Assessing Students' Numeracy through Item Response Theory

Nu Nu Nyunt [1], Win Ei San[2]

Abstract

The objectives of this study were to investigate the Grade 6 students' numeracy and to develop a numeracy test by using the two-parameter logistic Item Response Theory (IRT). Questionnaire survey method was used to measure students' numeracy. The data were collected in two regions and one state of Myanmar. Simple random sampling technique was used in this study. Alpha reliability for Numeracy Test for Grade 6 Students revealed at 0.85. In this study, gender related difference was found to be on number operation sense component, measurement and shape component, statistics component and the whole test of numeracy. In addition, results also revealed that the mean score of Grade 6 students from urban schools was higher than that of students from rural schools on each subcomponent of numeracy test as well as the whole numeracy test.

Key Terms: numeracy, gender, rural, urban, IRT

Introduction

The major purpose of this study is to investigate the students' numeracy. To develop a numeracy test to measure the numeracy of Grade 6 students in Myanmar is of next interest. And then, this study sought to construct a numeracy test by using the two-parameter logistic IRT model and also tend to get a wider knowledge of assessing students' numeracy through item response theory. Numeracy might define as the mathematical knowledge needed by the every human being to empower them for life in that society. Kemp (2005) argued that the term numeracy is used in a variety of ways in the literature. These range from defining numeracy as a set of basic mathematical skills through to a rich description of the use of mathematics in a whole range of different contexts. Bynne r & Parsons (1997), Gleeson (2005), Parsons & Bynner (2005) found that poor numeracy skills had more impact on an individual's life than poor literacy skills. People without numeracy skills suffered worse disadvantage in employment than those with poor literacy skills alone.

1. Dr, Assistant Lecturer, Department of Educational Psychology, Yangon University of Education (email: nyuntnn@gmail.com)

2. Junior Teacher, State High School No (2), Hinthada, Ayeyarwaddy Region

They left school early, frequently without qualifications, and had more difficulty in getting and maintaining full-time employment. The jobs entered were generally low grade with limited training opportunities and poor pay prospects (Bynner & Parsons, 1997). According to Steen (1990), numeracy is to mathematics as literacy is to language. Each represents a distinctive means of communication that is indispensable to civilized life (as cited in Kemp, 2005). According to Dossey (1997), numeracy may be defined as the ability to interpret and apply the aspects of mathematics such as date representation, number and operation sense, measurement, variables and relation's geometric shapes and spatial visualization (as cited in Kemp, 2005).

Literature Review

Definitions of numeracy have proliferated. One view equates numeracy with mathematics and computational skills, in much the same way that literacy is viewed as mastery of basic reading and writing. A much broader view of numeracy focuses on people's capacity and propensity to interact effectively and critically with the quantitative aspects of the adult world (Gal, 2002a). Maguire and O'Donoghue's (2002) organizing framework developed through discussions with researchers and practitioners in A Research Forum (ALM), Adults Learning Mathematics, offers a way of bringing some order into the conceptual confusion surrounding adult numeracy. In the framework, concepts of numeracy are arranged along a continuum of increasing levels of sophistication. In the formative phase, numeracy is considered to be basic arithmetic skills; in the mathematical phase, numeracy is "in context," with explicit recognition of the importance of mathematics in everyday life. The third phase, the integrative phase, views numeracy as a multifaceted, sophisticated construct incorporating the mathematics, communication, cultural, social, emotional, and personal aspects of each individual in context.

Ginsburg L, Manly M & Schmitt M.J (2006) proposed three major components; (1) context, (2) content, (3) cognitive and affective. Context is the use or purpose for which an individual takes on a task with mathematical demands. In societal contexts, family or personal is related to an individual's role as a parent, head of household or family member. The demands include consumer and personal finance, household management, family and personal health care, and personal interests and hobbies. Work place deals with the ability to perform tasks on the job and to adapt to new employment demands.

Community includes issues around citizenship, and other issues concerning the society as a whole, such as the environment, crime, or politics. Further learning is connected to the knowledge needed to pursue further education and training, or to understand other academic subjects. The focus on applying mathematics in a context or having a social purpose to the use and application of the mathematics provides motivation for learners to engage with and learn about mathematics. This leads to conclude that it is the focus on, and prioritization of, context that differentiates an adult numeracy framework from a formal school mathematics framework.

The content component of numeracy consists of the mathematical knowledge that is necessary for the tasks confronted. Numeracy content will also vary from context to context within the same time period. For example, a carpenter need a high level of practical understanding of measurement and geometry to ensure accurate fits and structural integrity; an office worker may need an understanding of the algebraic concepts of variables and equations to use spreadsheets effectively; and a factory worker may use statistical process control measures that require an understanding of what constitutes abnormal deviation in the quality of the output of a certain machine. At the beginning of the twenty-first century, general numeracy content organizes around four mathematical strands; namely, number and operation sense, patterns, function, and algebra, measurement and shape and data, statistics, and probability (Ginsburg L, Manly M & Schmitt M.J, 2006).

Number and operation sense is a sense of how numbers and operations work and how they relate to the world situations that they are represent. Patterns, functions and algebra is an ability to analyze relationships and change among quantities, generalize and represent them in different ways, and develop solution methods based on the properties of numbers, operations and equations. Measurement and shape is the knowledge of attributes of shapes, how to estimate and determine the measure of these attributes directly, or indirectly, and how to reason spatially. Data, statistics and probability is the ability to describe populations, deal with uncertainty, assess claims, and make decisions thoughtfully (Ginsburg L, Manly M & Schmitt M.J, 2006).

According to Hambleton, Swaminathan & Rogers (1991); Lord (1997), (1980); Lord & Stocking (1988), item response theory (IRT) models the relationship between a person's level on the trait being measured by a test and the person's response to a test item or question (as cited in Aye Aye Myint, 2001). Item response theory has two postulates: (1) the examinee performance can be predicted or explained by a set of factors called latent traits or abilities or θ and (2) the relationship between examinees' item performance and the set of traits underlying item performance can be described by a monotonically increasing function called an item characteristic curve (ICC). It provides the probability of examinees answering an item correctly for examinees at different points on the scale. The ICC is the basic building block of item response theory (Crocker & Algina, 1986).

Perhaps, the most important advantage of unidimensional item response models (Wright, 1968) is that is, given a set of test items that have been fitted to an item response model (that is, item parameters are known), it is possible to estimate an examinee's ability on the same ability scale from any subset of items in the domain of item have been fitted to the model. The domain of items needs to be homogeneous in the sense of measuring a single ability. If the domain of items is too heterogeneous, the ability estimates will have little meaning.

There are three primary advantages of item response theory models (Hambleton & Swamingthan 1985):

1. Assuming the existence of a large pool of items all measuring the same trait, the estimate of a examinee's ability is independent of the particular sample of test item, that are administered to the examinee,
2. Assuming the existence of a large population of examinees, the descriptions of test item (for example, item difficulty and discrimination indices) are independent of the particular sample of examinees drawn for the purpose of calibrating the items,
3. A statistic indication the precision with which each exminee's ability is estimated is provided. This statistic is free to vary from one examinee to another

On the other hand, IRT has some disadvantages (Lord & Stocking, 1988). It is currently not possible to completely check the accuracy with which the assumptions are met by the data. For data that appear to meet the assumptions, however, it is reassuring that predictions made from item response theory can often be independently verified. Applications of IRT are generally more expensive than similar applications of classical test theory, and many applications of IRT require the use computer (as cited in Aye Aye Myint, 2001).

Methodology

Research Design

Quantitative perspective was used in this study. Questionnaire survey method was used to measure students' numeracy. Four aspects of Numeracy such as number and operation sense, algebra, measurement and shape, and statistics were measured. Simple random sampling technique was used in this study. First of all, the sample for 14 High Schools and 6 Middle Schools such as 4 High Schools and 3 Middle Schools from Yangon Region, 6 High Schools and 3 Middle Schools from Ayeyarwaddy Region, 4 High Schools from Rakhine State were selected. A total of 1005 Grade 6 students participated in this study. Out of 1005 Grade 6 students, 464 (46.2%) are boys and 541 (53.8%) are girls and their ages range from 10 to 13 years. The socioeconomic status of the sample ranged from lower, middle and upper class families.

Instrument

In this study, researcher adapted from the content components of NCSALL Occasional Paper (2006). The instrument used in this study was constructed under the guidance of experts in educational test and measurement field and with the guidance of Teacher Guide Book and Grade 6 Mathematics Text Book to be suitable for Grade 6 students in Myanmar. The detail procedures for constructing a numeracy test for Grade 6 students were as follows. Base on the table of specification, 25 multiple-choice items for each sub-component such as (1) number and operation sense, (2) algebra, (3) measurement and shape and (4) statistics, totally 100 items. The instrument was reviewed by ten experts from Educational Psychology Department, two experts from Educational Methodology Department, one expert from Educational Theory Department of Yangon Institute of Education and one retired expert from Educational Psychology Department of Yangon Institute of Education. Next, revisions in item length, and the wording of items were made according to supervision and editorial review of these experts. Pre-pilot study was done with a sample of 50 Grade 6 students from Basic Education High School, Thuwana to test whether the wording of test items had clarity or not and items were appropriate and relevant to Grade 6 students. According to the pre-piloting result, numeracy test was modified. Finally, Numeracy Test for Grade 6 Students' was developed by applying the two-parameter logistic IRT model.

Table 1 Table of Specification for Numeracy Test of Grade 6 students

Content	Learning Outcomes						Total Numbers of Items	
	Knowledge		Comprehension		HOT			
	No. of Items	%	No. of Items	%	No. of Items	%	No. of Items	%
Number and Operation Sense	5	9	7	12	2	4	14	25
Algebra	1	2	11	19	6	10	18	32

Measurement and Shape	7	12	5	9	7	12	19	33
Statistics	3	5	2	4	1	2	6	10
Total	16	28	25	44	16	28	57	100

To construct numeracy test by using the two-parameter logistic IRT model was employed using BILOG-MG 3 and Microsoft Excel. As an input, the program provides Phase 1 output, Phase 2 output and Phase 3 output. Test and item identification and classical item analysis results of the test appear in Phase 1 file. The test is calibrated in Phase 2 output. In Phase 3, ability scores for all samples from the test are computed. According to Phase 2 output, the usual range for "a" is from 0 to 2 and high value of "a" indicates that the higher discrimination power of an item between high and low achievement of students (Hambleton etal, 1991). The values of "b" typically vary from about -2 to +2 and the negative sign indicates that easier item difficulty and positive sign indicates that the harder item difficulty. The values of the discrimination power of items were range from 0.192 (item 44) to 1.303 (item 22). So, it can be said that all items had good discrimination power. The range of difficulty parameter of items in the test was from -2.402 (item 47) to 4.654 (item 44). It was observed that item 14, 44, 45, 49,50, 51, and 53 were very difficult for the students due to b values of the items were greater than +2. It was found that item 43 and 47 were very easy for the students due to b values of the items were less than -2 and the remainders had good difficulty level.

In this study, the result of Phase 2 output was more emphasized than Phase 1 output because classical item statistics are dependent on the examinee sample in which they are obtained (Hambleton & Jones, 1991). According to Phase 2 output, item 14, 43, 44, 45, 47, 49, 50, 51 and 53 were not used in the test to assess Grade 6 students' numeracy.

Investigation of IRT Graphic Illustration for the Numeracy Test

The following figure indicates that the matrix plot of Item Characteristics Curves (ICC) for the test items. The item characteristics curve (ICC) is a graphical representation of the probability of choosing the correct answer to an item as a function of the level of the attribute being analyzed by the test. The item characteristics curve (ICC) serves as the foundation of item response theory. ICC also summarizes much of the information conveyed by item analysis and suggests how this information might be used to understand the relationship between the attribute being measures and test responses (Lord, 1997: Lord & Novick, 1968).

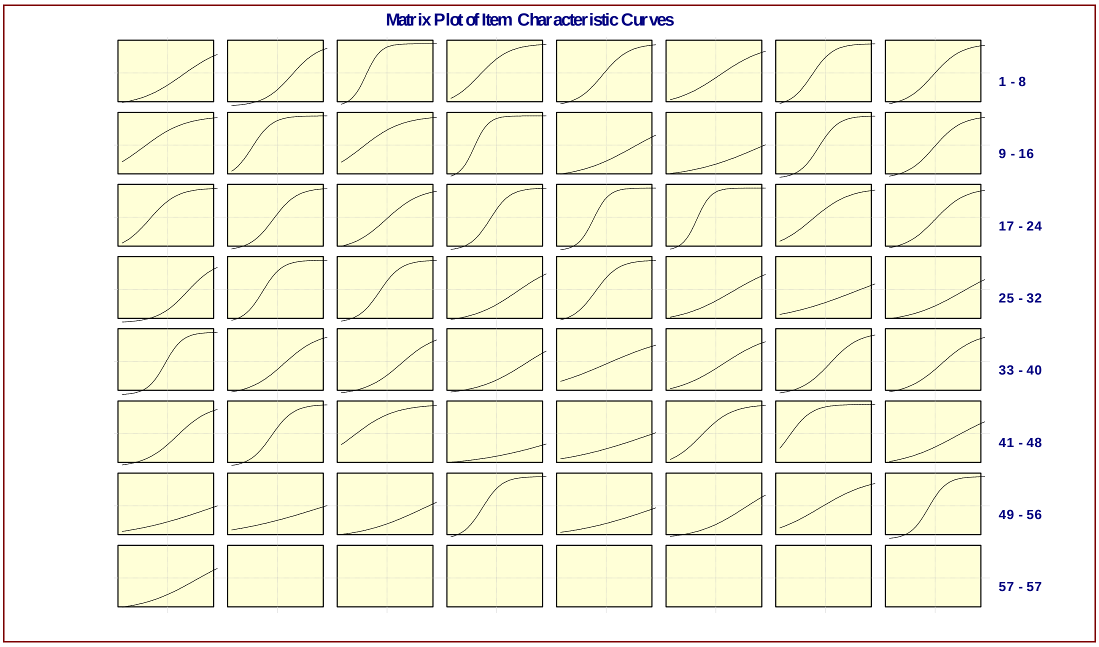
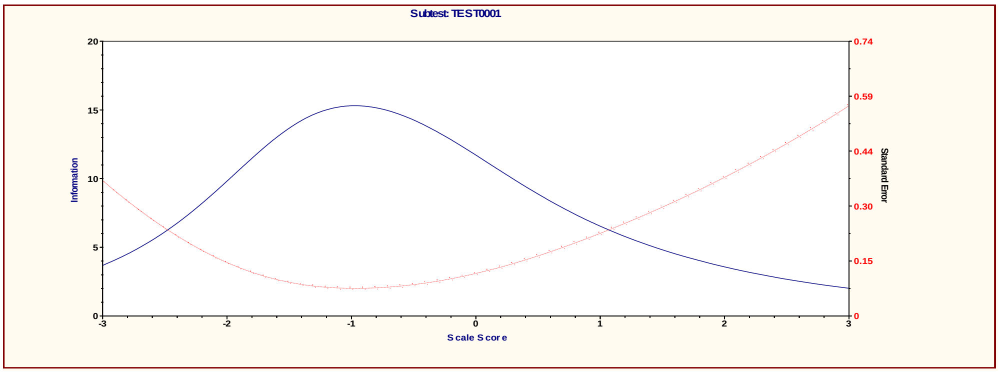


Figure 1 Matrix Plot of Item Characteristics Curves for the Test Items

The shape of the item information function is dependent on the item parameters. The higher the item's discrimination, the more peaked the information function will be, thus, higher discriminations parameters provide more function about individuals whose ability (θ) lie near the item's difficulty value.

The total information curve (TIC) gives the average probability or excepted proportion of the correct as a function of the underlying latent trait. TIC is used as a replacement for the



traditional concept of reliability and standard error of measurement (Samejima, 1977). The standard error of the test is the inverse of the square root of information, thus, the greater information causes the smaller the standard error and the greater the reliability (DeMars, 2010). Based on the results of results of parameter estimation of the test form, the total information curves (TIC) was plotted.

Figure 2 Total Information Curve of Test Form

Figure 2 illustrated that the test form had smaller standard error across the ability scale from -2.5 to +1.0 and larger standard error at the low and high ends of the scale. The maximum amount of information was I (θ) =15 at θ = -1.0. Ability estimates were more precise across the ability scale from -2.5 to +1.0 than at the high and low ends of the scale. Therefore, it may be interpreted that the test form can be suitable to measure for students whose numeracy is θ = -1.0.

Investigation of Phase 3 Output for the Test Form

Regarding the scoring output (Phase 3), raw scores of each student who participated in this study and their ability scores were investigated. The range of ability estimation were ranged from -2.56 (the lowest ability score) to +4 (the highest ability score) and the mean of ability was +0.0216. Moreover, it was observed that the same raw score could not show the same ability because the ability scores of students were different based on their response pattern. Raw scores of students and their ability scores are described in Appendix F. According to the following Table 3.8 and Figure 3, the expected ability distributions of the students were normally distributed across the ability scale.

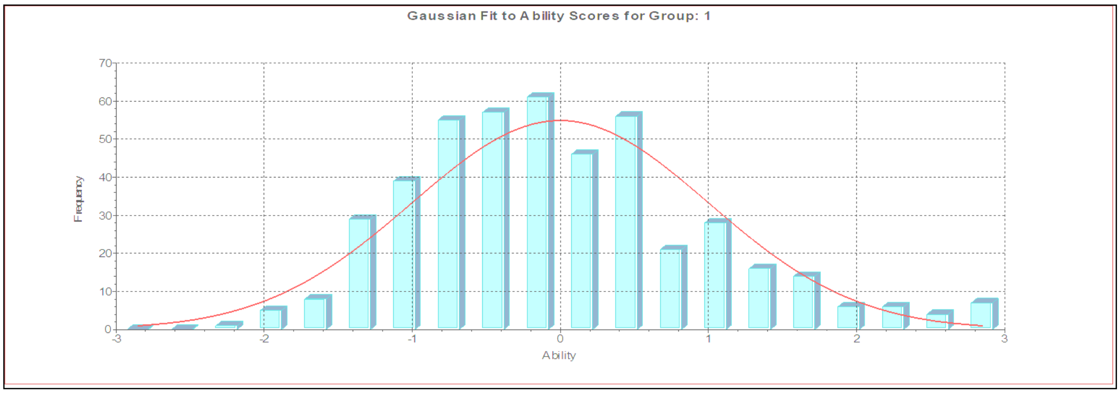


Figure 3 Ability Distributions of the Students

According to the purpose of this study, the numeracy test was constructed by using the two-parameter logistic IRT model. After constructing the test by using the two-parameter logistic IRT model, the difficult items as well as easy items (item 14, 43, 44, 45, 47, 49, 50, 51, and 53) were not used in the test. So, the revised numeracy test remained 48 items (see Table 2).

Table 2 Table of Specification for Revised Numeracy Test of Grade 6 students

<table>
<tr><td rowspan="3">Content</td><td colspan="6">Learning Outcomes</td><td colspan="2" rowspan="2">Total Numbers of Items</td></tr>
<tr><td colspan="2">Knowledge</td><td colspan="2">Comprehension</td><td colspan="2">HOT</td></tr>
<tr><td>No. of Items</td><td>%</td><td>No. of Items</td><td>%</td><td>No. of Items</td><td>%</td><td>No. of Items</td><td>%</td></tr>
</table>

Number and Operation Sense	5	10.42	7	14.58	1	2.08	13	27.08
Algebra	1	2.08	11	22.92	6	12.49	18	37.49
Measurement and Shape	5	10.42	3	6.25	4	8.33	12	25
Statistics	3	6.25	1	2.08	1	2.08	5	10.42
Total	14	27.17	22	45.83	12	25	48	100

Data Analysis and Result

Grade 6 Students' Numeracy Components by Standard Z Score

The standard z-scores for Grade 6 students' numeracy components are shown in Table 3. Numeracy test includes four components such as number and operation sense, algebra, measurement and shape and statistics.

Table 3 Grade 6 Students' Numeracy Components by Standard Z Score

Numeracy Components	No. of Items	Z Score
Number and Operation Sense	13	0.099
Algebra	18	0.065
Measurement and Shape	12	0.075
Statistics	5	0.109

Table 3 showed that the standard score of statistics component of the Grade 6 students was the highest in the four components and that of number and operation sense component was the second highest. The standard score of measurement and shape component of the Grade 6 students was the third highest and that of algebra component was the lowest on the whole numeracy test. Therefore, it can be said that Grade 6 students perform the best in statistics component than other components of numeracy. The standard score of algebra component of the Grade 6 students was found to be the lowest on the entire numeracy test. It can reasonably be concluded that students performed best in statistics component because items from this component are more concrete than other components. Students' performance on number and operation sense component was the second highest among all components. It can reasonably be said that number and operation sense items were more familiar with students since they had exposure since their early childhood mathematics learning. The standard score of measurement and shape component was the second last stand because students cannot thoroughly understand the relationships between different systems of units, identify equivalent period of unit within a system and carry out conversions with units of time, year, money, length, volume and weight. It can reasonably be concluded that students performed lowest in algebra component because the items from this component are more abstract than other components.

Comparison of Grade 6 Students' Numeracy by Gender

Whether there was gender difference, or not, in Grade 6 students' numeracy was worthwhile to explore. It was observed that the mean score of female students was higher than that of male students on the whole numeracy test. The mean scores for each subscale of female students' numeracy were also higher than that of male students.

Table 4 Results of Independent Sample t-test for Grade 6 Students' Numeracy by Gender

Subscales of Numeracy Test	Male	Female	*t*	*p*
Number and Operation Sense	8.27	8.60	-2.139	0.03
Algebra	10.17	10.39	-0.975	0.33
Measurement and Shape	5.20	5.49	-1.84	.048
Statistics	2.51	2.67	-2.11	.034
NUMERACY	26.15	27.14	-2.01	.04

Again, the independent sample t-test was used to examine whether these differences were significant or not. According to table 4, there was significant difference in Grade 6 students' numeracy by gender at 0.05 level. It may be concluded that female students were better than male students in numeracy. Moreover, there was significant difference in Number and Operation Sense by gender at 0.05 level and it can be interpreted that female students were better than male students in Number and Operation Sense. There was also significant gender difference in Measurement and Shape at 0.05 level and it can be interpreted that female students were better than male students in Measurement and Shape. Similarly, significant difference was found to be on Statistics by gender at 0.05 level and it can be interpreted that female students perform better than male students on Statistics (see Table 4). Female students perform better than male students on the whole numeracy test because girls are more concerned with helping their parents in buying groceries, commodities and stationery for their family.

Comparison of Grade 6 Students' Numeracy between Rural Schools and Urban Schools

In order to investigate whether there was significant difference in Grade 6 students' numeracy between rural schools and urban schools, descriptive analysis was done and the differences in mean scores of Grade 6 students' numeracy between rural schools and urban schools was presented. The difference of means can be seen in the following figure.

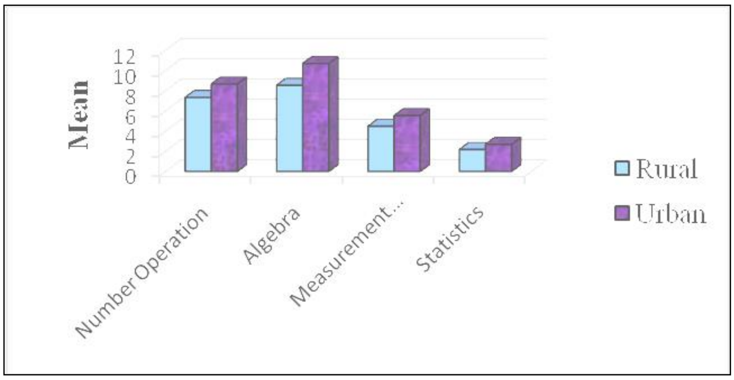


Figure 2 Mean Comparison of Grade 6 Students' Numeracy between Rural Schools and Urban Schools

Concerning the four subscales, statistics subscale was the lowest stand among on the subscales. In addition, algebra subscale was the second last stand among four subscales. Algebra subscale was the highest stand and number and operation sense subscale was the second highest stand. It can be said that Grade 6 students from urban schools perform better than Grade 6 students from rural schools on each subscale of numeracy test as well as on the whole numeracy test.

Again, the independent sample t-test was used to examine whether these differences were significant or not. According to the results of table 3, there was significant difference in Grade 6 students' numeracy between rural schools and urban schools at 0.05 level. It may be concluded that Grade 6 students from urban schools performed better than Grade 6 students from rural schools in the whole numeracy test. Moreover, there was significant difference on Number and Operation Sense between rural schools and urban schools at 0.05 level and it can be interpreted that Grade 6 students from urban schools performed better than Grade 6 students from rural schools in number and operation sense subscale.

There was also significant difference in algebra subscale between rural schools and urban schools at 0.05 level and it can also be interpreted that Grade 6 students from urban schools performed better than Grade 6 students from rural schools in algebra subscale. There was also significant difference in measurement and shape subscale between rural schools and urban schools at 0.05 level and it can also be interpreted that

Grade 6 students from urban schools performed better than Grade 6 students from rural schools in measurement and shape subscale. Similarly, there was also significant difference in statistics subscale between rural schools and urban schools at 0.05 level and it can also be interpreted that Grade 6 students from urban schools performed better than Grade 6 students from rural schools in statistics subscale.

There was also significant difference in the whole numeracy test between rural schools and urban schools at 0.05 level. It can be concluded that Grade 6 students from urban schools performed better than Grade 6 students from rural schools. The students from urban schools perform better than students from rural schools on each subscale of numeracy test as well as on the whole numeracy test. It can reasonably be said that students from urban schools had more opportunities to apply their numeracy skills in day by day experience from their environment than students from rural schools (see Table 5).

Table 5 Results of Independent Sample t-test for Grade 6 Students' Numeracy between Rural Schools and Urban Schools

Subscales of Numeracy Test	**Rural**	**Urban**	*t*	***p***
Number and Operation Sense	7.33	8.66	-5.080	**.000**
Algebra	8.53	10.69	-5.767	**.000**
Measurement and Shape	4.48	5.53	-3.914	**.000**
Statistics	2.18	2.69	-3.911	**.000**
Whole Numeracy Test	22.53	27.56	-6.098	**.000**

Note: NO =Number and Operation Sense, A=Algebra, MS =Measurement and Shape, S =Statistics

Comparison of Grade 6 Students' Numeracy by Socioeconomic Status

In order to test whether Grade 6 students' numeracy were depend on their socioeconomic status or not, checklists were prepared to get some information deal with students' socioeconomic status such as living status, father's occupation, father's educational level, mother's occupation, mother's educational level, family income, number of family members, number of graduated members in family, health care condition, usage of mobile at home, usage of computer at home, usage of internet at home, internet usage time, usage of vehicles at home, usage of electricity, electricity usage time, the number of reading time in library. To get the complete information of students' socioeconomic status, the factors were combined as follows. There are two factors for getting the students' living status, kinds of house and type of housing that is their house is hired or owned.

After assigning the above factors with their respective values, all the factors were combined and the range of values was from 13 to 67 and the mean of values was 36.62. And then the standard deviation of values was 11.2 and P25=28, P50=35, P75=45 and P99=67.Based on the percentile results, the values above P75 is defined as high socioeconomic status, the values between P50 and P75 is defined as middle socioeconomic status and the values below P50 is defined as low socioeconomic status .

Table 6 ANOVA result of Grade 6 Students' Numeracy by SES Level

Sub-scales	SES Level	N	Mean	S.D	F	P
NO TOTAL	Low SES	282	7.55	2.208	47.253	0.000
	SES Middle	131	9.02	2.182		
	SES High	133	9.65	2.263		
	Total	546	8.42	2.396		
A TOTAL	Low SES	282	8.82	3.149	67.853	0.000
	SES Middle	131	11.43	3.026		
	SES High	133	12.30	3.116		
	Total	546	10.29	3.473		
MS TOTAL	Low SES	282	4.55	2.087	36.472	0.000
	SES Middle	131	5.95	2.506		
	SES High	133	6.42	2.459		
	Total	546	5.34	2.431		
S TOTAL	Low SES	282	2.39	1.238	9.910	0.000
	SES Middle	131	2.77	1.042		
	SES High	133	2.88	1.135		
	Total	546	2.60	1.188		
NUMERACY	Low SES	282	23.30	6.515	72.993	0.000
	SES High	133	31.26	7.296		
	Total	546	26.65	7.678		

Note: NO =Number and Operation Sense, A=Algebra, MS =Measurement and Shape, S =Statistics

In order to investigate whether Grade 6 students' numeracy was different by their socioeconomic status, descriptive statistics was done. Based on the result of table 3.4, it was observed that the mean score of the students from high socioeconomic status was highest on each subscale as well as on the whole numeracy test (see Table 6).

To make the confirmation of the significant differences of Grade 6 students' numeracy by their level of socioeconomic status, ANOVA was executed. According to the results of table 6, there was significant difference in Grade 6 students' numeracy across different socioeconomic status at 0.05 level. It can reasonably be concluded that the students from the high socioeconomic status families were the best among the students from other groups (middle socioeconomic status families and low socioeconomic status families) on each subscale as well as on the whole numeracy test. To obtain more detailed information, the Post-Hoc Test carried out by Tukey method (see Table 7).

Table 7 Results of Tukey HSD Multiple Comparisons for Grade 6 Students' Numeracy by SES Level

Subscales	(I) SES Level	(J) SES Level	Mean Difference (I-J)	P
NO TOTAL	SES Middle	Low SES	1.473*	.000
	SES High	Low SES	2.104*	.000
A TOTAL	SES Middle	Low SES	2.605*	.000
	SES High	Low SES	3.478*	.000
MS TOTAL	SES Middle	Low SES	1.400*	.000
	SES High	Low SES	1.875*	.000
S TOTAL	SES Middle	Low SES	.384*	.006
	SES High	Low SES	.493*	.000
NUMERACY	SES Middle	Low SES	5.863*	.000
	SES High	Low SES	7.951*	.000
		SES Middle	2.088*	.035

Note: * The mean difference is significant at 0.05 level.

NO =Number and Operation Sense, A=Algebra, MS =Measurement and Shape, S =Statistics

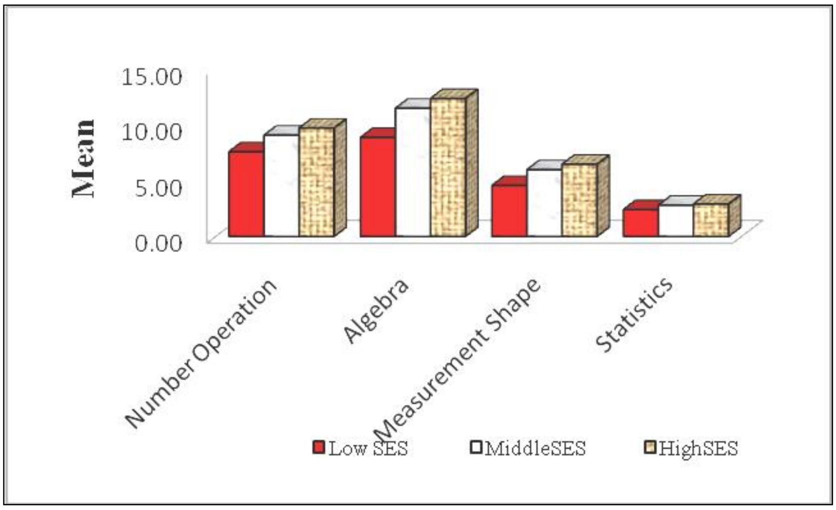


Figure 3 Mean Comparison of Grade 6 Students' Numeracy by SES

Concerning the whole numeracy test, the mean score of students from high socioeconomic status families was significantly higher than that of students from low socioeconomic status families and middle socioeconomic status families. With regard to number and operation sense subscale, the mean scores of students from middle socioeconomic status families and higher were higher than that of students from low socioeconomic status families. In regard to algebra subscale, the mean scores of students from middle socioeconomic status families and high socioeconomic status families were higher than that of students from low socioeconomic status families. In related to measurement and shape scale, the mean scores of students from middle socioeconomic status families and high socioeconomic status families were higher than that of students from low socioeconomic status families. Regarding the statistics subscale, the mean scores of students at middle socioeconomic status families and high socioeconomic status families were higher than that of students at low SES level (see table 7). The students from high socioeconomic status were highest on each subscale as well as on the whole numeracy test than students from middle and low socioeconomic status. It can reasonably be concluded that students from high socioeconomic status get many opportunities to enhance their numeracy because of their rich and conducive living environment for their learning.

Discussion

Since the earliest days, education has been highly regarded in Myanmar. Myanmar regards children as precious gems for a future community. The strong tradition of monastic education has contributed significantly to a high literacy level since the time of the Myanmar kings. Nowadays, education places more emphasis on the formal system with its schools and institutions at primary, secondary and tertiary levels. The Ministry of Education (MOE) and 12 other ministries provide varied and diverse courses for learning in higher education sector, but the MOE is also responsible for the basic education schools for all children.

Nowadays, interest in numeracy has been increasing in education because it can predict education and professional success. However, in Myanmar there was relatively rare awareness of the important of numeracy and there were relatively standardized numeracy test. Therefore, in this research numeracy test was using two-parameter IRT logistic model. Consequently, the numeracy test composed of 48 items was developed.

A number of studies have suggested that numeracy is grounded in number competence (such as recognizing the value of quantities and grasping the principles of counting) (Jordan, Kaplan, Rameni, & Locuniak, 2009), informal number sense (e.g., understanding terms such as "more", "less", "bigger" and "smaller"; knowing that numbers in a counting sequence refer to specific quantities and that higher numbers reflect greater quantities) (Griffin, 2004), and more general factors sometimes characterized as "working memory" (Raghubar, Barnes, & Hecht, 2010). Reid (2008) showed that an informal understanding of quantitative relationships provides the basis for developing formal mathematical knowledge. Numeracy test includes four components such as number and operation sense, algebra, measurement and shape and statistics. Since the numbers of items in four components were not equal, the marks for the items were transformed to standard score. The standard score of statistics component of the Grade 6 students was the highest in the four components and that of number and operation sense component was the second highest. The standard score of measurement and shape component of the Grade 6 students was the third highest and that of algebra component was the lowest on the whole numeracy test. Therefore, it can be said that Grade 6 students perform the best in statistics component than other components of numeracy. The standard score of algebra component of the Grade 6 students was found to be the lowest on the entire numeracy test. It can reasonably be concluded that students performed best in statistics component because items from this component are more concrete than other components. Students' performance on number and operation sense component was the second highest among all components. It can reasonably be said that number and operation sense items were more familiar with students since they had exposure during their early childhood mathematics learning. The standard score of measurement and shape component was the second last stand because students cannot thoroughly understand the relationships between different systems of units identify equivalent period of unit within a system and carry out conversions with units of time, year, money, length, volume and weight. It can reasonably be concluded that students performed lowest in algebra component because the items from this component are more abstract than other components.

Next, significant difference between gender, region and between different levels of socioeconomic status were also found on Grade 6 students numeracy test. There was significant difference in Grade 6 students' numeracy by gender at 0.05 level. It may be concluded that female students were better than male students in numeracy. Moreover, there was significant difference in Number and Operation Sense by gender at 0.05 level and it can be interpreted that female students were better than male students in Number and Operation Sense. There was also significant gender difference in Measurement and Shape at 0.05 level and it can be interpreted that female students were better than male students in Measurement and Shape. Similarly, significant difference was found to be on Statistics by gender at 0.05 level and it can be interpreted that female students perform better than male students on Statistics. Female students performed better than male students on the whole numeracy test because girls are more concerned with helping their parents in buying groceries, commodities and stationery for their home. There was also significant difference in the whole numeracy test between rural schools and urban schools at 0.05 level. It can be concluded that Grade 6 students from urban schools performed better than Grade 6 students from rural schools. The students from urban schools perform better than students from rural schools on each subscale of numeracy test as well as on the whole numeracy test. It can reasonably be said that students from urban schools had more opportunities to apply their numeracy skills in day by day experience from their environment than students from rural schools.

There was significant difference in Grade 6 students' numeracy across different socioeconomic status at 0.05 level. It can reasonably be concluded that the students from the high socioeconomic status families were the best among the students from other groups (middle socioeconomic status families and low socioeconomic status families) on each subscale as well as on the whole numeracy test. The students from high socioeconomic status were highest on each subscale as well as on the whole numeracy test than students from middle and low socioeconomic status. It can reasonably be concluded that students from high socioeconomic status get many opportunities to enhance their numeracy because of their rich and conducive living environment for their learning.

Suggestion for Future Research

This investigation highlights the need for a clearer operational definition of the construct of numeracy as well as additional research into other components of numeracy not included in this study. The limited study area pointed out the necessity to conduct a nationwide study to explore more detailed differences between rural and urban areas. In this study, the sample of students were chosen from Yangon Region, Ayeyarwady Region, and Rakhine State, so further research should be carried out by selecting students from other states and regions so that samples might be more representative.

REFERENCES

Aye Aye Myint. (2000). Development of the numerical reasoning ability test and the vocabulary comprehension ability test for Myanmar high school students. Unpublished Doctoral Dissertation. University of Tokyo: Japan.

Bransford, J.D, et al. (2003). How People Learn: Brain, Mind, Experience, and School (8th ed). Washinton: National Academy Press.

Bruner, J.T. (1993). School for Thought. Cambridge, MA: MIT Press.

Bynner, J. & S. Parsons (1997). Does numeracy matter: Evidence from the National Development Study on the impact of poor numeracy on adult life. London: Basic Skills Agency.

Crocker, L.M. & Algina, J. (1986). Introduction to classical and modern test theory. FL: Holt, Rinehart and Wiston.

Delvin, K. (2000). The four faces of mathematics. In Burke, M.J& Curcio, F.C (Eds), Learning Mathematics for a New Century 2000 Yearbook. Reston VA, USA: National Council of Teacher of Mathematics Incorporated.

Dossey, J.A. (1997). National indicators of quantitative literacy. In Steen, L.A (Ed),Why Numbers Count: Quantitative Literacy for Tomorrow's America. New York: College Entrance Board.

Gal, I. (1993). Issues and Challenges in Adult Literacy. Philadelphia: National Center on Adult Literacy, University of Pensylvania.

Gal, I. (2002a). Systemic needs in adult numeracy education. Adult Basic Education, 12(1), 20-33.

Gal, I., & Schmitt, M. J. (Ed.). (1994). Adult mathematical literacy conference proceedings (NCAL Brief). Philadelphia, PA: National Center on Adult Literacy, University of Pennsylvania. (ERIC Document Reproduction Service No. ED397219).

Ginsburg, L, Manly, M & Schimitt, M.J. (2006). The Components of Numeracy. Cambridge: Harvard Graduate School of Education.

Gleeson, L. (2005). Economic returns to education and training for adults with low numeracy skills. Adelaide.

Hambleton, R. K., Swaminathan, H. (1985). Item Response Theory: Principle and Applications. Boston : Kluwer.

Hambleton, R. K., Swaminathan, H.,& Rogers. H.J. (1991). Fundamentals of Item Response Theory. Newbury park, CA: Sage.

Heymann, H. W. (2003). Why Teach Mathematics: a focus on general education. Dordrecht: Kluwer Academic Publishers.

Jordan, N., Kaplan, D., Rameni, C., & Locuniak, M. (2009). Early math matters: Kindergarten number competence and later mathematics outcomes. Developmental Psychology, 45(3), 850–867.

Kemp, M. (2005). Developing Critical Numeracy at the Tertiary Level. Unpublished doctotral dissertation, University of Murdoch, Western Australia.

Lord, F. M. (1977). Practical applications of item characteristic curve theory. Journal of Educational Measurement, 14, 117-138.

Lord, F. M. (1980). Applications of item response theory to practical testing problems. Hillsdale, NJ: Lawrence Erlbaum.

Lord, F. M., & Stocking, M. L. (1988). Item response theory. In J. P. Keeves, (Eds.) Educational research. methodology, and measurement: An international handbook, pp.269-272. NY: Pergamon Press.

Maguire, T., & O'Donoghue, J. (2002). A grounded approach to practitioner training in Ireland: Some findings from a national survey of practitioners in Adult Basic Education. In L. Ø.Johansen & T. Wedege (Eds.), Numeracy for empowerment and democracy? Proceedings of the 8th International Conference of Adult Learning Mathematics - A Research Forum (ALM8) (pp. 120-132). Roskilde, Denmark, Roskilde University, Centre for Research in Learning Mathematics. Hent, UK: Avanti Books.

Ministry of Education. (1959). 15 to 18: A report of the Central Advisory Committee for Education (England). London: Department of Education and Science.

Parsons, S. & J. Bynner. (2005). Does Numeracy Matter More? London.

Raghubar, K., Barnes, M., & Hecht, S. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. Learning and Individual Differences, 20, 110–122.

Reid, K. (2008). Preschool children's informal understanding of discrete and continuous quantity concepts. (Unpublished doctoral dissertation). University of Melbourne, Melbourne.

Steen, L. A. (1990). Numeracy. Daedalus, 119(2): 211–231.

Su Wai Han. (2011). Basic concept attainment of preschool children: implications for teaching concepts. Unpublished Master Thesis, Yangon Institute of Education, Myanmar.

Talbert, J.E., & Mclaughlin, M.W.(in press). Understanding Teaching in Context. In Cohen, D.K, Mclaughlin, M.W& Talbert, J.E (Eds), Teaching for Understanding: Chllenges for Policy and Practice. San Francisco: Jossey-Bass.

Wright, B.D. (1968). Sample- Free Test Calibration and Person Measurement. Proceedings of the 1967 Invitational Conference on Testing Problems. Pirnceton, NJ: Educational Testing Services.