AN ANALYSIS OF THE DIMENSIONALITY OF THE SPM ADDITIONAL MATHEMATICS OPEN-ENDED ITEMS USING THE MULTIDIMENSIONALITY MODEL FOR DIF FRAMEWORK

LEONG CHEE KIN

UNIVERSITI SAINS MALAYSIA

2007

AN ANALYSIS OF THE DIMENSIONALITY OF THE SPM ADDITIONAL MATHEMATICS OPEN-ENDED ITEMS USING THE MULTIDIMENSIONALITY MODEL FOR DIF FRAMEWORK

by

LEONG CHEE KIN

Thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

MARCH 2007

ACKNOWLEDGEMENTS

The completion of this thesis was accomplished through the assistance of many significant people. The important and necessary support and guidance from the beginning to the end of this thesis that my supervisor Associate Prof. Dr. Mokthar Ismail and my co-supervisor, Dr. Ong Saw Lan are greatly appreciated. I would like to express my deepest gratitude and thanks to them. Their patience and contribution were critical in the completion of this thesis. Their professional expertise and many insightful and constructive comments contributed tremendously to the success of this endeavour in focusing my thoughts and ideas. No amount of words can express how touched and indebted I am by their commitment, kindness, and dedication in helping, encouraging, motivating and inspiring me. It was a great privilege to be under their supervision and guidance for they were ever so willing to give me generously their time and attention.

I would also like to sincerely thank Assoc. Prof. Dr. Yoong Suan for providing the DIFPACK software, specifically the Poly-SIBTEST which is a component bundle together in DIFPACK and relevant documentation. I am especially appreciative of his constructive suggestions and advice during the Research Proposal Seminar. Also, I am very grateful to all the members of the Research Proposal Seminar Committee for their constructive suggestions to enhance the quality of my research.

I am also grateful to the Malaysian Government, The Public Service Department Malaysia, The Ministry of Education Malaysia and every Malaysian taxpayer for directly and indirectly providing me the financial support to study for this degree.

I also deeply appreciate the invaluable assistance and cooperation rendered by the State Education Department of Penang and principals of schools that have participated in this study. My sincere thanks and appreciation to students who have responded to the study's instruments and furnished me the vital data and information for my study.

I am also indebted to Prof. Kenneth Tobin of City University of New York whom I greatly respect for his invaluable assistance through e-mails and consent to using his TOLT in my research. Also, related to TOLT, I would like to express my thank to Mr. Lim Bok Seng for his invaluable help, advice and suggestions in translation work.

I am very grateful to all my friends who have contributed to the success of this work. My heartfelt thanks to Mr. Tan Jen Jong who assisted me by proofreading and editing my thesis and Mr. Lee Lat Heng who assisted me supervising and administering the instruments to the students during the data collection period. Also, to Mr. Manogaran a/l Subramanian, thank you for your efforts in motivating and stimulating my intellect that expanded my mind, and for encouraging me to push myself further.

I am especially appreciative of the language and mathematics experts who gave generously their precious time in serving as raters who provided their expertise in their review as well as constructive comments at various stages of my study. Without their assistance, this study would not have been possible. They were Mr. Jasbir Singh, Mr. Khor Inn Seng, Ms. Lim Phaik Lean, Mr. Tan Jen Jong, and Ms. Yau Yik Yin.

In conclusion, I want to direct my thanks to family and friends who have supported and encouraged me through the trials, tribulations, and distractions of this process. In particular, I want to express my gratitude to Mariam, who has supported my work in countless ways. She is without question one of the primary reasons I was able to complete this thesis and she knew all along that I would finish even when I was not sure. Finally, I am especially grateful to my late parents for the values taught and I hope I do make them proud.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

SATU ANALYSIS DIMENSI ITEM-ITEM HUJUNG TERBUKA MATEMATIK TAMBAHAN SPM BERDASARKAN MODEL MULTIDIMENSI UNTUK RANGKA DIF

ABSTRAK

Dimensionaliti ujian merupakan satu isu penting yang berkaitan dengan kesahan dan kebolehpercayaan ujian. Dimensionaliti mempengaruhi anggaran kebolehpercayaan ujian, tafsiran, pentadbiran, penskoran, dan analisis data. Malahan, dimensionaliti ujian juga mempengaruhi kesahan keputusan yang diperoleh (Abedi, 1997). Kajian ini menyelidik dimensionaliti ujian item-item hujung terbuka Matematik Tambahan. Dimensionaliti dianalisa secara penerokaan dengan Analisis Faktor Komponen-Komponen Utama (Principal Components Factor Analysis) dan penyesuaian model (model fitting) dalam Analisis Faktor Pengesahan (Confirmatory Factor Analysis [CFA]). Ujian item-item hujung terbuka tersebut telah disahkan bersifat multidimensi melalui kaedah penerokaan. Kajian ini seterusnya mencadangkan penggunaan analisis "Differential Test Functioning (DTF)" Poly-SIBTEST yang berasaskan model multidimensi untuk rangka "Differential Item Functioning" yang disarankan oleh Shealy dan Stout (1993). Hal ini adalah bertujuan untuk mengesahkan bahawa ujian hujung terbuka Matematik Tambahan adalah multidimensi dan mengenal pasti dimensi-dimensi sekunder yang mungkin terlibat. Matematik Tambahan merupakan satu jenis matematik khusus yang ditawarkan kepada pelajar-pelajar Tingkatan 4 dan Tingkatan 5 yang cenderung dalam bidang sains dan teknologi serta sains sosial (Kementerian Pendidikan Malaysia, 2001). Item-item hujung terbuka ini dihipotesiskan untuk menilai dua dimensi kognitif yang berbeza. Hal ini adalah sebagai melengkapi trait terpendam utama yang diukur

iaitu kebolehan komunikasi matematik dan kebolehan penaakulan formal selain konstruk utama iaitu kebolehan matematik tambahan umum. Kajian ini menyelidik DTF antara pelajar-pelajar mahir dan tidak mahir dalam penulisan serta antara penaakul-penaakul formal dan tidak formal yang telah mempelajari semua topik Matematik Tambahan Tingkatan 4. Kajian ini juga meneroka caranya item-item hujung terbuka ini berfungsi secara "differentially" untuk kumpulan-kumpulan yang mempunyai aras kecekapan menulis dan berkebolehan penaakulan formal yang berlainan. Kejadian DTF dalam situasi-situasi ini sebenarnya mengukuhkan lagi kesahan konstruk ujian. Hal ini adalah kerana mengikut kurikulum dan sukatan Matematik Tambahan, kedua-dua dimensi sekunder ini merupakan komponen-komponen penting dalam kebolehan Matematik Tambahan. Kedua-dua kaedah iaitu penerokaan analisis faktor dan penyesuaian model dalam CFA serta analisis DTF Poly-SIBTEST menunjukkan bahawa ujian item-item hujung terbuka Matematik Tambahan adalah multidimensi dan multidimensionaliti ini adalah akibat daripada pemasukan dimensi-dimensi sekunder. Analisis-analisis ini menyarankan bahawa ujian item-item hujung terbuka Matematik Tambahan memperkenalkan dua dimensi tambahan, iaitu komunikasi matematik dan penaakulan formal di samping kebolehan matematik tambahan umum. Walau bagaimanpun kehadiran DTF bukan sahaja mempengaruhi kesahan inferens-inferens yang dibuat pada skor ujian malahan akibat daripada hanya berdasarkan satu laporan skor ujian akan mendorong kepada keputusan-keputusan dan inferens-inferens yang tidak tepat terhadap subjek-subjek yang diuji.

AN ANALYSIS OF THE DIMENSIONALITY OF THE SPM ADDITIONAL MATHEMATICS OPEN-ENDED ITEMS USING THE MULTIDIMENSIONALITY MODEL FOR DIF FRAMEWORK

ABSTRACT

Test dimensionality is an important issue as it is related to test validity and reliability. It affects the test reliability estimate, its interpretation as well as the administration, scoring, and analysis of the data and more importantly, it affects the validity of the results obtained (Abedi, 1997). This study investigates the dimensionality of the Additional Mathematics open-ended items test. It was analyzed exploratory utilizing the Principal Components Factor Analysis and model fitting in the Confirmatory Factor Analysis (CFA). The open-ended items test was verified exploratory to be multidimensional. This study then advocates the use of the Poly-SIBTEST's Differential Test Functioning (DTF) analysis, which is based on the multidimensional model for Differential Item Functioning framework proposed by Shealy and Stout (1993) to confirm that the Additional Mathematics open-ended test is multidimensional and to identify the possible secondary dimensions involved. Additional Mathematics, a specialized mathematics is offered to students in Form Four and Form Five who are inclined towards the field of sciences and technology as well as social sciences (Kementerian Pendidikan Malaysia, 2001). These open-ended items are hypothesized to assess two distinct cognitive dimensions to supplement the dominant latent trait being measured, namely the mathematical communication and formal reasoning abilities beside the dominant construct, the general additional mathematics ability. This study investigates the DTF between proficient and non-proficient writers and DTF between formal reasoners and non-formal reasoners who are in Form Four and have completed their Form

Four Additional Mathematics's topics. This study also explores how the open-ended items test functions differentially for the various groups with different levels of writing proficiency and formal reasoning ability. The occurrence of DTF in these situations actually enhanced the construct validity of the test because according to the Additional Mathematics' curriculum and syllabus, these two secondary dimensions are important components of Additional Mathematics ability. Both exploratory factor analysis and model fitting in the CFA as well as Poly-SIBTEST's DTF analyses showed that the Additional Mathematics open-ended test was multidimensional and the multidimensionality was due to the inclusion of the secondary dimensions. The analyses suggested that the open-ended items test introduced two extra dimensions, namely the mathematical communication and formal reasoning, in addition to the general additional mathematics ability. However, the presence of DTF influences the validity of inferences that is made on the test scores and having one test score reported leads to incorrect decisions and inferences being made about the examinees.

CHAPTER ONE
INTRODUCTION

This study investigated the Additional Mathematics open-ended items test dimensionality based on the multidimensional model of Differential Item Functioning (DIF) framework. Additional Mathematics is a specialized mathematics offer as an elective subject to Form Four and Form Five secondary schools students in Malaysia. This study focused on showing that the Additional Mathematics open-ended items test is multidimensional as a result of the inclusion of the mathematical communication and formal reasoning dimensions as the secondary dimensions besides the primary dimension, the general additional mathematics. This chapter presents the background of the study, specified the statement of the problem and the purpose of the study, describes its objectives and questions, presents the study theoretical framework and the significance of the study. And finally, the delimitations of the research will be presented.

1.1 Background of the study

Mathematics has always been considered as an important subject in our society. Mathematics has served nearly all the branches of the sciences and plays a vital role in the current information technology era. The technologies used in homes, schools, and workplaces are all built on mathematical knowledge. Mathematics is the language of all these technologies and the mastery of mathematics needs to be increased to prepare a mathematically literate workforce to deal with all these technologies processes aligned to the development and needs of a developed nation as we head towards the year

2020 (Kementerian Pendidikan Malaysia, 2001). This accounts for the great deal of concern with the success of mathematics education by our government and steps were taken such as reviewing our mathematics curriculum to meet this demand.

However, mastering mathematics is not easy and it requires various abilities. The reviewed mathematics curriculum require students to reason, think critically, make connections, make inferences, draw conclusions and to communicate mathematically (Kementerian Pendidikan Malaysia, 2001). The National Council of Teachers of Mathematics (NCTM) (1996) specified mathematical literacy as students having the ability to solve mathematical problems, to communicate and reason mathematically. Mathematical literacy equips students with logical reasoning skills, problem-solving skills, and the ability to think in abstract ways (National Council of Teachers of Mathematics, 1996).

As the curriculum changes so must the test, (Collis & Romberg, 1991; Kementerian Pendidikan Malaysia, 2000; National Council of Teachers of Mathematics, 1989; Romberg, 1992). Various studies have identified that mathematics tests require students to apply various skills (for example; Abedi, 1997; Muthen, Khoo, & Goff, 1997; Walker & Beretvas, 2000; 2001; Wu, 2003). This issue, defined from the multidimensional perspective, means different groups of examinees may have different multidimensional ability distributions due to the various abilities involved (Ercikan, 1998). In addition, if test items are capable of measuring these multiple dimensions, then using any unidimensional scaling procedure may produce item bias (Ackerman, 1992). From this perspective, a test item functioning differentially between two groups is an item measuring a secondary dimension that favours one of the groups after controlling the main dimension that the test is intended to measured (Camilli & Shepard, 1994). Hence, there are two major purposes in this study. Firstly, a specialized mathematics open-ended items test dimensionality was investigated exploratory using the Factor Analysis (FA) and model fitting analysis. Secondly, if these analyses revealed that the test had a multidimensional structure, the dimensionality was further investigated using a DIF approach, which utilized the framework that multidimensionality causes DIF. This study not only confirmed the multidimensionality structure of the test but investigated the possible secondary dimensions involved in the test using this framework.

1.2 Rationale of the Study

The main purpose of this study was to determine the dimensionality of the “Sijil Pelajaran Malaysia” (SPM) Additional Mathematics open-ended items test. If the test was revealed to be multdimensional the possible secondary dimensions involved were then identified. Also, the analyses involved in identifying these possible secondary dimensions utilized an appropriate multidimensional-based framework approach.

The issue of dimensionality is an important consideration because it affects the administration, scoring, data analysis and reporting of the results. Ignoring the underlying multidimensionality led to the inflation of reliability estimates and interpretation of the test reliability (Abedi, 1997). Importantly, it will affect the validity of the study's results obtained. There are studies that showed secondary dimensions which caused multidimensionality in tests involved in mathematics, for example, Walker and Beretvas (2000, 2001), Wu (2003), and Flores (1995). However, in Malaysia, there is a lack of published reports investigating the dimensionality of the Additional Mathematics test, specifically the recently revamped open-ended items test. More often then not, it is always assumed that such a test is unidimensional as evidenced by its reporting of the general additional mathematics ability using one single score only. However, there is substantive evidence indicating that the Additional Mathematics open-ended test is multidimensional. The following four paragraphs will further elaborate on this evidence and indications.

The implementation of the newly reviewed curriculum for Additional Mathematics has provided educators of Additional Mathematics a vision of what it means to know and understand additional mathematics. In the revamped curriculum and syllabus, the emphasis is on the need for additional mathematics students to spend more time on solving mathematics problems that require them to reason, think critically, creatively, make connections, make inferences, draw conclusions, and to communicate mathematical ideas (Kementerian Pendidikan Malaysia, 2001). This implies that the students do require various abilities in mastering the newly reviewed Additional Mathematics.

The reviewed curriculum also sees changes in assessing the students in Additional Mathematics. Currently, two instruments are used in SPM Additional

Mathematics namely Paper 1 (3472/1) which is an objective test and Paper 2 (3472/2) which is a subjective test with open-ended items to assess the students' additional mathematical abilities (Kementerian Pendidikan Malaysia, 2002). In the evaluation objectives of the new Additional Mathematics Items Format, it is stated, that the Additional Mathematics open-ended items are to evaluate students' mastery in identifying problems, planning and choosing solving strategies, implementing the strategies, and reviewing the solutions acquired for the problem:- all of which require various skills at the application, analysis, and synthesis stages (Kementerian Pendidikan Malaysia, 2002, p5). Problem solving is also the main focus in teaching and learning Additional Mathematics processes (Kementerian Pendidikan Malaysia, 2001). Studies on problem solving have shown that it requires various cognitive processes (see for example, Bryson, Bereiter, Scardamalia, & Joram, 1991; Chase & Simon, 1973; Chi, Feltovich, & Glaser, 1981; Forsyth & Spratt, 1980; Greeno, 1977; Simon & Hayes, 1976). Hence, the changes in assessing the students in Additional Mathematics open-ended items do indicate the need to evaluate various abilities.

Ackerman (1992) considered most cognitive ability tests to be multidimensional which required several composites of abilities that vary from item to item. Hamilton, Nussbaum, Kupermintz, Kerkhoven, and Snow (1995) and Kupermintz, Ennis, Hamilton, Talbert, and Snow (1995) who carried out research on the National Education Longitudinal Study of 1988, suggested that such tests should be treated as multidimensional. Abedi (1997) and Muthen, Khoo, and Goff (1997) reported multidimensional structures underlying the 1992

National Assessment of Educational Progress (NAEP) mathematics tests. Studies by Walker and Beretvas (2000, 2001) and Wu (2003) suggested that the Washington Assessment of Student Learning (WASL) mathematics test is multidimensional. Hence, the question arises as to whether our Malaysian Additional Mathematics open-ended items test which is used to assess students' performances in Additional Mathematics is also multidimensional.

Studies by Perkhounkova & Dunbar (1999), Walker & Beretvas (2000, 2001), and Wu (2003) have shown that the dimensionality of a test is related to item format. In their studies, it was shown that open-ended and multiple-choice item format do assess different constructs. According to Traub and MacRury (1990) there is evidence that these two formats appear to measure different abilities. According to Badger and Thomas (1992), open-ended items focus on students' understanding, their ability to reason and their ability to apply knowledge. In mathematics, open-ended items require students to reason and to offer evidence for their thinking, to communicate and present their ideas in mathematics, and find connections across mathematics (Foong, 2002; Sanchez & Ice, 2004).

This reasoning and arguments forwarded based on the curriculum requirements, assessment objectives, the various related previous research results and its open-ended format strongly suggests that the Additional Mathematics open-ended items test should be multidimensional. However, very few researches have been done on the Additional Mathematics open-ended items test dimensionality, specifically the newly reviewed Additional

Mathematics open-ended instrument in Malaysia. Although, the evidence indicated specifically the open-ended items test is multidimensional, such assumption should not be made without proof or without any studies being carried out. Hence, it is the purpose of this study to investigate the dimensionality of the SPM Additional Mathematics open-ended test. A study, such as this, will provide evidence and answers on the dimensionality structure of the open-ended Additional Mathematics instrument. This study will explore the dimensionality of the SPM Additional Mathematics open-ended items, utilizing both the Principal Component FA approach and Confirmatory Factor Analysis's (CFA) model fitting approach. If these exploratory analyses revealed multidimensionality, the test will be further analyzed using the confirmatory Poly-SIBTEST approach to determine the possible secondary dimensions involved in this multidimensional structure.

It is also the purpose of this study to identify the possible secondary dimensions involved if the multidimensional structure is present. There could be various possible secondary dimensions involved in the test. For example, based on the Additional Mathematics' curriculum and its assessment objectives, the secondary dimensions involved could be mathematical connection, mathematical reasoning, and mathematical communication besides the general additional mathematics dimension. Questions arise as to which of these abilities played a more important role and which of these abilities are also assessed by the multidimensional structure. Such information will be most valuable to mathematics educators as well as students. However, this study has undertaken to investigate two secondary dimensions only: the mathematical communication and formal reasoning which are believed to play a more prominent role in the open-ended items assessment. The following paragraphs will discuss why these two secondary dimensions are deemed to be more important in Additional Mathematics. The substantive reasons detailed below provide a prior for the poly-SIBTEST Differential Test Functioning (DTF) analyses to identify the secondary dimensions involved (Walker & Beretvas, 2001)

Regardless of the students' abilities, the content of mathematics is not taught without language. It is clear that although the focus of assessment is mathematics, the questions are presented through the medium of language, and the students too have to present their solutions through the medium of language, a factor by itself that may compromise the validity of the test if the language in any way obscures the mathematical demand in completing the tests (Miller, 1993). Under these circumstances, it is not only the mathematical competence that is being assessed, but also the communication aspects of the language in presenting their thoughts, reasoning, interpretations, and understanding of the question demands. This means that if students have to solve the items correctly, they need the communication ability to express their ideas, their thought and understanding of the demands of the questions.

The role of communication in developing mathematical understanding is emphasized in the reviewed Additional Mathematics' curriculum and syllabus. This role is also acknowledged in other parts of the world. For example, the National Council of Teachers of Mathematics in the Curriculum and Evaluation

Standards for School Mathematics (National Council of Teachers of Mathematics, 1989), Australian Educational Council in the National Statement on Mathematics for Australian Schools (1991) and the Professional Standards for Teaching Mathematics (National Council of Teachers of Mathematics, 1991), stressed the need for students to communicate mathematical ideas. According to the NCTM's curriculum standards, communication helps students "construct links between their informal, intuitive notions and the abstract language and symbolism of mathematics, and it also plays a key role in helping students make important connections among physical, pictorial, graphic, symbolic, verbal, and mental representations of mathematical ideas" (National Council of Teachers of Mathematics, 1989, p.26). For Mumme and Shepard (1990), effective communication about mathematics enhances students' comprehension and empowers them as learners. When students are presented with a mathematical problem, they need to describe, justify, explain and create a solution. Students' understanding of mathematics is dependent on their knowledge of both mathematics as a language and the language used to teach mathematics. Students need to make the cognitive links among familiar language, real-world concepts, formal mathematics language, and symbolic manipulation (National Council of Teachers of Mathematics, 1989). To empower students with the mathematical knowledge essential to participate fully in the society, students must be involved in the expressive aspects of language by having them speak and write about mathematics (Miller, 1993).

There has been a great deal of research and debate concerning the relationship between languages and learning in mathematics. Secada (1992) concluded that the studies he reviewed indicated a relationship between a student's language proficiency and his or her performances on measures of mathematics achievement. Shield and Galbraith (1998) investigated the mathematical understanding of the students as reflected in their writing. Their study supported the claim that writing enhances learning in mathematics. Hence, one can conclude that there is another aspect besides mathematical ability, which is the writing component, which influences one's mathematical achievement. Students who are not good writers have difficulty expressing themselves in writing (Miller, 1993). If students do not practice communicating their mathematical knowledge in writing, they will probably not improve their ability to write as well as to solve mathematical problems. In the educational context, students at many levels find writing difficult and may not communicate their mathematical thinking effectively (see MacNamara & Roper, 1992, on reports of investigative work in secondary school; Alibert & Thomas, 1991, on proofs at the undergraduate level). Research has shown that writing is a problem solving endeavour (Flower, Schriver, Carey, Haas, & Hayes, 1989) and as such it involves similar cognitive processes as in problem solving. It is an ability that is not easy to master (see for example, Husin & Abdul Aziz, 1979).

Basic to learning any language are the four components of listening, speaking, writing and reading (Kinmont, 1990). However, to understand the problem and to produce a written solution, only the reading and writing abilities are needed. Studies by Knifong and Holtan (1976), and Jerman (1973, 1974), show that the level of reading ability needed is low. Hence, for this study, the writing ability is assumed to contribute significantly and the written skill represents the mathematical communication ability. The writing ability is the skill the students apply to express their thoughts and understanding in writing their solutions to the Additional Mathematics open-ended items (Walker & Beretvas, 2001). Students lacking this skill, the writing ability, which represents the mathematical communication, will find the Additional Mathematics open-ended items function differentially against them if these items do required this additional dimension. Hence, this study hypothesized that the mathematical communication ability represented by the writing mode is assessed by the Additional mathematics open-ended items test as one of the secondary dimensions.

According to Dowden (1993), solving problems, making decisions, evaluating and assessing the quality of information, require the application of reasoning ability. Studies by Evans (2002), and Stanovich and West (2000) have identified that reasoning plays an important role in mathematics. They observed that not only students but also adults have great difficulty in solving mathematical reasoning problems. The failure of students to reason is a serious problem for both the theory and practice of education, and it is precisely the formal reasoning skills that are necessary for mastering academic subjects such as mathematics (Genovese, 2003). Lawson (1982) demonstrated that formal reasoning was related to mathematics. Researches conducted at the college level supported the idea that students who do not demonstrate formal reasoning abilities experienced difficulties in mathematics (Bunce & Hutchinson, 1993; Niaz, 1989).

“Successful students approach mathematical problems as they would in an investigation “ (Conley, 2003, p23). This report identifies key skills such as being able to think conceptually as well as procedurally, reason logically, use common sense to evaluate solutions, and construct arguments and proof using supporting evidence. Logical thinking can be defined as “the process of determining the authenticity, accuracy or value of something; characterized by the ability to seek reasons and alternatives, perceive the total situation and change one’s view based on evidence. This is not an ability with which one is born with but rather something that can and should be taught. Piaget’s theory supports the assumption that secondary students generally should be capable of reasoning at the level of formal thought, thereby demonstrating the cognitive abilities which are central to the secondary school level thinking.

However, several researchers reported that many secondary students have not reached the level of formal operational reasoning, but are still characterized by either the concrete level of reasoning or a transitional level between concrete and formal thought (Berenson, Carter, & Norwood, 1992; Lawson, 1982; Reyes & Capsel, 1986). Many researchers believe that students who cannot function at the formal level of thought experience more difficulties in mathematics courses (Bunce & Hutchinson, 1993; Niaz, 1989). For example, Niaz found that non-formal thinkers could not translate algebraic equations as effectively as formal thinkers.

Mathematics requires logical thinking for concepts to be understood clearly. Logic is the study of formal reasoning based upon statements or propositions (Price, Rath, & Leschensky, 1992). It encourages students to think critically and creatively about mathematics and gives them the possibility to articulate the expression of their opinions. The Third International Mathematics and Science Study (TIMSS) documented that mathematics teachers in U.S. focus on teaching students how to do mathematics and not on understanding what they do (NCES, 1996). There is a part of mathematics that's really about developing formal reasoning but teachers often focus on rote drilling. Epistemologically, reasoning is the foundation of mathematics. As science verifies through observation, mathematics relies on logic (Arthur, 1999).

Hence, this research also hypothesizes that students who are non-formal reasoners will be at a disadvantage when answering the open-ended Additional Mathematics items as well. Just like the mathematical communication aspect, groups who lack such ability would be disadvantaged if test items assess formal reasoning. Hence, it is hypothesized that answering the open-ended Additional Mathematics items correctly also require another additional ability, the formal reasoning ability. This means the Additional Mathematics open-ended items seem to assess the communication ability, formal reasoning ability, and the general additional mathematics ability. It needs to be noted that there would be other dimensions involved such as making mathematical connections as documented in the Additional Mathematics' curriculum (Kementerian Pendidikan Malaysia, 2001). However the three dimensions mentioned are investigated in this study as it is felt that these dimensions influence most the solutions of the examinees as well as their final score. In this study, the primary dimension will be the general additional mathematics ability whereas the mathematical communication and the formal reasoning ability are the secondary dimensions involved.

The presence of these two hypothesized secondary dimensions in the Additional Mathematics open-ended items inevitably introduced what Roussos and Stout (1996) termed as auxiliary or benign dimensions, where these secondary dimensions are intentionally assessed as part of the constructs on the test. Items that measure the secondary dimension should demonstrate a disproportionate difference between the reference and focal groups relative to what should be observed on the items that measure only the primary dimension, thereby producing group differences. This important source of group differences produces DIF.

DIF studies are designed to identify and interpret these secondary dimensions by using a combination of substantive and statistical analyses (Gierl, 2005). The statistical analysis of test unfairness at item-level is universally called DIF and at test level it is known as DTF. DIF is present when examinees from different groups have different probability or likelihood of answering an item or a bundle of items correctly (Shepard, Camilli, & Averill, 1981). Until recently, most of the analyses still focus on cultural differences using students grouped according to some inherent cultural attribute such as gender, ethnicity and race (Scheuneman & Gerritz, 1990; Schmitt & Dorans, 1990; Wang & Lane, 1996). Such analyses implicitly assumed that some test items measure some construct-irrelevant (Messick, 1989) abilities that favour certain cultures. Typically, such analyses have applied the unidimensional Item

Response Theory (IRT) models, which mean that only one ability of one composite of multiple abilities is measured by the test (Hambleton, Swaminathan & Rogers, 1991). However, this is a contradiction if the test measures several latent traits rather than a single one (Reckase, Ackerman & Carlson, 1988; Traub, 1983).

Similarly, although the format of standardized mathematics test has begun to change, the use of a unidimensional IRT model to describe the data is still common (Walker & Beretvas, 2001). Formats such as open-ended test types require students to write explanations, which involve communication skills (Foong, 2002; Sanchez & Ice, 2004) and formal reasoning. Some researchers in test and item bias research have recognized the necessary presence of multidimensionality underlying test items responses where bias is present. For example, Lord (1980, p.220) states “if many of the items in a test are found to be seriously biased, it appears that the items are not strictly unidimensional”. There is an overwhelming consensus in the testing community that the basic cause of DIF is the presence of multidimensionality in a test. However, rigorous mathematical multidimensional models of DIF have only been very recently been proposed (Kok, 1988; Shealy & Stout, 1993). A clear understanding of the practical implications of such models has only just begun to emerge (Douglas, Roussos, & Stout, 1996).

One such model was developed by Shealy and Stout (1993), the multidimensional model for DIF (MMD), which serves as the theoretical basis for understanding how DIF occurs. It is based on the premise that DIF is produced by multidimensionality. Based on the MMD, Roussos and Stout (1996) developed a multidimensional-based DIF analysis paradigm, linking substantive and statistical analyses so that researchers and practitioners can systematically identify and study the sources of DIF. This psychometric DIF paradigm shift from a totally reactive (removing unfair items after they have been constructed and pretested) and single-item-based approach to a partially proactive (i.e., also applied at the test design stage) and item-bundle-based approach to DIF. This approach stresses substantively interpreted latent-dimensionality-based explanations of causes of DIF (Stout, 2002). This proposed paradigm is a confirmatory approach (Gierl, 2005; Stout, 2002; Walker & Beretvas, 2001).

One such DIF approach that is based on both the MMD and the multidimensional DIF analysis paradigm is the Poly-SIBTEST. Poly-SIBTEST is the direct outgrowth of the Shealy-Stout multidimensional model for DIF (Shealy & Stout, 1993). When primary and secondary dimensions characterize item responses, the data are considered multidimensional. Poly-SIBTEST uses differences in the expected scores conditional on the primary dimension across groups to test for DIF. This study utilizes the Poly-SIBTEST in the analysis to determine if the Additional Mathematics open-ended items test is multidimensional and as well as testing the DIF hypothesis generated that the secondary dimensions involved are the mathematical communication and formal reasoning.

The main purpose of administrating test or examination is to evaluate the student’s ability in a subject in a fair and impartial manner. However, the results of such subject in the public examination are reported in a single overall score which is normally the sum of total number of correct items achieve. However, tests like Additional Mathematics open-ended items test mentioned above has always been assumed to measure a single ability. The psychometric properties obtained and research findings have also generally applied using unidimensional models or methodology. This gives rise to questions about the validity of such results or findings obtained. Can the research findings and results be reliable and trusted when the unidimensional model was applied if the data used were multidimensional? These are questions that arise when such a mismatch occurs.

This study has the intention to avoid such mistakes of using a unidimensional approach on a multidimensional structure data. The vast volume of literature, documents, reports and research findings has shown that generally mathematics items specifically problem solving items in fact require more than a single ability or cognitive process to obtain the solution to the problem. As mentiond earlier, this study will investigate the dimensionality of the open-ended Additional Mathematics items, which was not introduced by accident but by design using the Poly-SIBTEST. If the open-ended Additional Mathematics items are multidimensional, then it is possible that DIF will occur for groups of students with comparable Additional Mathematics content knowledge but differ in mathematical communication ability or / and formal reasoning ability. To improve the authentic diagnostic utility of the additional mathematics test, three scores should be reported, one for the writing ability, one for formal reasoning ability and the other for the ability in solving mathematical problem. This study is timely and relevant especially in the context of Malaysia because there have been few published reports on the investigation of the dimensionality of the reviewed Additional Mathematics open-ended terms as well as any reported studies on the possible secondary dimensions involved if the open-ended test is multidimensional. Also, much research study has been done based on unidimensional IRT models and not the appropriate multidimensional IRT model, and also dichotomous data has been employed instead of polytomous data.

1.3 Problem Statement

Researchers such as Reckase et al. (1988), Wang (1986), Ackerman (1992) and Yen (1985) and many others have argued that tests are multidimensional. Other studies too such as the National Education Longitudinal Study of 1988 have suggested that tests should be treated as multidimensional (Hamilton et al., 1995). Researches on open-ended formats showed that such formats required various abilities (see for e.g. Badger & Thomas, 1992; Perkhounkova & Dunbar, 1999; Walker & Beretvas, 2000, 2001; Wu 2003). However, there has been no published report of studies investigating the dimensionality of the newly reviewed SPM Additional Mathematics open-ended items. However, based on the reporting of a single total score or a single grade as used by the Lembaga Peperiksaan Malaysia in reporting students' Additional Mathematics achievement, it is always assumed that the score reflects only the general additional mathematics ability. Hence, are these open-ended items test unidimensional or multidimensional? Do these open-ended items measure single trait or multiple traits?

What if the open-ended items test is indeed multidimensional? Assuming a multidimensional structure is verified, what then are the additional abilities being measured? There are various reasons that cause multidimensionality in a test. Some studies done on mathematical items have suggested that it involved various abilities and cognitive processes. See for example, Forsyth and Spratt (1980), Grenno (1977) and Simon and Hayes (1976). Hence, what would be the possible secondary dimensions?

Studies by Berenson et al. (1992) and Lawson (1982) demonstrated that formal reasoning skill is related to mathematics and according to Genovese (2003) formal reasoning is necessary for mastering mathematics. Bunce and Hutchinson (1993) and Niaz (1989) demonstrated that students who are non-formal reasoners experienced difficulties in mathematics. Ironically, the majority of students in primary and high school grades are unable to utilize the formal reasoning as documented in Chiappetta's (1976) study. Renner and Lawson's (1973) study showed a significant proportion of tertiary students have yet to attain formal reasoning level. Hence, is lacking in the formal reasoning ability a disadvantage in answering the Additional Mathematics open-ended items? If it is, then the Additional mathematics open-ended items will be functioning differentially against this group of examinees who are weak in formal reasoning ability.

However, mastering the formal reasoning ability is not sufficient. Cai, Jakabcsin and Lane (1996) in their study discovered that students could not communicate about how they reason mathematically. Various research undertaken such as by Boaler (1997); Resnick et al. (1991); Steffe and Gale (1995); and Secada (1992) have shown the importance of communication in mathematics. Research by Walker and Beretvas (2000) and Wu (2003) on the WASL mathematical opened-items revealed that these items required the communication ability besides the general mathematics ability. More importantly, the examinees required the mathematical communication skill to express their thoughts and understanding in writing their solutions.

Students at many educational levels find writing difficult and can not communicate their mathematical thinking effectively (see MacNamara & Roper, 1992, on reports of investigative work in secondary school; Alibert & Thomas, 1991, on proof at the undergraduate level). According to Miller (1993) students who are not good writers have difficulty expressing themselves mathematically in writing. Husin and Abdul Aziz (1979) contented that it is also an ability that is difficult to master. Hence, the failure to master this communication ability means that the Additional Mathematics open-ended items test will function differentially against them. This is what teachers normally experience in their classroom where their students might be able to “do mathematics” but are unable to communicate what they know clearly or are unable to reason mathematically.

Though there could be other abilities required in the Additional Mathematics open-ended items, for this study it is hypothesized that the formal reasoning ability and communication ability would play a more prominent role in answering these open-ended items. Communication skills demand students to use a combination of components that consist of oral, reading, listening, and writing skills (Kinmont, 1990). However, according to Matz and Leier (1992) reading and writing components are only expected to influence the students' written solutions. Reading is needed to understand the items and writing skill is needed to reason and express their understanding in the written form. However, the level of reading ability required is low (Knifong & Holtan, 1976; Jerman, 1973, 1974). Hence, according to Walker and Beretvas (2001) writing is hypothesized to represent the communication ability. As such, this study will analyze whether the secondary dimensions involved are the mathematical communication ability as represented by the writing mode and the formal reasoning if the open-ended items test has a multidimensional structure.

However, most studies and analyses done on tests more often than not, applied models or approaches that assumed unidimensionality in their analyses (Wu, 2003). Some analyses carried out have even suppressed the multidimensional structures by making it fulfill certain assumptions so as to be able to function as unidimensional items. This was probably due to the approaches utilized being based on certain IRT models which hinged on the major assumption that performance on a set of items is unidimensional (Lord, 1980).

Such mismatch of utilizing a unidimensional approach on the multidimensional structure data has grave consequences. It affects the administration, scoring, data analysis and reporting (Abedi, 1997). According to Abedi, ignoring the underlying multidimensionality will inflate the reliability estimates and interpretation of the test reliability. It will also affect the fit of latent trait models in some detrimental way (Reckase, 1979) and might jeopardize the invariant feature of the IRT models (Ackerman, 1994; Reckase, 1985). More important, it will affect the validity of the results obtained in the results.

Hence, it is of great importance that analyses of multidimensional structure data should be done with an appropriate multidimensional-based framework approach. This study utilized the Poly-SIBTEST that not only works with polytomous scored data but is also based on the Shealy and Stout (1993) multidimensional model for DIF. Until now, there are no significant studies on multidimensionality of Additional Mathematics test or the utilizing of the MMD and Roussos and Stout (1996) multidimensional-based DIF analysis paradigm that have been published in Malaysia. Thus, the purpose of this study was to investigate the dimensionality of the Additional Mathematics open-ended test. If the open-ended test has a multidimensional structure, an appropriate multidimensional approach to the multidimensional data will be utilized to investigate if the mathematical communication and formal reasoning are the secondary dimensions assessed by the SPM Additional Mathematics open-ended items. Also, this study investigated how the Additional Mathematics open-ended items functioned differentially for groups of examinees chosen from the combinations of the two extremes along these two distinct cognitive dimensions.

1.4 Research Objectives

The first objective of this study was to investigate the dimensionality of the SPM Additional Mathematics open-ended items test. The dimensionality of this test was determined by performing both the exploratory and confirmatory approaches analyses, the Principal Components FA and the CFA’s model fitting. In short, is the Additonal Mathematics open-ended item test unidimensional or multidimensional? If a multidimensional structure was verified by these analyses, this study will reconfirm the multidimensional structure with the Poly-SIBTEST’s DTF analyses.

Secondly, if the underlying Additional Mathematics open-ended items data are verified to be multidimensional, this study will investigate the additional abilities involved in the open-ended test’s multidimensionality. This means that this study will determine the possible secondary dimensions besides the targeted primary dimension being assessed by these open-ended items test. The secondary dimensions being investigated are the formal reasoning ability and the mathematical communication ability. The primary dimension is the general additional mathematics ability. The formal reasoning dimension is represented by the correlation reasoning mode, controlling variables mode, combinatorial reasoning mode, probabilistic reasoning mode and proportional reasoning mode, whereas the mathematical communication dimension is represented by the writing mode

Finally, the study also explored how the Additional Mathematics open-ended items test functioned differentially for groups of examinees chosen from the combinations of two distinct extreme of these two secondary dimensions, formal reasoning and mathematical communication.

This study has also avoided these grave consequences due to mismatch of using a unidimensional approach in analyzing the multidimensional structure data. The analysis on the open-ended items test that has been verified to be multidimensional was conducted using an appropriate multidimensional-based framework. This study used the unique and novel application of DIF procedures which is based on Shealy and Stout (1993) multidimensional model for DIF framework as well as Roussos and Stout (1996) multidimensional-based analysis paradigm. The analyses which were conducted with Poly-SIBTEST utilized this multidimensional model of DIF framework and the multidimensional-based analysis paradigm. Poly-SIBTEST generated DIF hypotheses to enable it to make comparisons to explore what the open-ended items test may be measuring in addition to what they should be measuring. For example, in identifying if the mathematical communication ability is one of the secondary dimensions assessed by the open-ended items, DIF hypothesis will be generated by Poly-SIBTEST to compare if the open-ended items test is functioning differentially for the proficient or non-proficient writers. Similarly, Poly-SIBTEST will generate DIF hypothesis to compare if the open-ended items test is functioning differentially for the formal reasoners or non-formal reasoners in determining if the mathematical communication is the other secondary dimension involved.

1.5 Research Questions

This study guided by the research objectives outlined endeavoured to develop a framework that could provide answers to the following questions:

1. Are the Additional Mathematics open-ended items multidimensional?
2. Do the Additional Mathematics open-ended items function differentially for proficient and non-proficient writers?
3. Do the Additional Mathematics open-ended items function differentially for formal reasoners and non-formal reasoners?
4. Do the Additional Mathematics open-ended items function differentially for groups of examinees who are
   a. Both proficient writers and formal reasoners
   b. Proficient writers but non-formal reasoners
   c. Non-proficient writers but formal reasoners
   d. Both non-proficient writers and non-formal reasoners

1.6 Significance of the Study

Most of the studies done on DIF have focused on the application of a unidimensional IRT procedure although the data used were multidimensional (Ackerman, 1989, 1991; Ansley & Forsyth, 1985; Drasgow & Parsons, 1983; Folk & Green, 1989; Harrison, 1986; Reckase, 1979; Reckase et al., 1988; Way, Ansley, & Forsyth, 1988; Yen, 1984). This mismatch of utilizing a unidimensional approach on multidimensional data has many negative implications. According to Abedi (1997), this mismatch will affect the administration, scoring, data analysis and reporting. It will also inflate the reliability estimates, affect the validity of results obtained (Abedi, 1997) and affect the fit of latent trait models in some detrimental way (Reckase, 1979) as well as jeopardize the invariant feature of the IRT models (Ackerman, 1994).

This study avoids such mismatch between the analysis approach used and the data being analyzed.