-517. doi:10.1111/j.1468-5957.1987.tb00109.x

Financial Accounting Standards Board (FASB). (2013). *Statement of financial accounting standard No 117, financial statements of not-for profit organizations*. Stamford, CT: Financial Accounting Standards Board.

Flynn, P., & Hodgkinson, V. A. (2001). Measuring the Contributions of the Nonprofit Sector. Nonprofit and Civil Society Studies, 3–16. doi:10.1007/978-1-4615-0533-4_1

Forbes, D. (1998). Measuring the Unmeasurable: Empirical Studies of Nonprofit Organization Effectiveness from 1977 to 1997. *Nonprofit and Voluntary Sector Quarterly*, 27 (2), 183-202. doi:10.1177/0899764098272005

Ford, M. W., & Evans, J. R. (2001). Baldrige assessment and organizational learning: the need for change management. *Quality Management Journal, 8*(3), 9–25. ADD DOI

Foster, G. (1986). *Financial statement analysis.* Prentice-Hall.

Girden, E.R. (2001). Evaluating research articles from start to finish. Thousand Oaks, CA, Sage Publications.

Greenlee, J.S., Bukovinsky, D. (1998). Financial Ratios for Use in the Analytical Review of Charitable Organizations. *Ohio CPA Journal*. pp. 32-38.

Herman, R.D., Renz, D.O. (1998). Nonprofit Organizational Effectiveness: Contrasts Between Especially Effective and Less Effective Organizations. *Nonprofit Management and Leadership,* 9(1), 23-38. doi: 10.1002/nml.9102

Herman, R.D., Renz, D.O. (1999). Theses on Nonprofit Organizational Effectiveness. *Nonprofit and Voluntary Sector Quarterly*, 28(2), 107-125. doi: 10.1177/0899764099282001
Hoefer, R. (2000). Accountability in Action? Program Evaluation in Nonprofit Human Service Agencies. *Nonprofit Management and Leadership*, 11(2), 167-177. doi: 10.1002/nml.11203
Ijiri, Y., Simon, H.A. (1977). *Skew distributions and the sizes of business firms*. North-Holland.
Jackson, D.K., Holland, T.P. (1998). Measuring the Effectiveness of Nonprofit Boards. *Nonprofit and Voluntary Sector Quarterly,* 27, 159-182. doi: 10.1177/0899764098272004
Kanter, R.M., Summers, D.V. (1987). Doing Well While Doing Good: Dilemmas of Performance Measurement in Nonprofit Organizations and the Need for a Multiple Constituency Approach. In W.W. Powell (ed.). *The Nonprofit Sector: A Research Handbook.* New Haven, Conn.: Yale University Press.
Kline, R. (2005). *Principles and practice of structural equation modeling*. New York: The Guildford Press.
Murphy, G., Trailer, J., Hill, R. (1996). Measuring performance in entrepreneurship research. *Journal of Business Research* 36: 15-23. http://dx.doi.org/10.1016/0148-2963(95)00159-x
Murray, V., Tassie, B. (1994). Evaluating the Effectiveness of Nonprofit Organizations. In R.D. Herman (ed.), *Jossey-Bass Handbook of Nonprofit Leadership and Management*. San Francisco: Jossey-Bass.
Pallant, J. (2007). SPSS survival manual, a step by step guide to data analysis using SPSS for windows (3rd ed). Sydney: McGraw Hill.
Peter, J.P. (1979). Reliability: a review of psychometric basics and recent marketing practices. *Journal of Marketing Research*, 18. 133-45. http://dx.doi.org/10.2307/3150868
Public Company Accounting Oversight Board (PCAOB). (2003). Professionalism is primary. Remarks delivered by Douglas R. Carmichael at AICPA National Conference, Washington, D.C., (December 12): PCAOB.
Ramsey, P. (2006). Improved communality estimation in factor analysis. *Journal of Statistical Computation and Simulation*, 76,2, 93-101. doi: 10.1080/00949650412331320855
Rice, J. (1995). *Mathematical Statistics and Data Analysis* (Second ed.), Duxbury Press.
Ritchie, W.J., Kolodinsky, R.W. (2003). Nonprofit organization financial performance measurement: An evaluation of new and existing financial performance measures. *Nonprofit Management & Leadership*, 13(4), 367-381. doi: 10.1002/nml.5
Rojas, R.R. (2000). A Review of Models for Measuring Organizational Effectiveness Among For-Profit and Nonprofit Organizations. *Nonprofit Management and Leadership*, 11 (1), 97-104. doi: 10.1002/nml.11109
Ruppert, D. (2004). *Statistics and finance: an Introduction*. New York, NY: Springer-Verlag.
Schumacker, R. E., & Lomax, R. G. (2012). *A Beginner's Guide to Structural Equation Modeling, 3rd ed.* New York, NY: Taylor and Francis.
Siciliano, J.I. (1996). The Relationship of Board Member Diversity to Organizational Performance. *Journal of Business Ethics*, 15, 1313-1320. http://dx.doi.org/10.1007/bf00411816
Siciliano, J.I. (1997). The Relationship Between Formal Planning and Performance in Nonprofit Organizations. *Nonprofit Management and Leadership*, 7(4), 387-403. http://dx.doi.org/10.1002/nml.4130070405
Sori, Z., Hamid, M., Nassir, A, Mohamad, S. (2006). Some basic properties of financial ratios: Evidence from an emerging capital market. *International Research Journal of Finance and Economics*, 2, 71-87. http://dx.doi.org/10.2139/ssrn.923736
Stone, M.M, Bigelow, B., Crittenden, W. (1999). Research on strategic management in nonprofit organizations. *Administration and Society*, 31(3), 378-423. http://dx.doi.org/10.1177/00953999922019184
Tuckman, H.P., Chang, C.F. (1998). How Pervasive are Abuses in Fundraising Among Nonprofits? *Nonprofit Management and Leadership*, 9 (2), 211-221. doi: 10.1002/nml.9208

Venkatraman, N., Ramanujam, V. (1986). Measurement of business performance in strategy research: A comparison of approaches. *Academy of Management Review.* 11, 801-814 http://dx.doi.org/10.5465/amr.1986.4283976

Watts, R.L., & Zimmerman, J.L. (1986). *Positive accounting theory*, Englewood Cliffs, NJ: Prentice-Hall.