Developing and Validating an Instrument to Evaluate a Mathletes Training Program

Asia Pacific Journal of Education, Arts and Sciences
Vol. 4 No.1, 1-9
January 2017
P-ISSN 2362-8022
E-ISSN 2362-8030
www.apjeas.apjmr.com

Flordeliza Sorono-Gagani (MA)[1] and Amelia M. Bonotan (PhD)[2]
[1]University of San Carlos, [2]Cebu Normal University, Cebu City, Philippines
[1]sflorely@yahoo.com; [2]bonotana@cnu.edu.ph

Date Received: November 16, 2016; Date Revised: December 6, 2016

Abstract - *Most of the schools underserve high-ability and highly motivated students by denying them access to high-quality mathematics and giving only opportunities for the disadvantaged students through intervention programs. Hence, a Mathletes Training Program was initially implemented in order to expose these students to advance Mathematics and to bring their mathematical knowledge alive through engaging in outside school competitions and real life activities. However, the absence of an evaluation tool discontinued the program. This prompts the researchers to design an evaluation instrument to measure the effectiveness of the mathletes program in terms of attitude towards Mathematics and self-concept of the mathletes after undergoing the training. This study then details the development and validation of the Mathletes Training Program Evaluation Instrument. Samples of 60 elementary and junior high student- respondents were randomly selected from the total population of 211 Mathletes. Using Exploratory Factor Analysis, findings revealed that the three constructs-attitude toward math(ATM), mathlete training program(MTP) and self-concept (SC) were suited for factor analysis (KMO> 0.5) and were all reliable with Cronbach's Alpha values of 0.72, 0.73 and 0.85 respectively and overall reliability of Cronbach's alpha of 0 .84. From 20 ATM items, 61% had good factor loadings (>.4), 3 items were dropped and 5 factors were finally identified. Twenty-one (21) out of 25 MTP items captured 63% of the variance of the program effectiveness identified 7 factors after a reduction method. For the 12-item self-concept construct (SC), 3 factors were identified and captured approximately 61% of the total variance of the program effectiveness using the 3 factors. The results show that there is adequate evidence that the program evaluation tool designed is valid and reliable.*

Keywords: *Factor Analysis, Mathletes, Reliability, Validity*

INTRODUCTION

The 2000 National Council of Teachers of Mathematics (NCTM) President Lappan [1] once said:

> *"Our primary goal must be mathematical power for all students. We speak often about providing rich opportunities for disadvantaged students. But among the students we have in our mathematics programs are some that have either high abilities or high interest, or both. Our programs must include opportunities for these students as well. These students are likely to become significant users of mathematics as our future scientists, mathematicians, statisticians, engineers, technologists, and researchers. They deserve programmatic attention just as students with other kinds of special needs do."*

Hence, the University of San Carlos, Basic Education Department – South Campus ,Cebu City, Philippines had initially implemented a Mathletes Training Program for the *talented, gifted, interested and motivated students in Mathematics* who want to learn and appreciate more the beauty of Mathematics as the "queen of all sciences".

The Mathletes Training Program opened during school year 2015 – 2016 to the 211 qualified participants from Grades One to Ten with the objectives of honing the Mathematical skills of the students through exposing them to advance Math lessons, preparing them for outside school competitions, motivating and giving them opportunities to enjoy Math and see its real-life applications. Exposure to the topics that are beyond the general curriculum eventually would encourage them to pursue various topics that might be of interest to them. Further, this will give the students who have mastered the content to pursue the advanced topics in greater depth.

The program consists of a 2-hour session for 10 Saturday facilitated by the best Math teachers. A

qualifying examination patterned after MTAP (Metrobank Teachers Association of the Philippines) and PMO (Philippines Mathematical Olympiad) questions was given to select the mathletes.

The program went well. The mathletes were found to learn and enjoy the activities especially the Math Camp which serves as its culminating activity. In the succeeding school year, 2016-2017, many of the students asked if the Mathletes Training Program will be offered again. However, the Mathematics Department of the University *failed* to evaluate the program, due to the absence of an evaluation tool a requirement of the Administration for the continuation of the program. Hence, the program was temporarily discontinued as there was **no existing instrument** that would assess the level of effectiveness of the program.

In order to address the above-mentioned problem, the principal researcher was requested by the school administration to develop a valid tool to evaluate the training program as basis for the decision-making of the administration. The researchers accepted the challenge as they believe that the Math enrichment program is really needed by the students taking Science, Technology, Engineering and Mathematics (STEM) Strand as their K-12 curriculum career track.

Moreover, research studies also showed that these kind of enrichment programs increase students' success. Lubinski [2] mentioned that longitudinal studies of students with gifts and talents indicate that accelerated students who participated in talent search activities or other outside of school competitions achieved exceptional success and report high career and life satisfaction. In the same manner, Gavin Casa [3] also stressed that students who received a challenging math curriculum with focus on problem solving outperformed a comparison group of students of like ability from the same schools.

Hence, these thoughts prompted the researchers to develop and validate an instrument to evaluate the said program and to provide a model tool to other schools who may also want to engage in this type of enrichment program.

OBJECTIVES OF THE STUDY

This paper aims to develop and validate an instrument to evaluate the Mathletes Training Program. This includes discussing the process associated with the design and validation of an instrument.

Specifically, this study seeks to determine the construct validity and reliability of the (a) evaluation instrument – effectiveness of the program, measured also in terms of the other two (2) related constructs – Attitude towards Math (ATM) and Self-Concept (SP) of the Mathletes.

METHODS

Since this study is quantitative in nature, a descriptive-survey method through a researcher-made questionnaire measuring 3 constructs namely: (1) Mathletes' Attitude towards Mathematics (20 items), (2) Mathletes' Self-concept (12 items) (3) Effectiveness of Mathlete Training Program (25 items) were administered to the randomly selected 30 respondents in each of the elementary and junior high school level or a total of 60 out of the 211 University Mathletes. For the first two (2) constructs, a Likert – type scale that ranges from strongly agree to strongly disagree was utilized, while for the third construct, a semantic differential, was used where the respondents describe their feelings or attitude towards self and others before and after joining the program by simply checking the line from 1 to 5 where 1 is the lowest and 5 is the highest. Moreover, some five (5) open-ended questions were added to produce narratives of their experiences and were used to support the study through their significant statements.

The study was conducted in the University of San Carlos, Basic Education Department – South Campus, Cebu City, Philippines. , School Year 2016-2017. This is a private Catholic school managed by Society of the Divine Word (SVD) priests and is the only Level 3 Philippine Accrediting Association of Schools, Colleges and Universities (PAASCU) accredited Basic Education institution in Cebu City, Philippines.

In order to have a systematic view of the development of the instrument, a workflow of designing the Mathlete Training Evaluation Instrument is presented below:

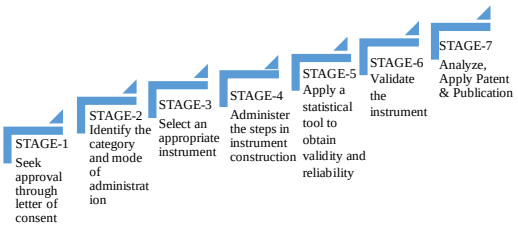


Fig. 1. Workflow of designing a Mathlete Training Program Evaluation Tool

The workflow started with the seeking of approval first from the authorized persons. Then, the researcher began to plan the development of the instrument

P-ISSN 2362-8022 | E-ISSN 2362-8030 | www.apjeas.apjmr.com

through identifying the category and its mode of administration, selecting the most appropriate instrument and administering the steps in instrument construction. Validity and reliability of the instrument was then determined using a multivariate statistical technique and the internal consistency was measured using Cronbach's alpha. The instrument was then validated by pilot testing it to a hundred respondents. Once evidence showed a valid and reliable instrument, application for patent and publication are the next in line.

According to Colton and Covert [4], "*an instrument is a mechanism for measuring phenomena, which is used to gather and record information for assessment, decision making, and ultimately understanding. An instrument such as a questionnaire is typically used to obtain factual information, support observations, or assess attitudes and opinions*".

Further, instruments are categorize in several ways. One way is based on a mode of administration which tells the one responsible in completing the instrument. This could be self-report, observation (observer-based) or a combination of both. In this study, the respondents used a self-report where through paper and pencil, they supplied the information directly. Although, self-reports are "*subject to respondent's personal interpretation of an item, which may or may not be what you, as instrument designer, had intended. In either case it is important to test and revise the instrument to minimize these potential problems*." [4]. There was no need for the study to apply external raters (intra or inter raters) to observe or to rate the respondents.

Instruments were also classified according to use or purpose. This study used a more friendly type of survey instrument such as a checklist. Checklists are commonly used for nominal data, something that is dichotomous in nature such as the yes-no answers, present or not present and open-ended questions. This type of instrument was selected due to the following reasons: purpose of the study which is to evaluate a program; research design which is a survey research; object of measurement which refers to the focus of the inquiry; data collection methodology and resources.

Further, a rating scale that ranged from strongly agree to strongly disagree were used in the checklist. As Aiken [5] notes that "*rating scales are a primary tool in contemporary assessment methodology, second only to teacher-made achievement tests in frequency of use for rating people, objects, and events*."

The undecided choice in the range was included by the researcher despite negative feedbacks of its value because there is no enough evidence in literatures that prove that including undecided in the rating scale minimize validity and reliability of the questionnaire.

The administered survey questionnaire contained the following: a title that convey the purpose of the instrument; introduction that sets the mode of the respondents by letting them know the reason of conducting the study and making them feel comfortable upon knowing the confidentiality of the result; directions or instructions that were stated clearly to avoid errors and misconceptions; items that were not ambiguous through avoiding double barreled questions; demographics that include respondents personal background and an ending/closing statement to thank the respondents for sharing their precious time.

The Process of Instrument Development

Background. Instrument construction is a science that deals with the systematic procedure of creating an instrument to obtain data and information to solve a certain phenomenon. The design of the instruments' constructs were based on the theory in which this study is anchored.

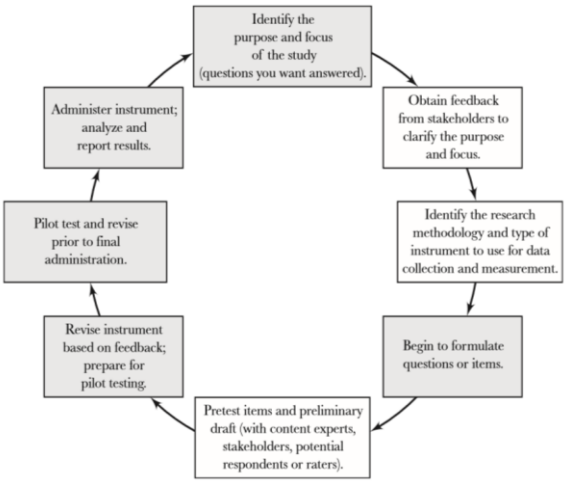
The social cognitive theory of Albert Bandura and the zone of proximal development of Lev Vygotsky prompt the researcher to include the constructs on attitude towards mathematics and self –concept aside from making a tool to evaluate the training program alone. This is done because the researcher hypothesize that some students just join the Mathlete Training program for the math camp purposes and not really on the intellectual trainings that they're supposed to gone through. Hence, the inter-correlations between attitude, the program and students' self-concept must be seen to be highly correlated with one another.

Furthermore, according to Bandura [6], positive outlook on oneself towards mathematical tasks lessen the idea that math is a difficult and a complicated subject. The more interested students are in math, the higher the achievement they may possibly attain. He further stated that students who have strong self-efficacy treated failure as insufficient efforts of deficient knowledge and skills which are acquirable [6].

Meanwhile, Vygotsky's [7] theory of Zone of Proximal Development (ZPD) supported this idea and stressed that social interaction plays a fundamental role in the development of cognition. Full cognitive development requires social interaction [8].

P-ISSN 2362-8022 | E-ISSN 2362-8030 | www.apjeas.apjmr.com

Process. Figure 2 formally details the process of instrument construction in this research study using Colton and Covert [4] research instrument design.



Source: Colton and Covert, 2007

Figure 2. Steps in the Instrument Construction Process

A letter of approval was secured first from the administration office for the planned development of an evaluation tool in order to answer the need of the Mathematics Department of the University of San Carlos.

A statement of purpose was made first in order identify the purpose and focus of the study. It included evidences of researchers' interest and reasons of conducting the study. Initial questions were framed as basis in item construction in the table of specifications (TOS). Presumptions and interpretations were done since it influenced the choice of topic and questionnaire items in the study. Through writing the statement of purpose, the researchers were able to identify the types of test and the kind of questionnaire to administer. In this way, an initial analysis and interpretations were crafted which were needed in the creation of the table of specification. Second, the TOS was then created to organize ideas so as not to miss any important points. It provided a systematic view of the constructs, its operational definitions, dimensions of the constructs, type of assessment/rater and sample items.

Feedbacks were then obtained through a meeting with the Math Area Coordinator and the rest of the Math teachers for clarity of purpose and focus of the study. A State University Professor who handled a post-graduate subject on Research Instrumentation was also asked to identify aspects to focus on or exclude from the study. Then from obtaining feedback, the methodology were determined as well as the type of instrument in collecting and measuring data.

Using the table of specifications, the researcher began to formulate the items to be included in the survey questionnaire. Different approaches in developing items for an instrument were done such as conducting a literature review, making use of existing processes like policies or regulations and procedures of the program. Brainstorming which happened to be the most common method of idea generation and asking a repetitive why just like a child were done in order to filter from generalities to specifics and to identify the right methods to better understand the situation. Other group approaches such as snowballing or pyramiding, nominal group technique (NGT), Delphi technique which solicits the input of content and methodological experts, employing item pools and Q sort and concept mapping were not done in the item development of this study.

After the item generation, the instrument was pretested to five (5) respondents. Feedbacks were obtained regarding the appropriateness of the language used in the items, the clarity of the instructions, accuracy of the format and the length of time in answering the questionnaire. Content experts such as the school's resident Psychometrician, State University Professor, Math Area Coordinator, elementary and junior high school Math teachers were asked to critic on the evaluation tool. Revisions of the instrument based on feedback were done immediately. At this point, the evaluation tool was submitted to the school principal informing that the tool was ready for pilot testing. However, feedbacks were again given that led to minimal revisions in the format and instructions.

With the approval of the school Principal, the tool was pilot tested to randomly selected 30 respondents in each elementary and junior high school levels. Feedbacks and observations were noted especially on the length of time for the respondent to answer the questionnaire, their comfortability and the appropriateness of time to administer the instrument. The appropriateness of the language in each item were taken into consideration too since test-takers also involve elementary students like grade 1. The manner of administration were discussed during the meeting before the final administration which was done by the teacher-trainers.

The data were then collected and analyzed using Microsoft excel and a statistical package for social

P-ISSN 2362-8022 | E-ISSN 2362-8030 | www.apjeas.apjmr.com

sciences (SPSS) software version 16.0. Reports were given to the Math Area Coordinator and to the school Principal. Based on the feedbacks of the school administration, the Math teachers gathered again during the Math Area meeting to discuss things related to the program.

The Process of Instrument Validation

Background. Validation refers to assembling the evaluative summary of evidence. Evidence based on content (TOS), responses, processes (item analysis), internal structure (Factor Analysis), relation with other variables, and based on the consequence of measuring.

The main purpose in any Social science research instruments is to obtain valid and reliable data and information in order to solve and gain better understanding of a certain phenomenon. The accuracy and truthfulness of a research study is nothing if there is no "good" instrument. An instrument of good quality produces an accurate, valid and reliable result while instruments of poor construction quality yield an invalid result. Hence, in any educational research studies, instruments are the "captain of the ship". It holds the control and taking shortcuts should be discouraged to any researcher.

In this study, the validity and reliability of the evaluation tool is dealt with utmost priority. According to Reston [9], validation of the test scores and responses generated from measuring instruments became one of the arising problem in researches. She further stated that "*unless one uses a standardized test and research instruments as data-gathering tools, the researcher has to take necessary measures to ensure validity of the data generated from their instruments. This problem is important in research because poor measurement can invalidate any scientific investigation."*

In measurement, validity and reliability are two important characteristics of a good measuring instrument. While validity is traditionally defined as the degree to which a test or instrument measures what it is designed to measure, reliability refers to the degree of consistency of a measure [10]. The traditional view of validity considers validity as a characteristic of a test or measuring instrument. Most people talk about a test being valid or not valid but with the modern unified view of validity, Reston [9] stressed that "*a test is just a test and any measuring instrument is just a data gathering tool; what needs to be validated are the scores and responses generated from these measuring instruments."* As Cronbach [11] puts it, "*One validates, not a test, but an interpretation of data arising from a specific procedure*." Moreover, Sheppard [12] pointed out that "*validity must be established for each particular use of a test*."

Hence, Messick [13] argues that the traditional view of validity is fragmented and incomplete especially because it fails to consider the evidence of the value implications of data interpretations arising from research instruments as basis for action and social implications.

The modern unified view of validity on the other hand, focuses on the appropriateness and justifications of decisions based on the data derived from tests and other measuring instruments. Validity is defined as "*the degree to which the evidence supports that the interpretations are correct and the manner on which they are used is appropriate.*" [14]

Furthermore, it involves the process of accumulating evidences that support the appropriateness of the inferences that are made of responses to instruments for specified uses. The more evidences, the better. There is no cut off score and it is best to triangulate to achieve more valid results and not just rely on one instrument.

The American Psychological Association (APA) identified four types of validity in 1954 according to purpose of the test or instrument: (1) content validity, (2) predictive validity, (3) concurrent validity, and (4) construct validity. [15]. However, these four types were reduced into 3 in 1966, fusing predictive validity and concurrent validity to a single category: criterion-related validity.

The Process. The instrument was subjected to processes related to establishing its validity according to purpose such as face validity, content validity and construct validity. Initial steps were done to establish face validity by letting the Math teachers see the questionnaire to check its appearance and items. Content experts such as the school's psychometrician, a state university professor who handled a post-graduate subject on Research Instrumentation, Math Area coordinator, Math teachers in both elementary and junior high school and the school principal were asked to check the tool for content validity.

While for construct validity, the researchers followed the definition given Heppner, Kivlighan and Wampold [16], as referring to the "*degree to which the measured variables used in the study represent the hypothesized constructs*." This means that in order for the instrument to have a valid construct it should answer the question: "*Does this test or instrument really measure what it intends to measure*?"

P-ISSN 2362-8022 | E-ISSN 2362-8030 | www.apjeas.apjmr.com

A construct refers to a characteristic which is presumed to exist but which cannot be directly measured such as intelligence, anxiety or self-esteem. Its presence and strength or amount is inferred from sample of behavior gathered by a measure designed to assess it [17].

Factor Analysis. In order to determine the construct validity of the three (3) constructs of this study, the researchers performed a multivariate statistical procedure called Factor Analysis which is "*a method for reducing a large number of measures to a smaller number, called factors, by discovering which ones measure the same thing and the relations between the clusters of measures that go together*." [17]. Further, a factor is defined as a construct, a hypothetical entity that is assumed to underlie tests, scales, items and measures of almost any kind.

Factor analysis is broad and wide in scope. It is classified according to purpose –exploratory or confirmatory. This study used the exploratory factor analysis (EFA) to identify the factor structure or model for a set of variables. The number of existing factors and the pattern of factor loadings were determined. EFA was chosen as the statistical tool for data analysis in this study since it is more of a theory – generating than a theory testing procedure which is a characteristic of confirmatory factor analysis (CFA).

In doing factor analysis, data were screened which includes looking at the table of descriptive statistics for the variables included in the study and the number of variables and cases with non-missing data. The r values between pairs of variables in the correlation matrix table were assessed too. Kaiser-Meyer-Olkin (KMO) and Bartlett's Test of Sphericity test were used to determine data's suitability for a factor analysis. As a rule of thumb, when the value of this KMO is less than 0.5, the matrix is said to be *not* suited to factor analysis; values from 0.5 to 0.7 are considered marginal; and values greater than 0.7 are adequate.

In addition, the study used a technique called principal component analysis, a method used which provides a unique solution so that the original data can be reconstructed from the results. It explains as much variance as possible with the first factor extracted. This variance was then removed, and a second factor was extracted to explain as much as possible of the remaining variance until 100% of the variance has been explained. This process is called the principal axis method and results in uncorrelated factors are called orthogonal factors. During the analysis, it was noted that a variable with a low communality (0.40 or lower) was understood to be less represented by combined factors. The table of the total variance explained and scree plot have helped in deciding the number of factors to be retained. The factor loadings on the rotated component matrix after applying a Varimax rotation method with Kaiser Normalization and the corresponding plot of the loadings were used to interpret the components. As a rule of thumb in interpreting the components, loadings of magnitude lower than 0.30 or 0.40 are considered low and according to Zwick and Verlicer [18] *"variables with these loadings are typically ignored",* while loadings in the 0.50 to 0.80 are considered substantial.

The study undertook the four basic steps in factor analysis, namely: (1) data collection and generation of the correlation matrix, (2) extraction of initial factor solution, (3) rotation and interpretation, and (4) construction of scales or factor scores.

To measure the internal consistency of the items, Cronbach's Alpha reliability test was used in the study.

RESULT AND DISCUSSION

Using Exploratory Factor Analysis, a Principal Axis Factor with a Varimax (orthogonal) rotation of 17 of the 20 Likert scale questions from the attitude towards mathematics (ATM) survey questionnaire was conducted on data gathered from 60 respondents.

An examination of the Kaiser-Meyer Olkin measure of sampling adequacy suggested that the sample was factorable (KMO = 0.608) and Bartlett's test of Sphericity is $x^2(190) = 465.864$. While the second and third constructs measuring the Mathletes Training Program (MTP) with 21 of the 25 Likert scale questions and self-concept (SC) with 12 questions in sematic differential form yield a KMO value of 0.526 with Bartlett's Test of Sphericity ofx^2 (300) = 522.315 and 0.780 with x^2 (66) = 263.626 which is considered marginal and adequate respectively. This means that the three (3) constructs were suited for factor analysis.

Table 1 shows the results of an orthogonal rotation of the solution. Based on exploratory factor analysis with Varimax rotation on initial 20 attitude items, when loadings less than 0.5 were excluded due to poor factor loadings, the analysis yielded a 6 factor solution with 17 retained items. However, for redundancy reasons, these factors were analyzed further through controlling the number of factors which resulted to five (5) final factors.

P-ISSN 2362-8022 | E-ISSN 2362-8030 | www.apjeas.apjmr.com

Table 1. Factor loading of each item based on Exploratory Factor Analysis on Attitude towards Mathematics (ATM)

Items	Cronbach's Alpha if item deleted	Component & Factor loadings 1	2	3	4	5
1.rewarding	.692	.776				
2.useful for daily living	.698	.714				
3.excites me	.686	.680				
4.interesting	.696	.678				
5. useful in daily act.	.701	.622				
6.enjoy a great deal	.697	.585				
7.sleepy thinking	.704		.753			
8.nervous thinking	.705		.696			
9.not sure of myself	.712		.671			
10.Never like math	.708		.652			
11.tired solving	.711		.525			
12.like no. like math	.720			.823		
13.like math more	.711			.601		
14.not challenging	.738				.880	
15.cant see challenges	.732				.772	
16.cant see much value	.696					.681
17.not practical	.712					.607

The five (5) identified domains/factors are namely (1) Positive experience, (2) Apprehensions, (3) Choice Preference, (4) Impact, and (5) Practicality. Each of the 17 items were analyzed using Cronbach's Alpha. Overall reliability statistics yielded a value of 0.721 which is a good and reliable value. The cumulative percentage of variance using six (6) factors was 66.68% but when reduced into five (5) factors, the combined factors accounted for approximately 61% of the combined variance of ratings of the five (5) factors which is an acceptable value in educational research.

On the other hand, Table 2 revealed the 21 from 25 item under Mathletes Training Program (MTP) construct.

Originally, nine (9) factors were found out after performing an extraction method using Varimax with Kaiser Normalization. Again, redundancy and similarities of factors were noticed which led to reducing it into seven (7) factors namely, (1) Offers bright opportunities, (2) Realized Provisions, (3) Appropriateness, (4) Efficient activities, (5) Satisfaction, (6) Good planning, and (7) Effective training. The total reliability measure using the Cronbach's Alpha found out to be 0.725 for 25 items with an estimated 63% of the cumulative variance of ratings for the 7 factors.

Table 2.Factor loading of items based on Exploratory Factor Analysis on Mathlete Training Program (MTP)

	Items	Cronbach's Alpha if item deleted	Factor loadings
Factor 1	1.Quiz bowl is memorable	.712	.698
Offers Bright	2.taught me to explain	.718	.636
Opportunities	3.opened opportunities	.708	.625
	4.max participation	.698	.618
	5.teacher had genuine interest	.700	.569
Factor 2	6. teacher provide materials	.710	.676
Apprehensions	7.venue is suitable	.710	.632
	8.made to realize enjoyable subject	.710	.537
Factor 3	9.math camp is appropriate	.709	.776
Choice Preference	10.provide basic knowledge	.713	.748
Factor 4	11.schedule is sufficient	.723	.838
Efficient Activities	12.classroom is right venue	.748	.691
	13.enough training session	.728	.603
Factor 5	14.good materials, sound system	.751	.662
Satisfaction	15.exam is best way to select	.742	.591
	16.math camp most awaited	.719	.532
Factor 6	17.acivities are well thought of	.715	.755
Good Planning	18.need to have orientation	.727	.745
	19.enough allotted time	.715	.631
Factor 7	20.effective training program	.723	.736
Effective Training	21.participation of members	.716	.635

As can be gleaned in the third construct on self – concept (SC), table 3 showed 3 factors namely (1) Highly motivated, (2) Positive outlook and (3) Happiness. The Mathletes Training Program revealed that students began to develop self-confidence, became highly motivated to study more and on their own and as a whole they were happy with what they

P-ISSN 2362-8022 | E-ISSN 2362-8030 | www.apjeas.apjmr.com

were doing and what they have undergone as shown by bigger factor loadings compared to the rest of the items.

Table 3. Factor loading of each item based on Exploratory Factor Analysis on Self-Concept (SC)

	Cronbach's Alpha if Item Deleted	Component		
		1	2	3
1.confident	.834	.805		
2.active	.819	.740		
3.friendly	.829	.721		
4.energetic	.833	.679		
5.proud	.835	.674		
6.outgoing	.828	.665		
7.interested	.835	.581		.505
8.relax	.848		.864	
9.optimistic	.829		.732	
10.cooperative	.831		.595	
11.happy	.863			.858
12.satisfied	.842			.574

Being confident (factor loading of 0.834) in the first factor is on the top rank. This means that students were able to cope up from their first baptisms of fire. They became satisfied and happy with a high 0.858 factor loading as evidence that they have improved a lot due to their contentment with the offerings of the program. As a result, they develop a positive outlook in life and considers joining the program again since it brought positive feelings and experiences that eventually led them to embrace the abstract nature of the subject and the difficulty it presents.

Using Cronbach's Alpha, the 12-item third construct obtained a reliability value of 0.847 which is highly reliable. The construct captured approximately 61% of the total variance using the three (3) factors. Overall reliability statistics revealed a highly reliable Cronbach's Alpha value of 0.843 among the 57 items. This means that there is enough evidence that the three (3) constructs are valid and reliable.

CONCLUSION AND RECOMMENDATION

There is sufficient empirical evidence that the evaluation instrument of the Mathletes Training Program is valid and reliable. It had undergone processes for face, content and construct validity. All the components of the instrument were checked by experts and the category and mode of administration were checked for accuracy and appropriateness. The steps of instrument construction as design by Colton and Covert [4] was strictly followed as well as the processes in the validation of instrument. Factor analysis as a multivariate statistical tool proved the constructs are valid. Furthermore, this valid tool served may serve as a model instrument in measuring the effectiveness of any program in the school as well as measuring attitude and self-concept of students as fruits of participation of the enrichment program.

Furthermore, the Mathletes Training Program creates a significant impact on the attitude and self-concept of the students as revealed by the factors of the constructs. A positive outlook in life, a highly motivated student and most importantly, a happy person came out to be the factors common among the Mathletes. This shows that the training program is effective as it creates a change on the students not only academically but socially as well. This is the most interesting part of this study since most intelligent and mathematically gifted persons (branded as "the nerds") are mostly seen spending their time reading Math books in libraries. In this study, the factors revealed that these students can be more cognitively successful if they engage themselves socially thus giving enough evidence to the constructivist and Banduras' way of making the learners construct their own knowledge by engaging them in meaningful and lasting learning experiences with others.

The new tool may also be used in evaluating an intervention or an enhancement programs in *any* subjects provided that it will undergo some revisions for the items to fit and will measure what it intends to measure. Developing and validating an instrument using group approaches is also highly recommended.

REFERENCES

[1] Lappan, G. March 1999. 'Mathematics for All' Must Include High-Ability and Highly Motivated Students. *NCTM News Bulletin*. Retrieved from https://goo.gl/5rY2BC on October 17, 2016.

[2] Lubinski, D., Benbow, C. P., Webb, R. M., & Bleske-Rechek, A. (2006). Tracking exceptional human capital over two decades. Psychological Science, 17, 194–199.

[3] Gavin, M. K., Casa, T. M., Firmender, J. M., & Carroll, S. R. (2013). The impact of advanced geometry and measurement curriculum units on the mathematics achievement of first-grade students. *Gifted Child Quarterly, 57*, 71–84.

[4] Colton, D. and Covert, R.,W. (2007). *Designing and constructing instruments for social Research and evaluation, 1st Ed.*, John Wiley & Sons, Inc. pp. 18.

[5] Aiken, L. R., Jr. (1975). Procedures and problems in designing a college course evaluation and questionnaire. *College and University, 50*, 247 – 253.

[6] Bandura, A. (1994). Self-efficacy. In V. S. Ramachaudran (Ed.), Encyclopedia of human behavior (Vol. 4, pp. 71-81). New York: Academic Press. (Reprinted in H. Friedman [Ed.], *Encyclopedia of mental health*. San Diego: Academic Press, 1998).Vygotsky, L.S. (1978). *Mind and society: The development of higher mental processes*. Cambridge, MA: Harvard University Press

[7] Kearsley, G. (1994e). Social development theory (L. Vygotsky). [Online]. Available: https://goo.gl/no0GZS [December 1, 1999].

[8] Reston, E. D. (2003). Factor Analysis as Tool in Construct Validation of Research Instruments, *USC Graduate Journal*, *20* (1), pp. 136-145.

[9] Chase, C. I. (1990). *Measurement for educational evaluation*. New York Macmillan Company.

[10] Cronbach, L.I. (1989). "Construct Validation after Thirty Years." *Intelligence: Measurement Theory and Public Policy*. In R.L. Linn (Ed.) Urbana University of Illinois Press.

[11] Shepard, L.A. (1993). "Evaluating Test Validity." *Research in Education*. In L. Darling-Hammond (Ed.). Vol. 19:405-450. Washington, DC: American Educational Research Association.

[12] Messick, S. (1996). "Validity." Educational Measurement. In R.L. Linn (Ed.). Third Ed. New York Macmillan.

[13] *American Educational Research Association*, (1999). Retrieved from https://goo.gl/Zgm8aM

[14] *American Psychological Association*, (1954).

[15] Heppner, P.P., D.M. Kevlighan, Jr. and B.F. Wampold. (1992). Research Design in Counselling. CA Books Cole.

[16] Kerlinger, F. & Lee, H. (2000). *Foundations of Behavioral Research*. Fourth Edition. New York Holt, Rhinehart and Winston, Inc.

[17] Zwick, W. R. & Velicer, W.F. (1982).Factors influencing four rules in determining the number of components to retain. *Multivariate Behavioral Research*, *17*, 253 – 269.

[18] Musa R., Aris, M., Draman, S., Abdullah K, Bujang. (2015). Designing and Validating a New Asian Family Scale.
Retrieved from https://goo.gl/N4BNsz, on October 27, 2016.