

主論文の要約

論文題目 Developing Multiple-Task Spatial Ability Tests for Myanmar Middle School Students Using Item Response Theory

(項目応答理論を用いたミャンマーの中学生のための複数課題空間能力テストの開発)

氏 名 Nu Nu Khaing

This dissertation presents a test construction of new multiple-task spatial ability tests. The purposes of the study are to develop new unique spatial ability tests in order to properly measure the spatial abilities of the Myanmar middle school students, and apply the tests into investigation of the spatial abilities of the students by ethnicity, gender and age factor.

Spatial ability involves an integrated use of visual spatial capacities to solve problems and it expands creative thinking. Moreover, spatial ability can predict educational and professional success. Therefore, interest in spatial ability has been increasing in education. However, until recent years, Myanmar, one of developing countries, had not yet become widely aware of the importance of spatial ability in education. In addition, when it is considered conducting a particular spatial research in Myanmar, the main issue is which spatial ability tests are appropriate for Myanmar students. Although a large number of spatial ability tests are known to exist, cultural factors play a big part in explaining differences in spatial abilities and spatial ability tests. For these reasons, this study considered developing new unique spatial ability tests for Myanmar students.

Actually, spatial ability itself is not a unitary construct. It includes a major factor which has certain sub-factors and each emphasizes different aspects of spatial ability. On the other hand, it can be observed that some researchers apply a single-task spatial ability test, some apply two or more single-task spatial ability tests simultaneously, and some utilize a multiple-task spatial ability test. Here if each spatial ability task can measure the major spatial factor to some extent, a multiple-task test can measure more aspects of spatial ability than a single-task test. Therefore, the aim of this study became to develop new multiple-task spatial ability tests.

Using different tests and comparing their results cause sample-dependent and test-dependent problem when the test construction procedure is undertaken by utilizing only the item analysis of the classical test theory (CTT). Item response theory (IRT) overcomes the shortcoming of CTT by providing information on how examinees at different ability levels

have performed on an item. Thus, in order to develop the intended tests systematically, an item analysis method of item response theory (IRT) model should be applied.

For the above reasons, this dissertation focuses a test construction procedure of new multiple-task spatial ability tests by item response theory (IRT). There are four steps: (1) developing a preliminary test in which items of multiple spatial tasks were included, (2) developing two equivalent spatial ability tests/forms that were composed of four spatial tasks by linking common items, (3) identifying potentially biased items in the two spatial ability tests by using three DIF analysis methods in order to ensure that no students will be unfairly panelized when taking the tests, (4) applying the two equivalent tests into investigation of spatial ability of Myanmar middle school students by ethnicity, gender and age factor.

Chapter 1 reviews the previous research on spatial ability and the importance of spatial ability. Then, reasons why unique spatial ability tests are necessary for Myanmar middle school students, issues related to definition of spatial ability, issues related to types of spatial ability tests, gender and age differences in spatial ability, ethnicity and spatial ability, purposes of the study, and organization of the dissertation are presented.

Chapter 2 includes an overview of statistical models and analysis methods used in this research. It is covered test-dependent problem and sample-dependent problem of classical test theory (CTT), concepts of item response theory (IRT) models, equating methods, and necessity of differential item functioning (DIF) analysis. Among many DIF approaches them, concerning Lord's Chi-Square Method, Logistic Regression Method, and Mantel-Haenszel (MH) Statistic are explained because the three methods were used in this research.

Chapter 3 presents a test development process of multiple-task spatial tests in this research. There are two studies. Study-1 concerns the preliminary study of constructing a multiple-task spatial test. It presents the way of how to draw item pictures for various types of spatial tasks and construct an item pool, and how to analyze systematically and statistically items, and test administration. From the preliminary study, a multiple-task test with 31 items could be developed, but some information and some limitations were observed. Study-2 presents how to develop two equivalent tests based on the result of the preliminary study. In this study, the two equivalent tests included four spatial tasks. The two tests were constructed by the IRT common item equating design. A factor analysis was conducted to confirm the major spatial factor and sub-factors of the tests. After that, a validity study was also performed.

Chapter 4 presents a study of DIF analysis into spatial ability tests in order to develop the spatial ability tests that better captures the spatial ability of diverse students when taking the tests. In this study, the DIF items were identified across gender and ethnic

groups by using three famous DIF methods; Lord's Chi-square method, Logistic Regression method, and Mantel-Haenszel method. Ethnic DIF items were removed from the two tests because these items were likely to be differential functioning due to the cultural and environmental differences among the ethnic groups. Regarding gender DIF items, no items were removed from the tests in this step because the gender based DIF items of the two tests might be impact items that could measure more accurately gender differences in spatial ability. After discarding DIF items, the two tests were revised and reconstructed without DIF items based on the result of the DIF analysis. 32 items remain in the Test A and 33 items in the Test B.

Chapter 5 mentions the study of investigation of spatial ability of Myanmar middle school students across gender, age and ethnic groups by applying the two equivalent spatial ability tests. In this study, two student groups were administered the two tests; Group A was administered Test A and Group-B was administered Test B. The spatial abilities of the students were investigated by ethnic group comparison, the gender group comparison, and age group comparison according to their ability scores. The finding showed that the cultural or ethnical differences affected the spatial ability, males performed better in spatial tasks than females, but gender difference emergence is highly dependent on the type of the spatial test, except Paper Folding task that measures spatial visualization. Then, it was found that spatial abilities were likely to develop from the level of the younger student group to that of the elder student group, especially females obviously develop from 13 to 15, males from 13 to 14.

Chapter 6 begins with a summary of the findings of this study. Following this summary, significances of the study, overall discussion and recommendations for future research, and conclusion are presented. Finally, it was concluded that by doing this research two new unique spatial ability tests for Myanmar middle school students were developed as two tests. These tests were four-task spatial ability tests which can measure more aspects of spatial ability than single-task spatial ability tests. Test items were analyzed systematically by unidimensional IRT analysis. Because of conducting the common item parameter estimation method, they can be used on the same ability scale. In addition, they had moderately good construct validity. As they have been revised without DIF items, the items can be better captures the spatial ability of diverse students of Myanmar middle school students.