Students' Mathematics Self-efficacy Scale: Item Development and Validation

Rhosechelle A. Riboroso[1]

[1]College of Arts and Sciences, University of Northern Philippines, Vigan City

rhosechelle.riboroso@unp.edu.ph

Abstract *– This research aimed to resolve the shortage of available resources that assess Filipino college students' expectations about their ability to perform mathematics skills successfully. This ability is measured using the Mathematics Self-efficacy Scale (MaSS). The scale started with 72-item statements derived from a literature review on mathematics competencies, adapted from existing mathematics self-efficacy questionnaires and surveys from mathematics teachers. The items underwent items analysis, exploratory analysis, confirmatory analysis. Likewise, it went through an internal consistency test to check the validity of the items. The results show that there are four factors which measure the mathematics self-efficacy of college students named as mathematical modeling, mathematic mathematical representation, mathematics communication and use of technology in mathematics. The four-factor MaSS exhibits validity and internal consistencies.*

Keywords*: mathematics self-efficacy, exploratory factor analysis, confirmatory factor analysis, scale development*

1. Introduction

Mathematics, often viewed as a challenging subject, is taught traditionally with just formulas, rules, and procedures. In recent years, mathematics education has become more holistic, relevant, and connected with the learners' everyday lives, their culture included.

Mathematics education aims to establish mathematically empowered citizenship as asserted by the Department of Science and Technology - Science Education Institute, and the Philippine Council of Mathematics Teacher Education, Inc. (SEI-Dost and MATTHED, 2011). Patena and Dinglasan (2013) conducted a study on the students' performance on mathematics departmental examination. They found that there is a poor performance with regards to College Algebra (Math 1) since "not met" obtained the highest frequency among all the colleges enumerated. Similarly, Trigonometry's poor performance on the four colleges as they could not meet the required criteria of above 83 percent. Furthermore, the result shows that there is still no progress in the students' performance during the second semester since the previous results were consistent in their performance.

Consequently, identifying the factors that yield adverse effects on students' mathematics performance and achievements have been initiated. Over the years, researchers have been interested in the relationship between personality and academic performance. Specifically, researchers investigated the predictive power of students' attributes on their predictive performance in mathematics, among the numerous mathematics self-efficacy variables. Zimmerman defined it as the "judgment of one's capabilities to successfully perform a particular task given," as cited in (Zimmermann, Bescherer & Spannagel, 2010). It is how confidently individuals believe they can accomplish specific tasks using their skills under certain circumstances.

Studies show the predictive and mediational role of self-efficacy beliefs in successfully performing mathematical tasks (Dullas, 2010; Galla & Wood, 2012; Guolao, 2014; Zimmerman et al., 2010). A high positive self-efficacy has a positive impact on mathematics achievement among students at different educational levels, whether upper elementary (Fast et al., 2010), senior secondary, or in engineering courses (Soleymani & Rekabdar, 2016). Also, poor self-efficacy in non-mathematics majors often leads to low achievement. However, self-efficacy being task-related, it tries to increase effort and perseverance towards the difficult tasks. Consequently, there is an increased chance of accomplishing the tasks. Because of the effect of mathematics self-efficacy on students' performance in mathematics, it is then essential for teachers' essential stand students' mathematics self-efficacy to be in a better position to help them boost their self-efficacy and use it in improving their mathematics performance. Baloglu and Zelhart (2007) suggested that the first step in developing appropriate and effective intervention strategies is to have a psychometrically sound and efficient measurement.

There are many existing measures of self-efficacy. Among these include the Mathematics Self-efficacy Scale (MSES; Betz & Hackett, 1983), Mathematics Self-efficacy Expectations of Future Teachers (MaSE-T; Zimmerman, Bescherer, & Spannagel, 2010), Mathematics Self-efficacy Questionnaire (MSEQ; May & Glynn, 2008), and Sources of Middle School Mathematics Self-efficacy Scale (Usher & Pajares, 2009). However, these developed scales on mathematics self-efficacy were from abroad. Murphy, Wood, and Carter (2007) argued that one set of findings in a local context might not apply to another group of students with different demographics and educational contexts. Thus, they proposed that to make accurate analyses and generalizations on individual students' self-efficacy, researchers need to be aware of the circumstances in carrying out the research, assessing the students' nature, and the components of the scales used. Likewise, Furr (2011) mentioned that one could not assume confidently that the reliability of a scale's scores in one study worth one sample of participants generalizes to all studies or participants. He added that assuming psychometric properties generalized across participants' samples is not always valid, particularly for scales used across cultural groups. The scale items do not necessarily represent the same latent construct across groups; hence, the accuracy of interpretations about group differences on the latent construct is compromised (Tucker et al., cited in Furr, 2011). Bond and Fox (2001) interpretation of data can only be as good as the quality of instruments used (cited in Li, Toland, & Usher, 2013).

The researcher reviewed local studies such as of Dela Rosa (2010), Ouano (2011), Yazon (2015), and Nipaz et al (2016). There were three of the four reviewed studies adapted and reworded existing foreign instruments to measure self-efficacy. While two tested for internal consistency using Cronbach alpha from the sample, the other must-have assumed the instruments' original reliability. Dela Rosa (2010) made a self-efficacy questionnaire, which measures students' confidence in working with fractions and had it tested for internal consistency but not its construct validity. They used questionnaires developed abroad. Though these studies tested the instrument's reliability, it is insufficient since an instrument "may be reliable but not valid" (DeVon et al., cited in Parsian and Dunning, 2009, p. 4).

In this light, there is a need for a reliable and valid instrument to measure mathematics self-efficacy among Filipino college students. Identifying students with low self-efficacy will be a step forward in understanding and addressing the factors that contribute to students' low mathematics performance. Hence, the researcher aimed to develop an instrument to measure the students' mathematics self-efficacy and establish this localized scale's psychometric properties.

1.1. Objectives of the Study

This study aimed to develop an instrument to measure college students' mathematics self-efficacy using exploratory factor analysis and confirmatory factor analysis. Specifically, it sought to create items that measures the mathematics self-efficacy of students, assess the factor loadings of the identified constructs, determine the best fitting model of the scale (MaSS) and establish the internal consistency of the developed instrument.

1.2. Conceptual Framework

According to the Social Cognitive Theory of Bandura, self-efficacy is one of the four interrelated processes of goal realization. It tries to increase effort and perseverance to execute difficult tasks such as various mathematics competencies. Hence, developing a valid and

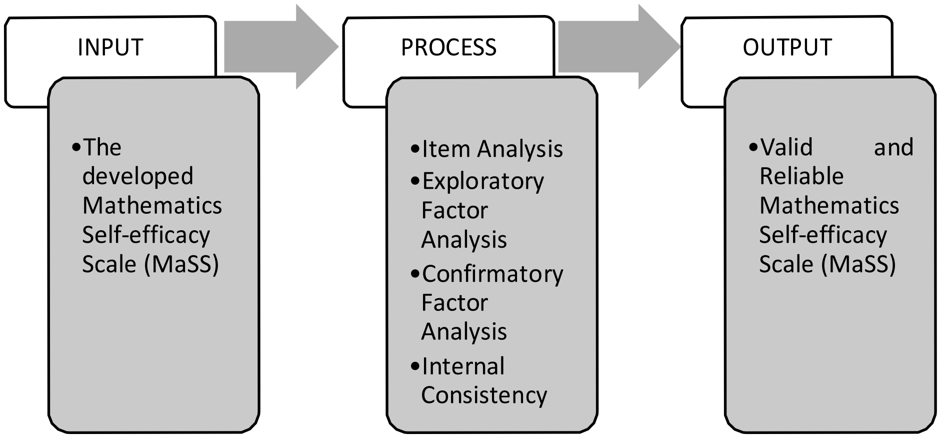

Figure 1. *Research Paradigm*

reliable instrument to measure college students' beliefs that they can do specific mathematics competencies successfully is deemed essential.

The paradigm in Figure 1 shows that an instrument measuring mathematics self-efficacy is developed and be called Mathematics Self-efficacy Scale (MaSS). Items of the constructed MaSS would undergo exploratory factor analysis, confirmatory factor analysis, and internal consistency test to establish its validity and reliability. Thus, the study's output would be a valid and reliable MaSS measuring the mathematics self-efficacy of college students.

2. Methodology

2.1. Research Design

This study used a descriptive-normative research design. This design combines two research methods: gathering information to describe the generated item statements of the MaSS as it is, has been, or is viewed (descriptive method) and evaluating the item statements to identify ways to improve it (normative method).

2.2. Subjects of the Study

The samples of the study were the 1030 first year college students. The students enrolled in academic programs such as elementary and secondary education, information technology, computer science, criminology, business administration, agriculture, hotel, restaurant, and tourism management, psychology, political science, English, social work, industrial technology, and public administration during the second semester of AY 2017-2018 from selected state universities and colleges in Ilocos Sur, La Union, Pangasinan, Abra, and Benguet. Furthermore, they are enrolled in at least one mathematics class (College Algebra, Plane Trigonometry, Contemporary Mathematics, Mathematics of Investment, Business Mathematics, or Statistics). They were the respondents since they have undergone a transition from middle school to university, which creates many experiences and challenges (van der Meer, 2012, cited in Ainscough et al., 2016). Students' first-year experience is crucial for overall academic success in higher education; that is, if they succeed in their first year of higher education, they are more likely to graduate from a university (Evan & Morrisson, 2011). Thus, self-efficacy could provide an essential foundation for university success.

In this study, there were 803 samples for EFA since, according to Izquierdo, Olea, and Abad (2014), the sample size for EFA should be greater than 300. Another 227 samples for CFA. Samples used for each analysis were selected using the Random Number Generation in Excel.

2.3. Data Gathering Tools

The study used the created pool of item statements about mathematics competencies wherein respondents indicate their self-belief level to do the task successfully from 0 (no confidence at all) to 10 (complete confidence). The items were content validated by six mathematics experts, and results of item validity indices range from .83 to 1.00 while the scale validity indices range from 0.92 to 0.96. Furthermore, all three language experts have agreed on the items' clarity and simplicity.

2.4. Data Gathering Procedure

The first step to developing an instrument that measures students' self-efficacy in mathematics is a need for a comprehensive review of literature, which is, understanding the self-efficacy theory, thorough examination of the guidelines for constructing self-efficacy scales, and exploration of mathematics competencies. Besides, it involves probing existing mathematics self-efficacy questionnaires and inquiry from experts. These led to a clear definition of the construct of interest, including its scope and subcomponents. Also, studying existing instruments that measure the same construct, related to the construct of interest or related instrument and experts' inputs, led to the development of pool items.

The experts initially validated the pool items and integrated their suggestions. Then the researcher asked permission from different authorities in the different state colleges and universities included in this study to run the MaSS to their first-year college students. The questionnaire was then collected from the participating universities and state colleges and tallied by the researcher.

Before performing factor analysis (EFA and CFA), items are screened if the skewness values have signs opposite to most items (Kim, 2009). Factor analysis was conducted to cluster items into common factors, interpret each factor according to the items having high loading on it, and summarize the items into a small number of factors (cited in Parsian & Dunning, 2009). Although the proposed Mathematics Self-efficacy Scales were based on a theoretical structure or model, the researcher applied exploratory factor analysis on the data to better understand the factor structure of the developed Mathematics Self-efficacy Scales. Moreover, it checks whether the theoretical model used as a basis for the items could be adopted for the Mathematics Self-efficacy Scales. Thus, factor analysis detects factors that best fit the data even if they depart from the original expectations.

Before undergoing exploratory factor analysis, the corrected item-total correlation (henceforth referred to as CITC) scores, which measure the extent of each item's association with all the scale items excluding itself scale items, filtered the items. A low CITC value (below 0.30) suggest an item to be discarded (Nunnally and Bernstein cited in Donaldson, 2015). Exploratory factor analysis checks the dimensionality of the scale developed (uni-dimensional or multi-dimensional). Besides, EFA reduced the number of items so that the remaining items can explain a more significant percentage of variance in the scale and maximize reliability (Netemeyer, Earden, & Sharma, 2003), and identify the underlying constructs with mathematics performance. The principal component analysis (PCA) was performed on the items since, as cited in Parsian & Dunning (2009), PCA was assumed to be entirely reliable and without error as it analyzes all the variance (both specific and common variance) of a variable. Likewise, the most commonly used orthogonal rotation, which is varimax, rotates the factors to maximize the loading on each factor and minimize the loading on other factors (cited in Cavas & Koc, 2013). In this study, EFA's goal is to reduce data and identify which items belong to which factors (components) and that the components are further analyzed; thus, the principal component analysis was used (May, 2009).

Confirmatory Factor Analysis aimed to test the significance of a specific factor loading. It also tests whether a set of factors are correlated or uncorrelated and assess the validity of a set of measures. Lastly, it compares the ability of two different models. AMOS yields several Fitness Indexes that reflect how fit the model is to the data. Yuet, Yusof, and Mohamad (2016) mentioned using at least one fitness model from each model fit category. The chi-square statistics, chi-square ratio to degrees of freedom, and Root Mean Square

Error Approximation (RMSEA) assessed the proper function estimated by the procedure.

Similarly, the degree of fit examined each item. Items will be good indicators of a particular dimension if the value of the t-test testing the null hypothesis that the actual value of the specified parameter is zero is significant at .05. The Akaike Information Criterion (AIC) and Expected Cross-Validation Index (ECVI) were compared to determine the best model to be adopted. AIC measures the relative quality of statistical models for a given set of data and ECVI for model comparison (University California Davis, 2016). The goodness of fit were indicated by the following: $2 \leq \chi2/df \leq 3$, $.05 < RMSEA \leq .08$, and that the lesser the value of AIC and ECVI, the better the model (Amora, 2017) and an $AVE \geq 0.5$ (Ab Hamid, Sami, and Mohmad Sidek, 2017).

Then, the scale's internal consistency was measured using Cronbach alpha and corrected item-total correlation.

2.5. Data Analysis

This study examined validity and reliability of the developed mathematics self-efficacy scales using item analysis, EFA, and CFA through the Statistical Package for the Social Sciences.

2.6. Ethical Considerations

This study underwent ethical review by the UNP Review Board Committee. An informed consent form which explains what the study is, what the study is for, who the participants are, what happens when they fill out the questionnaire, the benefits study, and the risks in the study was given to the participants before they answered the questionnaire. It also underscores to treat their identity and responses with the utmost confidentiality.

3.Results and Discussion

3.1. Development of MaSS

The suggestions on scale development and construction forwarded by DeVellis (2016), Germaine (2006), Furr (2011), and Bandura (2005) were used as basis in the development and validation of the Mathematics Self-efficacy Scale. Initially, Self-efficacy and Mathematics constructs are defined to ascertain that the scale items were valid. Self-efficacy in this study's context is the judgment of one's capability to execute given types of performances, the ability to attain a valued goal (Maddux & Kleinman, 2018). Mathematics was defined using Turner's Mathematics competency framework as a set of skills, knowledge, and attitudes fundamental to the development of "mathematical literacy" or applying mathematical knowledge to practical situations (Turner, 2010). Therefore, this study aimed to develop a Self-efficacy Scale regarding the six mathematical competencies based on the PISA problems in 2003: communication, mathematizing, representation, reasoning and argumentation, strategic thinking, and using symbols, formal and technical language, and operations.

The generated 72-item pool of the MaSS was adapted from existing instruments (13 items), extracted from literature (55 items), and elicited from Mathematics teachers (4 items). The 13 modified items came from instruments such as The Mathematics Self-efficacy for Secondary Schools by Marat (2005), Mathematics Self-efficacy and Anxiety Scale by May (2009), and the Self-efficacy for Learning Scale by Klobas, Renzi, and Nigrelli (2007). On the other hand, the items generated from literature are based on the definitions and discussions of the mathematics competencies by several authorities and from the descriptions of a mathematically proficient student by the Common Core State Standards for Mathematics.

The pool of items reflecting college students' mathematics self-efficacy was written in the first person statements with 11 response options (No Confidence at All=0 to Complete Confidence=10). The scale is a simple response format of 0–10, which is a stronger predictor of performance than a five-interval scale (Pajares, Hartley & Valiente cited in Bandura, 2005).

3.2. Assessment of Factor Loadings

Examining the mean, skewness, and kurtosis of the 72 items MaSS, the mean of items range from 4.77 to 6.35 with an average mean of 5.58. The range of the item means of the scale is close to the midpoint of the possible score range, five since it made use of an 11-point scale (0 to 10 scale). The results indicate that the students used a range of responses to arrive at this means, which implies an auspicious attribute of the items (DeVellis, cited in Donaldson, 2015). The variation of the students' responses for each of the items ranges from 1.85 to 2.21. The results show that the participants' responses to the mathematics self-efficacy items are diverse; that is, the students' variation of responses existed. Thus, the items of the scale are capable of discriminating students with different levels of mathematics self-efficacy. Furthermore, skewness values range from -0.31 to 0.20, and kurtosis values range from -0.73 to 0.09, which fall within the acceptable range of -2 to +2. The results support the multivariate normality of data distribution (Bobbio & Manganelli, 2009).

Furthermore, inspecting the sign of skewness values, 10 items were removed since these have skewness values with signs that are opposite to the majority of the items. The scale has only 62 items left. To determine whether the removal of these items had a significant effect on the mathematics self-efficacy scale, Cronbach's alpha was examined before and after removing the items. The computed Cronbach's alpha before and after the removal of the ten items are both 0.99. Thus, there is no difference in the reliability coefficient before and after removing the ten items indicating that the deleted items did not contribute significantly to the proportion of variance attributable to the valid Mathematics Self-efficacy score.

Evaluating the CITC values, which measure the extent of each item's association with all the scale items excluding itself, it surpassed the minimum value of 0.30, which indicates the 62 items could undergo EFA. Thus, the principal component analysis (henceforth referred to as PCA) using varimax rotation and replacing missing values with means so that none of the responses gets eliminated (Li, 2007) explored the 62 items of MaSS. The table shows the result of KMO and Bartlett's Test. On the first run of PCA on MaSS, the KMO coefficient is 0.981 which falls into the superb range according to Kaiser as cited in Parsian and Dunning (2009). It tells that the sample size is adequate to yield distinct and reliable factors (Muzaffar, 2016). The Bartlett's Test of Sphericity is highly significant (χ^2 (1891) = 32505.487, p=.000), which shows that the correlation matrix is significantly different from an identity matrix in which correlations between variables are all zero. The result implies that the items are correlated highly enough to provide a reasonable basis for factor analysis. However, Kaiser's criterion specifies only five factors with eigenvalues ≥1.0 in the first run of PCA. The total variance of the draft of MaSS factors is 69.87 percent, considered as a firm by Tabachnick and Fidell (2006). However, the results of the first run of EFA show that there are items that have a factor load of less than 0.50.

Meanwhile, three items have cross-loadings, indicating that students did not interpret them as they intend to. Revising the draft using EFA was done on a step-by-step basis, and that no more than one item is deleted in each step. This procedure resulted in the deletion of 16 items, which leaves 46 items, four-factor MaSS. The total variance explained with the four factors is presented in Table 2.

The exploratory factor analysis's final results reveal that MaSS has a four-factor solution accounted for 69.37 percent of the total variance. This amount of variance is adequate because, according to Beaver et al. (2013). Evident is the

Table 1. KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling	0.981
Bartlett's Approx. Chi-test of Square Sphericity	32505.487
df	1891
sig	.0000

Table 2. Total Variance Explained

Component	Initial Eigenvalues		
	Total	% of variance	Cumulative %
1	27.676	60.166	60.166
2	1.981	4.305	64.471
3	1.163	2.528	66.999
4	1.091	2.371	69.371

Table 3. Factor Loading and Communality of the Factors of MaSS

Item No.	Factor Loadings	Communality	Item No	Factor Loadings	Communality
Factor 1			66	0.72	0.72
31	0.58	0.66	67	0.68	0.69
32	0.60	0.70	68	0.72	0.72
33	0.61	0.71	69	0.73	0.67
35	0.61	0.67	70	0.71	0.68
36	0.57	0.63	Factor 2		
40	0.60	0.66	13	0.63	0.68
41	0.65	0.72	14	0.68	0.73
42	0.67	0.67	15	0.69	0.76
44	0.67	0.67	16	0.73	0.77
45	0.64	0.60	17	0.65	0.74
46	0.67	0.70	18	0.62	0.71
47	0.68	0.71	19	0.59	0.70
49	0.67	0.63	20	0.62	0.66
50	0.68	0.61	Factor 3		
51	0.70	0.69	1	0.72	0.66
52	0.68	0.69	2	0.76	0.73
53	0.71	0.71	3	0.72	0.74
54	0.68	0.71	4	0.71	0.71
55	0.70	0.69	5	0.69	0.66
57	0.72	0.72	6	0.62	0.66
58	0.72	0.74	Factor 4		
59	0.73	0.68	22	0.76	0.80
64	0.71	0.62	23	0.74	0.77
65	0.74	0.70	24	0.54	0.65

fact that the first factor is higher than the rest of the factors extracted. The first factor accounts for 60.166 percent of the total variance. The four factors extracted represent the four levels of specificity to measure students' mathematics self-efficacy (Ling, 2016).

Table 3 exhibits the factor loading and communality of the four-factor 46 items MaSS.

3.2.1. Factor 1 – Self-efficacy in Mathematical Modeling

According to DeVellis (2012), the highest loadings are the ones most similar to the latent variable. He added that these items provide a window into the nature of the factor in question. The 29 items of this factor consists of statements As can be seen from the table, although the factor contributes to a diverse set of items, all the items pertain to mathematical modeling. According to Bliss and Libertini (2016), modeling is a competency using mathematics to represent, analyze, and make predictions or otherwise provide insight into real-world phenomena. This study adapted the construct given by Bliss and Libertini (2016) and therefore renamed modeling. These items represent the basic modeling cycle as described by the Common Core State Standard. There were two on identifying variables in the situation and selecting those that represent essential features; four items on formulating a model by creating and selecting geometric, graphical, tabular, algebraic, or statistical representations that describe relationships between the variables; eight items on analyzing and performing operations on these relationships to conclude; three items on interpreting the results of the mathematics and in terms of the actual situation; five items on validating the conclusions by comparing them with the situation and then either improving the model or if it is acceptable; and two items on reporting on the conclusions and the reasoning behind them. In addition, there are two items that are about making and evaluating assumptions are present throughout the modeling cycle. Thus, the label Mathematical Modeling appropriately represents the underlying dimension of the factor, that is, the 29 items measure *self-efficacy in Mathematical Modeling*. It involves the students' confidence in their capability to turn real-world experiences into a mathematical problem (mathematizing), applying various symbols in mathematical relationships, formulas, expressions, and operations (Symbolic Language and Operations). The factor loadings of the 29 items of Self-efficacy on Mathematical Modeling range from 0.57 to 0.74, which are more significant than 0.50, the most strict requirement for factor loading (Hair et al., 2010). These results indicate High convergent validity (Hair et al., 2010) since the loadings ascertained the high degree to which the factor contributes to the items.

Furthermore, the communality values of the items range from 0.61 to 0.77. These values are higher than the recommended minimum value of communality, which is 0.40. It means that the percentages of variance explained in each item range from 61 to 77 percent. It signifies that the items explain the factor, self-efficacy in mathematical modeling well enough.

3.2.2. Factor 2 - Self-efficacy in Mathematical Representation

Factor 2 comprises eight items, which are all mathematical representation measures; that is, referring to devising or using depictions of mathematical objects or relationships, equations, formulas, graphs, tables, diagrams, and textual descriptions. Thus, this factor is named as *Self-efficacy in Mathematical Representation*. It has factor loadings ranging from 0.59 to 0.73, which are higher than the benchmark value of 0.40 and even the most strict benchmark value of 0.50. The results of the factor loadings recognized the amount to which factor 2 contributes to the eight items. It means that the extracted factor represents well by the items. The communality values, the estimates of the variance in each item accounted for by the factor, range from 0.66 to 0.77.

3.2.3. Factor 3 - Self-efficacy in Mathematics Communication

The hypothesized underlying construct measured by six items is named *Self-efficacy in Mathematics Communication.* It measures the students' ability to read and making sense of and interpreting mathematical statements and information. Students should be comfortable using the language of mathematics to express mathematical ideas and arguments precisely, concisely, and logically (Kaur & Toh, 2012). College students must read mathematical language expressed in tables, graphs, charts, and other data representations. Similarly, they must have the ability to understand mathematical problems by grasping, uncovering, and extracting relevant information in order to create conclusions or generalizations from these data. Factor 3 has factor loadings ranging from 0.62 to 0.76, which is greater than the strictest benchmark of 0.5. These are the amount of which items that add to the factor. It denotes that the items satisfactorily represent the extracted factor. The communality values, the estimates of the variance in each item accounted for by the factor, ranges from 0.66 to 0.74.

3.2.4. Factor 4 - Self-efficacy on the Use of Technology in Mathematics

There are three items loaded in this factor. These items deal with the utilization of technology in mathematics. Thus, *Self-efficacy on the Use of Technology in Mathematics* would be the name of this factor. The results signify that the use of technology is a proficiency that contributes to the mathematics self-efficacy of college students. This result supports the claim of Andrade-Arechiga et

Table 4. Reliability of the MaSS after Exploratory Factor Analysis

Factors	No. of Items	Cronbach Alpha	Range of CITC
Mathematical Modeling	29	0.98	0.75 – 0.84
Mathematical Representation	8	0.94	0.77 – 0.82
Mathematics Communication	6	0.91	0.71 – 0.79
Use of Technology in Mathematics	3	0.88	0.70 – 0.83
Overall	46	0.985	

al. (cited in the study of Nicolescu, 2015) that technology has a beneficial impact in motivating students, which in turn may increase their confidence about their mathematical abilities. The factor loadings of factor 4 ranges 0.54–0.76 are more significant than the accepted value.

Moreover, the items' communality values range from 0.65–0.80, which connote that the factor does a good job accounting for the variation of the items (May, 2009). Furthermore, results indicate that the extracted three items represent the factor satisfactory.

Table 4 poses the reliability of the four-factor, 46-item MaSS. The internal consistencies of the factors, namely: Self-efficacy on Mathematical Modeling, Self-efficacy on Mathematical Representation, Self-efficacy on Mathematics Communication, and Self-efficacy on the Use of Technology in Mathematics, are 0.98, 0.94, 0.91, and 0.88, respectively. These values exceeded the minimum recommended value of alpha, which is 0.70 for a new instrument (DeVon et al., 2007) in order to be valid. The high degree of internal consistency suggests that the Mathematics Self-efficacy Scale's sub-scales provide support to a reliable scale. Similarly, the corrected item-total correlations coefficients range from 0.70 to 0.84, greater than the accepted value of 0.50. It implies that the individual item is measuring the same thing as the rest of the sub-scale items. Furthermore, findings indicate that the items under each factor are near associated, which is highly indispensable in scale development.

3.3. Best fitting model of the Scale

The resulting four-factor model of MaSS from the EFA was assessed using a) statistical significance and the feasibility of parameter estimates along with the appropriateness of standard errors; b) Squared multiple correlations for each indicator observed and c) model fit indices as suggested by Bryne (cited in Kim, 2009). The initial four-factor MaSS model, together with its correlation values, standardized factor loadings, and SMC, is shown in Figure 2.

The correlation coefficients between the factors range from 0.63 to 0.87, which indicates a high relationship. By scanning the values of the standardized factor loadings, the values range from 0.72 – 0.87 (self-efficacy on mathematical modeling), 0.78 – 0.86 (self-efficacy on mathematical representation), 0.71 - 0.86 (self-efficacy on mathematics communication), and 0.80 – 0.94 (self-efficacy on the use of mathematics technology). The findings denote strong loadings since all values are higher than 0.6 (Garson, 2010).

Table 5 shows the goodness of fit of the initial four-factor a model of MaSS. The initial four-factor model of MaSS rejects the model fit since χ^2 is 2739.36 at 983 degrees of freedom. However, since χ^2 is affected by sample size, the χ^2/df of 2.79 is less than 3.0, which indicates a good fit according to the recommended rule of thumb for χ^2/df by Kline (cited in Mishra et al., 2016). On the other hand, the RMSEA value of .09 indicates poor fit. Thus, the initial estimation of the four-factor model of MaSS reveals that the model did not provide satisfactory results. The

Table 5. Comparison of the Goodness of Fit Evaluation

Fit Measure	Initial Four-factor Model	Modified Four-factor Model
χ^2(df)	2739.36 (983)	2577.49 (981)
χ^2/df	2.79	2.63
RMSEA	.09	.08
AIC	2935.36	2777.49
ECVI	12.99	12.29

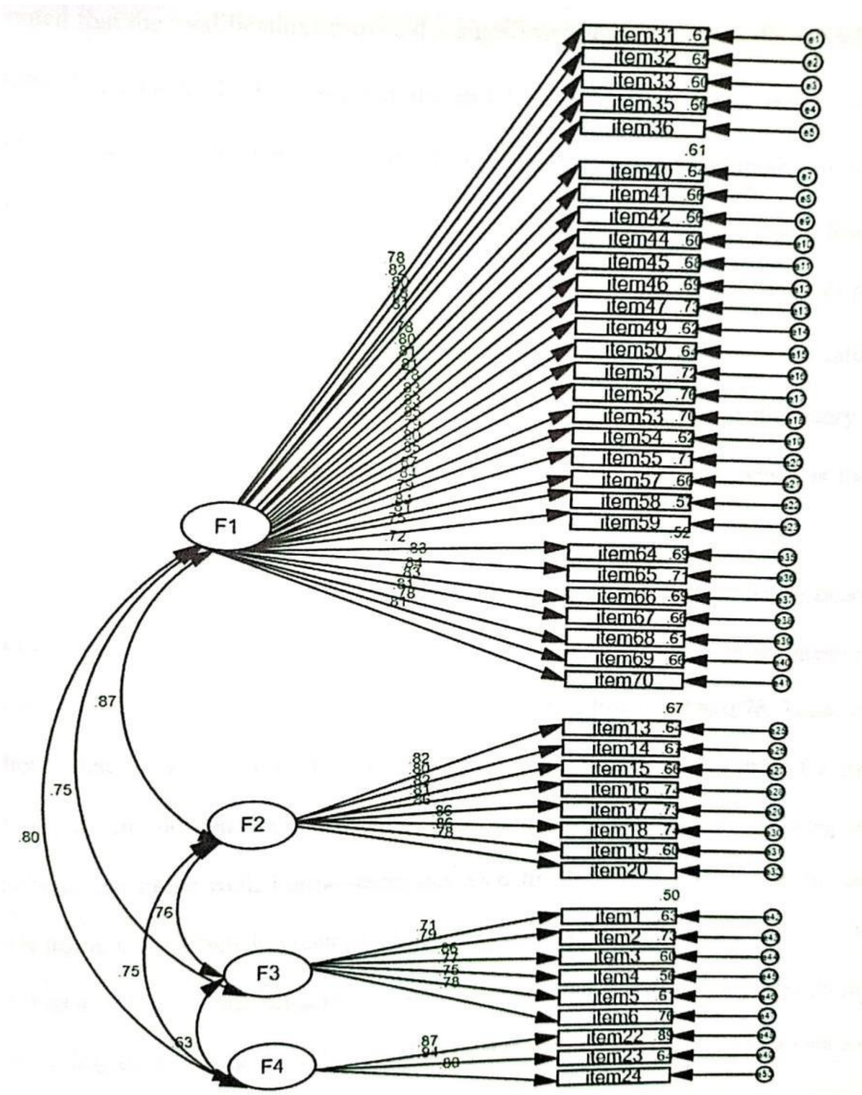

***Figure 2**. Initial Four-factor Model of MaSS*

modification indices were examined and suggested by freeing the corresponding error covariance parameter between the error terms for the measured variables, 67 and 68, and 31 and 32. These can be done since the items' load on the same factor (Alkan, 2016). Likewise, both items 31 and 32 are related pertains to making or evaluating critical assumptions in the estimation, modeling, and data analysis, while items 67 and 68 concerns modifying a model and demonstrating a function by a model. Thus, the modified CFA was repeated to examine the improved model fit, as shown in Figure 3.

Results of the CFA of the modified four-factor model show that the χ^2 /df, which is 2.63 indicates acceptable fit and the RMSEA = .08, which indicates a mediocre fit. Thus, the four-factor MaSS model is plausible. These results are similar to the findings of Sanjaya Mishra et al. (2016). It can be noted that the modifications provided a significant contribution to the model fit and support the fact that it has four-factor structures.

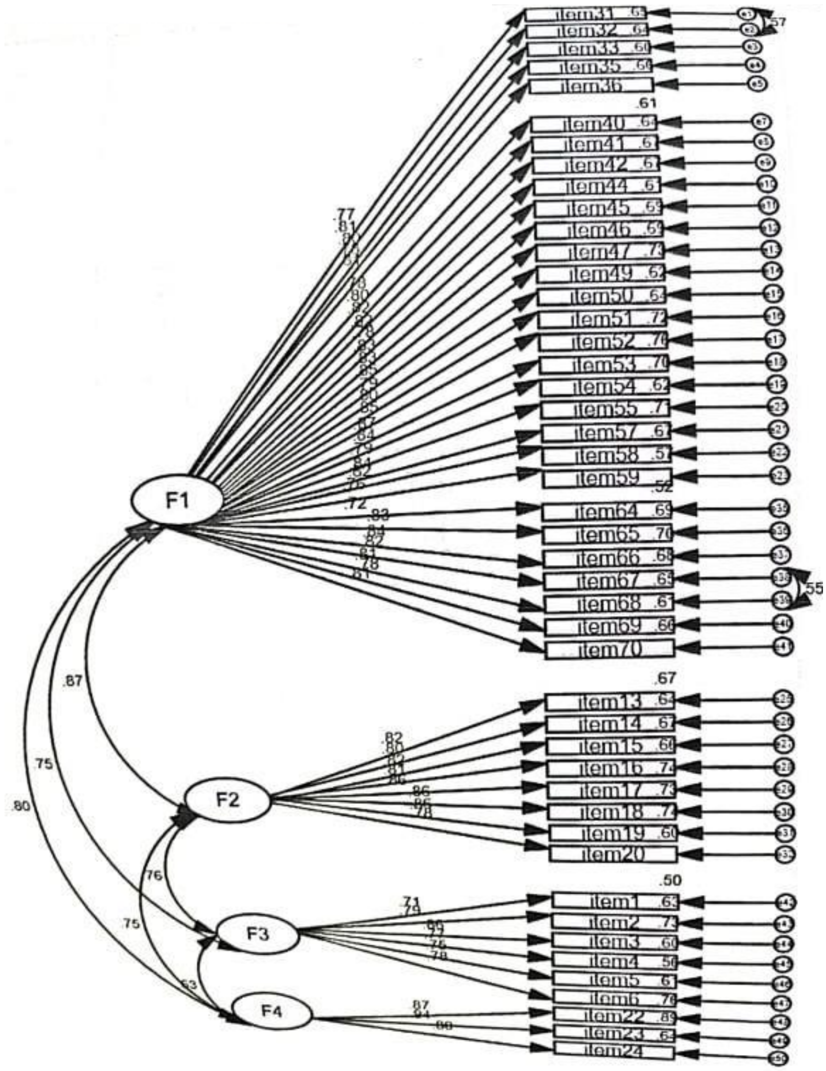

Figure 3. *The modified four-factor model of MaSS*

Thus, the construct validity of the model is achieved. Also, the modified model shows better fit than the original model as indicated by the AIC and ECVI values, that is, lower AIC (from 2935.36 to 2777.49) and ECVI (from 12.99 to 12.29) reflect the better fit (Vahedi & Farrokhi, 2011).

The standardized factor loading and squared multiple correlations and average variance extracted on each of the factors of MaSS are presented in Table 6.

The confirmatory factor analysis output shows that the standardized factor loadings range from 0.71 to 0.94, more significant than the threshold value of 0.50. Also, the squared multiple correlation (SMC) values range from 0.50 to 0.88. These results indicate that 50 to 88 percent of each item's variance is accounted for by the factors they belong to and that the extracted factors represent the items well. The Average variance extracted are 0.65, 0.68, 0.61, and 0.76 for self-efficacy on mathematical modeling, self-efficacy on

Table 6. Results of the Confirmatory Factor Analysis of the MaSS

Item No.	Standardized Factor Loading	Squared Multiple correlations	Item No.	Standardized Factor Loading	Squared Multiple correlations
Factor 1			68	0.81	0.66
31	0.77	0.59	69	0.78	0.61
32	0.81	0.66	70	0.81	0.66
33	0.80	0.64	AVE	0.65	
35	0.78	0.61	Factor 2		
36	0.81	0.66	13	0.82	0.67
40	0.78	0.61	14	0.80	0.64
41	0.80	0.64	15	0.82	0.67
42	0.82	0.67	16	0.81	0.66
44	0.82	0.67	17	0.86	0.74
45	0.78	0.61	18	0.86	0.74
46	0.83	0.69	19	0.86	0.74
47	0.83	0.69	20	0.78	0.61
49	0.85	0.72	AVE	0.68	
50	0.79	0.62	Factor 3		
51	0.80	0.64	1	0.71	0.50
52	0.85	0.72	2	0.79	0.62
53	0.87	0.76	3	0.86	0.74
54	0.84	0.71	4	0.77	0.59
55	0.79	0.62	5	0.75	0.56
57	0.84	0.71	6	0.78	0.61
58	0.82	0.67	AVE	0.61	
59	0.76	0.58	Factor 4		
64	0.72	0.52	22	0.87	0.76
65	0.83	0.69	23	0.94	0.88
66	0.84	0.71	24	0.80	0.64
67	0.82	0.67	AVE	0.76	

mathematical representation, self-efficacy on mathematics communication, and self-efficacy on the use of technology in mathematics, respectively. These values suggest high convergent validity of the four factors; that is, the items encompassing in each of the factors are interrelated.

3.4. Establishing the internal consistency

Table 7 displays the reliability results of the MaSS. The overall reliability of the Mathematics self-efficacy scale is 0.985, wherein it is higher than 0.70. The result indicates that MaSS is a good

Table 7. Reliability of the Mathematics Self-efficacy Scale (MaSS)

Factors	No of Items	Cronbach Alpha	Range of CITC
1 Mathematical Modeling	29	0.98	0.72 - 0.85
2 Mathematical Representation	8	0.94	0.74 - 0.83
3 Mathematics Communication	6	0.90	0.67 - 0.80
4 Use of Technology in Mathematics	3	0.90	0.73 - 0.88
Overall	46	0.985	

scale. The four subscales' reliabilities based on Cronbach's alpha are 0.98, 0.94, 0.90, and 0.90, signifying excellent reliability (Stephanie, 2017). The results imply that the items within each mathematics self-efficacy sub-scale measure the same construct since it tends to pull together the items. This means that if a parallel test is given to a student, the same rating will be given by the student on the same item. Similarly, the correlations between the individual item and the total subscale score range: 0.72–0.86 (self-efficacy on mathematical modeling), 0.74–0.83 (self-efficacy on mathematical representation), 0.67–0.80 (self-efficacy on mathematics communication), and 0.73–0.88 (self-efficacy on the use of technology in mathematics). The corrected item-total correlation shows that the scale is reliable since all items significantly correlate (CITC > 0.3) with the corresponding subscale's total score.

4. Conclusions and Recommendations

The validated instrument which measures mathematics self-efficacy of college students is valid and reliable as sustained by substantial shreds of evidence drawn from the findings.

An initial 72-item statements reflecting college students' mathematics self-efficacy was written in the first person statements with 11 response options. These items were adopted from existing instruments (13 items), extracted from literature (55 items), and elicited from Mathematics teachers (4 items)

There are four constructs that measures the mathematics self-efficacy of college students named as mathematical modeling, mathematical representation, mathematics communication, and use of technology in mathematics;

The best fit model is a four-factor MaSS, with the error terms of items 67 and 68, and of items 31 and 32 correlated; and the MaSS demonstrates internal consistencies.

The Mathematics Self-efficacy Scale (MaSS) for first-year college students remedies the need for a locally available instrument to measure their beliefs in their ability to perform mathematics competencies. This scale is suggested for use by mathematics teachers and guidance counselors for diagnostic purposes and planning and implementing instructional intervention strategies. Other researchers could further validate the results using larger number of college students specially those who are end results of the Senior High School program since the group of students utilized in this study were not yet graduates of the K to 12 curriculum. Lastly, a set of norms for the MaSS should be established to interpret the students' mathematics self-efficacy.

REFERENCES

Ab Hamid, MR., Sami, W. & Mohmad Sidek, MH. (2017). Discriminant validity assessment: use of fornell & larcker criterion versus htmt criterion. Journal of Physics: Conf. Series 980. doi:10.1088/1742-6596/890/1/012163.

Ainscough, L., Foulis, E., Colthrope, K., Zimbardi, K., Robertson-Dean, M., Chunduri, P. & Lluka, L. (2016). Changes in biology self-efficacy during a first year university course. CBE Life Sci Educ, 15(2). doi:10.1187/cbe.15-04-0092.

Alkan, F. (2016). Development of chemistry laboratory self-efficacy beliefs scale. Journal of Baltic Science Education, 15(3), 350 - 359.

Amora, J. (2017). Training Manual on structural equation modeling using Amos. Eins Review & Statistical Consultancy Center.

Baloglu, M. & Zelhart, PF. (, 2007). Baloglu and Zelharta (2007) suggested that the first step in developing appropriate and effective intervention strategies is to have a psychometrically sound and efficient measurement. The Psychological Record, 57, 593–611.

Beavers, AS., Lounsbury, JW., Richards, JK., Huck, SW., Skolits, GJ., & Esquivel, SL. (2013). Practical considerations for using exploratory factor analysis in educational research. Practical Assessment, Research & Evaluation, 18(6).

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. 84(2), pp. 191 - 215. doi:10.1037/0033-295X.84.2.191.

Bandura, A. (1991). Social cognitive theory of self-regulation. Organizational Behavior and Human Decision Processes, 50(2), 248 - 287. Retrieved February 11, 2017, from http://search.proquest.com.ezaccess.libraries.psu.edu/docview/61800683? accountid=13158.

Bandura, A. (1997). Self-efficacy: The exercise of control. Retrieved February 11, 2017, From http://search.proquest.com.ezaccess.libraries.psu.edu/docview/619147930? accountid=13158.

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. P. Urdan, Self-efficacy beliefs of adolescents (pp. 307 - 338). Greenwich, CT: Information Age.

Bliss, K., & Libertini, J. (2016). What is mathematical modeling? In S. G. (Eds.), Guidelines for Assessment and Instruction in Mathematical Modeling Education (GAIMME) (pp. 7 -22). Bedford, MA: Consortium for Mathematics and its Applications (COMAP) and Society for Industrial and Applied Mathematics (SIAM). Retrieved November 23, 2017, from http://www.siam.org/reports/gaimme-full_color_for_online_viewing_pdf.

BrckaLorenz, A., Chiang, Y., & Nelson Laird, T. (2013). Internal consistency. Retrieved March 22, 2017, from FSSE Psychometric Portfolio: ffse.indiana.edu.

Cavas, B. & Koc, SA. (2013). The development of laboratory safety questionnaire for middle school science teachers. Retrieved August 15, 2018, from www.icaseonline.net/journal/index.php/sei/article/view/34/49.

Cipora, K., Szczygiel, M., Willmes, K., & Nuerk, H. (2015). Math anxiety assessment with the abbreviated math anxiety scale: Applicability and usefulness: Insights from the polish adaptation. Frontiers in Psychology, 6, 18-33. doi:10.3389/fpsyg.2015.01833.

Clutts, D. (2010). Mathematics self-efficacy of community college students in developmental mathematics courses. (Doctoral Dissertation). Retrieved February 1, 2017, from http://digitalcommons.liberty.edu/cgi/viewcontent.cgi?article=1416 &context=doctoral.

DeVellis, R. (2012). Scale Development theory and applications. Sage Publications, New York.

DeVon, H. A., Block, M. E., Moyle-Wright, P., Ernst, D. M., Hayden, S. J., Lazzara, D. J. et al. (2007). A psychometric toolbox for testing validity and reliability. Journal of Nursing Scholarship, 39(2), 155-164.

Donaldson, D. (2015). Development of a Scale for Measuring Perceptions of Trustworthiness for Digitized (Dissertation). Retrieved January 7, 2018, from https://deepblue.lib.umich.edu/bitstream/handle/2027.42/111489/devand_1.pdf?.

Dullas, A. (2010). Academic performance and self-efficacy of filipino science high school students on mathematics and english subjects. Retrieved December 27, 2016, from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2152791.

Fast,LA., Lewis, JL., Bryant, MJ., Bocian, KA.;,Cardullo, RA., Rettig, M., & Hammond, KA. (2010). Does math self-efficacy mediate the effect of the perceived classroom environment on standardized math test performance? Journal of Educational Psychology, 102(3), 729 - 740. Retrieved from http://dx.doi.org/10.4135/9781446287866.n8.

Furr, M. (2011). Scale construction and psychometrics for social and personality psychology. http://dx.doi:10.4135/9781446287866.n8.

Galla, B. & Wood, J. (2012). Emotional self-efficacy moderates anxiety related impairments in math performance in elementary school-age youth. Personality and Individual Differences, 52(2), 118 - 122. doi:10.1106/j.paid.2011.09.012.

Garson, G. (2010). Factor analysis. Retrieved August 6, 2017, from http://faculty.chass.ncsu.edu/garson/PA765/factor.htm.

Germaine, M. (2006). stages of psychometric measure development: The example of the generalized expertise measure (GEM). Retrieved August 6, 2017, from https://archive.org>details>ERIC_ED492775.

Guolao, M. (2014). The relationship between self-efficacy and academic achievement in adults' learners. Athens Journal of Education, 1(3), 237 - 246. Retrieved February 1, 2017, from www.atiner.gr/journals/education/2014-1-3-4-Guolao.pdf.

Hall, M. & Ponton, M. (2005). Mathematics self-efficacy of college students. Journal of Developmental Education, 28(3), 26 - 32.

Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., Tatham, R.L. (2010), Multivariate Data Analysis. 7th ed. New York: Pearson.

Izquierdo, I., Olea, J., & Abad, FJ. (2014). Exploratory factor analysis in validation studies: Psicothema, 26(3), 395-400. doi:10.7334/psicothema2013.349.

Kaur, B., Toh, TL. (2012). Reasoning, Communication and Connections in Mathematics: An Introduction. 1-10. doi:10.1142/9789814405430_0001.

Kim, J. (2009). Development of a scale to measure memorable tourism experiences (Dissertation). Retrieved November 23, 2017, from http://scholarworks.iu.edu/dspace/bitstream/handle/2022/8786/Kim_indiana_0093A_10479.pdf;sequence=1.

Klobas, JE., Renzi, S. & Nigrelli,, MS. (2007). A scale for the measurement of selfefficacy for learning (SEL) at university. Working Papers 002, Carlo F. Dondena Centre for Research on Social Dynamics (DONDENA). Retrieved from https://ideas.repec.org/p/don/donwpa/002.html.

Li, Z., Toland, M., & Usher, E. (2013). Psychometric examination of the mathematics skills self-efficacy scale. Retrieved January 27, 2017, from http://sites.education.uky.edu/motivation/files/2013/08/LiTolandandUsher.pdf.

Ling, S. (2016). Development and validation of the singapore secondary mathematics self-efficacy scale. Dissertation. University of California. Proquest. Retrieved August 1, 2017.

Maddux, JE. & Kleiman, E. (2018). Self-efficacy. In R. B.-D. Diener, Noba textbook series: Psychology. Champaign, IL: DEF Publishers. doi:nobaproject.com.

Marat, D. (2005). Assessing mathematics self-efficacy of diverse students from secondary schools in Auckland: Implications for academic achievement. Issues in Educational Research, 15.

May, D. (2009). Mathematics self-efficacy and anxiety questionnaire. Doctoral Dissertation. Retrieved July 22, 2016, from https://getd.libs.uga.edu/pdfs/may_diana_k_200908_phd.pdf.

May, D. & Glynn, S. (2008). A mathematics self-efficacy questionnaire for college students. Retrieved December 27, 2016, from http://sigmaa.maa.org/rume/crume2008/Proceedings/May%20SHORT.pdf.

Mishra,S., Sharma, M., Sharma, RC., Singh,A., & Thakur, A. (2016). Development of a scale to measure faculty attitude towards open educational resources. Open Praxis, 8(1), 55 - 69. doi: http://dx.doi.org/10.5944/openpraxis.8.1.236.

Murphy, P., Wood, L., & Carter, L. (2007). A critical review of research in self-efficacy in mathematics education. Delta Special Edition. Retrieved January 27, 2017, from http://directorymathsed.net/montenegro/Murphy.pdf.

Muzaffar, B. (2016). The development and validation of a scale to measure training culture: The TC scale. Journal of Culture, Society and Development: An International Peer-reviewed Journal, 23, 49 - 58.

Nicolescu, R. (2015). The analysis of self-efficacy for students enrolled in calculus I course at a community college. (Dissertation). University of Oklahoma, Norman, Oklahoma.

Nipaz, JGG., Belecina, RR., & Garvida, MD. (2016). Language of encouragement: Effects on mathematics anxiety, self-efficacy and mathematics performance of college students in the Philippines. World Journal of Research and Review, 2(5), 9 - 14.

Nwosu, KC., & Okoye, RO. (2014). Students' Self-Efficacy and Self-Rating Scores as. Journal of Educational and Social Research, 4(3). doi:10.5901/jesr.2014.v4n3p223.

Ouano, J. (2011). Changes in self-efficacy among filipino adolescents as moderated by performance goal orientation. Philippine Journal of Counseling Psychology, 13(1), 65 - 75.

Parsian, N. & Dunning, T. (2009). Developing and validating a questionnaire to measure spirituality: A psychometric process. Global Journal of Health Science, 1(1), 111. doi:10.5539/gjhs.v1n1p2.

Patena, A. & Dinglasan, BL. (2013). Students' performance on mathematics departmental examination: basis for math intervention program. Asian Academic Research Journal of Social Science and Humanities, 1(14). www.asianacademicresearch.org.

SEI-DOST and MATTHED. (2011). Mathematics framework for Philippine basic education. Manila: SEI-DOST & MATHTED.

Soleymani, B. & Rekabdar, G. (2016). Relation between mathematics self-efficacy and mathematics achievement with control of math attitude. Applied Mathematics, 6(1), 16 -19. doi:10.5923/j.am.20160601.03.

Stephanie. (2017). Cronbach's alpha: Simple definition, use and interpretation. Retrieved December 22, 2017, from Statistics How to Theme: www. StatisticsHowToTheme.

Turner, R. (2010). Exploring mathematical competencies. ACER's Research Conference 2010: Teaching Mathematics? Make It count. Retrieved October 22, 2017, from http://www.acer.edu.au/conference.

Usher, E. & Pajares, F. (2009). Sources of self-efficacy in mathematics: A validation study. Contemporary Educational Psychology, 34, 89 - 101. doi:10.1016/jcedpsych.2008.09.002.

Yazon, A. (2015). Self-esteem, self-efficacy and academic performance of the college of teacher education students at laguna state polytechnic university, los banos campus, AY 2013-2014. Proceeding of the 3rd Global Summit on Education GSE 2015. Kuala Lumpur, Malaysia. Retrieved from http://WorldConferences.net.

Yuet, FKC., Yusof, H. & Mohamad, SIS. (2016). Development and validation of the teacher leadership competency scale. Malaysian Journal of Learning and Instruction, 13(2), 43 - 69.

Zimmermann, M., Bescherer, C., & Spannagel, C. (2010). A questionnaire for surveying mathematics self-efficacy expectations of future teachers. Retrieved August 2, 2017, from http://www.cerme7.univ.rzeszow.pl/WG/14/CERME7-WG14-PaperZimmermann-Bescherer-&-Spannagel-REVISED-Dec2010.pdf.