

DEVELOPMENT OF A SCALE TO MEASURE POTENTIALITY TOWARD BECOMING A GREAT EDUCATIONAL LEADER

ARNEL A. USMAN
Researcher

*Doctor of Philosophy
Major in Educational Management*

Abstract

A sequel to the researcher's thesis that generated a grounded theory, the present study explored the construct validity of the scale to measure the potentiality of school heads toward becoming great educational leaders. Seven out of the 13 original factors and 20 out of the 56 initial set of items made the final cut after undergoing Confirmatory Factor Analysis (CFA) using responses to the scale administered to 595 public school heads in Region III in the Philippines. The Influence of Teachers (IT) factor edged all others by registering the highest Average Variance Extracted (AVE) at .677 and topmost composite reliability at .861. Data analysis resulted in the hypothesized hierarchical model transfiguring into two oblique models with the latter being adjudged as the best and preferred among the three competing models; albeit the goodness of fit is not acceptable statistically (p -value of Chi-square is less than .05), all the fit indices (e.g. RMSEA, SRMR, NFI, GFI, ECVI) are within the acceptable range. Implications for school management and leadership are discussed amid the plethora of concepts that abound in literature.

Keywords: construct validity, potentiality, great educational leaders, Confirmatory Factor Analysis

INTRODUCTION

"Some are born great, some achieve greatness, and some have greatness thrust upon them."

(Shakespeare, 1601)

Are great leaders born or are they made? Di Giulio and Giulio (2015) examined various definitions and theories of leadership in answer to this age-old question. For instance, the authors defined leadership, citing a previous annotation by Northouse, as a process of influencing people in a new way to achieve a common goal. This description was resonated by Sorensen who traced the origin of the term "leader" in the 1300's in which it meant "to travel" or "to show the way."

The same paper quoted Bennis raising a word of caution by saying that the most dangerous leadership myth is that leaders are born – that there is a genetic factor to leadership. This myth asserts that people either have certain charismatic qualities or not.

Di Giulio and Giulio's research offers interesting insights on this old yet engaging issue on whether leadership is a pre-programmed trait in the human deoxyribonucleic acid (DNA) or something that is acquired and developed through time.

The study brings to mind The Great Man Theory popularized by Thomas Carlyle in the 1840s, which suggested that leadership traits are intrinsic, that great leaders are born and will emerge when confronted with the appropriate situation. In the 1860s, however, Herbert Spencer formulated a counter-argument that remained influential throughout the 20th century, asserting that such great men are the products of their societies and that their actions would be impossible without the social conditions built before their lifetime (Spector, 2015).

Surprisingly, this contentious issue on whether leaders are born or made did not escape the attention even of medical doctors who published an article in the American Journal of Pharmaceutical Education on the role of genetics in leadership in the hope of providing a physiological answer to the debate.

The American Association of Colleges of Pharmacy (2005) investigated the heritability of leadership by presenting studies conducted on twins to evaluate both genetic and environmental influences on leadership. As indicated in the article, these studies employed different methodologies such as the use of questionnaires and mathematical models to evaluate the genetic and shared or unique environmental experiences of twins in leadership roles.

The medical group cited the work of DeNeve, which tested what traits determine leadership role occupancy. Results revealed that the shared environment of the twins was not significant with only a 10% correlation; whereas, the genetic component was found to be

24%. The leadership role occupancy, the study pointed out, may be associated with the rs4950 genetic marker, a single nucleotide polymorphism (SNP) that resides on the neuronal acetylcholine receptor gene on chromosome 8, which can be related to personality traits, suggesting a link between leadership role occupancy and personality traits.

In the academic field, the birth of research on leadership has been attributed to James MacGregor Burns, regarded as the father of leadership studies, after his 1978 lecture on the subject (Sorenson, 2000). However, with the rapid rise of studies produced on leadership, the debate on whether it is nature or nurture, whether a leader is born or developed, continues and remains to be the topic of such undertakings.

A number of studies, for instance, looked into the traits of principals as possible determinants to the success of schools. Crawford (2002) sought to distinguish between the myth and fact of being a charismatic school leader. De Pree (2008) examined the essential elements of a great leader. Leithwood and Jantzi (2005) conducted an in-depth review of the characteristics of a transformational leader.

Moreover, the factors that influence school leadership and management have been the focus of numerous research. Amanchukwu, Stanley, and Ololube (2015) found that leadership theories affect leadership styles. Camburn, Rowan and Taylor (2003) scrutinized how distributed leadership shaped a model school. Tirrozi and Kappan (2001) described the evolving role of the secondary school principal in the face of the challenges brought about by the changing times.

In the Philippines, where education is highly regarded not only as an achievement but also as wealth (Doyle 2005), several studies on educational leadership have been explored primarily with the aim of proposing solutions to the problems and challenges being faced by school leaders.

San Antonio (2008), for example, used experimental research in probing how democratic school leadership creates better schools. Subjecting the participants in a one-year observation, the research yielded valuable inputs on how school leaders should utilize the indicators of this type of leadership toward school improvement. This was collaborated by the findings of Sindhvad (2009) which emphasized the role of principals as instructional leaders.

School leadership in the Philippines, Sutherland and Brooks (2013) explained, has historical, cultural and policy dynamics underpinnings. The influences of foreign colonizers shaped Philippine education in terms of system and policy. This in turn, according to the study, is translated on how school leaders manage schools. Kinship, for example, considered as the nucleus of the Filipino social organization, induces leadership practice in schools.

Notwithstanding the external factors that seemingly define and describe the plight of Philippine education, countless research have been carried out with the school leader as the focus. De Guzman and Guillermo (2007), for instance, centered on the role of the principals' tacit knowledge derived from their experiences, day-to-day dealings with the stakeholders, observations, insights, and reflections as vital inputs in understanding the dynamics of school leadership. Quiambao (2013) described in an autoethnography the lived experiences of a secondary public school principal assigned in a coastal community. In higher education institutions, Nicdao (2015) employed mixed methods research in examining the transformational leadership practices of presidents of state universities and colleges.

While there is a myriad of knowledge on educational leadership that abound in literature, the interest in educational leadership continues considering that the same problematic situation persists in schools. As can be gleaned from the aforementioned studies, the most important factor of the school's success is not the curriculum nor perennial issues such as class size, but, as enunciated in the findings of these works, and to borrow the words of Collins (cited in Pierce, 2013), having a great leader at the helm.

In fact, the Department of Education, through a series of seminars conducted by the National Educators Academy of the Philippines (2015) share the same view with regard to the pivotal role of school heads in the success or failure of the delivery of quality education in the public school system. As part of its Principal's Development Framework and Guidelines, the DepEd described the hindering and facilitating factors that principals face, citing stories on school leadership and management ranging from difficulties on the use of information and communication technology to misuse of funds, among others, in leading and managing schools. The inputs served as bases for the identification of leadership competencies expected of school heads.

Moreover, the results of the Principals' Test in the last two years, which registered a record low in the passing rate, 11 percent and two percent respectively, with domains in school leadership and school management operations at the bottom of the breakdown of scores, (D.M. 11, s. 2016 & D.M. 145, s. 2017), reflect the alarming situation on the plight of educational leadership in the country considering that most of the test-takers are designated already as school heads, officers-in-charge, pending their permanent appointment as full-fledged principals whose main requirement is the passing of the said examination. In other words, if the qualifying test is the main measure, specifically the low scores in school leadership and operations, it can be inferred that said school heads appear to lack the needed leadership competencies.

Clearly, the sad reality of educational leadership in the country, particularly in basic education institutions, is primarily an issue on the effectiveness of school heads, which the DepEd integrated into its competency-based standards. While it is true that effectiveness is a requisite to headship positions such as the principalship, and this is concretely enumerated

by the required competencies, it cannot be undermined that transformational leadership, ascribed in this study as great leadership, is of equal, if not, of higher importance.

In his best-selling book, Covey (2004) professed that effectiveness comes first but the one truly important facet of leadership is what succeeds it, and that is greatness. Also called transformational leadership, greatness means not only being effective by acquiring all the needed leadership competencies in order to attain success and quality, but also by leading others to do the same, which in the words of Covey is “finding your voice and helping others find their voices, too.” The author’s words entail competence, integrity, and humility, which in the end would be a leader’s lasting legacy.

The present study stemmed from the master’s thesis of the researcher (Usman, 2010) which generated a grounded theory on the making of great educational leaders. Purely qualitative in nature, the previous study, which had Covey’s mantra on effectiveness and greatness as part of its theoretical bases, sought to investigate what makes great educational leaders by looking through the significant experiences of the participants in the various developmental stages of life and by eliciting their personal insights on the factors that they think may have contributed to their being acknowledged as great educational leaders.

Findings of the study revealed that the participants had various experiences in the developmental stages of life that significantly contributed to their being acknowledged as great educational leaders. In childhood and adolescence, the significant experiences of the participants revolved in three areas: (1) home, (2) school, and (3) Church. Foremost of the findings were the participants’ humble beginnings, average academic standing, and active parish life. Other factors to which they attributed their being acknowledged as great educational leaders included faith in God, love for reading and the influence of teachers.

Embedded in the grounded theory are the categories, which emerged from the significant experiences of the participants in the developmental stages of life, and the other factors (leadership qualities and attributes) to which they credit their being acknowledged as great educational leaders.

Consequently, the present study seeks to develop a scale to measure the potentiality of school heads toward becoming great educational leaders. Potentiality, the variable under scrutiny, is the probability of something happening (Premuzic, 2016) which in the case of leadership is possible because it is partly dependent on genetic and childhood experiences. Most of the commonly used indicators to gauge leadership potential include educational achievement, emotional intelligence, ambition and intellectual quotient.

The present study, however, centered on significant human experiences in the developmental stages of life, including childhood experiences as indicated by Premuzic, as well as leadership qualities and attributes, which were based on the grounded theory, as possible determinants of potentiality toward becoming a great educational leader.

Accordingly, the conceptual paradigm of the present study may be described graphically as:

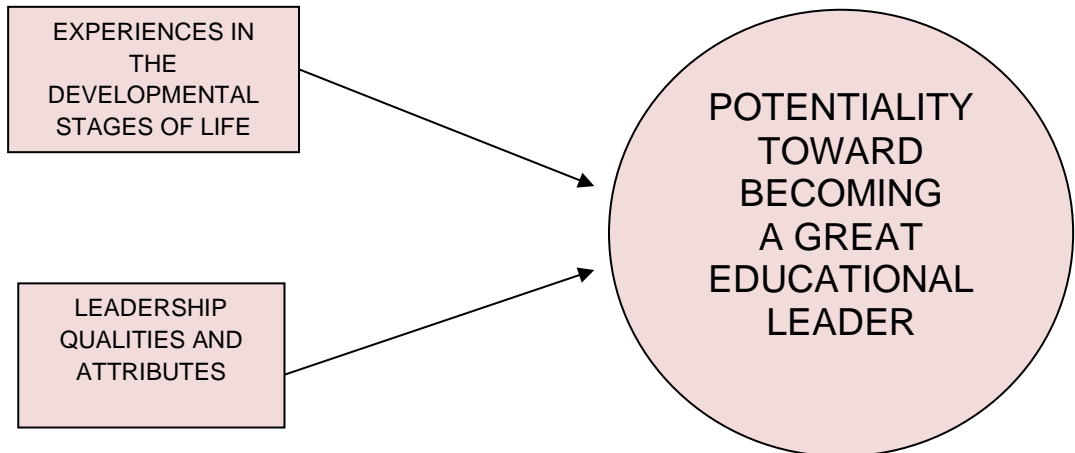


Figure 1. The Conceptual Paradigm: Hypothesized Model Scale to Measure Potentiality toward Becoming a Great Educational Leader

Confirmatory Factor Analysis was used in realizing the objective of the present study. Malo (2016) prescribed the existence of a theory as one of the requirements before conducting CFA. A theory, such as the aforementioned grounded theory, is the systematic set of causal relationships that provide the comprehensive explanation of a phenomenon (Statistics Solutions, 2013). The other requirement, according to Malo, is the scale which could either be newly developed or come from prior research. Undeniably, these two qualifications were met by the present study.

CFA is one of the two components of Structural Equation Modelling (SEM) that deals specifically with measurement models --- that is, the relationships between observed variables and latent variables (Malo).

Also called measured variable and manifest variable, an observed variable can be measured directly. It is represented by a square or rectangle graphically in a measurement or structural model (Suhr, 2017). In the present study, it refers to item or indicator.

A latent variable cannot be measured directly. It is also known as unobserved variable, latent construct, and latent dimension. It is depicted graphically with a circle or oval (Suhr). In the present study, it refers to factor or construct.

Unlike its counterpart, Exploratory Factor Analysis, CFA requires the researcher to prespecify all aspects of the model, which is the illustration of the theoretical relationships among the observed and unobserved variables.

To prespecify all aspects of the model, the researcher must have a firm priori sense, based on past evidence and theory, of the number of factors that exist in the data, of which indicators are related to which factors, and so forth (Brown, 2015). Again, these required conditions were satisfied in the present study.

While there is a prevalence in literature and relevant studies on the use of Exploratory Factor Analysis in instrument development, there appears to be a limited number, with the exception of a few such as Magno (2013) which assessed students' critical thinking skills and approaches to learning, that used Confirmatory Factor Analysis as the primary research method.

The growing popularity of SEM addresses this observation, as CFA is one of its two components; the measurement model, which shows the relationships between latent variables and their indicators. The other component is the structural model, which shows the potential causal dependencies between endogenous and exogenous variables (Prudon, 2013).

The present study is limited, however, to the measurement component of SEM as it focused only on scale development. It did not pursue the structural component, thereby, deferring the testing of the grounded theory. In particular, the present study focused on potentiality as one possible determinant toward becoming a great educational leader. There could be other factors or variables that may likewise serve as determinants and this could be answered in a full SEM study.

Given this limitation, the present study is, nevertheless, worth pursuing because of the following reasons, namely: (1) the continuing problematic situation in Philippine education with regard to the lack of great educational leaders; (2) the apparent scarcity on the use of CFA as a research method in scale development; and (3) the novelty of the present study being anchored on a previous study which generated a grounded theory.

Undoubtedly, these underlying reasons point to the gap in knowledge that this study humbly aims to address and the contribution that it may offer to the body of knowledge on educational leadership and future research undertakings.

Specifically, results of the study would be of significance to educational agencies and institutions, particularly the Department of Education, in assessing and strengthening their programs on human resource and development, which includes recruitment, hiring, placement and induction, as well as learning and professional development.

Likewise, school administrators and teachers, particularly those holding school head positions such as the test-takers mentioned earlier, who have aspirations of becoming principals, may benefit from the outcome of the study as this would provide them inputs on the essential factors, from the needed experiences and indispensable qualities, in order to be great educational leaders themselves in the future.

In addition, researchers who are engaged in instrument development may find the results of the present study useful in conducting similar studies. As an offshoot of this study, a full SEM may be pursued as a type of theory-testing undertaking.

METHOD

Anchored on the previous study which generated a grounded theory, this is a quantitative study that used Confirmatory Factor Analysis in developing a scale to measure the potentiality of school heads toward becoming great educational leaders.

Furr (2013) described CFA as a means to evaluate “measurement hypotheses” regarding a scale’s internal structure. It enables researchers to evaluate the degree to which their measurement hypotheses are consistent with actual data produced by respondents using the scale. Three key sets of results (parameter estimates, fit indices, and modification indices) are examined to test measurement hypotheses, which can be later modified to be more consistent with the actual structure of participants’ responses to the scale.

The steps in conducting CFA and reporting the results that were followed in this study were patterned after that of Brown (2015); Furr and Bacharach (2013); Jackson, Gillaspay, and Stephenson (2009); Malo (2016); Schreiber, Nora, Stage, Barlow and King (2006); and Suhr (2017). Adhering to the works of these experts lend credence to the manner by which the study was conducted and further solidify the results.

Preliminary Steps

At least two important preliminary steps are required before conducting CFA. The initial development of a scale is the first of these steps. For the present study, a 55-item scale was initially constructed based on the existing grounded theory. Schreiber et.al., described CFA as a confirmatory statistical technique, which should be theory-driven. As articulated, the present study was the major result of a previous study that generated a theory with plenty of interrelated variables.

The initial scale was composed of 13 constructs, referred to in this study as factors, which were the same emergent categories in the theory, namely: (1) family upbringing; (2) academic foundation; (3) church involvement; (4) college formation; (5) joy in teaching; (6) administrative and leadership preparation; (7) faith in God; (8) support of family; (9) values; (10) passion for work; (11) love for reading; (12) continuous learning; and (13) influence of teachers. Indicators, referred to in this study as items, for each factor were drawn from the definitions/statements in the said categories of the theory.

The instrument used a six (6) point Likert Scale where option one (1) represented the “very untrue of me” category while option five (5) referred to the “very true of me” category.

Likert's scale six (6) points had a higher trend of discrimination and reliability than Likert's scale five (5) points (Chomeya, 2010).

The initial draft of the questionnaire was subjected to content validation by five (5) experts with specialization on instrument development and validation. Each expert was asked to rate each item as to content and language with the rating scale set to one (1) for "not manifestive" and four (4) for "highly manifestive".

Content validity concerns the degree to which a sample of items, taken together, constitute an adequate operational definition of a construct (Polit & Beck, 2006), which in the case of the present study, resulted in the modification of the draft particularly on item 17 which was split into two specific indicators, hence, resulting in a final draft consisting of 56 items.

The second preliminary step in CFA is the collection of responses to the scale. The respondents for this study consisted of 595 public school heads from two (2) provincial Schools Division Offices (SDOs) and one city Schools Division Office (SDO) under the Department of Education Regional Office III. The respondents were selected based on the following criteria: (1) shall have been issued a permanent appointment by the Department of Education; (2) shall have been appointed or designated as school head in any of the following positions, namely: principal, head teacher, school officer-in-charge, or teacher-in-charge; (3) shall have been appointed or designated as school head in any of the following levels: elementary, junior high school, senior high school or integrated school; and (4) shall have had at least three years of experience as school head. Weindling (2000) described the third year of headship as the "refinement" stage characterized by the accumulation of experiences beginning with the preparation and entry in the first year and going to the consolidation and plateau years in the succeeding years.

The sample size collected through random sampling was deemed sufficient as Worthington and Whittaker (2006) argued that one common rule of thumb is to ensure a person-to-item ratio of 10:1; another rule of thumb is that $N = 300$ is usually acceptable. Malo (2016) posited that measurement models with a large number of constructs require a minimum sample of 500.

The questionnaire was pilot tested to a group of 46 school heads coming from 14 out of the 20 schools division offices (SDOs) in Region III.

A pilot test can be defined as a small study to test research protocols, data collection instruments, sample recruitment strategies, and other research techniques in preparation for a larger study. A pilot study is one of the important stages in a research project and is conducted to identify potential problem areas and deficiencies in the research instruments and protocol prior to implementation during the full study (Hassan, Schatnner & Mazza, 2006).

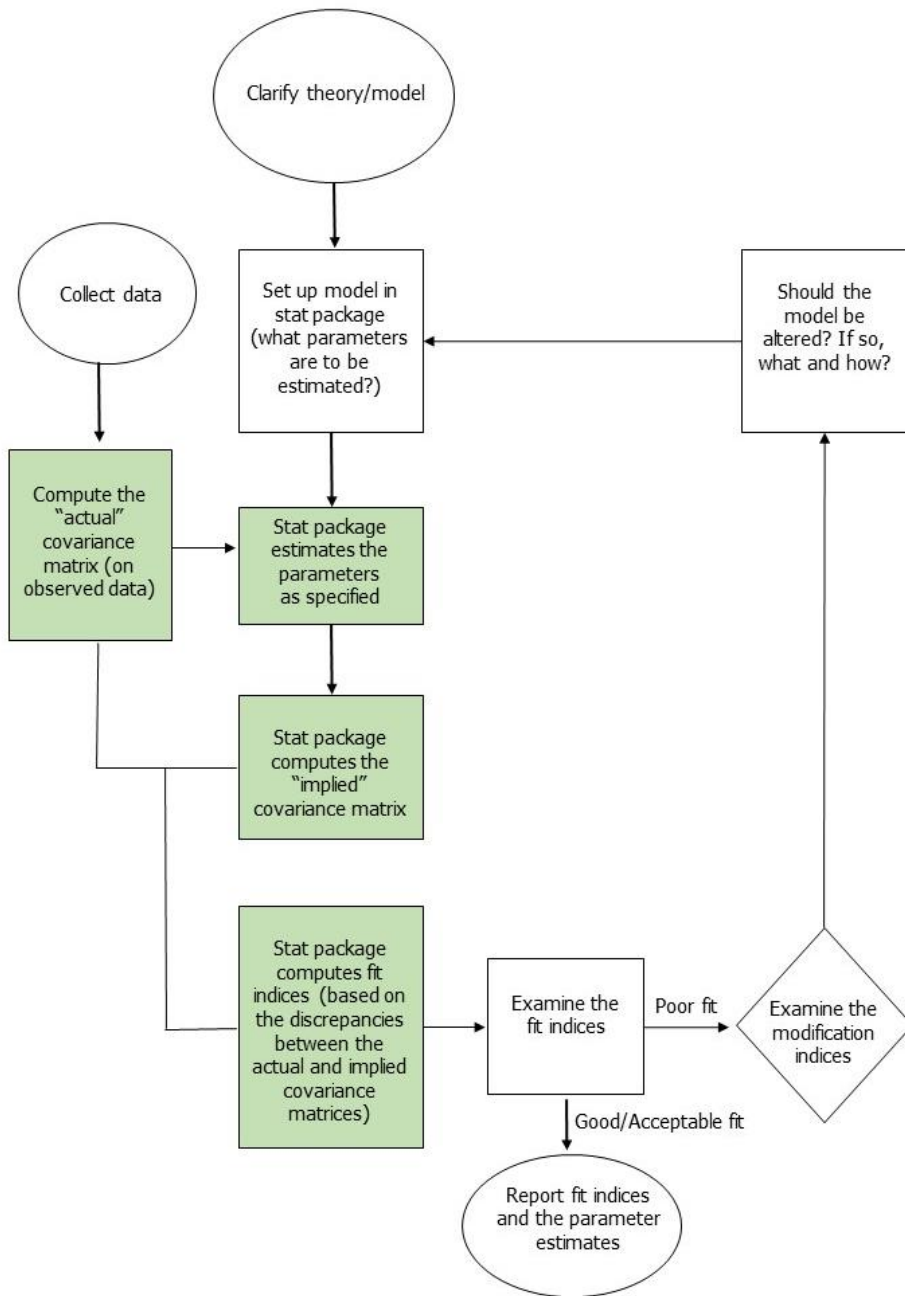
To address the issues that came out during the pilot study, such as certain items being skipped and the average time in filling out the responses took longer than expected, it was decided that an online version of the instrument be created to facilitate the process.

Online surveys are cost-efficient, practical, and results are gathered and processed easily and quickly (De Bois, 2016). However, some of the disadvantages include lack of conscientious responses from some of the respondents, lack of personal encounter, and accessibility issues. It was deemed necessary by the researcher that two versions of the questionnaire, the online and hardcopy, be made available for the respondents to choose which method suits them.

Coordination with concerned authorities was made through formal communication in order to secure their permission in administering the instrument. The objective of the study was explained to the respondents prior to the administration of the instrument. Their right to withdraw from the study was also discussed as well as the provision of confidentiality. Respondents who opted to answer the questionnaire online were provided a supplementary note regarding this clause.

After the questionnaires have been collected, the data were prepared for statistical treatment. Codes were used to represent the 13 factors for easy reference: FU (Family Upbringing); AF (Academic Foundation); CI (Church Involvement); CF (College Formation); JT (Joy in Teaching); AL (Administrative and Leadership Preparation); FG (Faith in God); SF (Support of Family); V (Values); PW (Passion for Work); LR (Love for Reading); CL (Continuous Learning); and IT (Influence of Teachers). The overall construct was coded PGEL (Potentiality towards becoming a Great Educational Leader).

Figure 2 illustrates the process flowchart in conducting Confirmatory Factor Analysis as suggested in Furr (2013) which was adopted in the present study. The figure differentiates steps in which researchers have active roles (the unshaded boxes) from those performed by the statistical software (in shaded boxes). The figure also shows that CFA can be an iterative process in which the hypothesized measurement model is articulated, evaluated, revised, and re-evaluated. Ultimately, researchers report information about the model-testing process, including any revisions to the model(s), with primary attention to the model best matching the scale's actual internal structure.



Legend: ○ start/end of the process □ step performed by researcher
 ◇ decision point ■ step performed by statistical software

Figure 2. Flowchart of Confirmatory Factor Analysis

Specification of Measurement Model

The preliminary steps in CFA set the stage for the first step in the CFA proper, which was the articulation of the measurement model with the help of CFA-capable software. Statistical software packages, such as IBM SPSS Amos v.23, which was used in the present study, have lessened, if not eradicated, this tedious process, which used to demand familiarity with matrix algebra and esoteric programming syntax. Researchers now draw diagrams depicting their measurement models, with the software translating diagrams into appropriate statistical notation for processing.

At least three important facets of the measurement model were specified: (1) the number of factors or latent variables (represented by ovals) hypothesized to underlie the scale's items (represented by rectangles); (2) the items linked to (i.e., load on) each factor, with at least one item linked to each factor (though this is not necessarily true for hierarchical measurement models); and (3) the possible associations between factors.

In the present study, Model 1 served as the hypothesized measurement model (See Appendix E). This model reflects a hierarchical measurement model, which can be described as follows: PGEL, being the overall construct, causes the 13 factors and each factor causes the corresponding set of items.

In CFA, the software estimates the precise association of the items and factors. In addition, when working with multi-dimensional models, researchers specify that particular items are not associated with one or more latent variables. Items have a zero loading on one or more latent variables. For example, Model 1 hypothesized items 1 to 5 load only on Family Upbringing (FU) factor and that items 6 to 8 load only on Academic Foundation (AF) factor. This is illustrated by means of a pathway (i.e., arrow) between an item and a factor, which indicates that there is a hypothesized non-zero association between the two. On the other hand, the lack of a pathway indicates a hypothesized zero association. Model 2 (See Appendix F), on the other hand, implies that all factors are hypothesized to be correlated with each other to some degree and each factor causes the corresponding set of items.

Analysis

Analysis of the model and the data from the answers of the respondents to the scale served as the next step after model specification. This involves a four-phase process with the statistical software performing the following functions: (1) computes the collected data to get the items' actual variances and covariances; (2) uses the item's actual variances and covariances to estimate item parameters and to gauge the model's accuracy; (3) uses the estimated parameters to create "implied" item variances and covariances; and (4) generates information reflecting the overall adequacy of the hypothesized model.

Covariance measures the direction and magnitude of the relationship between two variables/factors. It is similar to variance, but where variance tells how a single variable /factor varies, covariance tells how two variables/constructs vary together (Hall, 2018).

The information gathered in the analysis of the model and the data from the answers of the respondents to the scale is useful in gauging the model's accuracy. In the second phase of the analysis, for instance, the software used the items' actual variances and covariances to estimate parameters as specified by the researcher. The software used the actual association between two items to estimate the factor loadings that they might have on a common factor. Such information was needed to estimate all factor loadings, inter-factor correlations, error variances, and so on. In the third phase, the software used the estimated parameters to create "implied" item variances and covariances. Since the hypothesized model in the present study was poor, it could be surmised that the implied values differed greatly from the actual values. In the fourth phase, the software generated information reflecting the overall adequacy of the hypothesized model. The large discrepancies in the indices, as in the case of Model 1, showed that the hypothesized measurement model did not adequately account for the scale's data. The "modification indices" provided specific ways in which the hypothesized measurement model might be modified.

Interpreting and Reporting Output

After conducting the analyses, the results were interpreted and reported. To test the overall adequacy of the model, fit indices were examined with their acceptability range as a guide. For the present study, the following sources on model fit indices were consulted, namely: Bagozzi and Fornell (1982); Baron and Kenny (1986); Hair, Ringle, and Sarstedt (2011); and, Hooper, Coughlan, and Mullen (2008).

A good fit indicates that the hypothesized measurement model is consistent with observed data while a poor fit reveals the opposite and is interpreted as evidence against the adequacy of the model.

Regarded as the most commonly used index, the Chi-square was computed to show the degree of mis-fit of the model. Small values indicate support for the model or good fit while large values indicate strong evidence against the model or poor fit. However, due to the fact that sample size affects Chi-square (large samples produce large Chi-square values), as in the case of the present study, other fit indices were explored to obtain better results.

Stapleton (1997) explained that the Chi-square statistic is very sensitive to sample size, rendering it unclear in many situations whether the statistical significance of the Chi-square statistic is due to poor fit of the model or to the size of the sample.

The following fit indices were used to evaluate the acceptance or rejection of the measurement models in the present study: Root Mean Square of Approximation (RMSEA), Standardized Root Mean Square Residual (SRMR), Normed Fit Index (NFI), Goodness of Fit

Index (GFI), Adjusted Goodness of Fit statistic (AGFI), Comparative Fit Index (CFI), Akaike Information Criterion (AIC), Consistent AIC (CAIC), and the Expected Cross-Validation Index (ECVI). (For the range of the acceptability of each of the said fit indices, please refer to the aforementioned sources on fit indices. As an example, a discussion on model fit indices by Hooper, Coughlan, & Mullen is provided in Appendix I).

After examining the overall fit of the hypothesized measurement model, parameter estimates were examined, particularly the items' factor loadings, inter-factor associations, and error variances. These parameters were vital in evaluating the scale's factorial structure and psychometric properties.

Indicator/Item loading is the relationship between an item and construct. P values associated with the item loadings should be equal to or lower than .05; and that the loadings be equal to or greater than .50 (Hair et al., cited in Kock, 2013). Items/indicators for which these criteria were not satisfied were removed to improve the model.

Aside from fit indices and factor loadings, other statistical tools were used in this study to lend validity and reliability to the results: (1) Average Variance Extracted (AVE) which is the proportion of variance in the items that is explained by the construct/latent variable. An AVE threshold frequently recommended for validity is .50 (Fornell & Larcker, cited in Kock, 2013); and (2) Composite Reliability and Cronbach's Alpha which are both measures of internal consistency. The rule of thumb about composite reliability and the Cronbach's alpha are discussed in Kock as follows: "More conservatively, both the composite reliability and the Cronbach's alpha should be equal to or greater than 0.7 (Fornell & Larcker; Nunnally; Nunnally & Bernstein, cited in Kock). The more relaxed version of this criterion, which is widely used, is that one of the two coefficients should be equal to or greater than 0.7. This typically applies to the composite reliability coefficient, which is usually the higher of the two (Fornell & Larcker, cited in Kock). An even more relaxed version sets this threshold at 0.6 (Nunnally & Bernstein, cited in Kock). If a latent variable does not satisfy any of these criteria, the reason will often be one or a few indicators that load weakly on the latent variable. These indicators should be considered for removal" (p.1).

Model Modification and Re-Analysis (if necessary)

Since the hypothesized measurement model of the present study revealed poor fit indices, weak and non-significant factor loadings, and low validity and reliability, the researcher deemed it necessary to modify and re-analyze the model.

In CFA, more than one model may accurately describe the data and that a number of fit indices should be used to determine the fit of the various models.

Guided by the software, the researcher analyzed the modification indices of the hypothesized measurement model. Each modification index referred to a specific parameter that was set to zero in a measurement model.

The size of modification index reflected the benefit of revising the relevant parameter, which eventually showed the improvement of the overall fit of the model. Thereafter, software conducted re-analysis. This resulted in revised models with entirely new outputs --- new fit indices and new parameter estimates, among others.

Comparing Models

With the fit indices of the hypothesized measurement model, referred to in this study as Model 1, landing outside the acceptable range, the said model was modified and re-analyzed as described in the previous CFA step.

It is important to note that in interpreting the findings of CFA analysis, more than one model can be determined that will adequately fit the data (Biddle & Marlin; Thompson & Borrello, cited in Stapleton, 1987) and other ways can be evaluated in order to improve the model.

Model 2 was tested as an alternative to the first model. However, Model 2 was still problematic in terms of fit indices, validity, and reliability.

Since the two CFA models (Models 1 and 2) were considered poor and problematic, an improved model (called Model 3) was proposed which emerged as the best and preferred model with relatively strong fit to the observed data.

Statistical Package for Social Sciences (SPSS) v. 23 was utilized to carry out the computations and analyses.

RESULTS

As articulated in the previous section, preliminary steps were undertaken prior to the specification of the hypothesized measurement model: (1) the initial development of the scale and (2) the collection of responses to the scale.

The initial survey questionnaire was composed of 55 items, which was later improved to a 56-item scale after content validation by experts. The instrument was administered to the randomly selected respondents composed of 595 public school heads from two (2) provincial Schools Division Offices (SDOs) and one city Schools Division Office (SDO) under the Department of Education Regional Office III.

Results of Confirmatory Factor Analysis

Data gathered were prepared for statistical treatment using computer software. Codes were used to represent the 13 factors for easy reference: FU (Family Upbringing); AF (Academic Foundation); CI (Church Involvement); CF (College Formation); JT (Joy in Teaching); AL (Administrative and Leadership Preparation); FG (Faith in God); SF (Support of Family); V (Values); PW (Passion for Work); LR (Love for Reading); CL (Continuous Learning); and IT (Influence of Teachers). The overall construct was coded PGEL (Potentiality towards becoming a Great Educational Leader).

As described in the CFA flowchart, the first step was the specification of the hypothesized measurement model referred to in this study as Model 1. Briefly, the model being confirmed can be described as follows: PGEL (Potentiality toward becoming a Great Educational Leader), being the overall construct, causes the 13 factors and each factor causes the corresponding set of items. (See Appendix E). The model is a hierarchical model which is characterized by the use of single-headed arrows to imply direct effect of the factors to the indicators.

After subjecting the model to data analysis via statistical software, it was revealed that most of the fit indices for Model 1 were outside the acceptable range. The construct validity (both convergent and discriminant validity) for Model 1 was problematic. For example, factor loadings of many items were less than the .5 threshold. (See Column 2 of Table 1 for the fit indices of Model 1 and Appendix E for the statistical software analysis outputs of Model 1).

Since the specified hypothesized measurement model was considered problematic, model modification and re-analysis ensued which resulted in Model 2, which was tested as an alternative measurement model. Model 2, an oblique model, can be described as follows: The 13 factors are assumed correlated and each factor causes the corresponding set of items. An oblique model is characterized by the use of double-headed arrows to denote correlation among the factors.

However, Model 2 was still problematic in terms of fit indices, validity, and reliability. (See Column 3 in Table 1 for the fit indices of Model 2 and Appendix F for the statistical software analysis outputs of Model 2).

Considering that the two measurement models (Models 1 and 2) were problematic, an improved model (called Model 3) was proposed. This oblique model is composed of seven (7) factors which are assumed to be correlated and each factor causes the corresponding set of items (See Appendix G for the statistical software outputs of Model 3).

After comparing the three models, which was the last step in the CFA process, Model 3 emerged as the best and preferred model with relatively strong fit to the observed data.

Accordingly, this model shall be the focus of the presentation and interpretation of the results of the present study.

Results of Goodness-of-Fit and Fit Indices

As can be gleaned from the results, Model 3 emerged as the best and preferred model because all the fit indices (e.g., GFI, RMSEA, SRMR) are within the acceptable range although the goodness of fit is not acceptable statistically (p -value of Chi-square is less than .05).

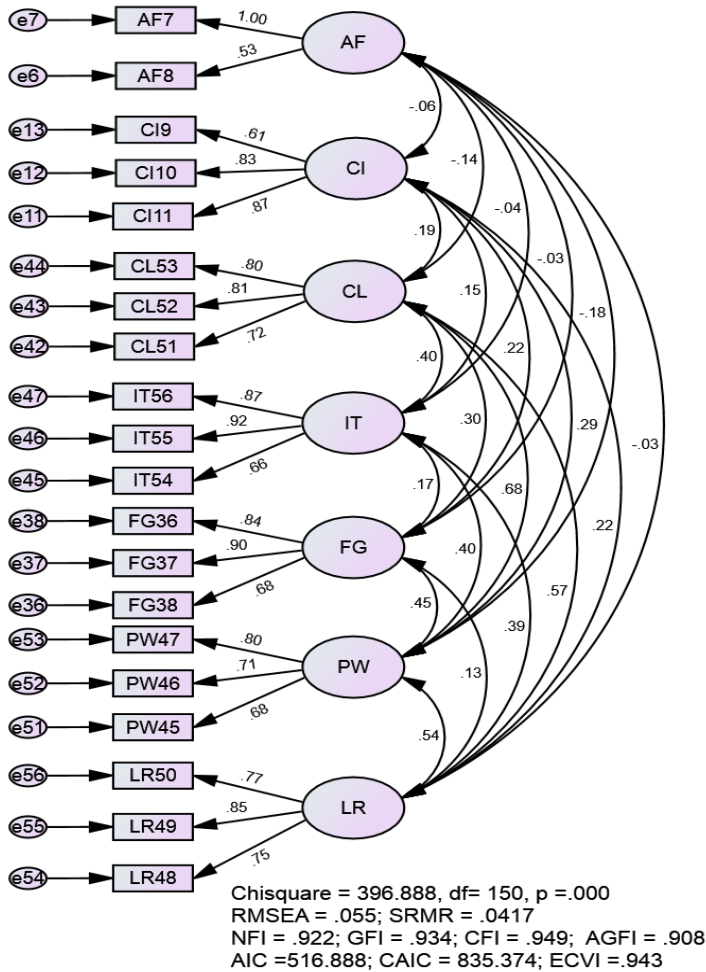
Table 1 presents the goodness-of-fit and fit indices of the three models which shows that Model 3 emerged as the best among the three competing models because of the strong statistical evidence that the estimates are acceptable: the Root Mean Square Error of Approximation (RMSEA) = 0.055 (acceptable within the range 0.05 to 0.08); Standardized Root Mean Square Residual (SRMR) = .0417 (acceptable within the range 0 to 0.05) ; Normal Fit Index (NFI) = .922 (acceptable within the range .90 to .95); Goodness of Fit Index (GFI) = .934 (acceptable within the range .90 to .95); Adjusted Goodness of Fit Index (AGFI) = .908 (acceptable within the range .90 to .95); Comparative Fit Index (CFI) = .949 (acceptable within the range .95 to .97).

Table 1
Goodness of fit and fit indices

Goodness-of- Fit and Fit Indices	Model 1	Model 2	Model 3	Remarks
Chi-square P value	3273.36 0.000	2933.396 0.000	396.888 0.000	Model 3 has the smallest chi-square; hence, it is better than Models 1 and 2
Chisquare/df	2.225	2.086	2.646	All three models within the acceptable fit range.
RMSEA	.047	.048	.055	Model 3 within the acceptable fit range.
SRMR	.062	.055	.042	Model 3 is the best among the three models.
NFI	.720	.726	.922	Model 3 is the best among the three models.
GFI	.810	.831	.934	Model 3 is the best among the three models.
AGFI	.794	.809	.908	Model 3 is the best among the three models.
CFI	.822	.849	.949	Model 3 is the best among the three models.
AIC	3523.36	3313.396	516.888	Model 3 is the best among the three models.
CAIC	4186.87	4321.935	835.374	Model 3 is the best among the three models.
ECVI	6.429	6.046	.943	Model 3 is the best among the three models.

Legend: RMSEA= Root mean square error of approximation; SMSR= Standardized Root mean square residual; NFI= Normal Fit Index; GFI = Goodness of Fit Index; AGFI = Adjusted Goodness of Fit Index; CFI= Comparative Fit Index; AIC= Akaike Information Criterion; CAIC= Consistent AIC; ECVI= Expected Cross Validation Index.

Figure 3 illustrates Model 3 which emerged as the best and preferred measurement model composed of seven (7) factors with each factor consisting of a set of items, namely: Academic Foundation (2 items); Church Involvement (3 items); Faith in God (3 items); Continuous Learning (3 items); Influence of Teachers (3 items); Passion for Work (3 items); and Love for Reading (3 items).



Legend: AF (Academic Foundation); CI (Church Involvement); CL (Continuous Learning); IT (Influence of Teachers); (FG) Faith in God; (PW) Passion for Work; LR (Love for Reading)

Figure 3. Measurement Model 3

Results of Convergent and Reliability Statistics

After establishing the final structure of the scale/instrument, construct validity such as convergent validity and discriminant validity along with reliability were determined.

The Average Variance Extracted (AVE), regarded as the proportion of variance in the items that is explained by the construct/latent variable, was computed. The AVE threshold frequently recommended for validity is 0.50 (Fornell & Larker, cited in Kock, 2013). It can be deduced from Table 2 that all the AVE's of the constructs/factors are above 0.5, hence the scale has good convergent validity.

In terms of internal consistency, composite reliability and Cronbach's alpha were computed. The rule of thumb about composite reliability and Cronbach's alpha suggests that, more conservatively, both should be equal to or greater than 0.7 (Fornell & Locker; Nunnally; Nunnally & Bernstein; cited in Kock). All the composite reliability values are all above .7 suggesting that the scale has internal consistency.

Table 2
Convergent validity and reliability statistics

Constructs/Items	Item Loadings	Average Variance Extracted	Composite Reliability
A. Academic Foundation (AF)		.642	.766
1. I was not 'conscious' of academic awards when I was in elementary and high school.	0.532		
2. I preferred to be just in the corner with no active involvement in extra or co-curricular activities when I was in elementary and high school.	1.000		
B. Church Involvement (CI)		.605	.821
3. My family exposed me to parish/Church activities, which characterized my very active parish/Church life at a young age.	0.871		
4. I joined in parish/Church organizations and served in simple acts such as cleaning the chapel or assisting in the Eucharistic celebrations/worship service during my childhood and adolescence.	0.829		
5. I was a regular churchgoer when I was growing up.	0.611		

Constructs/Items	Item Loadings	Average Variance Extracted	Composite Reliability
C. Faith in God (FG)		.662	.853
6. I have a personal relationship with God.	0.682		
7. I lift up and attribute everything to God.	0.901		
8. God is at the center of my life.	0.841		
D. Continuous Learning (CL)		.605	.821
9. I continue to seek out for new knowledge.	0.716		
10. I view learning as a continuous process.	0.813		
11. New learning excites me.	0.801		
E. Influence of Teachers (IT)		.677	.861
12. I had (a) teacher/s who greatly influenced my life.	0.657		
13. My former teacher/s inspired me.	0.917		
14. I was motivated by my former teachers.	0.872		
F. Passion for work (PW)		.540	.778
15. I see to it that I always exert my best effort in any work or project.	0.682		
16. I got interested in managing schools because I have a passion for education.	0.714		
17. I have a passion for excellence.	0.805		
G. Love for Reading (LR)		.627	.834
18. I love reading books and content materials.	0.754		
19. The books that I read help me in managing my school.	0.850		
20. I regularly read books on management and other fields.	0.768		

Results of Discriminant Validity Statistics

Table 3 shows the discriminant validity statistics. Discriminant validity is a measure of the quality of a measurement instrument and is computed by getting the square root of the AVE of each latent variable. The values should be higher than the correlations with all the other latent variables (Mora, Steenkamp & Raisinghani, 2012). Obviously, the scale has discriminant validity since the diagonal elements are all larger than the off-diagonal elements.

Table 3
Discriminant validity statistics

	AF	CI	FG	CL	IT	PW	LR
AF	0.801						
CI	-0.061	0.779					
FG	-0.029	0.221***	0.814				
CL	-0.144**	0.194***	0.298***	0.778			
IT	-0.04	0.154**	0.170***	0.404***	0.823		
PW	-0.181***	0.286***	0.455***	0.675***	0.404***	0.735	
LR	-0.03	0.217***	0.132**	0.575***	0.385***	0.535***	0.792

Note: Diagonal elements are the square root of AVE between constructs.

For discriminant validity, the diagonal elements should be larger than the off-diagonal elements.

In conclusion, after having satisfied all the foregoing validity and reliability indicators, it could be said that Model 3, the emerged scale for measuring the potentiality of school heads toward becoming great educational leaders, is deemed valid and reliable.

The Validated Scale

After undergoing Confirmatory Factor Analysis, which resulted in Measurement Model 3 as the best and preferred measurement model for the development of a scale to measure the potentiality of school heads toward becoming great educational leaders, the following validated instrument composed of 20 items serve as the result of the present study.

The first five (5) items are from the two (2) factors in Measurement Model 3, which were taken from the statements in the categories of the grounded theory that answered the first objective of the previous study, which pertained to the childhood experiences of the participant-great leaders.

The remaining 15 items are from the other five (5) factors in the CFA model which were taken from the statements in the categories of the said theory that answered the second objective of the study, which pertained to the participants' insights as to what other reasons to which they attribute their being acknowledged as great educational leaders.

**VALIDATED SCALE TO MEASURE POTENTIALITY TOWARD BECOMING
A GREAT EDUCATIONAL LEADER**

	INDICATORS	1	2	3	4	5	6
1	I preferred to be just in the corner with no active involvement in extra or co-curricular activities when I was in elementary and high school.						
2	I was not 'conscious' of academic awards when I was in elementary and high school.						
3	I was a regular Churchgoer when I was growing up.						
4	I joined in parish/Church organizations and served in simple acts such as cleaning the chapel or assisting in the Eucharistic celebrations/worship service during my childhood and adolescence.						
5	My family exposed me to parish/Church activities, which characterized my very active parish/Church life at a young age.						
6	God is at the center of my life.						
7	I lift up and attribute everything to God.						
8	I have a personal relationship with God.						
9	I see to it that I always exert my best effort in any work or project.						
10	I got interested in managing schools because I have a passion for education.						
11	I have a passion for excellence.						
12	I love reading books and content materials.						
13	The books that I read help me in managing my school.						
14	I regularly read books on management and other fields.						
15	I continue to seek out for new knowledge.						
16	I view learning as a continuous process.						
17	New learning excites me.						
18	I had (a) teacher/s who greatly influenced my life.						
19	My former teacher/s inspired me.						
20	I was motivated by my former teachers.						

DISCUSSION

The results show that from the original 13 factors and 56 items in the scale, seven (7) factors consisting of 20 items, namely: (1) Academic Foundation; (2) Church Involvement; (3) Continuous Learning; (4) Influenced of Teachers; (5) Faith in God; (6) Passion for Work; and (7) Love for Reading were confirmed after undergoing Confirmatory Factor Analysis.

The statistical outcome, illustrated in the measurement model, referred to in this study as Model 3, implies that the aforementioned factors and items met the optimal combination of the following: Goodness-of-fit and fit indices, Convergent Validity (e.g., factor loading, AVE), Discriminant Validity and Reliability.

The primary goal of CFA is to make sure that the values of such indices are at least within the acceptable range. As presented in the results, the two prior models failed in satisfying this goal. The process of CFA is to improve the model by removing items. For example, convergent validity can be improved if both item factor loading and AVE are not less than .50. Even if item factor loadings and AVE are already .50 or higher, but the criteria for discriminant validity is not yet met, the researcher continues to remove items with the smallest factor loadings (Stapleton, 1997).

It is thus safe to state that the deleted six (6) factors, namely: (1) Family Upbringing; (2) College Formation; (3) Joy in Teaching; (4) Academic and Leadership Preparation; (5) Support of Family; and (6) Values, consisting of a total of 36 items, did not meet these statistical criteria and do not contribute to the model. Consequently, the said deleted factors and items are statistically insignificant to the development of the scale.

It is to be recalled that the deleted factors and items, except for Values, were from the emergent categories that elicited the significant experiences of great educational leaders in the different developmental stages of life, which served as the first objective of the grounded theory, which was the basis of this study. On the contrary, the confirmed factors and items that surfaced from the CFA were from the emergent categories in the same theory which asked for the insights of the participant-great leaders, the other research objective, as to what other reasons aside from experiences (which turned out to be leadership qualities and attributes) contribute to their being acknowledged as great educational leaders.

With this striking observation, it may be misconstrued that the results of the CFA may debunk the existing grounded theory. However, as highlighted in the previous chapters, the present study is not meant to test the said theory, as this would require full Structural Equation Modelling (SEM). Rather, it solely seeks to develop a scale based on one aspect of the theory, that is, potentiality.

Confirmatory Factor Analysis, as Furr (2013) opined, is a means to facilitate theory testing, theory comparison, and theory development in a measurement context. CFA is the

measurement component; the structural component is the other part needed to make it a SEM and to test the existing theory. What is being tested in CFA is the scale's internal structure -- the measurement hypotheses that can be modified to be more consistent with the actual structure of participants' response to the scale.

Since the focus of CFA is on the factors and items that were confirmed, it would be good to examine their relationships, which in the words of Brown (2015) is the relationship between observed measures of indicators (referred to in this study as items) and latent variables (referred to in this study as factors). Likewise, Malo (2016) posited that in examining the relationship between the factors and items, CFA validates the measurement model on three accounts: (1) whether the factors are unidimensional and valid; (2) the number of items that should be used for each factor; and (3) the ability of the measures to portray the factor or explain it.

Among the items under Academic Foundation (AF), item 7, which is the "preference of the respondents to take no involvement in extra and co-curricular activities when they were in elementary and high school", exhibited the highest loading of 1.000 not only inside this factor but also in the entire measurement model.

This item runs contrary to the popular and widely accepted belief that extra and co-curricular activities, such as that advocated in Billah (2017) and previously affirmed in the studies of Daniyal, Nawaz, Hassan and Mubeen (2012), complement academic courses in the education of the child.

Contradictory as it may seem, this particular item does not discount the importance of extra and co-curricular activities in the academic growth and in the future of the learner. As noted in the previous chapters, the scale under study is just a part of the existing theory, which is to validate the scale. Whether the particular item is true or not remains to be proven in a full SEM study.

Aside from the item on extra and co-curricular activities, the item on "not being conscious (mindful) of academic awards in elementary and high school" was confirmed through CFA as an indicator of the Academic Foundation factor.

Although academic honors are viewed generally as a form of motivation for students to aspire for academic excellence and achievement, there have been studies that support this item on the possible adverse effects of academic awards on the education of the child. For example, Bursztyrn and Jensen (n.d.), in a natural experiment involving high school learners, found that fear of social sanctions was one of the effects of public recognition by means of awards and honors.

Likewise, it appears unusual that one original item in the Academic Foundation factor, item 6, which stated "I was just an average student in terms of academic performance when I was in elementary and high school," was deleted due to low item loading among other

reasons. Unlike the other remaining six (6) factors, which had their original set of items still intact, the AF factor lost this original item. As is the case in CFA, which is likewise the explanation for the deleted six (6) factors and 36 indicators, items with the smallest factor loadings (Malo, 2016) are deleted and removed because they do not meet the optimal combination of statistical measures such as Goodness-of-Fit, fit indices, convergent validity, discriminant validity, and reliability.

Suhr (2017) stated that CFA requires specification of the number of factors and which items load on each other; hence, it is confirmatory in nature. On the possibility of the deleted items or factors loading to other items or factors, cross loading or secondary loading, this is not part of the output (Prudon, 2013). CFA does not determine whether a certain item would perhaps have been better assigned to another cluster/factor especially if its primary loading is low. Instead, modification indices for each item are part of the output. These indices give an impression of how the GOF indices would improve if the item would be deleted from its predicted factor.

The second factor that was confirmed by CFA in the scale under study is Church Involvement (CI). "Exposure to parish/Church activities at a young age" scored the highest loading in this cluster followed by "membership in organizations" and "regular attendance to Church". This implies that this factor contributes significantly to the scale, suggesting that this particular experience in the early developmental stages of the participant-great educational leaders' lives is a determinant of the potentiality of the respondent-school heads toward becoming great educational leaders themselves.

The role of the youth in the life of the Church cannot be undermined in the same breath as the impact of the Church in the holistic development of the youth and ultimately in the future of any individual. Clarke (2014) declared that the youth are the future of the Church. Being a youth leader herself, the author encouraged members of youth organizations to keep an active participation in their activities, as this is an essential ingredient in character formation, which would ultimately contribute to their success in the future. Similarly, regular attendance to Church service such as the Eucharist is an essential element of a Christian. Pope Francis (cited in Glatz, 2017) used simile to underscore the importance of this indicator by saying, "Just like a plant needs sun and nourishment to survive, every Christian needs the light of Sunday and the sustenance of the Eucharist to truly live" (p.1).

Scoring the second best among the confirmed seven (7) factors in terms of composite reliability, the Faith in God (FG) factor is also the second when it comes to Average Variance Extracted. The composite reliability estimates the extent to which a set of latent construct indicators/items share in their measurement of a construct/factor, while the AVE is the amount of common variance among latent construct indicators/items (Hair, Black, Bain, Anderson, & Tatham, 1998).

In terms of the said statistical measures, this result indicates that the items for the FG factor, which include "having a personal relationship with God", "lifting up and attributing

everything to God”, and “God being at the center of one’s life” are internally consistent in measuring the said factor. This observation also denotes that the responses of the respondents on this factor and items make this cluster a good measure for the scale.

Robertson (2012) emphasized the significance of spirituality, religiosity, and faith in her study, which intended to link these three dimensions to educational leadership. The same study cited the work of Clugston which observed that “most people identify themselves as religious and/or spiritual, and for many, their faiths call them to live in ways that respect and care for all life, present and future, and to focus on being more, not having more, after basic needs are met” (p. 4).

Continuous Learning (CL) is another factor, including its corresponding items, which was confirmed as a good factor to achieve the objectives of the scale. The items “seeking for new knowledge”, “viewing learning as a continuous process”, and “learning as exciting” all met the required points for item loadings. This result connotes that this factor and its items contribute to the overall fitness of the scale and measurement model.

Bennis and Nanus (cited in Robbins, 2013) described successful leaders as taking responsibility for their own development and are perpetual learners. The same book cited the paper of Schlechty, which stressed that in order for school principals to help their teachers improve what they do, the principals must continuously learn how to improve what they are doing themselves. In the same manner, Senge (cited in Robbins, 2013) pointed to the desire to learn as an important characteristic of a great school leader.

Registering the highest score among the seven confirmed factors with regard to composite reliability and Average Variance Extracted, the Influence of Teachers (IT) factor, together with its set of items, is the top determinant of the validation of the scale. This finding presupposes that this factor is a good measure in assessing the potentiality of the participant-school heads toward becoming great educational leaders.

The items “having inspiring former teachers”, “teachers who were motivational” and “teachers who influenced one’s life” all point to the role of teachers in shaping the lives of their students and the legacy that they leave that transcends the four walls of the classroom. These phrases are best captured in the famous quote by Henry Adams, “A teacher affects eternity; he can never tell where his influence stops” (Albom, 1997, p.23).

Bennett (2012) reported the study conducted by Raj Chetty and John N. Friedman of Harvard University and Jonah E. Rockoff of Columbia University, which tracked 2.5 million students over 20 years from a large urban school district from fourth grade to adulthood that made the research as one of the largest educational studies to date. The investigation focused on the long-term impact of teachers based on their “value-added” ratings. Findings affirmed that teachers are second to parents as the most important influence in a child’s education.

Great teachers make a great difference; poor teachers hurt a child's life chances surfaced as the conclusion of the study.

Passion for Work (PW) is another factor, which was established through CFA as a valid factor for the scale. The items in this cluster include "exerting one's best effort in any work or project," "having a passion for education," and "having a passion for excellence." The latter item scored the highest in terms of item loading which means that this item has met the highest score in this particular criterion.

Sutcliffe (2013) conducted a national survey where the best school heads of the United Kingdom were asked as to what makes a great school leader. Having a passion for education emerged as one of the essential qualities which came out in the interviews. Specifically, this refers to being passionate about teaching and learning and exhibiting commitment to children.

Passion is the force and energy that move a school leader to succeed (McConnell, 2016). The author described passion as a strong emotion that one feels when he/she is doing a meaningful work that makes other people also feel the same. Further, passion is the key that will create a big impact in moving the school toward achievement and success. The study recommended concrete steps that a principal can do to be passionate about his/her work such as surrounding oneself with equally passionate people, connecting with other school leaders, and believing in one's ability to make a difference.

The remaining factor that was attested by CFA to be a valid construct for the scale is Love for Reading (LR). The items "reading books help in managing the school," "regularly reading books on management and other fields," and "the love for reading books and content materials" are all good indicators of this factor. Consequently, the said factor and items are significant in the validation of the said scale.

US President Harry S. Truman (cited in Faiola, 2012 p.35) postulated, "Not all readers are leaders, but all leaders are readers" in underscoring the value of reading to the success of leaders. Reading is an essential ingredient to leadership. With books, a leader, including school leaders, gets to discover ideas, which he/she can adopt in the workplace. Many successful leaders attribute their success to the people that they have read in books.

In particular, many school leaders profess that their leadership styles and management approaches are often influenced by the authors and books that they read. Rosenthal (2017) asserts that a good school leader thinks about innovative ways of doing things. According to the author, innovation can be achieved by building a "library of career," that is, through reading leadership and management books.

Overall, the seven (7) confirmed factors and 20 items were proven to be well-defined and internally consistent measures of the scale. With the use of CFA, the study generated

three measurement models where Model 3 emerged as the best and preferred model in terms Goodness-of-Fit, fit indices, factor loadings, validity and reliability. Evidently, this model is valid and reliable.

For the recommendations of the study, to test further the validity of the scale, the researcher recommends undertaking CFA in another sample to support the generalizability of the instrument.

Another research using a different method such as Exploratory Factor Analysis or qualitative analysis may be pursued to refine and revalidate the grounded theory that served as the basis of the study.

It is also recommended that a full Structural Equation Modelling (SEM) be performed to determine the interrelationships between and amongst the variables of the theory.

In conclusion, educational agencies and institutions, particularly the Department of Education, may utilize the validated scale in assessing and strengthening their programs on human resource and development. Potentiality could be linked to two specific areas of the HRD System: (1) RSPI (Recruitment, Selection, Placement and Induction) and (2) L and D (Learning and Development).

Likewise, researchers may learn on the use of Confirmatory Factor Analysis as a research method in instrument development, taking into account its strengths and weaknesses, based on the actual steps taken by the researcher.

In the end, the school leaders, as well as teachers who have aspirations of becoming school heads, are the main beneficiaries of the results of the present study. The validated scale may serve as a yardstick in their introspection on whether or not they are making a difference in the lives of the learners and in the society, which is the hallmark of potentiality, regardless that their being a leader is something that is innate or inborn or something that is made or acquired through time and experience.

REFERENCES

- Albom, M. (1997). *Tuesdays with Morrie: An old man, a young man, and life's greatest lesson*. New York: Doubleday.
- Amanchukwu, R., Stanley, G. & Ololube, N. (2015). *A review of leadership theories, principles and styles and their relevance to educational management*. Scientific and Academic Publishing Journal 5 (1): 6-14.
- American Association of Colleges of Pharmacy (2005). *Point/counterpoint: Are outstanding leaders born or made?* American Journal of Pharmaceutical Education Volume 1-13, 1937-49.
- Bagozzi, R. & Fornell, C. (1982). *Theoretical Concepts, Measurement, and Meaning. A Second Generation of Multivariate Analysis*: Praeger, pp. 5-23.
- Baron, R. & Kenny, D. (1986). *The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations*. Journal of Personality and Social Psychology, 51, 1173-1182.
- Bennett, W. (2012). *The lasting impact of good teachers*. Retrieved February 13, 2018 from <https://edition.cnn.com/2012/01/11/opinion/bennett-good-teachers/index.html>.
- Billah, M. (2017). *Importance of extra and co-curricular activities*. Retrieved February 13, 2018 from <http://www.theindependentbd.com/post/124836>.
- Brown, T. (2015). *Confirmatory factor analysis for applied research: Second Edition*. Guilford Publications.
- Bursztyn, L. & Jensen, R. (n.d.). *Should schools recognize or award?* Retrieved February 13, 2018 from <https://pdfs.semanticscholar.org/e06f>
- Camburn, E., Rowan, B. & Taylor, J. (2003). *Distributed leadership in schools: The case of elementary schools adopting comprehensive school reform models*. Educational Evaluation and Policy Analysis Vol 25, Issue 4, pp. 347 – 373.
- Chomeya, R. (2010). *Quality of psychology test Between Likert Scale 5 and 6*. Journal of Social Sciences 6 (3): 399-403, 2010 ISSN 1549-3652.
- Clarke, J. (2014). *Youth are the hope of the church*. Retrieved February 13, 2013 from www.mdpi.com/2077-1444/2/4/628.
- Covey, S. (2004). *The 8th Habit*. NY: Free Press/ Simon & Schuster, Inc.
- Crawford, M. (2002). *The Charismatic school leader--Potent myth or persuasive effect?* School Leadership & Management v. 22. 273-287.
- Daniyal, M., Nawaz, T., Hassan, A. & Mubeen, I. (2012). *The effect of co-curricular activities on the academic performance of students: A case study of the Islamia*

- University of Bahawalpur, Pakistan. Bulgarian Journal of Science Education and Policy (BJSEP) Vol.6 November 2, 2012.
- DeBois, S. (2016). *9 Advantages and disadvantages of questionnaires*. Retrieved February 13, 2018 from <https://surveyanyplace.com/questionnaire-pros-and-cons/>.
- De Guzman, A. & Guillermo, M. (2007). *The serendipity of principalship: Meaning-making of a Filipino secondary school principal*. Asia Pacific Education Review, 2007, Volume 8, Number 2, p. 216.
- Department of Education. (2016). *Results of 2015 principal's test*. Department Memorandum No.11, s. 2016
- Department of Education. (2017). *Results of 2016 principal's test*. Department Memorandum No.45, s. 2017.
- De Pree, M. (2008). *Leadership jazz - Revised edition: The essential elements of a great leader*. Crown Business, USA.
- Di Giulio, J. & Giulio, D. (2014). *Are leaders born or made?* Retrieved February 13, 2018 from <https://www.researchgate.net/publication/270684591>.
- Doyle, M. (2005). *The long walk out of poverty*. Retrieved February 13, 2018 from <http://news.bbc.co.uk/1/hi/world/asia-pacific/426948.stm>.
- Failoa, A. (2012). *All leaders are readers: Six must-read books to inspire success*. Retrieved February 13, 2018 from <https://forbesbooks.com/all-leaders-are-readers-six-must-read-books-to-inspire-success>.
- Furr, R. & Bacharach, V. (2013). *Psychometrics. An Introduction Second Edition*. Wake Forest University and Appalachian State University.
- Glatz, C. (2017). *Pope Francis: We need to go to Sunday mass*. America The Jesuit Review. Retrieved February 13, 2018 from <https://www.americamagazine.org/faith/2017/12/13/>.
- Hair, J., Black, W., Babin, B., Anderson, R. & Tatham, R. (1998). *Multivariate data analysis*. Berlin-New York: Mouton de Gruyter.
- Hair, J., Ringle, C., & Sarstedt, M. (2011). *PLS: Indeed a silver bullet*. Journal of Marketing Theory and Practice, vol. 19, no. 2 (spring 2011), pp. 139–151.
- Hall, M. (2018). *What is the difference between variance and covariance?* Retrieved February 13, 2018 from <https://www.investopedia.com>
- Hassan, Z., Schattner, P. & Mazza, D. (2006). *Doing a pilot study: Why is it essential?* Malays Family Physician Journal. 2006; 1(2-3): 70–73.

- Hooper, D., Coughlan, J., Mullen, M. (2008). *Structural equation modelling: Guidelines for determining model fit*. Electronic Journal of Business Research Methods Volume 6 Issue 1 2008 (53-60).
- Jackson, D., Gillaspay, J., & Stephenson, R. (2009). *Reporting practices in confirmatory factor analysis: An overview and some recommendations*. American Psychological Association: Psychological Methods Vol. 14, No.1, 6-23.
- Kock, N. (2013). *WarpPLS 4.0 user manual*. Laredo, TX: Script Warp Systems.
- Leithwood, K. & Jantzi, D. (2005). *The effects of transformational leadership on organizational conditions and student engagement with school*. Journal of Educational Administration. 38. 112-129.
- Magno, C. (2013). *Assessing students' critical thinking and approaches to learning*. The International Journal of Educational and Psychological Assessment January 2013, Vol. 12(2).
- Malo, P. (2016). *Structural equations modeling – Part 1: Confirmatory factor analysis*. Aalto University School of Business.
- McConnell, S. (2016). *The secret to a successful principal*. Retrieved February 13, 2018 from <http://www.principalprinciples.net/2016/06/the-secret-to-successful-principal.html>.
- Mora, M., Gelman, O., Steenkamp, A., & Raisinghani, M. (2012). *Research methodologies, innovations and philosophies in software systems engineering and information systems*. USA: Information Science Reference.
- National Educators Academy of the Philippines (2015). *Principal's development framework and guidelines*. Department of Education, Pasig City.
- Nicdao, R. (2013). *Transformational leadership practices of SUC presidents: A triangulation study*. (Unpublished Dissertation). University of the Assumption, Pampanga.
- Pierce, D. (2013). *Good to great: How to have great schools*. Retrieved February 13, 2018 from <https://www.schoolnews.com/2013/02/22/good-to-great-author-how-to-have-great-schools>.
- Polit, D. F. & Beck, C. T. (2006). *The content validity index: Are you sure you know what's being reported? Critique and recommendations*. Res. Nurs. Health, 29: 489–497. doi:10.1002/nur.20147.
- Premuzic, T. (2016). *What science tells us about leadership?* Retrieved February 13, 2018 from <https://hbr.org/2016/09/what-science-tells-us-about-leadership-potential>.
- Prudon, P. (2013). *Confirmatory factor analysis: A brief introduction and critique*. Retrieved February 13, 2018 from <https://www.researchgate.net/publication/273003228>.

- Quiambao, R. (2013). *A metamorphosis of a coastal school community amidst adversities: An autoethnography*. (Unpublished Thesis). University of the Assumption, Pampanga.
- Robbins, R. (2013). *Leader as learner*. Retrieved February 13, 2018 from <https://uk.sagepub.com/sites/default/files/upm-binaries/61031>.
- Robertson, D. (2012). *Faith values and leadership*. Retrieved February 13, 2018 [www.acu.edu.au/data/assets/pdf_file/0006/470454/Deborah Robertson](http://www.acu.edu.au/data/assets/pdf_file/0006/470454/Deborah_Robertson).
- Rosenthal, L. (2017). *9 Best books on management for school leaders*. Retrieved February 13, 2018 from <https://schoolleadersnow.weareteachers.com/best-school-management-books>.
- San Antonio, D. (2008). *Creating better schools through democratic school leadership*. International Journal of Leadership in Education Vol. 11, Iss. 1, 2008.
- Schreiber, Nora, Stage, Barlow & King (2006) *Reporting structural equation modeling and confirmatory factor analysis results: A review*. The Journal of Education Research: July –August 2006 vol.99 no.6.
- Shakespeare, W. (2003). *No Fear Twelfth Night*. (Crowther, J., Ed.). New York: Washington Square-Pocket.
- Sindhvad, S. (2009). *School principals as instructional leaders: An investigation of school leadership capacity in the Philippines*. University of Minnesota Digital Conservancy. Retrieved February 13, 2018 from <http://hdl.handle.net/11299/56986>.
- Sorenson, G. (2000). *An intellectual history of leadership studies: The role of James MacGregor Burns*. Annual Meeting of the American Political Science Association, Washington DC.
- Spector, B. (2015). *Carlyle, Freud, and the great man theory more fully considered*. SAGE Journals. Volume: 12 issue: 2, page(s): 250-260.
- Stapleton, C. (1997). *Basic concepts and procedures of confirmatory factor analysis*. Retrieved February 13, 2018 from <http://ericae.net/ft/tamu/Cfa>.
- Statistics Solutions. (2013). *Confirmatory factor analysis*. Retrieved February 13, 2018 from <http://www.statisticssolutions.com/academic-solutions/resources/directory-of-statistical-analyses/confirmatory-factor-analysis>.
- Suhr, D. (2017). *Exploratory or confirmatory factor analysis?* Retrieved February 13, 2018 from [http://tx.liberal.ntu.edu/DataAnalysis/FactorAnalysis](http://tx.liberal.ntu.edu>DataAnalysis/FactorAnalysis).

- Sutcliffe, J. (2013). *The eight qualities of successful school leaders*. Retrieved February 13, 2018 from <https://www.theguardian.com/teacher-network/2013/sep/24/eight-qualities-successful-school-leaders>.
- Sutherland I., & Brooks, J. (2013). *School leadership in the Philippines: Historical, cultural, and policy dynamics*. *Collective Efficacy: Interdisciplinary Perspectives on International Leadership*. 2013, 199-213.
- Tirozzi, G. & Kappan, P. (2001). *The artistry of leadership: The evolving role of the secondary school principal*. *SAGE Journals Vol 82, Issue 6*, pp. 434 – 439.
- Usman, A. (2010). *A grounded theory on the making of great educational leaders*. (Unpublished Thesis). University of the Assumption, Pampanga.
- Weindling, D. (2000). *Stages of headship: A longitudinal study of the principalship*. Retrieved February 13, 2018 from <https://www.researchgate.net/publication/234702931>.
- Worthington, R. & Whittaker, T. (2006). *Scale development research. A content analysis and recommendations for best practices*. *Counseling Psychologist*, 34, 806-838.