A Development of the Multiple Intelligences Measurement of Elementary Students

Chaiwat Waree

Abstract — This research aims at development of the Multiple Intelligences Measurement of Elementary Students. The structural accuracy test and normality establishment are based on the Multiple Intelligences Theory of Gardner. This theory consists of eight aspects namely linguistics, logic and mathematics, visual-spatial relations, body and movement, music, human relations, self-realization/self-understanding and nature. The sample used in this research consists of elementary school students (aged between 5-11 years). The size of the sample group was determined by Yamane Table. The group has 2,504 students. Multistage Sampling was used. Basic statistical analysis and construct validity testing were done using confirmatory factor analysis. The research can be summarized as follows; 1. Multiple Intelligences Measurement consisting of 120 items is content-accurate. Internal consistent reliability according to the method of Kuder-Richardson of the whole Multiple Intelligences Measurement equals .91. The difficulty of the measurement test is between .39-.83. Discrimination is between .21-.85. 2). The Multiple Intelligences Measurement has construct validity in a good range, that is 8 components and all 120 test items have statistical significance level at .01. Chi-square value equals 4357.7; p=.00 at the degree of freedom of 244 and Goodness of Fit Index equals 1.00. Adjusted Goodness of Fit Index equals .92. Comparative Fit Index (CFI) equals .68. Root Mean Squared Residual (RMR) equals 0.064 and Root Mean Square Error of Approximation equals 0.82. 3). The normality of the Multiple Intelligences Measurement is categorized into 3 levels. Those with high intelligence are those with percentiles of more than 78. Those with moderate/medium intelligence are those with percentiles between 24 and 77.9. Those with low intelligence are those with percentiles from 23.9 downwards.

Keywords — Multiple Intelligences, Measurement, Elementary Students.

I. INTRODUCTION

NARINEE Reungnoo have done a long-term research on intelligence of students aged between 6 and 12 years. It was found that IQ average was 91.2. This shows that the intelligence of Thai students by average is at a lower level than normal. Students with IQ of more than 110 or geniuses were less than 10%. This may be because parents do not know how to improve their children’s brain potentials. Teaching that does not emphasize in various intelligence developments may also be one of the reasons [10].

Measuring learners’ intelligences will help those involved in the education system to apply results from the measurement to design consistent and effective curriculum and learning experience for learners according to learning standards comprising of eight subject areas. These factors will result in different abilities to learn and work. Having good teaching and learning managements positively affect the intelligences of learners. Learners can have their missing intelligences filled and improved. Intelligence Tests are usually based on various intelligence theories for instance Multiple Factor Theory of Thurstone, Wechsler Intelligence Scale, Standford-Binet L-W Form Test and Kaufman Assessment Battery for Children: K-ABC. These tests cannot comprehensively measure intelligences as they only measure aspects of linguistics, logic, mathematics and some parts of visual-spatial relations. They are also culturally biased [12].

There have been a few researches on intelligence test development in Thailand whereas the research on development of aptitude testing is more popular. This type of test is usually based on Multiple Factor Theory of Thurstone and Guildford’s Theory of Intelligence Structure. These tests only measure sub-components of intelligence in the aspects of linguistics, visual-spatial relations and mathematics [11], for instance, Mathematics Aptitude Test [6], Mathematics Test [3], Intelligence Test on Logic and Mathematics [13], and Aptitude Tests on linguistics, arithmetic and reasons [1]. Hence, they are not enough to tell that a child is clever or the extent of a child’s capability. There are other aspect of cleverness or intelligence that current tests cannot measure for instance music, sports and arts [15]. Hence, in order to comprehensively measure intelligence, multiple intelligences measurement is to be developed.

Theory of multiple intelligences (MI) by Gardner is based on a basic belief that each human being has a different way of learning [2]. Each person has different strengths and weaknesses. Intelligence, according to this theory, can be categorized into eight aspects namely: 1.) Linguistic intelligence which means the ability to use language proficiently and the ability to communicate through reading, writing, listening or speaking; 2.) Logical-mathematical intelligence which refers to the ability to understand the principle of reasons, calculations and mathematical operations; 3.) Visual-spatial intelligence which means the ability to correctly perceive visually; 4.) Bodily-kinesthetic intelligence which means the ability to use bodily organs such as hands and feet in relation to things which the person wishes to express in order to generate work or solve a problem; 5.) Musical intelligence which means the ability in music, being sensitive to sound, being absorbed in rhythm, level of sound, melody and differences of sound; 6.) Interpersonal intelligence which means the ability in understanding the emotions of others; 7.) Interpersonal intelligence which means the ability in understanding the emotions of others; 8.) Naturalist intelligence which means the ability to understand and solve problems concerning the natural environment.
intelligence refers to the ability to understand and relate to other people, ability to adjust oneself appropriately to society; 7.) Intrapersonal intelligence which means the ability to understand oneself, being able to control appropriate expressions according to time and occasions and 8.) Naturalist intelligence which means the ability to deeply realize and understand nature, being all-rounded in science and the environment.

There are various tools to measure multiple intelligences of learners for instance two-way response record, behavior observing record, Likert Scale, PMI Chart, Open-ended and guided questionnaires and tests [7]. The popular means is testing as it is an easy and convenient way to realize and understand the intelligence levels of learners allowing one to accurately understand the ability of learners. The data received from using a quality test with high validity and reliability is more meaningful than data received from other means apart from testing. Results can be used to predict the success of learners, analyze problems, categories learners and allow parents or those involved such as teachers, guiding teachers and school executives to accurately understand the level of ability and success of learners [12]. Other tools of measurement of multiple intelligences are limited by factors such as time, labor, budget and prejudice. They are suitable for measurement and assessment in actual classes with multiple intelligences with multiple measurements and assessments over the term or the academic year [20].

From the significance mentioned above, the researcher argues that development of multiple intelligences measurement qualified for standard and appropriate to the context of Thai society will be beneficial to the education reform which aims to develop human resources according to their potentials. Learners will be developed to become intelligent, moral and with lasting content.

II. RESEARCH OBJECTIVES  
- Develop the multiple intelligences measurement for elementary students
- Test construct validity
- Establish normality

III. SCOPE OF THE STUDY  
- The population used in this research consists of 5,762,700 elementary students (aged between 5-11 years). Data has been retrieved from the National Statistical Office.
- Sample group used in this research consists of elementary students (aged between 5-11 years). The sample size was determined using Yamane Table. The number of students in the group was 2,504. Multistage sampling was used; 1.) Cluster Sampling from 4 regions. Two provinces were selected from each region. One educational service area was selected from each province. Hence there were 8 educational service areas in total. 2.) Stratified Sampling – each educational service area was divided according to school size. Three schools were selected from each school size. Each educational service area was to have 3 schools per size. Hence there were 72 schools in total. Of these, there were 1,004 male students and 1,500 female students. The total number of those who answered the questionnaires was 2,504. Criteria for school size are as follows; Small- The number of students is between 1 and 120; Medium- The number of students is between 121 and 600 Large- The number of students is from 601.

- The multiple intelligences measurement according to Gardner developed by the researcher consists of 8 aspects of intelligence namely; Linguistic Intelligence, Logical and Mathematical Intelligence, Visual-spatial Intelligence, Bodily-kinesthetic Intelligence, Musical Intelligence, Interpersonal Intelligence, Intrapersonal Intelligence, Naturalist Intelligence

IV. CONCEPTUAL FRAMEWORK  
This research is the development of multiple intelligences measurement of elementary students. The researcher uses the basis of Gardner’s multiple intelligences theory which categorizes intelligence into 8 aspects namely; linguistic, logical and mathematical, visual-spatial, bodily-kinesthetic, musical, interpersonal, intrapersonal and naturalist intelligences [2]. In order to develop the test to be able to measure each aspect of intelligence accurately, the researcher has produced operational definition for each aspect of intelligence to be used as guidelines in generating the test and as a check whether the test is able to measure properties accurately according to multiple intelligences theory of Gardner. Construct validity and normality establishment are as shown in Fig. 1;

![Fig. 1 Conceptual framework](image-url)

V. RESEARCH PROCESS/METHODOLOGY  
For this research, the researcher has used research tools to collect data from the sample group. Details are as follows;
- Publicized online through the website to ask for the collection of data. The sample size consists of a group of students who voluntarily accepted and registered on the website.
- The collection of data was done with students who formed the sample group between 4 June and 31 August
B.E.2555 (2012). The number of students in the sample group was 2504.

- Marked and recorded scores.

VI. SUMMARY OF THE RESEARCH
The research results can be summarized as follows:
Quality findings of 120 items in the Multiple Intelligences Measurement for elementary students are as follows; 1). Content validity of the Multiple Intelligences Measurement- there is content validity with index of item objective congruency (I.O.C.) between the test and operative definitions between 0.67 and 1.00. 2). Difficulty for each item in the Multiple Intelligences Measurement is between 0.39 and 0.83. It can be categorized according to components as follows; linguistic intelligence has the difficulty between 0.57 and 0.83; logical and mathematical intelligence has the difficulty between 0.52 and 0.61; visual-spatial intelligence has the difficulty between 0.61 and 0.73; bodily-kinesthetic intelligence has the difficulty between 0.54 and 0.66; musical intelligence has the difficulty between 0.39 and 0.56; interpersonal intelligence has the difficulty between 0.64 and 0.78; intrapersonal intelligence has the difficulty between 0.53 and 0.62 and naturalist intelligence has the difficulty between 0.55 and 0.61. 3). Itemized discrimination value of the Multiple Intelligences Measurement is between 0.21 and 0.85. It can be categorized according to components as follows; linguistic intelligence has the discrimination value between 0.21 and 0.85; logical and mathematical intelligence has the discrimination value between 0.45 and 0.85; visual-spatial intelligence has the discrimination value between 0.47 and 0.81; bodily-kinesthetic intelligence has the discrimination value between 0.51 and 0.78; musical intelligence has the discrimination value between 0.23 and 0.65; interpersonal intelligence has the discrimination value between 0.53 and 0.72; intrapersonal intelligence has the discrimination value between 0.39 and 0.71 and naturalist intelligence has the discrimination value between 0.22 and 0.72. 4). Reliability of the Multiple Intelligences Measurement- the overall internal consistency according to the method by Kuder-Richardson formula KR20 for the whole measurement is 0.91.

Construct Validity of the Multiple Intelligences Measurement for elementary students.
- Confirmatory factor analysis in order to test for construct validity of the Multiple Intelligences Measurement can be categorized into 8 components which are summarized as follows; 1). The result of the test for construct validity of the Multiple Intelligences Measurement in terms of linguistic intelligence with 15 items is that every item is statistically significant at the level 0.01 with factor weight of between 0.20 and 0.94. 2). The result of the test for construct validity of the Multiple Intelligences Measurement in terms of logical and mathematical intelligence with 15 items is that every item is statistically significant at the level 0.01 with factor weight of between 0.21 and 0.86. 3). The result of the test for construct validity of the Multiple Intelligences Measurement in terms of visual-spatial intelligence with 15 items is that every item is statistically significant at the level 0.01 with factor weight of between 0.19 and 0.60. 4). The result of the test for construct validity of the Multiple Intelligences Measurement in terms of bodily-kinesthetic intelligence with 15 items is that every item is statistically significant at the level 0.01 with factor weight of between 0.27 and 0.44. 5). The result of the test for construct validity of the Multiple Intelligences Measurement in terms of musical intelligence with 15 items is that every item is statistically significant at the level 0.01 with factor weight of between 0.29 and 0.39. 6). The result of the test for construct validity of the Multiple Intelligences Measurement in terms of interpersonal intelligence with 15 items is that every item is statistically significant at the level 0.01 with factor weight of between 0.25 and 0.41.
- Second confirmatory factor analysis in order to test construct validity of the Multiple Intelligences Measurement in 8 components with 120 items overall has been found that factor weight for each component is statistically significant at the level of 0.01. Factor weight value is between 0.18 and 0.67. This means that the overall Multiple Intelligences Measurement has construct validity in a good level. Chi-square value equals to 4357.7; p = 0.00 at the degree of freedom of 244. Goodness of Fit Index (GFI) equals to 1.00. Adjusted Goodness of Fit Index (AGFI) equals 0.92. Comparative Fit Index equals 0.68. Root Mean Squared Residual (RMR) equals 0.064 and Root Mean Square Error of Approximation equals 0.82.

Normality of the Multiple Intelligences Measurement for elementary students- the researcher divides the criteria used to consider the level of intelligence into 3 levels; high, medium and low. Those with high intelligence are those from percentile 78 onwards and they have raw score between 68 and 99. Those with medium intelligence are those in percentile between 24 and 77.9 and they have raw score between 43 and 67. Those with low intelligence are those in percentile between 0 and 23.9 and they have raw score between 0 and 42. These can be categorized according to components as follows; 1). High linguistic intelligent students- those with high intelligence have raw scores between 11 and 15. Those with medium intelligence have raw scores between 6 and 10. Those with low intelligence have raw scores between 0 and 5. 2). High logical and mathematical...
intelligent students—those with high intelligence have raw scores between 11 and 15. Those with medium intelligence have raw scores between 6 and 10. Those with low intelligence have raw scores between 0 and 5. 3). High visual-spatial intelligent students—those with high intelligence have raw scores between 12 and 15. Those with medium intelligence have raw scores between 5 and 11. Those with low intelligence have raw scores between 0 and 4. 4). High bodily-kinesthetic intelligent students—those with high intelligence have raw scores between 10 and 14. Those with medium intelligence have raw scores between 6 and 9. Those with low intelligence have raw scores between 0 and 5. 5). High musical intelligent students—those with high intelligence have raw scores between 11 and 15. Those with medium intelligence have raw scores between 6 and 10. Those with low intelligence have raw scores between 0 and 5. 6). High interpersonal intelligent students—those with high intelligence have raw scores between 11 and 15. Those with medium intelligence have raw scores between 6 and 10. Those with low intelligence have raw scores between 0 and 5. 7). High intrapersonal intelligent students—those with high intelligence have raw scores between 12 and 15. Those with medium intelligence have raw scores between 5 and 11. Those with low intelligence have raw scores between 0 and 4. 8). High naturalist intelligent students—those with high intelligence have raw scores between 10 and 14. Those with medium intelligence have raw scores between 6 and 9. Those with low intelligence have raw scores between 0 and 5.

is more than 0.50 for every item. Itemized difficulty for the whole test/measurement is between 0.39 and 0.83. Itemized discrimination for the whole test/measurement is between 0.21 and 0.85. The reliability of the whole test/measurement is 0.91. Discussions are as follows; 1). Result of content validity analysis taking into account the Index of Item Objective Congruence between the test and the operative definition is that every item has Index of Item Objective Congruence between that item and operative definition of more than 0.50. That means the researcher has produced the test in line with the operative definition. Thus this Multiple Intelligences Measurement has content that conforms to the Multiple Intelligences Theory of Gardner. 2). In terms of itemized difficulty analysis, it has been found that itemized difficulty for the whole test/measurement is between 0.39 and 0.83. According to the general criteria on difficulty, good test should have the difficulty level between 0.21 and 0.85 to be considered appropriate. If the difficulty level reaches 0.95, it will mean that the test is very easy. If the difficulty level falls between 0.81 and 0.95, it will mean the test is easy. If the difficulty level falls between 0.20 and 0.80, it will mean that the test is appropriately difficult. If the difficulty level is less than or equal to 0.19, it will mean that the test is hard or too hard. The whole test/measurement should have the average difficulty level of 0.50 [17]. 3). In terms of itemized discrimination of the Multiple Intelligences Measurement, it has been found that the Multiple Intelligences Measurement produced by the researcher has itemized discrimination value of between 0.21 and 0.85. In general, good test should have positive discrimination value and should be more than 0.20. If the discrimination value is from 0.40 onwards, it will be considered a test with very good discrimination. If the discrimination value is between 0.30 and 0.39, 0.20 and 0.29 and below 0.20, it will be considered a test with good, satisfactory and no discrimination respectively. The test used should be a test with discrimination of 0.20 onwards [17]. With regards to the Multiple Intelligences Measurement, it has been found that all items have discrimination value of at least 0.20 each. Hence the test is able to discriminate the intelligence level of elementary students according to Gardner’s Multiple Intelligences Theory. Hence the Multiple Intelligences Measurement is a quality test and is also a test with concurrent validity [16]. 4). In terms of the overall reliability of the Multiple Intelligences Measurement, it has been found that the overall reliability is 0.91. This means that a test has very good reliability. In general, a norm-referenced test of good quality should have the reliability value of at least 0.80 [14]. Generally, a test with validity coefficient of at least 0.71 is regarded as a test with high validity. If the coefficient falls between 0.41 and 0.70, the test will be considered as a test with medium validity. If the coefficient is less than or equals to 0.40, the test will be considered as a test with low validity. Popham mentions criteria in assessing validity coefficients of the tests [19]. He states that if the coefficient is between 0.85 and 0.95, the test will be
considered with a high validity. Moreover, Sucheea Patrayuttawat refers to various criteria in assessing validity coefficients of psychological tests [9]. In 1991, Murphy and Davidshorfer specified that if the coefficients are 0.95 and 0.90, the test will be considered with high and medium validity respectively [5]. In 1994, Williams specified criteria for assessing validity coefficient and stated that tests with coefficients 0.80, 0.70 and 0.60 will be considered as good, acceptable and just acceptable [18]. Fisher and Corcoran have stated that tests with coefficients 0.00-0.20, 0.21-0.40, 0.41-0.70 and 0.71-1.00 will be considered with very low, rather low, average and high validities respectively [4]. Hence, it can be said that the whole Multiple Intelligences Measurement produced by the researcher is one with high validity. It is a quality tool that gives consistent and certain results under the same circumstance and conditions.

- Construct validity of the Multiple Intelligences Measurement for elementary students- according to the confirmatory factor analysis done to check the construct validity of the Multiple Intelligences Measurement that the researcher has produced from the Multiple Intelligences Theory of Gardner consisting of 8 aspects namely: linguistic; logical and mathematical; visual-spatial; bodily-kinesthetic; musical; interpersonal; intrapersonal and naturalist aspects, the researcher has produced 160 testing items. 150 items are with content validity. After a trial test, there remains 120 quality testing items which can be categorized into 15 items in linguistic aspect; 15 items in logical and mathematical aspect; 15 items in visual-spatial aspect; 15 items in bodily-kinesthetic aspect; 15 items in musical aspect; 15 items in interpersonal aspect; 15 items in intrapersonal aspect and 15 items in naturalist aspect. After confirmatory factor analysis in order to check for construct validity, it has been found that all 120 test items have statistical significance which gives rise to the following points of discussion: The results from the confirmatory factor analysis in order to check for construct validity of 8 aspects of the Multiple Intelligences Measurement totaling 120 testing items have shown that factor weight of every item has statistical significance. Factor weight of all 8 factors according to the Multiple Intelligences Measurement of Gardner is at 0.01. Statistical values used to assess the consistency between the model and empirical evidence are namely: factor weight which lies between 0.18 and 0.67 implying that the whole Multiple Intelligences Measurement has a very good level of construct validity; Chi-square value which equals 4357.7; p<0.00 at the degree of freedom of 244; Goodness of Fit Index (GFI) which equals to 1.00; Adjusted Goodness of Fit Index (AGFI) which equals 0.92; Comparative Fit Index which equals 0.68; Root Mean Squared Residual (RMR) which equals 0.064 and Root Mean Square Error of Approximation (RMSEA) which equals 0.82. In this regard, Chatsiri Piyapimonsit has explained the criteria in assessing the consistency between the model and empirical evidence [8]. Chi-square must not have any statistical significance. Goodness of Fit Index (GFI) and Adjusted Goodness of Fit Index (AGFI) must be over 0.90. Comparative Fit Index (CFI) must be over 0.95. Root Mean Squared Residual (RMR) must be below 0.08 and the Root Mean Square Error of Approximation (RMSEA) must be below 0.06. Back to the statistical results from the confirmatory factor analysis of the researcher, it has been found that statistical values used to assess the consistency between the model and empirical evidence are in line with the criteria. This shows that the model of results of the confirmatory factor analysis of the researcher is consistent with the empirical evidence. That is the Multiple Intelligences Measurement for elementary students aiming to assess 8 aspects of intelligence is able to accurately measure according to Gardner’s Multiple Intelligences Theory. This shows that this Multiple Intelligences Measurement has construct validity. It is the tool that can measure 8 Multiple Intelligence aspects of elementary students accurately and comprehensively. These result in reliability and confidence when translating the meanings of the test.

- Normality of the Multiple Intelligences Measurement- This Multiple Intelligences Measurement for elementary students is a norm-referenced measurement. The researcher has specified the standard score by creating normality of the Multiple Intelligences Measurement from 2,504 elementary students. This normality can be used to compare intelligence levels of students by taking into account raw scores and compare them against percentile locations. The researcher has categorized intelligence into 3 levels namely high, medium and low intelligences. The normality results from norm referenced measurement which states that the scores from the test will be meaningful when scores of those doing the test have been adjusted into any one form of standard scores and then measured against the group [9]. In other words, when the school asks students to take the Multiple Intelligences Measurement then compare raw scores with the normality generated by the researcher, the school will know levels of intelligence of students. The normality of this Multiple Intelligences Measurement is considered appropriate to use because it is a proxy of group and is modern. Being a proxy refers to a proxy of a sample group used to generate the normality. The sample group comes from the targeted population and members were selected at random. The targeted population is elementary students. The sample group used to generate normality consists of 2,504 elementary students which had been through various stages of random selection. Hence, it can be said that the normality of this Multiple Intelligences Measurement is a proxy of the group. With regards to modernity, it can be said that this measurement aims to assess various aspects of intelligence of students. When the measurement is used to test students and the results are compared with the level of intelligence, students’ levels of intelligence will be known allowing teachers to adjust the learning process to better suit learners. This will encourage learners to develop naturally with their full potentials which in turn leading to happy learners. These
conform to the National Education Act B.E.2542 (1999) and the amended version (2nd Version) B.E.2545 (2002) which aims to develop learners so that they are complete in physical, mental, intelligent, knowledge and moral aspects. Learners should lead their lives with morality and they should develop a culture to live with other people happily. These are consistent with the core curriculum of basic education B.E.2551 (2008) which aims at having learners who are moral, intelligent and happy with potentials to further their education and work. Due to these reasons, the normality of the Multiple Intelligences Measurement generated by the researcher is thus modern.

VIII. RECOMMENDATIONS FOR FURTHER RESEARCH

- Generate other tools to measure and assess Multiple Intelligences. These include Likert Scale, Behavior Observing Form and PMI Chart for example.
- Study and compare levels of intelligence of students in various provinces or regions by comparing between different genders, economic and social statuses.
- Generate normality of the measurement by region and by country.
- Generate normality of the measurement from the equation of scores of 8 component aspects of Multiple Intelligences retrieved from confirmatory factor analysis.
- Test predictive validity of the measurement and use results to predict students' future.
- Test construct validity of the questionnaire by analyzing all components of the measurement simultaneously.

REFERENCES