Error Analysis of Children with Mathematics Learning Difficulties in Tibet

Lei RONG

Master’s Dissertation
Department of Special Needs Education
Faculty of Educational Sciences
UNIVERSITY OF OSLO

This dissertation is submitted in part fulfilment of the joint degree of MA/Mgr. SIE – Erasmus Mundus, University of Roehampton, University of Oslo and Charles University

Autumn 2016
Error Analysis of Children with Mathematics Learning Difficulties in Tibet
Error Analysis of Children with Mathematics Learning Difficulties in Tibet

Lei Rong

http://www.duo.uio.no/

Trykk: Reprosentralen, Universitetet i Oslo
Abstract

Children with mathematics learning difficulties (MLD) in Tibet are in need of educational support, and error analysis is regarded as a prior step to imply better teaching design and to plan intervention programmes. This study focuses on two aspects: identification of error types and analysis of error patterns.

A Mathematics Error Pattern Identification Test (MEPIT) was developed in this study, which aimed to identify mathematical error types among children with MLD. Two primary school graduates in Lhasa participated the pilot test, thirty first-year seventh graders from two secondary schools in Tibet took the official test. Through analysing the test results, MEPIT showed a satisfactory validity and reliability. The analysis of error patterns consisted of three parts: the frequencies and structures of mathematical errors, the relationships between error types with gender and school types, and a model fit test between error types and cognitive stages in solving mathematics problems.

Based on the analysis, fact errors and comprehension errors happened most frequently, relevance errors were avoided in most of cases. Girls seemed to be more vulnerable to mathematical errors than boys. Children from rural school tended to make more comprehension errors compared to those from urban school, there were no differences between other error types. Most correlations between error types were weak or modest, which indicated the independence of each error type. Among them, comprehension error, fact error, procedure error, and measurement error had a stronger effect on the score of MEPIT. The model fit test suggested that eight types of mathematical errors did not fit into four stages of cognition in solving mathematics problems very well.

Keywords: error analysis, mathematics learning difficulties, Tibet
Preface

This study is dedicated to all children with mathematics learning difficulties in Tibet. It is also devoted to teachers, administrators, researchers, policy makers, and other people who care about and want to improve mathematics education in Tibet.

Working on this master’s thesis is more like a learning process to me. I’ve always been interested in figuring out how do some students experience serious difficulties in learning mathematics, and I am so happy to find that the field of mathematics learning difficulties matches my interest so well. With the help of artificial intelligence in the future, human beings could have a deeper understanding of the nature of learning, and this field will become increasingly exciting.

In the first place, I would like to thank the Erasmus+ Programme, an amazing programme funded by the European Commission, without which I would not be able to study in these three unforgettable universities all across Europe. Here, I want to express my gratitude to Leda Kamenopoulou, who convenes this Erasmus Mundus MA/Mgr. Special and Inclusive Education Programme. She is a wonderful scholar and her feedback of assignments enlightens me every time. I would like to thank Jorun Buli-Holmberg, another convenor of this programme. She is knowledgeable and kind, she is always willing to support. I also want to thank Sarka Kanova, who helps me to settle down, to learn, and to experience life in Prague. Thanks to Nicolai Mowinckel-Trysnes as well as other administrative staff of the programme.

In addition, I am thankful to all the teachers and children in the schools involved in this study, for all their co-operation and help. I am thankful to my classmates and friends in this cohort, we are the last cohort, we’ve been experienced a lot together, and we will never forget this journey. I am thankful to my parents, sister, relatives, and my friends. Special thanks to Lang LIU, Long-long CAI, Shi-qin SHENG, and Bin-chuan DOU, for their generous support and valuable assistance with the research, I am very grateful.

Finally, I want to thank my supervisor Riikka-Maija Mononen for the great help and inspiration. Her open-minded guidance, comprehensive comments, as well as the pertinent books and journal papers she generously provided have highly improved the quality of this thesis. I am very lucky to have her supervision.
# Table of Contents

Error Analysis of Children with Mathematics Learning Difficulties in Tibet ............ III

Abstract ........................................................................................................................................ V

Preface ........................................................................................................................................ VII

Table of Contents ........................................................................................................................ IX

1 Introduction ............................................................................................................................... 1

2 Review of the Literature .......................................................................................................... 5

2.1 Mathematics Learning Difficulties (MLD) ........................................................................... 5

2.1.1 Introduction of Learning Disabilities ............................................................................... 5

2.1.2 Definition and Prevalence of MLD ............................................................................... 6

2.1.3 Characteristics of Children with MLD ........................................................................... 9

2.1.4 Support for Children with MLD .................................................................................... 10

2.1.5 Assessment of Children with MLD ............................................................................ 11

2.2 Mathematics Errors ............................................................................................................. 13

2.2.1 Mathematics Framework ............................................................................................... 13

2.2.2 Types of Mathematical Errors ....................................................................................... 14

2.2.3 Cognitive Stages in Solving Mathematics Problems .................................................... 17

2.3 Research Questions .............................................................................................................. 20

3 Research Methodology and Ethical Considerations .............................................................. 21

3.1 Research Methodology ......................................................................................................... 21

3.1.1 Research Design ........................................................................................................... 21

3.1.2 Research Participants ................................................................................................... 22

3.1.3 Research Tools ............................................................................................................. 23

3.1.4 Data Analysis Methods ................................................................................................. 24

3.2 Ethical Considerations .......................................................................................................... 26

4 Data Analysis and Results ....................................................................................................... 28

4.1 Result I: Mathematics Error Pattern Identification Test (MEPIT) ........................................ 28

4.1.1 Development of MEPIT .............................................................................................. 28

4.1.2 Pilot Test and Official Test ......................................................................................... 40

4.1.3 Validity and Reliability ............................................................................................... 44

4.2 Result II: Error Pattern Analysis ......................................................................................... 47

4.2.1 Coding of Error Types ............................................................................................... 47
4.2.2 Error Pattern Analysis ................................................................. 62

5 Discussion and Conclusion ............................................................. 70
  5.1 Discussion .................................................................................. 70
  5.2 Practical Implication ................................................................. 71
  5.3 Research Limitation ................................................................. 72
  5.4 Future Directions .................................................................... 73
  5.5 Conclusion ............................................................................... 74

References ...................................................................................... 75

Appendices ..................................................................................... 82
  Appendix 1 Norwegian Social Science Data Services ..................... 82
  Appendix 2 MEPIT Type I ............................................................... 83
  Appendix 3 MEPIT Type II ............................................................. 89
List of Tables

Table 1: The estimated prevalence of children with MLD .......................................................... 9
Table 2: Content Domains of Mathematical Basic Ability Diagnostic Test.......................... 29
Table 3: Target Percentages of the TIMSS 2015 Mathematics Assessment Devoted to Content
Domains in Each Framework........................................................................................................ 31
Table 4: Target Percentages of the TIMSS 2015 Mathematics Assessment Devoted to
Cognitive Domains in Each Framework..................................................................................... 31
Table 5: Constitution of the first draft of MEPIT ...................................................................... 32
Table 6: Constitution of the final draft of MEPIT ..................................................................... 33
Table 7: Details of the final draft of MEPIT ............................................................................. 34
Table 8: Participants’ demographic background information.................................................... 42
Table 9: The correlations between previous score and score of MEPIT .................................... 45
Table 10: Mean, standard deviation, and 0.05 level significance test of error types and gender
..................................................................................................................................................... 63
Table 11: Mann–Whitney U test of error types and gender ....................................................... 64
Table 12: Mean, standard deviation, and 0.05 level significance test of error types and school
types ............................................................................................................................................... 65
Table 13: Mann–Whitney U test of error types and school types .............................................. 65
Table 14: Using Pearson’s r correlation coefficient to look at relationships between eight error
types .............................................................................................................................................. 66
Table 15: Using multiple linear regression to look at the relationships between eight error
types and score of MEPIT ........................................................................................................... 67

List of Figures

Figure 1: Mathematics Framework of Singapore......................................................................... 13
Figure 2: Durations of four cognitive stages on mathematics trials............................................ 18
Figure 3: Frequencies of all possible error types ........................................................................ 36
Figure 4: Flow chart of official testing ...................................................................................... 43
Figure 5: Factor analysis of relationships between error types and children’s error patterns .. 45
Figure 6: Frequencies of error types - absolute number............................................................ 62
Figure 7: Frequencies of error types - accumulated percentage.............................................. 63
Figure 8: Path diagram of model fit between the types of errors and the stages of cognition in
solving mathematics problems ............................................................................................... 68
## List of Abbreviations and Symbols

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BOSTES</td>
<td>Board of Studies Teaching &amp; Educational Standards</td>
</tr>
<tr>
<td>CSI</td>
<td>Cognitive Strategy Instruction</td>
</tr>
<tr>
<td>DfE</td>
<td>Department for Education (United Kingdom)</td>
</tr>
<tr>
<td>DfES</td>
<td>Department for Education and Skills (United Kingdom)</td>
</tr>
<tr>
<td>ICD</td>
<td>International Classification of Disease</td>
</tr>
<tr>
<td>IDEA</td>
<td>Individuals with Disabilities Education Improvement Act</td>
</tr>
<tr>
<td>IEA</td>
<td>International Association for the Evaluation of Educational Achievement</td>
</tr>
<tr>
<td>IQ</td>
<td>Intelligence Quotient</td>
</tr>
<tr>
<td>MBADT</td>
<td>Mathematical Basic Ability Diagnostic Test</td>
</tr>
<tr>
<td>MEPIT</td>
<td>Mathematics Error Pattern Identification Test</td>
</tr>
<tr>
<td>MLD</td>
<td>Mathematics Learning Difficulties</td>
</tr>
<tr>
<td>MoE</td>
<td>Ministry of Education</td>
</tr>
<tr>
<td>MRI</td>
<td>Magnetic Resonance Imaging</td>
</tr>
<tr>
<td>NJCLD</td>
<td>National Joint Committee for Learning Disabilities</td>
</tr>
<tr>
<td>OALD</td>
<td>Oxford Advanced Learner’s Dictionary</td>
</tr>
<tr>
<td>OECD</td>
<td>Organisation for Economic Co-operation and Development</td>
</tr>
<tr>
<td>PISA</td>
<td>Programme for International Student Assessment</td>
</tr>
<tr>
<td>TIMSS</td>
<td>Trends in International Mathematics and Science Study</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
</tbody>
</table>

- **B**: unstandardized coefficients
- **Beta**: standardized coefficients
- **CFI**: comparative fit index
- **d**: Cohen’s d
- **df**: degrees of freedom
- **GFI**: goodness of fit index
- **M**: mean
- **p**: probability
- **r**: correlation or size effect of Mann–Whitney U test
- **RMSEA**: Root Mean Square Error of Approximation
- **$R^2$**: error variances
- **SD**: standard deviation
- **t**: t-test
- **Z**: Z-test
1 Introduction

“Mathematics as an expression of the human mind reflects the active will, the contemplative reason, and the desire for aesthetic perfection. Its basic elements are logic and intuition, analysis and construction, generality and individuality. Though different traditions may emphasize different aspects, it is only the interplay of these antithetic forces and the struggle for their synthesis that constitute the life, usefulness, and supreme value of mathematical science.”

(Courant & Robbins, 1996, p.1)

As a fundamental discipline, mathematics is important in many fields. The significance of mathematics lies not only in its indisputable weight occupying school curriculum, in its essence of beauty, but also in its practical value pervading individual’s everyday life. Mathematics often deals with elements and relationships of quantity. A good master of mathematics could enable individuals make effective use of the quantitative information available in their environment. Galileo Galilei (1564–1642) said, “The universe cannot be read until we have learned the language and become familiar with the characters in which it is written. It is written in mathematical language, and the letters are triangles, circles and other geometrical figures” (BBC, 2010).

Daily activities such as shopping, cooking, managing money, and surfing the internet cannot be done without mathematics. Further, many jobs require a strong background in mathematics, like banker, engineer, architect, accountant, just to name a few. Mathematics is vital to finance and economics, to software engineering and information technology, as well as to artificial intelligence which is believed to be the future direction of human society (Urban, 2016).

Learning mathematics improves problem-solving skills, which involves mathematically modelling a problem and developing the skills needed to solve that problem. Mathematical skills are abstract skills. Piaget (1953) explored the way children develop mathematical skills across various stages of cognitive development. He found that acquisition of mathematical skills is an active process, children learned the skills (such as addition, subtraction, multiplication) through interactions with the world around them. The International
Association for the Evaluation of Educational Achievement (IEA) addressed that “all children can benefit from studying and developing strong skills in mathematics” (IEA, 2015, p.11).

In this “technologically advanced and information-based world” (IEA, 2015, p.11), lack of mathematical skills would cause severe emotional, educational, and social problems. People with difficulties in mathematics are susceptible to negative emotions such as anxiety and depression, very often they show a distinct lack of persistence and perseverance in solving problems. When a high percentage of a country’s population lack of basic mathematical skills, its long-term economic growth will be negatively affected (OECD, 2016). Adults with low numeracy skills are not likely to gain full-time jobs and their employment options are often limited, as a result they are most often classed as receiving low wages (Dowker, 2005). Children with mathematics learning difficulties (MLD) suffer from mathematics classes constantly, very often this could lead to emotional and behavioural problems and the worst-case scenario, dropping out of the school (Jena, 2013). Especially those at the age of 15, they face a higher risk of quitting school (OECD, 2016). In addition, evidence shows that most of the adults with difficulties in mathematics tend to have already demonstrated mathematics learning difficulties in their childhood (Dowker, 2005; Jena, 2013). In other words, children’s mathematics learning difficulties can persist into their adulthood (Desoete, Roeyers, & Clercq, 2004). Therefore, it is essential to conduct research on children with MLD.

Nevertheless, from 1974 to 1997, only 28 studies on mathematics disabilities were cited in PsycInfo, meanwhile there were 747 researches could be found on reading disabilities (Noel, 2000, as cited in Desoete et al., 2004). As recently as 2007, studies on dyslexia outnumbered those on mathematics disability by 14:1 (Gersten, Clarke, & Mazzocco, 2007, as cited in Price & Ansari, 2013). It seems that MLD has been chronically understudied, the problems of children with MLD are underestimated and neglected (Hanich, Jordan, Kaplan, & Dick, 2001; Desoete et al., 2004; Price & Ansari, 2013). The global trend of inclusive education aims at reconstructing education system to meet the needs of children who are at risk (Malinen, 2013), children with MLD should not be ignored in this process and studies in this field should not lag behind.

It is especially imperative, necessary and valuable to conduct researches on children with MLD in Tibet. China seems to have a good reputation for its mathematics education. Take Shanghai as an example, based on the mathematics results of the Programme for International Student Assessment (PISA), Shanghai was ranked the first in 2009 and the first in 2012
consecutively among all the participating economies, making it the best performer (OECD, 2009, 2012). However, there is a big achievement gap between Tibetan students and other Chinese students in mathematics. The gap is verified by researcher’s classroom observation, four years’ teaching experience in Tibet as well as communications with other mathematics teachers. It is also endorsed by standardized test results and research findings (e.g., Zhang, 1995; Basang & Shi, 2006; Fang, Wang, Lou, Li, & Nima, 2008; Tibetan Examination Yuan, 2015). Geary (2013) estimates that the prevalence of MLD in school children should be 4% to 14%, to add low-achieving students, this number would range from 9% to 24%. By contrast, Wang (2008) claims that about 33% of Tibetan children experience some form of difficulties in learning mathematics.

There is an urgent need of educational support and intervention programmes for children with MLD in Tibet. However, teachers have been facing barriers to support and to intervene effectively. Generally speaking, mathematics teachers in Tibet can be divided into two types: the local Tibetan teachers and Han Chinese teachers. The number of Han Chinese teachers is nevertheless more than the number of local Tibetan teachers (MoE, P.R.C., 2015). The reasons might be two-fold. Firstly, many Tibetan students experience difficulties in learning mathematics, this may be caused by unique geographical circumstance, distinct cultural background, or lack of preliminary education (Basang & Shi, 2006; Fang et al., 2008; Wang, 2008). As a result, they are not likely to study mathematics in higher education, let alone to be a mathematics teacher. This causes a lack of qualified Tibetan mathematics teachers. To solve the problem, Ministry of Education encourages mathematics teachers all over China to support Tibet, especially those student teachers, and then there is increasingly more Han Chinese teachers in Tibet. This is the second reason which is often called the “yuánzàng” (Supporting Tibet) movement. The movement has dramatically improved the quality of mathematics education in Tibet, meanwhile it has been facing many challenges (Wang & Zhao, 2012; MoE, P.R.C., 2015). The major issue is that Han Chinese teachers do not have direct learning experiences in Tibet, very often they are not aware of the actual mathematical understandings of Tibetan students, especially potential errors that Tibetan children would make. As a result, their mathematics teaching is not always effective and sometimes they would set up misunderstandings in classes, especially for children with MLD, they might be confused and totally lost.
Bearing this in mind, in order to imply better teaching design and inform better intervention programmes, it is necessary to identify error patterns among children with MLD. Error analysis, or error pattern analysis, is the most appropriate method to fulfil this goal. Error analysis is a way to analyse learners’ errors manifested in their work with a perspective to infer possible explanations underneath these errors. Error analysis focuses on the pervasive errors that children make due to their lack of certain mathematics abilities (Ketterlin-Geller & Yovanoff, 2009; Herholdt & Sapire, 2014). Error analysis not only includes the analysis of children’s mathematical errors, but also suggests the better intervention programmes. Analysing errors enables to identify difficulties that children may have with facts, concepts, strategies and procedures. Identifying the types of errors allows the teacher to address learner needs more efficiently (Hansen, 2005; McGuire, 2013).

This study aims to identify and analyse mathematical error patterns among children with MLD in Tibet. The main focus is on two aspects: identification of error types and analysis of error patterns.
2 Review of the Literature

In this chapter, literature concerning mathematics learning difficulties will be reviewed, which covers definition, prevalence, characteristics, support and assessment. Types of mathematical errors will be discussed next. Research questions will be presented in the end of the chapter.

2.1 Mathematics Learning Difficulties (MLD)

2.1.1 Introduction of Learning Disabilities

Children with Mathematics Learning Difficulties (MLD) are always included under the definition of Learning Disabilities (Garnett, 1998), it is therefore reasonable to look at learning disabilities briefly. The International Classification of Disease (ICD), which is maintained by the World Health Organization (WHO) adopts the term “learning disorders” to describe difficulties in learning:

“Disorders in which the normal patterns of skill acquisition are disturbed from the early stages of development. This is not simply a consequence of a lack of opportunity to learn, it is not solely a result of mental retardation, and it is not due to any form of acquired brain trauma or disease.”

(ICD-10, Version 2016, F81)

This description focuses more on responding what factors cannot explain learning disorders. Meanwhile, in USA, the National Joint Committee for Learning Disabilities (NJCLD) currently uses the definition that was formally adopted in 1990:

“Learning disabilities is a general term that refers to a heterogeneous group of disorders manifested by significant difficulties in the acquisition and use of listening, speaking, reading, writing, reasoning, or mathematical abilities. These disorders are intrinsic to the individual, presumed to be due to central nervous system dysfunction, and may occur across the life span.”

(NJCLD, 1990, p.1)
The NJCLD uses the term “learning disability” to indicate a discrepancy between children’s potential to learn and their actual achievement. While adapting the achievement-intelligence discrepancy, some researchers suggested a measuring score to measure the discrepancy between intelligence age and achievement age, some others employed sophisticated statistical analyses and calculated an expected achievement through a combination of weighted factors (e.g., IQ, socio-economic status, age) to reflect individual’s potential achievement (e.g., Hopkins & Sitkei, 1969; Rutter & Rule, 1970; Myklebust, 1975).

On the other hand, the achievement-intelligence discrepancy is under criticism. Stanovich (1991) criticized the achievement-intelligence discrepancy as having badly misled the researchers and specialists, he claimed that defining a hypothetical construct like learning disability with another hypothetical construct like intelligence was problematic. Further, Seigel (1988, 1989) questioned the long-term association of learning disabilities and IQ in diagnosing process on the following grounds: 1) IQ is not valid in measuring the cognitive processes involved in reading, spelling, language, and memory tasks of individuals regardless of whether they have learning disability or not; 2) it is verified by empirical researches that the construct of learning disability is irrelevant to the concept of intelligence, any relationship between two concepts is questionable. Cole (1993) responded to Seigel’s revisionist theory and argued that there was a better conceptual clarification by using five basic parameters to define learning disability: 1) level of deficit based on criteria which have been widely accepted; 2) level of intellectual potential which can be measured by individually administered standardized IQ test; 3) relationship between the criterion and the intellectual potential; 4) assessment of discrepancy, for example to assess the discrepancy between IQ and achievement scores; 5) exclusion clause: to excludes the potential case of learning disability helps in arriving at a relatively clear diagnosis of the specific learning disability.

2.1.2 Definition and Prevalence of MLD

Many different terms are used in this field to describe a deficit in mathematics learning, such as mathematics learning difficulties, mathematics learning disabilities, mathematics disorder, dyscalculia, developmental dyscalculia (e.g., Hanich et al., 2001; Kuo, Hsu, Liu, Chang, & Fan, 2001; Desoete et al., 2004; Geary, 2004; Price & Ansari, 2013). This is primarily due to different severeness levels of the difficulties, but also reflects that it is a new field in academics. The main difference is the using of disability, difficulty, disorder, or dyscalculia.
The following is the meaning of these words from Oxford Advanced Learner’s Dictionary (OALD): Disability, “[countable] a physical or mental condition that means you cannot use a part of your body completely or easily, or that you cannot learn easily” (OALD, 2016). This definition suggests that children with mathematics learning disabilities can still learn, however they cannot learn easily. Disorder, “[countable, uncountable] (medical) an illness that causes a part of the body to stop functioning correctly” (OALD, 2016). According to this definition, mathematics disorder focuses more on medical model. Difficulty, “[countable, usually plural, uncountable] a problem; a thing or situation that causes problems” (OALD, 2016).

Concerning dyscalculia, there is a distinction between primary and secondary developmental dyscalculia: the former focuses on the deficits caused by impaired development of brain to process mathematical information, while the latter refers to difficulties resulted of external factors such as poor teaching, economically deprived situation (Price & Ansari, 2013). Nonetheless, dyscalculia is not defined by the dictionary (OALD, 2016). In general, the term mathematics learning difficulties is used to describe a broad spectrum of difficulties in mathematics learning (Geary, 2013; Karagiannakis, Baccaglini-Frank, & Papadatos, 2014). In this study, this term – mathematics learning difficulties (MLD), will be adopted to refer to characteristics of the targeted children.

The Diagnostic Statistic Manual of Mental Disorders (DSM-5, 2013) states MLD as specific leaning disorder, which means impairment in mathematics. Mathematics achievements of these children are notably below their chronological age. This is shown in the early school years and will be lasting for at least half a year, and the impairment is not attributed to intellectual disabilities, developmental disorders, or neurological or motor disorders. Globally, International Classification of Disease defines specific disorder of arithmetical skills as:

“Involves a specific impairment in arithmetical skills that is not solely explicable on the basis of general mental retardation or of inadequate schooling. The deficit concerns mastery of basic computational skills of addition, subtraction, multiplication, and division rather than of the more abstract mathematical skills involved in algebra, trigonometry, geometry, or calculus.”

(ICD-10, Version 2016, F81.2)
In America, Individuals with Disabilities Education Improvement Act of 2004 (IDEA 2004) defines MLD as an underachievement in mathematical computation and mathematical problem solving. Other underachievement includes listening comprehension, reading fluency, written expression, etc. They are all categorised under the term Learning Disability. In England, DfES (2001) regards MLD as “a condition that affects the ability to acquire arithmetical skills”. One thing to notice is that difficulties in mathematics is not written into Code of Practice for children with special educational needs (DfE, 2015).

The *Diagnostic Statistic Manual of Mental Disorders* (DSM-IV-TR, 2000) provides the following diagnostic criteria for mathematics disorder:

a) Mathematical ability, as measured by individually administered standardized tests, is substantially below than expected given the person’s chronological age, measured intelligence and age appropriate education.

b) The disturbance in Criterion A significantly interferes with academic achievement or activities of daily living that require mathematical ability.

c) If a sensory deficit is present, the difficulties in mathematics are in excess of those usually associated with it.

(DSM-IV-TR, Section 315.1)

In its latest version, children with MLD are described as “…with impairment in mathematics: number sense, memorization of arithmetic facts, accurate or fluent calculation, accurate math reasoning” (DSM-5, Section 315.1). The changes from DSM-IV-TR to DSM-5 highlight the difficulties in diagnosing mathematics learning difficulties, as learning deficits in the areas of reading, written expression, and mathematics commonly occur together.

In order to identify children with MLD, varying diagnostic criteria have been used by different researchers, this has resulted in a variation in estimating the prevalence rates. Many of them used a mathematical age-chronological discrepancy of two years as standards to screen children with mathematics learning difficulties (e.g., Gross-Tsur, Manor, & Shalev, 1996; Geary, 2004). Some others adopted the lowest quartile or standard score of less than 90 as the criteria to distinguish the targeted children (e.g., Butterworth, 2002; Desoete et al., 2004). Table 1 shows a few prevalence studies that have been done by investigators mainly in the west:
Table 1: The estimated prevalence of children with MLD

<table>
<thead>
<tr>
<th>Location</th>
<th>Term</th>
<th>Prevalence</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Norway (Ostad, 1998)</td>
<td>Mathematical disorders</td>
<td>10.9%</td>
<td>Children involved in special mathematics teaching programme</td>
</tr>
<tr>
<td>England (Lewis et al., 1994, as cited in Butterworth, 2002)</td>
<td>Specific arithmetic difficulties</td>
<td>3.6%</td>
<td>&lt; 85 on arithmetic test, &gt; 90 on non-verbal IQ test</td>
</tr>
<tr>
<td>Israel (Gross-Tsur et al., 1996)</td>
<td>Dyscalculic</td>
<td>6.4%</td>
<td>Two grades below Chronological Age</td>
</tr>
<tr>
<td>Germany (Bzufka et al., 2000, as cited in Dowker, 2005)</td>
<td>Mathematics difficulties</td>
<td>6.6%</td>
<td>&gt; 50th percentile in spelling, but &lt; 25th percentile in mathematics</td>
</tr>
<tr>
<td>Belgium (Desoete et al., 2004)</td>
<td>Mathematics learning disabilities</td>
<td>2.27% to 7.70%</td>
<td>2 SD below norm</td>
</tr>
<tr>
<td>General estimation (Geary, 2004)</td>
<td>Mathematics learning disabilities</td>
<td>5% to 8%</td>
<td>Performance that deviates from age-related norms</td>
</tr>
</tbody>
</table>

2.1.3 Characteristics of Children with MLD

Although the diagnostic criteria and the main character of MLD are very poor arithmetic skills, children with MLD often exhibit problems in other areas of math as well. Such as problems in identifying spatial relationships, performing visual-perceptual and visual-motor tasks, conceptualizing time and direction, understanding of number symbols and other mathematics symbols, communicating mathematics concepts, carrying out mathematics calculations (Jena, 2013). A big proportion of these difficulties is perhaps due to developmental delays, faulty learning and faulty reasoning can also contribute to these difficulties (Cawley & Miller, 1989).

On neurobiological level, a damage or underdevelopment in specific part of brain could cause the most profound MLD (e.g., angular gyrus), this is often labelled as dyscalculia. For example, dyscalculia appears as a part of the syndrome called Gerstmann syndrome, which is a neuropsychological disorder characterized by four primary symptoms and could cause difficulty in learning or comprehending mathematics. Gerstmann syndrome is often suggested by the presence of a lesion in a particular area of the brain (Vallar, 2007). Price & Ansari (2013) further distinguishes primary and secondary developmental dyscalculia. The primary
developmental dyscalculia is related to impaired development of brain, and the secondary developmental dyscalculia focuses on mathematical deficits caused by other factors.

On cognitive level, working memory (WM), IQ, and specific language skills are believed to have played key roles in developing mathematics skills. A group of researchers use meta-analysis to investigate the relationship between working memory components (central executive, phonological loop, and visual-spatial sketchpad) and mathematical performance, and the results suggest that all these components are related to mathematical performance in 4-12 years old children with the highest correlation between mathematics and verbal updating (Friso-van den Bos, van der Ven, Kroesbergen, & van Luit, 2013). The relations between WM and difficulties in mathematics have been revealed in many studies as well (e.g., Passolunghi & Siegel, 2001). Regarding IQ, it is suggested that low intelligence level could result in poor mathematical performance (e.g., Li & Geary, 2013). In general, many researchers have recognized that specific language impairment could result in poor early mathematical competence, and this would be lasting to school age (e.g., Mononen, Aunio, & Koponen, 2014).

On behavioural level, lack of number sense (a cognitive mechanism that supports the representation and processing of numerical magnitudes, see e.g., Price & Ansari, 2013), sense of directionality, and lack of basic arithmetical skills would result in mathematics learning difficulties. Such as having difficulties in automatically retrieving arithmetical facts (for children around 9 should be atomised to retrieve number range from 1 to 20), slow and error prone strategies (counting with fingers), poor understanding of relational concepts which forms the foundation of directionality (e.g., up-down, upper-lower, nearer-farther, longer-shorter, bigger-smaller), difficulties in using decimal places of “carrying” (Hansen, 2005; Jena, 2013; Mononen, 2016).

2.1.4 Support for Children with MLD

Children with MLD have been designed a number of teaching strategies, the procedures of which is adapted in accordance with grades and levels of difficulty. Such as Cognitive Strategy Instruction (CSI), which is believed to be able to improve the academic performance of children with MLD, especially in solving mathematics word problems (e.g., Montague & Dietz, 2009; Montague, Enders, & Dietz, 2011). Zhu (2015) designed an experiment in a
Chinese primary school in order to investigate the impact of CSI on children with MLD in dealing with mathematics words problems. Results showed that students at all ability levels in treatment group outperformed their peers in comparison group, the intervention effect was strongest for children with MLD. This study suggested CSI to be a contextually and pedagogically appropriate model and an effective intervention programme for children with MLD.

There are some other educational support programmes for children with MLD. For example, explicit instruction and visual representation. The former is a problem dependent approach that enables teacher to solve the problem step by step, and children are motivated to follow the procedure step by step. The latter is beneficial to raise the interest, to relate mathematics to the real world, and to solve mathematics problems in a vivid way through visualize mathematical concepts (Mononen, 2016). In addition, brain-compatible instruction, the self-monitoring tactic, strategy training approach, a peer-tutoring system are also found to be helpful in improving children’s mathematics abilities (Bender, 2008).

Two computer-assisted training tools has been especially designed to intervene children with severe MLD. The first programme is “the Number Race”, which specialises in improving the precision of numerical magnitude representations. The second is “Graphogame”, aiming to build bridge linking exact amount to number symbols (Arabic digits). These two programmes are developed based on cognitive neuroscience evidence and have been found helpful to improve number-comparison performance, however overall effects to arithmetic are still under evaluation (Price & Ansari, 2013).

2.1.5 Assessment of Children with MLD

In research studies, many methods have been utilized to assess children with MLD. These methods can be divided into two categories: formal assessment and informal assessment. In addition, the formal assessment techniques can be further categorised into two groups: standardised achievement test and diagnostic arithmetic test (Jena, 2013). A standardized achievement test is an achievement test that is administered and scored in a consistent or standard manner. More specifically, a standardized achievement test requires participants to answer the same questions or a group questions from common bank of questions in the same way. The test conditions, scoring procedures, and interpretations are consistent, this enables
people to compare the relative performance of individual students or selections of students (Popham, 1999). One of the major benefits of standardized testing is that the test scores can be empirically documented. The results can then be suggested to have a certain degree of validity and reliability, which are therefore generalizable and replicable (Kuncel & Hezlett, 2007). Examples of such tests include the Programme for International Student Assessment (PISA) (OECD, 2016), the Trends in International Mathematics and Science Study (TIMMS) (IEA, 2015), in which mathematics is assessed as one of the abilities. These two tests are conducted globally and served as benchmark tests in many countries.

Other than standardized achievement tests, mathematics diagnostic tests are designed with the purpose to determine the exact nature of deficits in preforming mathematical skills. Such as Diagnostic Test of Self-Helps in Mathematic (Grades 3-8), Enright Diagnostic Inventory of Basic Arithmetic Skills (Grades 6-Adult), Key Math-Revised (Grades K-6), Sequential Assessment of Mathematics Inventory (Grades K-8), Test of Early Mathematics Ability-2 (Grades Preschool-3), Test of Mathematical Abilities 2 (Grades 3-12), WISC Arithmetic, and BAS Number Skills Tests, in which mathematics is assessed as part to determine learning disabilities (Jena, 2013; Dowker, 2005). In Taiwan, a Mathematical Basic Ability Diagnostic Test (MBADT) has been developed (Kuo et al., 2001), which serves particularly for diagnosing mathematical errors of children with MLD.

Apart from using standardized and diagnostic tests, informal inventories should also be adopted in the assessing process, which are especially effective for school teachers (Jena, 2013). A teacher could build up informal assessment techniques to identify children with MLD according to the mathematics curriculum used in the classroom. For example, a teacher can assess the process of metal operations through observing students’ responses in performing arithmetic calculations. Mathematical errors like incorrect labelling of axis, use of inappropriate units, over or under generalisation errors in the calculations, errors concerning place value and so on, can be easily identified through marking students’ school assignments. When formal assessment cannot be implemented due to appropriateness or other reasons, informal assessment can serve as a reliable replacement. Informal assessment inventories, especially teacher’s observation, errors manifested in school assignments, and worksheets used in the classroom, are useful indicators for identifying children with MLD (Jena, 2013). Therefore, it is always a better choice to combine formal assessment and informal assessment together.
2.2 Mathematics Errors

2.2.1 Mathematics Framework

The purpose of mathematics framework is to guide implementation of mathematics standards. It is important to have a well-developed mathematics framework before discussing types of mathematical errors. Mathematics framework sets scope and standards of mathematics knowledge, within which the identification of error types and analysis of error patterns are significant and reasonable. Otherwise, the general taking of errors in mathematics is boundless and pointless.

![Mathematics Framework Diagram]

*Figure 1: Mathematics Framework of Singapore (MoE, Singapore, 2013)*

Singapore has always got a high reputation in its mathematics education (OECD, 2009, 2012). Figure 1 shows the mathematics framework that is currently being used in Singapore. This framework emphasizes conceptual understanding, skill proficiency, and mathematical processes, and places equal emphasis on attitudes and metacognition. All these five factors are correlated (MoE, Singapore, 2013).
Mathematical concepts include numerical, algebraic, geometric, statistical, probabilistic, and analytical concepts. These concepts are related and interdependent. Children should be exposed to various learning experiences such as hands-on activities and computer-aided teaching, which would help them relate the abstract mathematical concepts with concrete experiences. In this way, children are expected to develop a deep understanding of mathematical concepts.

Mathematical processes refer to the process skills involved in the process of acquiring and applying mathematical knowledge, which includes reasoning, communication and connections, applications and modelling, thinking skills and heuristics. Children should have opportunities to apply mathematics to deal with various questions, including open-ended and real-world problems.

Metacognition, or thinking about thinking, includes the awareness of, and the ability to control one’s thinking process. Through solving non-routine and open-ended questions, discussing their solutions, thinking aloud, and reflecting on what they are doing, children can better develop metacognition awareness and strategies.

Attitudes is the affective aspects of mathematics learning, which includes beliefs about mathematics and its usefulness, interest and enjoyment in learning mathematics and so on. Children’s attitudes towards mathematics are shaped by their learning experiences. Children should be provided with activities that are designed to make the learning of mathematics fun, meaningful. They should be taught to build confidence and develop appreciation for mathematics.

Mathematical skills can be grouped into numerical calculation, algebraic manipulation, spatial visualisation, data analysis, measurement, use of mathematical tools, and estimation. In modern classroom, these skills also include manipulate computer software to learn and use mathematics. Children should be taught these skills with an understanding of the underlying mathematical principles (MoE, Singapore, 2013).

**2.2.2 Types of Mathematical Errors**

The mathematical errors refer to those pervasive errors that children make, based on the difficulties they have experienced when dealing with mathematical problems (Ketterlin-Geller
There are many types of mathematical errors and it is important to identify the types of errors before error analysis.

In Mathematical Basic Ability Diagnostic Test, nine types of mathematical errors are identified, including mathematical concepts errors, computation errors, errors caused by perceptual deficit (Kuo et al., 2001). In Australia, the Board of Studies Teaching & Educational Standards (BOSTES) encourages teachers to use evidence of learning to analyse, diagnose, and remediate errors. The following steps are therefore suggested: 1) collect evidence of learning; 2) examine children’s responses; 3) look for error patterns; 4) search exceptions to error patterns; 5) analyse error types; 6) further investigation (BOSTES, 2016).

Types of errors are categorised into two broad groups: errors for arithmetic and errors for word problem. For example, errors for arithmetic refer to types of errors in mental and written computation, which include fact errors, operation errors, and procedural errors (placement errors, incorrect steps, and missing steps). Hansen (2005) edited a book that especially investigated children’s errors in primary mathematics. In this book, errors are categorised into four groups according to the related mathematical knowledge: Number, Shape and Space, Measures, and Handling Data. Take number as an example, it covers ten subtypes of errors including counting error, place value error, ratio and proportion error. On the other hand, Desoete, Roeyers, & Clercq (2004) suggest several skills that are important for solving mathematics problems and lack of these skills would result in mathematical errors. These skills are numerical comprehension and production skills, operation symbol comprehension and production skills, number system comprehension and production skills, etc.

Based on the literature review and empirical knowledge, the following eight types of mathematics errors are recognized in this study:

**Visual-spatial errors**: errors due to lack of visual-spatial ability; the child cannot recognize numbers, expressions, quantities, and shapes; the child cannot read the time on a clock (e.g., Hansen, 2005; Desoete et al., 2004; Kuo et al., 2001). For example, the child may mistake + as ×, 3 as 8, and 6 as 9.

**Comprehension errors**: children understand the words but do not understand the whole problem or specific terms within the problem; the child cannot grasp the mathematical meaning of the problem; the child cannot retrieve information from graphs, tables, texts, or other sources (e.g., Desoete et al., 2004; Hansen, 2005; BOSTES, 2016). For example: In the
classroom, there are 5 tables with 6 students at each table, how many students are there in the classroom? The child is expected to respond 5 groups of 6 is 30, however this child drew a picture showing 2 tables, with 5 students at one table and 6 students at the other, to obtain an answer of 11. This kind of errors would be categorised as comprehension errors (adapted from BOSTES, 2016).

**Transformation errors**: children understand the problem but cannot solve the problem; the child cannot implement strategies or operations to deal with problems which can be solved by commonly used mathematical methods, simple mathematical concepts and procedures (e.g., Kuo et al., 2001; Hansen, 2005; BOSTES, 2016). For example: *Tony is thinking of a number, if he doubles the number and adds 4 he gets 18, what is the number?* The child is expected to answer the number equals 7. If the child nevertheless calculated and gave the answer 5, which indicates that this child failed in transforming the information into calculation process, then it would be called a transformation error (adapted from BOSTES, 2016).

**Relevance errors**: children cannot exclude irrelevant information (e.g., Desoete et al., 2004; Hansen, 2005). For example: *Jack has 1 brother and 3 sisters, Bill has 2 two more brothers than Jack, and how many brothers does Bill have?* Here, 3 sisters are irrelevant information. For many children with MLD, they cannot exclude irrelevant numbers. Very often they would add all the numbers together and create a wrong answer.

**Fact errors**: errors concerning mathematical facts; the child cannot recall mathematical definitions, terms, properties of numbers, geometric properties, and notations; the child does not know the order of operations in computing a combination of +, −, ×, ÷; the child does not know the functions of brackets in altering the order of operations, like “()”; the child has some incorrect beliefs regarding the magnitude of numbers; the child cannot recognize entities that are mathematically equivalent (e.g., Hansen, 2005; Mazzocco, Devlin, & McKenney, 2008; MoE, Singapore, 2013). For example, calculate 21+43, the child is expected to respond 21+43=64, if the child responds 21+43=65, which may illustrate that this child believes 1+3=5. If this kind of error happens persistently, it will be concluded as a fact error.

**Procedural errors**: incorrect operation, wrong algorithm, placement errors, incorrect steps, missing steps; the child cannot compute +, −, ×, ÷, or a combination of these with whole numbers, fractions, decimals, and integers; the child makes mistakes in the process of carrying out algorithmic procedures (e.g., Kuo et al., 2001; Hansen, 2005; MoE, Singapore,
For example, in calculating $6.325 + 13.56$, the child is expected to answer

$$
\begin{array}{c}
6.325 \\
+ 13.56 \\
\hline
19.885
\end{array}
$$

if the child responds $6.325 + 13.56 = 73.381$, then it would be identified as a procedural error (adapted from BOSTES, 2016).

**Measurement errors**: errors in terms of the measurement; the child cannot recall the relationship between measurement units; the child cannot choose appropriate units of measurements or use measuring tools; the child always omits measurement units in real-life problems (e.g., Hansen, 2005; MoE, Singapore, 2013). For example, $1$ kilometre = ( ) metres, the child is expected to respond $1$ kilometre = (1000) meters, if the child does not know or just fills the blank with a random number, this kind of error will be categorised as measurement error.

**Presentation errors**: children cannot encode the answer in a correct mathematical manner; the child cannot create equations, inequalities, geometric figures, or diagrams that model problem situations correctly; the child cannot generate equivalent representations for a given mathematical entity or relationship (e.g., Kuo et al., 2001; Hansen, 2005; MoE, Singapore, 2013). For example: A class of 30 students is to be divided into three equal-sized teams. How many students will there be in each team? The child is expected to answer 10 students.

However, if the child writes like $\frac{10}{30} \div 3$, then it would be recognized as a presentation error (adapted from BOSTES, 2016).

### 2.2.3 Cognitive Stages in Solving Mathematics Problems

Many researchers have tried to construct a cognitive model to describe the hidden stages in solving mathematics problems (e.g., Sternberg, 1969; Mayer, 2004; Anderson et al., 2016). If such endeavours can be utilized in accordance with types of mathematical errors successfully, it would benefit children with MLD a lot. For example, if a specific type of mathematical error is identified, it can then be traced back directly to deficit of the certain cognitive stage. In this way, the development of more precise and effective educational interventions can be expected.
Among these studies, Mayer (2004) claimed the four fundamental cognitive-behavioural process involved in solving mathematics problems: 1) *translating*. Children with MLD often face difficulties in translating mathematics problems into mental representation; 2) *integrating*. This stage requires integrating all mental representations of the problem as a whole; 3) *planning and monitoring*. Underachievers in mathematics always find it difficult in generalize a specific solution to broad situations; 4) *executing*. Procedural knowledge is essential at this stage and many children fail in this part.

Most recently, Anderson, Pyke, & Fincham (2016) developed a new method which employed functional MRI (Magnetic Resonance Imaging) brain activation to recognize when individuals were engaged in different cognitive stages on mathematics trials. This method was a combination of multivoxel pattern analysis and hidden semi-Markov models, the former was used to identify cognitive stages and the latter were used to recognize the durations of these stages. Four distinct stages were identified in applying this method to problem-solving tasks: *encoding, planning, solving, and responding*.

![Figure 2: Durations of four cognitive stages on mathematics trials (Cited from Anderson et al., 2016)](image)

The illustration in Figure 2 demonstrates the durations of the four distinct stages in solving mathematics problems. The arrow in the four sample questions represents the new mathematical operator the research designed and taught the participants. In each stage, the axial slice highlights brain regions in which activation in that stage is significantly greater.
than the average activation during problem solving. The results revealed that there was a significant variation of duration in problem-solving among participants and among problems.

In accordance to these stages of cognition in solving mathematics problems, the eight types of mathematical errors can be categorised into four groups based on its features: 1) encoding stage (visual-spatial errors, comprehension errors). Encoding stage requires to read and comprehend the mathematical information; 2) planning stage (transformation errors, relevance errors). Planning stage requires mental transformation and exclude irrelevant information; 3) solving stage (fact errors, procedural errors). Solving stage requires fact knowledge and correct calculation process; 4) responding stage (measurement errors, presentation errors). Responding part requires the appropriate measurement unit and the right presentation of answers.

However, if the eight types of mathematical errors can actually match with this newly developed cognitive stages in solving mathematics problems remains disputable.
2.3 Research Questions

Children with MLD are in need of educational support. However, there have been far fewer intervention programmes for children with MLD than, for example, for children with reading difficulties (Dowker, 2005). Error analysis is beneficial to provide effective educational support and to plan better intervention programmes (McGuire, 2013), but there is a lack of relevant tests which can identify error patterns (Kuo et al., 2001). Bearing this in mind, one of the main purpose of this study is to develop a test identifying error patterns. With the understanding of remarkable higher prevalence of children with MLD in Tibet and barriers teachers have been facing to provide effective support (Wang, 2008; Fang et al., 2008), this test will be utilized to analyse error patterns among children with MLD in Tibet. This is another main purpose of the study.

Therefore, the research questions this study aims to answer are:

1. How to develop a test identifying the types of mathematical errors?

2. What are error patterns among children with MLD in Tibet?
   2.1. What are the frequencies and structures of mathematical errors?
   2.2. What are the relationships between error types with gender and school types?
   2.3. Does the types of mathematical errors fit in with the stages of cognition in solving mathematics problems?
3 Research Methodology and Ethical Considerations

In this chapter, research methodology will be discussed, including research design, research participants, research tools, and data analysis methods. In addition, the ethical issues concerning this study will be examined.

3.1 Research Methodology

3.1.1 Research Design

As the nuts and bolts of a study, research method should be addressed in the first place, it can be qualitative, quantitative or a mix of both. Nowadays, mixed research methods is becoming increasingly popular in educational science because researchers could choose and combine both qualitative and quantitative methods. However, a mixed research approach is always time-consuming and requires more resources (Johnson & Onwuegbuzie, 2004). In this study, a quantitative approach will be adopted.

The underlying philosophy and worldview of quantitative research is positivism, which takes the view that the aim of research is to uncover an existing reality (Muijs, 2010). The truth is already there, the researcher needs to be as detached from the phenomena he/she is observing as possible. The methods used should maximise objectivity and minimise the bias and emotional involvement of the researcher in the research. It is therefore suggested to adapt the methods used in natural science (e.g., mathematics, physics, biology) to the social research settings (e.g., education). However, it is problematic to assume the absolute reality that can be measured completely objectively. As part of the world, it is impossible for researchers to isolate themselves from what they are studying. The social/political/economical background of the researcher, including personal beliefs, would inevitable affect the research. This has been verified by historical studies (e.g., Tobin, 1993). As a result, there are a number of revisions to extreme positivism (Robson, 2002; Muijs, 2010; Remler & Van Ryzin, 2010).

Post-positivists focus on confidence rather than absolute truth, they ask questions like “to what extent are we certain about our findings”. Experiential realists believe in the limitation
to subjectivity with the understanding that we observe by interacting with the world through our bodies. The major contention of pragmatists is that the meaning and the truth of any idea is a function of its practical outcome, the use of methods depends on what kind of research questions we want to answer.

Robson (2002) claims three types of quantitative research design, which are: 1) true experimental aims to control variables and identify cause and effect relationship by intervening and manipulating some variables, random sample required; 2) quasi-experimental has the same goal as the first one but less strong, random sample required; and 3) non-experimental collects descriptive responses and provide tentative explanations and possible hypothesis for future research, sample allocation is flexible. Some other researchers argue that the fundamental distinction should be causal design and descriptive design (Remler & Van Ryzin, 2010). The purpose of causal design is to answer “if we change X, how Y will be affected” questions, while descriptive design focuses on mirroring the real world, reflecting the real situation.

In order to answer the first research question, a test aiming for identifying mathematics errors needs to be developed. The second research question aims to analyse error patterns among children with MLD in Tibet, which covers the frequencies and structures of mathematical errors, the relationships between gender, school types and error types, and a model fit test. The only dependent variable in this study is children’s mathematical errors. The two independent variables in this study are gender and school types. Through describing and analysing these errors, patterns are expected to be discovered and relationships between dependent variable and independent variables will be revealed. A test of model fit will be done afterwards. Children’s mathematical errors already exist without researcher, the whole analysis process is non-intrusive and without controlling variables. Therefore, it is appropriate to adopt quantitative descriptive design in this study.

3.1.2 Research Participants

In the real world research, it is impossible or very difficult to measure the whole population. Take this study as an example, I cannot investigate mathematics errors of children with MLD all over Tibet. A sample is thus needed, which means taking a small number of participants from a larger population to represent that population (Mann, 2015). Two factors play key
roles through sample selection process in quantitative research, they are representativeness of the sample and size of the sample (Robson, 2002). This is because researchers usually want to generalise the results they find in their sample to the whole target population. Representativeness measures how much a sample can represent the target population, the higher the better. In achieving this, a probability sample is suggested (Remler & Van Ryzin, 2010). The size of the sample matters because if the sample is too small, it may lead a biased result and cannot reflect the target population. A larger sample could reduce the risk of biased results (Coughlan, Cronin, & Ryan, 2007).

There are several different types of sampling, including *simple random sampling*, *stratified random sampling*, *quota sampling*, *cluster sampling*, *multistage sampling*, *volunteer sampling* and *convenience sampling*. The former five sampling methods are generally regarded as probability sampling, the latter two are not (Robson, 2002; Muijs, 2010). Due to descriptive design, probability sampling is not necessarily required. Convenience sampling, as the fast, inexpensive and probably the most commonly used sampling technique, will be used in this study. The target population of this study is the children with MLD in Tibet. Two schools in Tibet are chosen according to geographical location and convenience: one is an urban school located in Lhasa (the capital city of the Tibet Autonomous Region), and another is a rural school in the countryside of Lhasa region. In each school, 15 seventh graders with MLD are selected based on formal assessment and informal assessment. Seventh grade is the transition period from elementary school to secondary school, studying children with MLD in this grade is especially valuable (Kuo et al., 2001).

### 3.1.3 Research Tools

Selecting appropriate research tools can be crucially important, as neither a high-quality research design nor sophisticated statistical analyses can make up for bad measurement. Obviously, the quality of data depends on the quality of research tools. Many instruments can be used in quantitative research to collect data, these include survey, test, systematic observations (Mann, 2015). In this study, diagnostic arithmetic test will be deployed, which aims to determine the exact nature of mathematical skill deficits (Jena, 2013). After determining which instrument to use, the following work is to design the instrument in order to answer research questions. There are two choices to achieve this goal: design a new diagnostic test or adopt a previously designed one. Developing a new test is time-consuming
and difficult, sometimes it is lack of reliability and validity (Coughlan et al., 2007). Reliability tells if a test consistently reveals the same results, and validity indicates if a test actually measures what it is intended to measure (Mann, 2015). Due to financial and time constraints, it is not suitable to design a brand new test in this study. Adopting a well-designed and proven test is a better choice. Two such tests were chosen, which are presented next.

Firstly, Kuo et al. (2001) developed a Mathematical Basic Ability Diagnostic Test especially for diagnosing mathematics errors of children with MLD in Taiwan. One thousand and ten students were tested and revealed satisfactory reliability and validity. The test is in Chinese which overcomes language barrier to use it in Tibet (the official language in Tibet is Standard Chinese). Given the similar mathematics curriculum and education system (Chiang, 2006), this test is believed to be suitable to answer the research questions. Secondly, some items from TIMSS test will be used. As one of the popular international benchmark tests, TIMSS Mathematics test has a high degree of validity and reliability, the test items are therefore generalizable and replicable (Kuncel & Hezlett, 2007). The reason to use TIMSS Mathematics rather than PISA Mathematics is that the content of TIMSS Mathematics is more suitable to the target population. Finally, several question items adapted from a journal paper (Desoete et al., 2004) as well as some items designed by the researcher will also be used.

Sullivan (2011) reminds the researchers that the identification process we used to identify children with learning disabilities must be fair and equitable so that we do not inadvertently overidentify children from some racial and ethnic groups. In this study, items used from those tests and the journal paper will be revised to be more suitable to Tibetan context. Revisions will be made based on researcher’s knowledge gained through training in mathematics pedagogy and teaching experience in Tibet. In addition, some experienced mathematics teachers in Tibet will be advised and opinions of the specialist in MLD will be addressed. Further, a pilot test will be conducted after the test is revised, a few amendments are expected to be made in order to further improve the quality of the test.

3.1.4 Data Analysis Methods
Descriptive statistics describes numerical data and inferential statistics uses data to make predictions about the target population (Mann, 2015). Descriptive statistics will be used in this study for measuring central tendency and dispersion, such as calculating mean, percentage, standard deviation, etc. This kind of descriptive information can give useful information about variables and research questions. The results of descriptive statistics will be presented via graphs and tables, which can make the results easy to read and understand. Inferential statistics, as a way to generalise findings from sample data to the whole population, will be used more often in this study. Inferential statistics serves to make judgments of the probability that an observed difference occurring in the sample if there is no relationship in the population and of how large the strength of relationship is (Muijs, 2010). For example, it will be used to calculate if there is a statistical significance between error types and gender, what the correlations between error types are, and so on.

The raw data collected by test will be carefully scored and then coded into eight error type categories. The same criteria will be applied to every participant, and children’s test papers will be scored consistently and recorded into an Excel document (Microsoft Office 2013). Each error that a child has made will be coded into one or more specific error type categories. This will be strictly done according to the definition and scope of each error type. In addition, a random recoding process will be conducted, which means the researcher will pick up several test papers, mark, score, and record error types again. In other words, some children’s work will be coded more than once and to check if there is any inconsistency. This is to guarantee all the data have been analysed in a correct and consistent manner. The data will then put into SPSS (IBM SPSS Statistics 19) and LISREL (SSI LISREL 9.2 Students) for further analysis.

SPSS is probably the most common statistical data analysis software package used in educational research (Muijs, 2010). The t-test, Mann–Whitney U test, Pearson’s r correlation coefficient, and multiple linear regression will be conducted using SPSS.

LISREL enables social science researchers to empirically assess their theories. “If data are collected for the observed variables of the theoretical model, the LISREL program can be used to fit the model to the data” (SSI, 2016). LISREL specializes in factor analysis and structural equation modelling. In this study, factor analysis and the test of model fit will be conducted using LISREL.
3.2 Ethical Considerations

Ethical considerations can be addressed at individual and at societal levels. The way that participants are affected by the research needs ethical considerations. The protection of human genetic information in Australia (ALRC, 2003) notes the ethics as:

“An accumulation of values and principles that address questions of what is good or bad in human affairs. Ethics searches for reasons for acting or refraining from acting; for approving or not approving conduct; for believing or denying something about virtuous or vicious conduct or good or evil rules.”

Four essential ethical rules are emphasized by Beauchamp and Childress (2001): 1) autonomy, which states that one should be fully informed of information about the research and feel free to choose whether or not he/she wants to take part. In this study, children with MLD were given an introduction about the study along with the meaning and purpose of the test, the information was explained in a way that children could easily understand and could relate to. This introduction was presented by the researcher, and children’s class advisors were also there in order to answer other questions that children may be concerned with. Children’s willingness were respected and their oral consent were acquired before conducting test; 2) non-maleficence, which requires the study should be harmless to participant both physically and psychologically. There was no ground to cause physical harm in this study. A few children expressed their concerns about results of the test, and they were made sure that all their personal data would be safely protected, the test papers would be coded and analysed anonymously. Another issue was about the time of taking test. The participants took the test during their self-study time (basically all seventh graders in Tibet had 40 minutes’ self-study time every day, during which there was no formal teaching activities). The test did not engage their lesson time and thus would not cause them to miss out on other learning activities; 3) beneficence, which suggests that the study should be beneficial to participants and society. Obviously, mathematics teachers could use this study to inform better teaching design for participants. The results of the test could suggest better intervention programmes for children with MLD and provide useful information to improve mathematics education in Tibet; 4) justice, which demands all participants receive same treatments during the research, no one is treated superior than others. In this study, all children were treated equally regardless of their
gender, religion and other backgrounds. There was no sign or evidence that a child had been treated unfairly.

In addition, at the beginning of this study, the research took the notification test from the Norwegian social science data services. Since there was neither directly or indirectly identifiable personal data would be registered in the project, this study was subject to notification. Further, before the data collection, local authorities in Tibet such as principals of two schools were communicated, and their oral consent were gained. For children participants and especially those with special educational needs or disabilities, parental consents are very important (Remler & Van Ryzin, 2010). Therefore, the parents of participants were communicated by class advisors and their oral consent was gained.
4 Data Analysis and Results

In this chapter, two main results of this study will be presented to respond the research questions. The first result is the Mathematics Error Pattern Identification Test (MEPIT), which is developed for identifying the types of mathematical errors. The second result is the error patterns among children with MLD in Tibet, which covers the frequencies and structures of mathematical errors, the relationships between gender, school types and error types, and a model fit test.

4.1 Result I: Mathematics Error Pattern Identification Test (MEPIT)

4.1.1 Development of MEPIT

The Mathematics Error Pattern Identification Test is mainly a combination of the MBADT and TIMSS Mathematics, and they will be introduced in the following.

Introduction of Mathematical Basic Ability Diagnostic Test (MBADT)

In order to develop a test identifying the types of mathematical errors, some of test items from MBADT will be used. The purpose of MBADT is to diagnose mathematics errors of children with MLD in Taiwan (Kuo et al., 2001), which is suitable for children who are in the transition period from primary school to secondary school. MBADT is designed based on Taiwan National Primary School Mathematics Curriculum, the content of the MBADT test is presented in Table 2.

This test uses primarily three question formats: computation items, fill-up items and constructed-response items. Choice or multiple-choice items are not deployed in the test in order to avoid guessing and casual responses. Since it is a paper-based test aiming for identifying mathematics errors, using of calculators is not allowed. The skill of using calculators is thus not assessed. In each content domain of MBADT, there are two similar questions with the same degree of difficulty. Because of this kind of duplicate design, the whole test is split into two parts – type A and type B, which enables several advantages.
Firstly, MBADT has totally 64 question items, the motivation of children to finish the whole test is doubtable. After splitting each type has 32 question items, it is thus an appropriate amount. Secondly, this design requires children to take the test twice, the results will be more reliable. Finally, the design of Type A and Type B enables the researcher to compare the result of each test, this could help to justify a solid reliability.

**Table 2: Content Domains of Mathematical Basic Ability Diagnostic Test (MBADT)**

(Adapted from Kuo et al., 2001)

<table>
<thead>
<tr>
<th>Content</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Integers</td>
<td>Compute (+, −, ×, ÷) with integers.</td>
</tr>
<tr>
<td>Decimals</td>
<td>Compute (+, −, ×, ÷) with decimals.</td>
</tr>
<tr>
<td>Fractions</td>
<td>Compute (+, −, ×, ÷) with fractions.</td>
</tr>
<tr>
<td>Elementary arithmetic operations</td>
<td>Compute a combination of four arithmetic operations.</td>
</tr>
<tr>
<td>Comparison</td>
<td>Compare the magnitude of numbers, including integers, decimals and fractions.</td>
</tr>
<tr>
<td>Problem solving</td>
<td>Solve problems set in contexts, including those in real life situations.</td>
</tr>
<tr>
<td>Measurements</td>
<td>Convert the measurement units of time, length and volume.</td>
</tr>
<tr>
<td>Geometry</td>
<td>Measure physical attributes such as length, angle, area, and volume; and to use simple formulas to calculate areas and perimeters of squares and rectangles.</td>
</tr>
<tr>
<td>Ratio</td>
<td>Identify and find equivalent ratios.</td>
</tr>
<tr>
<td>Statistical charts</td>
<td>Read, compare, and represent data from tables, pictographs, bar graphs, line graphs and pie charts.</td>
</tr>
<tr>
<td>Chance</td>
<td>Judge chances of outcomes as certain, more likely, equally likely, less likely, or impossible in general terms.</td>
</tr>
<tr>
<td>Velocity</td>
<td>Apply mathematics to solve physics problems involving velocity and distance.</td>
</tr>
<tr>
<td>Coordinate</td>
<td>Use informal coordinate systems to locate points in a plane.</td>
</tr>
</tbody>
</table>

The pre-test of MBADT was conducted in seven schools all across Taiwan (Kuo et al., 2001). The total number of participants was 257. After the pre-test, the score of each question item and the score of whole test were compared using statistical t-test. According to the t-test results and responses of the tested children, several revisions were made. Then came the official test, 1010 students in 29 schools all across Taiwan were chosen to take the test as the norm sample. The split-half reliability was acquired up to as high as 0.93 by testing subjects’ scores of Type A and Type B. The criterion validity, up to 0.74, was acquired by testing 148
sixth graders’ scores on MBADT and their mathematic achievements of first semester. In terms of content validity, the items of MBADT were developed by analysing the curriculum and content materials of six graders. A robust content validity was thus secured.

With this high reliability and validity, MBADT is believed to have a good generalizability and can be used in other contexts. Chiang (2006) compared mathematics curriculum between Taiwan and China Mainland from the first grade to the twelfth grade, the result revealed a noticeable similarity in both content domain and cognitive domain. As the curriculums seem to be similar, this curriculum based assessment tool is therefore appropriate to applied in Tibet.

**Introduction of TIMSS Mathematics Assessment**

With the purpose to develop a high-quality test identifying mathematical errors, some test items will be taken from TIMSS Mathematics. The Trends in International Mathematics and Science Study (TIMSS) is an international assessment of mathematics and science at fourth and eighth grades. First held in 1995, and then TIMSS assessment took place every four years – 1999, 2003, 2007, and 2011. TIMSS 2015 has been the latest one of the TIMSS series, which involves more than 70 countries and economies from all over the world. There are three assessment frameworks in TIMSS 2015 mathematics assessments, including TIMSS Mathematics – Fourth Grade, TIMSS Numeracy, and TIMSS Mathematics – Eighth Grade. TIMSS Numeracy is a new, less difficult mathematics assessment in TIMSS 2015, which is designed “to assess mathematics at the end of the primary school cycle (4th, 5th, or 6th grades) for countries where most children are still developing fundamental mathematics skills” (IEA, 2015, p.17). Each of these three assessment frameworks is established according to two dimensions: content dimension (which specifies in assessing the subject matter) and cognitive dimension (which specifies in assessing the thinking processes).

Content domains and the target percentages of testing time allocated to each domain of TIMSS Mathematics – Fourth Grade, Numeracy and Eighth Grade are presented in Table 3. The variation of content domains in each framework reflects different focuses in mathematics teaching. At the fourth grade, number is the priority, which constitutes half of the test items. The geometry in this grade emphasizes on shapes and measures. By contrast, number and algebra comprise a large portion at the eighth grade. It is because the pre-algebra topics taught at the fourth grade are included as part of numbers. While the data domain at the fourth grade
emphasizes on reading and displaying data, it highlights more on interpretation of data and the preliminary concept of probability at the eighth grade. The content domain of TIMSS Numeracy focuses on the fundamental mathematical knowledge and skills. About two-thirds of the test items are numbers, including whole numbers, fractions, and decimals. The remaining one-third is geometric shapes and measures.

Table 3: Target Percentages of the TIMSS 2015 Mathematics Assessment Devoted to Content Domains in Each Framework (Adapted from IEA, 2015)

<table>
<thead>
<tr>
<th>Content Domains</th>
<th>Percentages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fourth Grade</td>
<td></td>
</tr>
<tr>
<td>Number</td>
<td>50%</td>
</tr>
<tr>
<td>Geometric Shapes and Measures</td>
<td>35%</td>
</tr>
<tr>
<td>Data Display</td>
<td>15%</td>
</tr>
<tr>
<td>Numeracy</td>
<td></td>
</tr>
<tr>
<td>Whole Numbers</td>
<td>50%</td>
</tr>
<tr>
<td>Fractions and Decimals</td>
<td>15%</td>
</tr>
<tr>
<td>Shapes and Measures</td>
<td>35%</td>
</tr>
<tr>
<td>Eighth Grade</td>
<td></td>
</tr>
<tr>
<td>Number</td>
<td>30%</td>
</tr>
<tr>
<td>Algebra</td>
<td>30%</td>
</tr>
<tr>
<td>Geometry</td>
<td>20%</td>
</tr>
<tr>
<td>Data and Chance</td>
<td>20%</td>
</tr>
</tbody>
</table>

TIMSS mathematics assessments assess a range of mathematics problem-solving questions, the majority of test items require children to use cognitive applying and reasoning skills. Cognitive domains and the target percentages of testing time allocated to each domain of TIMSS Mathematics – Fourth Grade, Numeracy and Eighth Grade are presented in Table 4. In each framework, the cognitive domains remain the same, the focus is nevertheless slightly different. For instance, TIMSS Numeracy emphasizes mainly on knowing skills while TIMSS Eighth Grade stresses more on applying and reasoning skills.

Table 4: Target Percentages of the TIMSS 2015 Mathematics Assessment Devoted to Cognitive Domains in Each Framework (Adapted from IEA, 2015)

<table>
<thead>
<tr>
<th>Cognitive Domains</th>
<th>Percentages</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fourth Grade</td>
</tr>
<tr>
<td>Knowing</td>
<td>40%</td>
</tr>
<tr>
<td>Applying</td>
<td>40%</td>
</tr>
<tr>
<td>Reasoning</td>
<td>20%</td>
</tr>
</tbody>
</table>
To sum up, TIMSS series are globally recognized assessment framework that have been conducted six times on a four-year cycle. With each assessment linked to the one that preceded it, TIMSS mathematics provides regular and timely data on trends in students’ mathematics achievement. TIMSS assessment has included dozens of countries and economies all over the world, which indicates its generalizability to serve Chinese, especially Tibetan context. Further, TIMSS assessment items are valid and reliable (IEA, 2015). One important purpose of the TIMSS study is “diagnosis of common learning difficulties in mathematics and science as evidenced by misconceptions and errors” (IEA, 2015, p.93). Therefore, TIMSS mathematics items can be deployed to identify mathematics errors among Tibetan students.

**Development of Mathematics Error Pattern Identification Test**

The first draft of the new test developed in this study is mainly a combination of MBADT and TIMSS Mathematics. MBADT is a useful and proven test with high reliability and validity. It is nevertheless a little bit out of date since it was designed in 2001. Consequently, several revisions are made based on researcher’s professional knowledge and teaching experience. On the other hand, TIMSS assessment has been conducted all over the world and its test items are valid and reliable. That gives good reason to deploy TIMSS Mathematics items into the new test. In addition, according to teaching experience in Tibet, the researcher developed a few question items as well. The constitution of the first draft is shown in Table 5. The new test will be called Mathematics Error Pattern Identification Test, referred to as MEPIT. MEPIT consists of two types of test with similar difficulty: Type I and Type II. Each type has 30 items and children are expected to finish the test in 40 minutes.

*Table 5: Constitution of the first draft of MEPIT*

<table>
<thead>
<tr>
<th>MEPIT Types</th>
<th>Percentages</th>
<th>Developed by the researcher</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MBADT</td>
<td>TIMSS Mathematics</td>
</tr>
<tr>
<td>Type I</td>
<td>73.3%</td>
<td>23.4%</td>
</tr>
<tr>
<td>Type II</td>
<td>73.3%</td>
<td>23.4%</td>
</tr>
</tbody>
</table>

After the completion of the first draft of MEPIT, the copies of Type I and Type II were sent to two seasoned mathematics teachers in Tibet for comments and suggestions. The expert in the field of MLD was also advised. Their suggestions were concluded as followings:
1) The layout of the test should be redesigned in order to make it easier for children to respond. The blank should be big enough to record all the information that children have written.

2) Some expressions in test questions should be revised to fit into Tibetan context. For instance, it would be better to adopt Tibetan names into mathematical word problems.

3) Few mathematics terms used in test items should be replaced with most commonly used mathematics terms in China. For example, instead of using “甲、乙、丙、丁、戊……” (A traditional Chinese way to represent a sequence), using of “A, B, C, D, E …” would be easier for children to understand.

The defects mentioned above were amended, question items about ratios were deleted because plenty of targeted children had not learned about this concept. In addition, some question items from a journal paper (Desoete et al., 2004) were used in the test. These question items were cleverly designed and purposeful in identifying certain types of mathematical errors among children with MLD.

The Final Draft of MEPIT

After several revisions, the final draft of MEPIT has been developed. As presented in Table 6, items of the final draft of MEPIT are from four sources.

Table 6: Constitution of the final draft of MEPIT

<table>
<thead>
<tr>
<th>MEPIT Types</th>
<th>MBADT</th>
<th>TIMSS Mathematics</th>
<th>Adopted from journal paper</th>
<th>Developed by the researcher</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type I</td>
<td>62.7%</td>
<td>18.7%</td>
<td>9.3%</td>
<td>9.3%</td>
</tr>
<tr>
<td>Type II</td>
<td>62.7%</td>
<td>18.7%</td>
<td>9.3%</td>
<td>9.3%</td>
</tr>
</tbody>
</table>

Note. Journal paper refers to *Children with mathematics learning disabilities in Belgium* (Desoete et al., 2004)

The details of the final draft of MEPIT are presented in Table 7. All test items of Type I and Type II are designed to be identical to each other in identifying the same types of mathematical errors. Eight types of mathematics errors are recognized in this study: ① visual-spatial error; ② comprehension error; ③ transformation error; ④ relevance error; ⑤ fact error; ⑥ procedural error; ⑦ measurement error; ⑧ presentation error.
Table 7: Details of the final draft of MEPIT

<table>
<thead>
<tr>
<th>Items</th>
<th>Description</th>
<th>Sources</th>
<th>Revisions</th>
<th>Most Possible Error Types</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. 1</td>
<td>Integers (addition)</td>
<td>MBADT</td>
<td>change of numbers</td>
<td>①⑤⑥</td>
</tr>
<tr>
<td>No. 2</td>
<td>Integers (multiplication)</td>
<td>MBADT</td>
<td>change of numbers</td>
<td>①⑤⑥</td>
</tr>
<tr>
<td>No. 3</td>
<td>Integers (subtraction)</td>
<td>MBADT</td>
<td>intact</td>
<td>①⑤⑥</td>
</tr>
<tr>
<td>No. 4</td>
<td>Integers (division)</td>
<td>MBADT</td>
<td>intact</td>
<td>①⑤⑥</td>
</tr>
<tr>
<td>No. 5</td>
<td>Integers (division, more digits)</td>
<td>MBADT</td>
<td>intact</td>
<td>①⑤⑥</td>
</tr>
<tr>
<td>No. 6</td>
<td>Integers (a combination of four arithmetic operations)</td>
<td>Adopted from journal paper</td>
<td>intact</td>
<td>①⑤⑥⑧</td>
</tr>
<tr>
<td>No. 7</td>
<td>Decimals (addition)</td>
<td>MBADT</td>
<td>intact</td>
<td>①⑤⑥</td>
</tr>
<tr>
<td>No. 8</td>
<td>Decimals (subtraction)</td>
<td>MBADT</td>
<td>intact</td>
<td>①⑤⑥</td>
</tr>
<tr>
<td>No. 9</td>
<td>Decimals (multiplication)</td>
<td>MBADT</td>
<td>intact</td>
<td>①⑤⑥</td>
</tr>
<tr>
<td>No. 10</td>
<td>Decimals (division)</td>
<td>MBADT</td>
<td>intact</td>
<td>①⑤⑥</td>
</tr>
<tr>
<td>No. 11</td>
<td>Fractions (addition)</td>
<td>MBADT</td>
<td>intact</td>
<td>①⑤⑥</td>
</tr>
<tr>
<td>No. 12</td>
<td>Fractions (division)</td>
<td>MBADT</td>
<td>intact</td>
<td>①⑤⑥</td>
</tr>
<tr>
<td>No. 13</td>
<td>Problem Solving (integers)</td>
<td>Adopted from journal paper</td>
<td>intact</td>
<td>①②③</td>
</tr>
<tr>
<td>No. 14</td>
<td>Problem Solving (real-life problems containing irrelevant information)</td>
<td>Developed by the researcher</td>
<td>none</td>
<td>②③④</td>
</tr>
<tr>
<td>No. 15</td>
<td>Comparison (integer, faction, decimal)</td>
<td>MBADT</td>
<td>intact</td>
<td>③⑤⑧</td>
</tr>
<tr>
<td>No. 16</td>
<td>Word Problems (Type I: buying pens in the stationer; Type II: age problem)</td>
<td>MBADT</td>
<td>change of amounts and cultural background</td>
<td>②③④⑧</td>
</tr>
<tr>
<td>No. 17</td>
<td>Word Problems (Type I: shopping; Type II: conversion of numbers)</td>
<td>MBADT</td>
<td>change of amounts and cultural background</td>
<td>②③④⑧</td>
</tr>
<tr>
<td>No. 18</td>
<td>Real-life Problems (Type I: read the clock; Type II: calculate the time)</td>
<td>Developed by the researcher</td>
<td>none</td>
<td>①④⑦⑧</td>
</tr>
<tr>
<td>No. 19</td>
<td>Measurements (converting the measurement units: length)</td>
<td>MBADT</td>
<td>intact</td>
<td>⑦</td>
</tr>
<tr>
<td>No. 20</td>
<td>Measurements (converting the measurement units: weight)</td>
<td>MBADT</td>
<td>change of amounts</td>
<td>③⑦</td>
</tr>
<tr>
<td>No. 21</td>
<td>Measurements (converting the measurement units: time)</td>
<td>MBADT</td>
<td>change of amounts</td>
<td>③⑦</td>
</tr>
<tr>
<td>No. 22a</td>
<td>Geometry (measure physical attributes: angles)</td>
<td>TIMSS Mathematics</td>
<td>intact</td>
<td>①④⑤⑦⑧</td>
</tr>
<tr>
<td>No. 22b</td>
<td>Geometry (compare physical attributes: angles)</td>
<td>TIMSS Mathematics</td>
<td>intact</td>
<td>①④⑦⑧</td>
</tr>
<tr>
<td>No. 23</td>
<td>Geometry (draw geometric figures: triangle, rectangle, square, and circle)</td>
<td>MBADT</td>
<td>intact</td>
<td>⑤⑦⑧</td>
</tr>
<tr>
<td>No. 24</td>
<td>Geometry (Type I: measure circumference of a triangle; Type II: measure circumference of a circle)</td>
<td>Developed by the researcher</td>
<td>none</td>
<td>⑤⑦⑧</td>
</tr>
<tr>
<td>No. 25</td>
<td>Geometry (Type I: measure area of a rectangle; Type II: measure area of a triangle)</td>
<td>MBADT</td>
<td>change of amounts</td>
<td>⑤⑦⑧</td>
</tr>
<tr>
<td>No. 26</td>
<td>Problem solving (arithmetic problem)</td>
<td>Adopted from journal paper</td>
<td>change of amounts and cultural background</td>
<td>②③④⑧</td>
</tr>
<tr>
<td>No. 27a</td>
<td>Statistical charts (read data from bar graphs)</td>
<td>TIMSS Mathematics</td>
<td>change of cultural background</td>
<td>①②④⑧</td>
</tr>
<tr>
<td>No. 27b</td>
<td>Statistical charts (compare data from bar graphs)</td>
<td>TIMSS Mathematics</td>
<td>change of cultural background</td>
<td>①②④⑧</td>
</tr>
</tbody>
</table>
When designing the MEPIT, one primary rule was to have all error types distributed equally in the test. Figure 3 illustrates the frequencies of all possible error types, which are calculated based on the most possible error types of each item as shown in Table 7. However, the absolute equal distribution is impracticable. In addition, when a child is taking the test, he/she might make some errors beyond the pre-determined hypothesis. Therefore, this frequency of all possible error types is just an experience-based estimation.
Note. This is the estimated frequencies of all possible error types. In fact, the errors children made during the test can be various and unexpected.

Sample Question Items of MEPIT

Some sample question item will be presented in the following. These questions come from four sources: MBADT, TIMSS mathematics, a journal paper and some items developed by the researcher.

Test item 6 is adapted from MBADT (Kuo et al., 2001). In type I, The question requires to calculate “4 × 2 − 30 ÷ 5 + 8 = ____”. Several types of mathematics errors can be identified in solving this kind of questions. Firstly, if the child cannot read numbers composed of one or more digits (e.g., read 8 as 3, 30 as 13) or does not recognize operation symbols (e.g., confuse × as −, + as ×), it will be recognized as a visual-spatial error. This kind of error often leads to mistakes such as 15 × 9 = 24. Secondly, if the child consistently believes that “8 − 6 = 3” or “30 ÷ 5 = 7” and thus leads to incorrect answer, it will be recognized as a fact error. Another fact error would be that the child does not know the order of operations. For instance, after calculating the multiplication and division, the child made a mistake like this, “8 − 6 + 8 = 8 − 14”, this would also be included as a fact error. Thirdly, if the child always makes mistakes with the place of a number on a number line, has difficulties in dealing with carries, borrows and units, it will be categorised as procedural errors. Finally, if the child has difficulties in representing the answer in the correct mathematical manner. For instance, if all the procedures are correct and the answer is wrong, it would be concluded as a presentation error. In type II, The question requires to calculate “72 ÷ 6 − 4 × 2 + 9 = ____”. The targeted error types are the same.

Test item 13 is adapted from an article by Desoete et al. (2004). In type I, the question is: 15 is 9 more than ____? A metal transformation skill is required to solve this problem. The key words in the problem (e.g., “more”) should be transformed into calculation procedure (e.g., “addition”). The correct answer is “6”, because 15 is 9 more than 6 (“9 + 6 = 15”). Without transformation ability would lead to blind calculation or number crunching. This superficial approach brings about the wrong answer like “24”. This kind of error will be categorised as a transformation error. In type II, the question is: 27 is 3 less than ____? Which aims for the same type of error.
Test item 14 was developed by the researcher. In type I, the problem is: *Sonam has 1 sister and 2 brothers, Dorjee has 3 more sisters than Sonam does, how many sisters do Dorjee have?* The answer is pretty simple, “1 + 3 = 4”, thus Dorjee has 4 sisters. It is noticeable that “2 brothers” in this question is irrelevant information, which means it is useless in creating the correct answer. Ordinary children can easily exclude the irrelevant information and get the right answer. However, for children with MLD, they can have difficulty ignoring and not using information (e.g., “2 brothers”) in the problem. They might add up all the numbers in the question (“1 + 2 + 3 = 6”) and create the incorrect answer. This is because children with MLD often think that they should use all the numbers appeared in the question to solve a mathematical problem. This question is used especially to identify relevance errors. The problem in type II is similar to this one.

Test item 18 was also developed by the researcher. The question in type I requires the child to read the time on the clock. It is a real-life problem. Ordinary children can directly answer the question. However, children with MLD may experience two kinds of difficulties. Firstly, they might have difficulties in retrieving information from the clock and thus they cannot read the time on the clock. This will be categorised as a visual-spatial error. Secondly, they might know the time on the clock but they cannot express it in the correct way. This kind of error will be categorised as a presentation error. The question in type II is also concerned with the time and aims for the same types of errors.

Test item 27 is adapted from TIMSS mathematics, which has two sub questions. The question items in Type I and Type II are very similar to each other. Both question items require the child to recognize bar chart in the first place. If the child fails at this step, it will be a visual-spatial error. And then, the child is asked to answer a question concerning the quantities on the bar chart. If the child cannot retrieve information from bar chart and thus makes a mistake in responding the question, it will be a comprehension error. In addition, since there are several bars on the bar chart, if the child uses irrelevant information or try to add all the bars together, it will be categorised as a relevance error. At last, if the child solves the problem but makes an error in presenting the answer, it would be a presentation error.

Test item 29 is also adapted from TIMSS mathematics. In type I, the child is required to locate a point on the plane according its coordinate. In type II, the child is asked to find the coordinate of a point on the plane. Both questions are concerned with points and coordinates, visual-spatial ability is thus essential in the whole problem-solving process. In addition, if the
child has difficulties in comprehending mathematical information from the coordinate, point, and plane, it will be a comprehension error. Further, if the child understands the information but cannot transform and solve the problem, it will be a transformation error. At last, if the child comes up with the correct answer but presents it in a wrong way, it will be categorised as a presentation error.
4.1.2 Pilot Test and Official Test

Pilot Test

A pilot test is a small-scale trial, which serves as a rehearsal for the official test. In pilot testing, a few participants from targeted group take the test and give their feedback on the test (Thabane et al., 2010). Schade (2015) suggests several benefits to go through a pilot test. Firstly, it is a check to see if the study has been properly prepared. Are the test papers well printed? Are there any printing errors (especially in printing mathematics equations, graphs, tables)? Is testing place ready and functional? A single mistake could cause serious problems in the official test. In this way, pilot testing will help identify potential practical problems in the following research procedure. Secondly, test items which may be completely clear to researcher can be confusing or misleading to participants. Some test items may even contain errors that researcher is unconscious of. Identifying and remedying problems with test items before official test will help guarantee a successful test and trustworthy results. Finally, even though the length of time for participants to take the test is critical for a successful study, determining the appropriate amount of time is always not an easy task. Pilot testing can help researcher to have a better estimation regarding how long participants may need to take the test, allowing the test results to be more effective.

Procedures of Pilot Testing

After the completion of the second draft of MEPIT, a pilot test was conducted in early August. Two primary school graduates in Tibet (they would go to secondary school in late August as seventh graders), were contacted to participate the pilot testing. These two children, one boy and one girl, were recommended by their mathematics teacher who had been explained the purpose of study and the requirement of participants in detail. They did quite well in Chinese literacy, but always ranked behind the rest of the class in mathematics. They were communicated afterwards and showed a lot of interest. Oral permissions from their parents were also obtained.

Before pilot testing, the participants were carefully explained the purpose of the study as well as the requirements of taking the test. The most important requirement was to keep records of all their procedures during the test. To do this, the two kids were instructed to write every procedure onto the blank part of test papers because all the details of their responses were
essential in error analysis. After that, they took the test in a quiet classroom at the same time. First to take Type I of MEPIT, two days later to take Type II of MEPIT. When designing the test, researcher estimated that children should finish the test in 40 minutes. With the aim to have a better estimation, children were given flexible time to take the test in pilot testing. The researcher was present and instructed all the process of pilot testing.

Results of Pilot Testing

After pilot testing, children’s feedback were collected. The content of MEPIT test was within their knowledge spectrum. No test item had been reported as new knowledge to them. No major problems with the test instructions were pointed out. Item 14 in Type I was reported to be unclear, this item was revised to be more straightforward after careful consideration. In Type I, item 18, the picture of clock was not very clear after printed out, this picture was then replaced by a higher quality one. The boy said that there was no enough space for certain questions, the blank space for those questions were widened. No other formatting and typographical errors and/or issues were identified. The researcher found that the quantities of item 14 in Type II and item 27 in Type I were not identifiable enough to recognize related mathematics errors, these items were then improved after piloting testing. All the issues with the test items and forms were addressed in the later revisions. The researcher also observed that both girl and boy finished the test in around 40 minutes, just as estimated. It was therefore suggested to set 40 minutes as the standard length of time for MEPIT test. Another thing to mention was that these two children would not attend official test.

Official Test

Choosing the Participants

Participants were chosen from two schools in Tibet. One is an urban school located in Lhasa (the capital of the Tibet Autonomous Region), and another is a rural school in the countryside of Lhasa region. The following criteria were used in the selecting process: 1) participants were all first year seventh graders; 2) participants’ mathematics scores from previous examination (Secondary School Entrance Examination, organized by local government) were 2 or more SD below norm; 3) participants’ Chinese literacy scores from previous examination were average or above average; 4) participants were recognized by their mathematics teachers as being having difficulties in learning mathematics; 5) participants themselves were willing
to take the test. In each school, the researcher advised the mathematics teachers, calculated the mean and standard deviations of mathematics and Chinese literacy scores from the previous examination, talked with the targeted children, and then selected the participants. The participants’ parents were also informed via phone calls and their oral permissions were gained. The participants’ demographic background information is shown in Table 8.

Table 8: Participants’ demographic background information

| Age           | 11 years old (n = 1), 3.33% |
|              | 12 years old (n = 3), 10.00% |
|              | 13 years old (n = 14), 46.67% |
|              | 14 years old (n = 10), 33.33% |
|              | 15 years old (n = 2), 6.67% |
|              | M=13.30, SD=0.88           |
| Grade        | first year of seventh grade (n = 30), 100% |
| Gender       | girls (n = 20), 66.67%     |
|              | boys (n = 10), 33.33%      |
| School type  | urban school (n = 15), 50.00% |
|              | rural school (n = 15), 50.00% |
| Ethnicity    | Tibetan (n = 29), 96.67%   |
|              | Han Chinese (n = 1), 3.33% |

Conducting the Test

The whole process of the official test took more than three weeks. The test was first done in an urban school and then in a rural school. In each school, participants were gathered together and instructed by the researcher before the test. In the following testing, a local mathematics teacher was there to invigilate. The researcher was not present in order to guarantee objectivity. Participants first took the Type I test, and then Type II test. The procedure of official testing is illustrated in Figure 4.

Muijs (2010) claims that if children are given too little time between the first test and the second test, they may remember how they solved the problem last time and will solve in the same way this time. This would lead to an overestimation of test’s reliability. On the other hand, if children are left too much time, they might have learned and improved in this period. This could lead to underestimating the reliability of the test. It is often suggested to use one to two weeks as test interval. Two days was obviously too short in pilot testing. Therefore, the one-week interval between two tests was used in the official test.
After test papers were gathered, they were stored safely in big envelopes. When the data collection in two schools completed, the researcher started to code the data anonymously. All the test papers, both Type I and Type II, were carefully marked. The scores of the whole test items were calculated and saved. Every time when a specific error type was identified, it would be analysed and recorded. The raw data was first stored in an Excel document, and then put into SPSS and LISREL for further analysis. The computers used were private and safe.
4.1.3 Validity and Reliability

“Reliability and validity are ways of demonstrating and communicating the rigour of research processes and the trustworthiness of research findings” (Roberts, Priest, & Traynor, 2006, p.41). These are two most important features of any instrument. Take test as an example. If it is a well-established test which has been used many times and the new study will not change it in any way, then the reliability and validity has been pre-determined. Nonetheless, if this test will be adapted or combined with other test to better suits the new study then the previous reliability and validity are not applicable, the researcher has to construct the reliability and validity again (Coughlan, Cronin, & Ryan, 2007). This is the case of MEPIT test, the following will show how the reliability and validity is established.

Validity

Validity measures if a test actually shows what it is intended to show (Mann, 2015). Or in Muijs’ (2010) words, are we measuring what we want to measure? He also points out that there are three types of validity: content validity, criterion validity and construct validity.

Content validity is concerned with whether or not the content of the manifest variables (e.g., items of MEPIT test) is appropriate to measure the latent variables (e.g., children’s error types), it refers to the relevance and representativeness of test items (Muijs, 2010; Roberts et al., 2006). Firstly, eight types of errors were developed through extensive literature review. The test items were adapted from reliable and trustworthy resources (Kuo et al., 2001; IEA, 2015; Desoete et al., 2004) and some of which had been revised to better identify error types. Secondly, the researcher had four years mathematics teaching experience in Tibet, several test items were developed by the researcher himself. The draft of MEPIT test had been scrutinized by experienced mathematics teachers in Tibet. Opinion of relevant scholar in this field was also advised. Finally, a pilot test was conducted. Responses from targeted children suggested a satisfactory content validity.

Criterion validity makes an assumption that scores of the newly-developed test should be able to predict or at least be related to scores of other test which measures the similar concept or phenomenon (Muijs, 2010; Roberts et al., 2006). Therefore, the scores of MEPIT test should be correlated to other mathematics test. During official test, the scores from participants’ previous mathematics examination (Secondary School Entrance Examination, results in
points) were collected. As shown in Table 9, using Pearson’s r correlation coefficient to calculate the relationship between previous score and score of MEPIT (the total score of correct answers, results in points), a satisfactory criterion validity is achieved.

Table 9: The correlations between previous score and score of MEPIT

<table>
<thead>
<tr>
<th>Previous Score</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Score of Type I</td>
<td>0.45*</td>
</tr>
<tr>
<td>Score of Type II</td>
<td>0.42*</td>
</tr>
<tr>
<td>Score of MEPIT</td>
<td>0.48**</td>
</tr>
</tbody>
</table>

* correlation is significant at the 0.05 level (2-tailed), ** correlation is significant at the 0.01 level (2-tailed). 0.3 < r < 0.5, moderate correlation.

Construct validity is related to demonstrating the relationship between the internal structure of a test and the concept it is measuring (Muijs, 2010; Roberts et al., 2006). In statistics, factor analysis is a way to demonstrate construct validity. Using confirmatory factor analysis to measure relationships between error types and children’s error patterns, the results are illustrated in Figure 5.

Figure 5: Factor analysis of relationships between error types and children’s error patterns
The value of Chi-Square suggests that the data fits the model to some extent. However, RMSEA (Root Mean Square Error of Approximation), another index of factor analysis, is not small enough to guarantee a good model fit (a good fit requires RMSEA < 0.05). This suggests that the construct validity of MEPIT test remains improvable. Another reason could be that factor analysis is sensitive to the size of sample, a small sample in this study ($n = 30$) would cause a poor model fit (Albright, 2006).

**Reliability**

Reliability measures if a test consistently reveals the same results, which refers to the repeatability and internal consistency of research tools (Mann, 2015; Roberts et al., 2006).Muijs (2010) puts it more specifically, reliability is the extent to which test scores are free of measurement error. There are two main types of reliability: repeated measurement reliability and internal consistency reliability.

Repeated measurement reliability refers to the ability of test to measure the same thing at different times (Muijs, 2010). Best effort had been made to guarantee that each item of Type I and Type II was on the same knowledge and difficulty level in a way to measure the same types of mathematics errors. In order to reduce carryover error, the participants were given one week’s interval to take two tests. The correlation coefficient between two test scores was 0.726 ($r > 0.7$, significant at the 0.01 level), which offered reasonable reliability for research purposes.

Internal consistency reliability measures how homogeneous the items of a test are or how well they measure a single construct (Muijs, 2010). In statistics, split half reliability and coefficient alpha are often used to calculate internal consistency reliability. The same statistics were applied to MEPIT, the results suggested a strong internal consistency reliability (split-half coefficient = 0.84, Cronbach’s alpha = 0.84).
4.2 Result II: Error Pattern Analysis

4.2.1 Coding of Error Types

In this section, the coding process of each error type will be illustrated by presenting errors excerpted in children’s work, describing these errors in detail, and categorising errors to the corresponding error type. Examples will be given to each error type.

1) Visual-spatial errors

Visual-spatial errors occur when children cannot recognize or draw geometric shapes accurately, or read the time on a clock correctly. Visual-spatial errors also involve confusions among numbers like 3 and 8, letters like C and G, operations like + and ×, and so on. During the coding process, the researcher found that nearly half of children’s visual-spatial errors were made due to carelessness.

Example 1:

<table>
<thead>
<tr>
<th>MEPIT Type II, item 5. This child mistook 29 as 25 in setting up the written algorithm, which indicates that the kid might have some difficulties in recognizing 9 and 5 in some situations.</th>
<th>MEPIT Type II, item 29. The problem requires to locate the position of C8, this child actually located G8. This error suggests a confusion of C and G in this child’s cognition.</th>
</tr>
</thead>
</table>
| 5. $145 \div 29 = 5.8$ | 29. 这是卓玛家所在的乡镇地图。下面标出学校的位置在 A5，市场的位置在 F4。桑旦家的位置在 C8，请在图中标出桑旦家的位置。

25/145

\[ \overline{125} \]

\[ \overline{200} \]

\[ \overline{200} \]

a | 8

\| 7

\| 6

\| 5

\| 4

\| 3

\| 2

\| 1

学校

市场

A B C D E F G H I | \[ C \]

\[ \circ \]
Example 2:

| MEPIT Type I, item 18. This child could not recognize the time on the clock. | MEPIT Type I, item 23. The problem requires to draw a rectangle and a circle. This child drew a rectangle that resembles a square and could not draw a circle at all. |

2) Comprehension errors

Comprehension errors occur when children with MLD understand the words in question items but do not understand the overall problem. When children cannot grasp the overall mathematical meaning of the problem, or cannot retrieve quantitative data from figures or tables, they would make errors. And this kind of errors will be categorised as comprehension errors.

Example 1:

| MEPIT Type I, item 18. It is a mathematics word problem requiring an equation for calculating the total price in a stationery shop. According to the written response, this child knew the words, but did not understand overall what this problem is asking. | MEPIT Type II, item 18. It is a two-step mathematics word problem asking for an equation and then an answer. Apparently, this child did not understand the requirement of this problem. |
Example 2:

MEPIT Type I, item 27. The bar chart illustrates 4 pens, 5 pencils, and 8 gel pens. This child nevertheless read 3 pens and 4.5 pencils on the chart, which indicates she actually could not acquire mathematics information correctly from the bar chart.

MEPIT Type I, item 29. The problem requires to find the coordinate of the given position. This child failed to give the correct coordinate, which suggests she could not retrieve mathematics information from this kind of grid graph effectively.

Note. This will not be categorised as visual-spatial error because this child did quite well in reading the time on the clock, drawing geometric graphs, and locating a specific point on the grid graph (MEPIT Type II, item 29).

Example 3:

MEPIT Type II, item 30. The problem asks how the graph will look like if it gets rotated \( \frac{1}{4} \) circle anticlockwise. And then there are four choices. Children are expected to choose an answer from four choices directly. This child however made a description to each choice in his own language. Although all these descriptions were correct, he did not actually answer this question. The most possible reason was that he had no idea about the terminology - “anticlockwise”.
3) **Transformation errors**

Transformation errors happen due to the poor transformation and planning abilities in solving mathematics problems. When children have grasped the mathematical meaning of the problem however cannot decide right methods to solve the problems, they will make this kind of errors. These errors will be categorised as comprehension errors.

These errors also manifest when children cannot implement strategies or operations to deal with problems which can be solved by commonly used mathematical methods, simple mathematical concepts and procedures. In one word, children understand the problem but cannot solve the problem correctly.

Example 1:

**MEPIT Type I, item 13.** 15 is 9 more than ___? The key word “more” in the problem should be transformed into planning calculation process – addition or subtraction, and be cautious about the subtle information – which is more than which? Apparently this child failed in this procedure.

**MEPIT Type II, item 13.** 27 is 3 less than ___? The key word “less” in the problem should be transformed into planning calculation process – subtraction or addition, and be cautious about the subtle information – which is less than which? This child failed in this procedure.

*Note.* These two examples are excerpted from two different children.
Example 2:

MEPIT Type I, item 26. Although the child knew what this question was asking, she failed in transforming the information into the appropriate calculation process. She gave the wrong answer.

MEPIT Type II, item 26. The child understood the overall problem. However she failed in determining the appropriate operations to solve the problem.

Note. These two examples are excerpted from two different children.

Example 3:

MEPIT Type II, item 18. The child understood that he was required to calculate the time. However he failed to transform the time into numbers and thus to make the effective calculation.

MEPIT Type I, item 30. The child understood that she was supposed to find the length of the curved line. However she could not find an effective strategy to solve the problem. Thus she turned to guess a familiar number.

Note. These two examples are excerpted from two different children.
4) **Relevance errors**

Relevant error belongs to “small minority” in the family of error types. In primary and secondary education, most children can easily ignore the irrelevant information and get the right answer. However, some children with MLD nonetheless have no idea to exclude the irrelevant information. They would always put all the numbers in the equation and thus create the wrong answer. These errors will be categorised as relevance errors.

Example 1:

<table>
<thead>
<tr>
<th>MEPIT Type I, item 14.</th>
<th>MEPIT Type II, item 14.</th>
</tr>
</thead>
<tbody>
<tr>
<td>This question contains irrelevant information (“2 brothers”) which should be excluded in the calculation process. This child however added all the numbers together and created the wrong answer.</td>
<td>This question involves irrelevant information (“10 cattle”) which is useless in calculation. This child put all the numbers in the equation and thus got an incorrect answer.</td>
</tr>
</tbody>
</table>

*Note.* These two examples are excerpted from two different children.

Example 2:

<table>
<thead>
<tr>
<th>MEPIT Type I, item 27.</th>
<th>MEPIT Type I, item 27.</th>
</tr>
</thead>
<tbody>
<tr>
<td>This child read correctly the information on the bar chart (4 pens, 5 pencils, and 8 gel pens). She also understood the meaning the problem, she got second part of the problem right. However, when dealing with the first part of the problem, she did not exclude the irrelevant information. She added all the numbers together and created the wrong answer.</td>
<td>This child read correctly the information on the bar chart (4 pens, 5 pencils, and 8 gel pens). She also understood the meaning the problem, she got second part of the problem right. However, when dealing with the first part of the problem, she did not exclude the irrelevant information. She added all the numbers together and created the wrong answer.</td>
</tr>
</tbody>
</table>
5) **Fact errors**

Fact errors occur due to the incorrect beliefs regarding mathematical facts, or even unaware of these facts. These errors happen when children fail to retrieve mathematics definitions, terms, properties of numbers, geometric properties, or notations from their memory. Through analysing children’s work these errors can be identified, such as wrong beliefs concerning the magnitude of numbers, unaware of the functions of brackets in altering the order of operations (e.g., “()” and “[”), arbitrary use of formulas in calculating circumference and area of geometric graphs. They will then be categorised as fact errors.

Example 1:

<table>
<thead>
<tr>
<th>MEPIT Type II, item 12. Apparently, this child did not know the meaning of division and multiplication. Another possibility is that she did not know the mathematical fact: when she wanted to divide a fraction, she needed to multiply its reciprocal rather than itself.</th>
<th>MEPIT Type I, item 11. This child succeeded in transforming fractions into common denominator. However, she could not recall the whole procedure to implement fraction addition. Another possibility is that she had the belief that $\frac{3}{6} + \frac{4}{6} = 7$.</th>
</tr>
</thead>
<tbody>
<tr>
<td>12. $\frac{4}{7} \div \frac{2}{3} = \frac{8}{21}$</td>
<td>11. $\frac{1}{2} + \frac{2}{3} = 7$</td>
</tr>
<tr>
<td>$\frac{4}{7} \times \frac{3}{3} = \frac{8}{21}$</td>
<td>$\frac{3}{6} + \frac{4}{6} = 7$</td>
</tr>
</tbody>
</table>

Example 2:

| MEPIT Type II, item 15. The response suggests that this child had wrong beliefs in measuring magnitude of numbers, or she did not know the meaning of “>”.
<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>15. 比较 $3$, $3.5$, $3\frac{1}{3}$ 的大小。</td>
</tr>
<tr>
<td>$3 &gt; 3\frac{1}{3} &gt; 3.5$</td>
</tr>
</tbody>
</table>
MEPIT Type I, item 15. Base on the written response, this child had no sense regarding which number is bigger and which number is smaller.

Note. These two examples are excerpted from two different children.

Example 3:

MEPIT Type II, item 24. According to the written response, this child did not know the formula for calculating circumference of the circle. It is not a comprehension error because he did know what he needed to calculate, he just forgot the formula.

MEPIT Type II, item 22. The first part of the question asks which one is the acute angle, the second requires to arrange these angles in ascending order. This child did know the magnitude of angles, however she failed in the first part. This suggests that she did know the concept of “acute angle”.

Note. These two examples are excerpted from two different children.

6) **Procedural errors**

When children with MLD doing calculations, you can always notice incorrect operation, incorrect orders of operations, wrong algorithm, placement errors, incorrect steps, and missing steps. They may experience severe difficulties in computing +, −, ×, ÷, or a combination of these calculations. The errors made during the process of carrying out algorithmic procedures will be categorised as procedural errors.
<table>
<thead>
<tr>
<th>Example 1:</th>
<th>Example 2:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MEPIT Type I, item 9.</strong> This child made a placement error when dealing with multiplication of decimals.</td>
<td><strong>MEPIT Type I, item 10.</strong> When calculating division of decimals, this child made another placement error.</td>
</tr>
<tr>
<td><strong>Example 2:</strong></td>
<td></td>
</tr>
<tr>
<td><strong>MEPIT Type II, item 5.</strong> This child used wrong algorithm to calculate the division of integers.</td>
<td><strong>MEPIT Type I, item 12.</strong> When dealing with division of fractions, this child implemented incorrect operation and leaded to a weird answer.</td>
</tr>
</tbody>
</table>
Example 3:

| 7. 5.36+4.57=10.03 | 6. \[ \begin{array}{c}
\frac{12}{6} - \frac{4}{2} - \frac{12}{0} - \frac{12-4+9}{8+9} \\
\end{array} \]

MEPIT Type I, item 7. This child made a mistake (a carrying error from low-order to high-order) in the process of carrying out adding procedures. It is not a fact error because he knew how to do every single addition.

MEPIT Type II, item 6. When dealing with a combination of +, −, ×, ÷, this child got every single step correctly. However she made a procedural mistake in creating the final answer.

Note. These two examples are excerpted from two different children.

7) Measurement errors

Measurement errors occur due to poor knowledge in measurement units. When children do not know the relationship between measurement units, cannot choose appropriate units of measurements or use measuring tools, these errors will happen. Some children always misuse or omit measurement units in real-life problems. These errors will be categorised as measurement errors.

Example 1:
MEPIT Type I, item 19, 20, and 21. These questions focus on converting units of measurement: meter – centimetre, kilogram – gram, hour – minute. This child failed all these questions.

According to the written response, this child failed in neither converting meters to centimetres or hours to minutes. She succeeded in converting kilograms to grams.

Note. These two examples are excerpted from two different children.

Example 2:

MEPIT Type I, item 19, 20, and 21. According to the written response, this child failed in neither converting meters to centimetres or hours to minutes. She succeeded in converting kilograms to grams.

Example 3:

MEPIT Type I, item 19, 20, and 21. According to the written response, this child failed in neither converting meters to centimetres or hours to minutes. She succeeded in converting kilograms to grams.

Note. These two examples are excerpted from two different children.
8) **Presentation errors**

Presentation errors occur because children are experiencing difficulties in encoding the calculation procedures and results in the correct mathematical manner. When children cannot present quantitative information on tables or figures, cannot create mathematical representations (including equations, inequalities, geometric figures, and diagrams) to model the problem correctly, or when they have the right procedure and calculation but write the wrong answer, they will make this kind of errors. These errors will be categorised as presentation errors.

Example 1:

<table>
<thead>
<tr>
<th>4. $427 \div 7= \boxed{61}$</th>
<th>8. $8.75-6.44= \boxed{2.31}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$7\frac{427}{42}$ $7\frac{61}{1}$</td>
<td>$\frac{8.75}{6.44}$ $\boxed{2.31}$</td>
</tr>
</tbody>
</table>

**MEPIT Type I, item 4.** This child knew how to implement the calculation, and he got the correct answer. However he made a mistake when he was presenting his division procedures.

**MEPIT Type II, item 8.** This child calculated the subtraction of decimals successfully. She nevertheless omitted the minus sign when she presented the whole calculation.

*Note.* These two examples are excerpted from two different children.

Example 2:

| 26. 次想了一个数字，这个数字的两倍再加4等于18，问这个数字是多少？ | 26. This child comprehended the problem and she solved the problem with correct answer. She however did not create correct equation to model the problem.
| $18-4=14$ $2\times7=14$ | $14=2\times7$ |

**MEPIT Type I, item 26.** This child comprehended the problem and she solved the problem with correct answer. She however did not create correct equation to model the problem.
MEPIT Type II, item 18. This child did not know how to encode time into an appropriate mathematical manner and then implement calculation. The child could not correctly generate equivalent mathematical representations for time relationship.

Note. These two examples are excerpted from two different children.

Example 3:

MEPIT Type II, item 22. Apparently, this child were not aware of using notations and symbols to represent mathematical entity. The second part of the problem asks for the relationship between angles. She could not generate equivalent representations for the given mathematical relationship.

9) **Multiple errors**

According to written responses from children with MLD, very often more than one error types can be identified in a single test item. It is understandable because children with MLD may experience multiple difficulties in dealing with one mathematics problem. Even though these test items aim to identify error types as accurate as possible, sometimes it is really difficult to categorise an error to just one specific error type. The errors made by children with MLD are various and unexpected. In this case, inference is used when coding the errors. The researcher needed to pretend that he was the child and to assume how he was going to deal with the problem, therefore to infer reasons underneath the error. The subsequent error analysis is based on this assumption. As a result, the absolute subjectivity will be affected.
Example 1:

MEPIT Type II, item 6. Apparently, this child made a presentation error, the way he presented the calculation is incorrect. Further, this child also made a procedural error by implementing wrong order of operations.

MEPIT Type II, item 11. At first sight, it is a visual-spatial error, this child read “+” as “÷”, and then used reciprocal rule. Following that, he began to make more procedural errors and got a weird answer.

Example 2:

MEPIT Type I, item 24. The question asks for the circumference of an equilateral triangle with 8cm side. At first sight, this child may be confused with “+” and “×” and write $8 + 8 + 8$ as $8 \times 8 \times 8$, then it would be a visual-spatial error. Another possibility is that he did not know the formula for calculating circumference of triangles, then it would be a fact error. In either case, he omitted the measurement units, it is a measurement error.
Example 3:

 MEPIT Type I, item 28. In brief, velocity is 30 km/h, time is half an hour, the question asks for distance. According the written response, this child first translated half an hour to 30 minutes and then he used 30 km/h to multiply 30 minutes. It is a measurement error. The measurement of final answer was in metres rather than kilometres, it is another measurement error. Further, when he did the multiplication, he believed $30 \times 30 = 1600$, it is a fact error.
4.2.2 Error Pattern Analysis

In this section, the second research question will be responded. The frequencies of mathematical errors types will be calculated, the relationships between error types, gender and school types will be determined, the model fit test between mathematical errors and cognitive stages will be discussed.

Frequencies of error types

The frequencies shown in Figure 6 were calculated according to the number of each type of errors that identified among all the participants \(n = 30\) through analyzing the raw test data.

![Frequencies of error types - absolute number](image)

*Figure 6: Frequencies of error types - absolute number*

Fact error constitutes the largest proportion among all eight types of errors while relevance error constitutes the least proportion. There might be two reasons to explain this phenomenon. One is that more children made fact errors and less children made relevance errors during the test. Another reason would be that in the MEPIT test, as well as other tests (e.g., MBADT, TIMSS tests, state-mandated tests), there are fewer question items concerning with relevance errors compared to other errors. As a result, the major deficiency of this frequency chart is that all eight types of possible errors are not equally distributed in the MEPIT test (see Table 7). One error type occupying a higher proportion may be because there are more question items aiming to identify this error. As shown in Figure 7, a new frequency chart is introduced in order to reduce this bias.
For one error type, the accumulated percentage frequency is calculated using the following formula: \( \sum \left( \frac{\text{number of identified errors}}{\text{number of all possible errors}} \times 100\% \right) \). In this way, it would be more accurate to see which type of error children made most and which type children made least. This time, comprehension error occupies the largest proportion and relevance error occupies the least proportion. Children made basically same amount of transformation error, measurement error, fact error, and procedural error. Presentation error and visual-spatial error are the third and second least error types that children have made during the test.

**Error types and gender**

In the following, the relationships between gender and error types will be discussed.

**Table 10: Mean, standard deviation, and 0.05 level significance test of error types and gender**

<table>
<thead>
<tr>
<th>Error Types</th>
<th>Girls</th>
<th>Boys</th>
<th>Sig. Level (t-test)</th>
<th>Effect Size (Cohen’s d)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Visual-spatial Error</td>
<td>1.30</td>
<td>1.17</td>
<td>0.60</td>
<td>0.52</td>
</tr>
<tr>
<td>Comprehension Error</td>
<td>2.60</td>
<td>1.54</td>
<td>2.00</td>
<td>1.41</td>
</tr>
<tr>
<td>Transformation Error</td>
<td>1.95</td>
<td>1.36</td>
<td>1.30</td>
<td>0.82</td>
</tr>
<tr>
<td>Relevance Error</td>
<td>0.55</td>
<td>0.51</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Fact Error</td>
<td>3.25</td>
<td>2.47</td>
<td>1.40</td>
<td>0.97</td>
</tr>
<tr>
<td>Procedural Error</td>
<td>2.10</td>
<td>1.29</td>
<td>1.20</td>
<td>1.23</td>
</tr>
<tr>
<td>Measurement Error</td>
<td>2.10</td>
<td>1.45</td>
<td>1.20</td>
<td>1.48</td>
</tr>
<tr>
<td>Presentation Error</td>
<td>1.25</td>
<td>0.91</td>
<td>1.50</td>
<td>0.85</td>
</tr>
</tbody>
</table>

*Note. n = 30 (20 girls, 10 boys), * \( p < 0.05 \)
As shown in Table 10, according to the means, girls seem to have made more errors compared to boys among all types of errors except presentation error. Boys made slightly more presentation errors than girls did. In addition, the standard deviations of girls are higher than those of boys except measurement error. It is thus suggested that the errors made by girls are more dispersed than those made by boys. Using the t-test for independent samples, a significant difference has been found between girls and boys regarding fact error ($t = 2.27, df = 28, p < 0.05$). Using Cohen’s $d$ to measure the size effect of this relationship, a strong effect is suggested ($d > 1$) (Field, 2009). Another significant difference is found between girls and boys in terms of relevance error ($t = 2.27, df = 28, p < 0.05$), the size effect is also strong ($d > 1$). One thing worth notice is that only girls have made relevance errors among all the participants despite the fact that none of boys have made this error. The differences between girls and boys regarding the remaining types of errors are not statistically significant according to the results of a series of t-tests.

However, since there are more girls than boys and the small sample size in this case, the assumption of normal distributions of t-test may have been violated. It is therefore suggested to adopt the Mann–Whitney U test, which is a nonparametric test that could guarantee the valid results without the assumption of normal distributions. The results of Mann–Whitney U test is presented in Table 11.

<table>
<thead>
<tr>
<th>Error Types</th>
<th>Sig. (p)</th>
<th>Z</th>
<th>Effect Size (r)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visual-spatial Error</td>
<td>0.06</td>
<td>-1.84</td>
<td></td>
</tr>
<tr>
<td>Comprehension Error</td>
<td>0.26</td>
<td>-1.12</td>
<td></td>
</tr>
<tr>
<td>Transformation Error</td>
<td>0.21</td>
<td>-1.26</td>
<td></td>
</tr>
<tr>
<td>Relevance Error</td>
<td>0.00</td>
<td>-2.90</td>
<td>0.53</td>
</tr>
<tr>
<td>Fact Error</td>
<td>0.02</td>
<td>-2.38</td>
<td>0.43</td>
</tr>
<tr>
<td>Procedural Error</td>
<td>0.61</td>
<td>-1.88</td>
<td></td>
</tr>
<tr>
<td>Measurement Error</td>
<td>0.45</td>
<td>-2.01</td>
<td></td>
</tr>
<tr>
<td>Presentation Error</td>
<td>0.51</td>
<td>0.65</td>
<td></td>
</tr>
</tbody>
</table>

*Note. The significance level is 0.05*

The results of Mann–Whitney U test consolidate the previous results. Significant differences have been found between relevance error and gender, fact error and gender respectively. The size effect of relevance error and gender is strong with $r > 0.5$, another one is medium with $0.3 < r < 0.5$ (Field, 2009).

**Error types and school types**

64
In the following, the relationships between school types and error types will be discussed.

Table 12: Mean, standard deviation, and 0.05 level significance test of error types and school types

<table>
<thead>
<tr>
<th>Error Types</th>
<th>Urban M</th>
<th>Urban SD</th>
<th>Rural M</th>
<th>Rural SD</th>
<th>t</th>
<th>p</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visual-spatial Error</td>
<td>1.07</td>
<td>1.22</td>
<td>1.07</td>
<td>0.88</td>
<td>0.00</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Comprehension Error</td>
<td>1.80</td>
<td>1.61</td>
<td>3.00</td>
<td>1.13</td>
<td>-2.36</td>
<td>0.03*</td>
<td>0.88</td>
</tr>
<tr>
<td>Transformation Error</td>
<td>1.53</td>
<td>0.92</td>
<td>1.93</td>
<td>1.49</td>
<td>-0.89</td>
<td>0.38</td>
<td></td>
</tr>
<tr>
<td>Relevance Error</td>
<td>0.27</td>
<td>0.46</td>
<td>0.47</td>
<td>0.52</td>
<td>-1.12</td>
<td>0.27</td>
<td></td>
</tr>
<tr>
<td>Fact Error</td>
<td>2.27</td>
<td>1.44</td>
<td>3.00</td>
<td>2.85</td>
<td>-0.89</td>
<td>0.38</td>
<td></td>
</tr>
<tr>
<td>Procedural Error</td>
<td>1.67</td>
<td>1.23</td>
<td>1.93</td>
<td>1.44</td>
<td>-0.55</td>
<td>0.59</td>
<td></td>
</tr>
<tr>
<td>Measurement Error</td>
<td>1.33</td>
<td>1.11</td>
<td>2.27</td>
<td>1.71</td>
<td>-1.77</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>Presentation Error</td>
<td>1.13</td>
<td>0.92</td>
<td>1.53</td>
<td>0.83</td>
<td>-1.25</td>
<td>0.22</td>
<td></td>
</tr>
</tbody>
</table>

Note. $n=30$ (15 urban school children, 15 rural school children), * $p < 0.05$

As shown in Table 12, the standard deviations between children from urban school and children from rural school fluctuate from one error type to another error type. The means of all types of errors made by children from rural school seem to be higher than those made by children from rural school, which indicates that children from rural school have made more errors. However, there are no statistically significant differences found between children from urban school and children from rural school in terms of these error types with one exception. According to the result of t-test, a significant difference has been suggested between children from urban school and children from rural school with respect to comprehension error ($t = -2.36, df = 28, p < 0.05$). The size effect of this difference is moderate as reported by Cohen’s d ($0.51 < d < 1$).

Table 13: Mann–Whitney U test of error types and school types

<table>
<thead>
<tr>
<th>Error Types</th>
<th>Sig. (p)</th>
<th>Z</th>
<th>Effect Size (r)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visual-spatial Error</td>
<td>0.67</td>
<td>0.43</td>
<td></td>
</tr>
<tr>
<td>Comprehension Error</td>
<td>0.02</td>
<td>2.42</td>
<td>0.44</td>
</tr>
<tr>
<td>Transformation Error</td>
<td>0.61</td>
<td>0.51</td>
<td></td>
</tr>
<tr>
<td>Relevance Error</td>
<td>0.26</td>
<td>1.12</td>
<td></td>
</tr>
<tr>
<td>Fact Error</td>
<td>0.73</td>
<td>0.34</td>
<td></td>
</tr>
<tr>
<td>Procedural Error</td>
<td>0.69</td>
<td>0.41</td>
<td></td>
</tr>
<tr>
<td>Measurement Error</td>
<td>0.12</td>
<td>1.57</td>
<td></td>
</tr>
<tr>
<td>Presentation Error</td>
<td>0.19</td>
<td>1.32</td>
<td></td>
</tr>
</tbody>
</table>

Note. The significance level is 0.05
The results of Mann–Whitney U test is presented in Table 13, which conform to the previous results. There is a significant difference between comprehension error and school type, and the size effect is medium with $0.3 < r < 0.5$.

**Relationships between error types**

As shown in Table 14, the correlations between eight error types will be discussed.

*Table 14: Using Pearson’s $r$ correlation coefficient to look at relationships between eight error types*

<table>
<thead>
<tr>
<th>Pearson’s r correlation coefficient (F-test)</th>
<th>Visual-spatial Error</th>
<th>Comprehension Error</th>
<th>Transformation Error</th>
<th>Relevance Error</th>
<th>Fact Error</th>
<th>Procedural Error</th>
<th>Measurement Error</th>
<th>Presentation Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visual-spatial Error</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comprehension Error</td>
<td>0.36 (0.54)</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transformation Error</td>
<td>0.26 (0.17)</td>
<td>0.28 (0.13)</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relevance Error</td>
<td>0.29 (0.13)</td>
<td>0.26 (0.16)</td>
<td>0.28 (0.13)</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fact Error</td>
<td>0.61** (0.00)</td>
<td>0.29 (0.12)</td>
<td>0.52** (0.00)</td>
<td>0.35 (0.06)</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Procedural Error</td>
<td>0.33 (0.07)</td>
<td>0.15 (0.44)</td>
<td>-0.29 (0.12)</td>
<td>0.28 (0.14)</td>
<td>0.31 (0.10)</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Measurement Error</td>
<td>0.21 (0.27)</td>
<td>0.01 (0.98)</td>
<td>0.01 (0.97)</td>
<td>0.15 (0.43)</td>
<td>0.35 (0.06)</td>
<td>0.19 (0.32)</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Presentation Error</td>
<td>0.01 (0.95)</td>
<td>0.21 (0.27)</td>
<td>0.24 (0.20)</td>
<td>0.27 (0.16)</td>
<td>0.20 (0.28)</td>
<td>0.06 (0.76)</td>
<td>-0.10 (0.58)</td>
<td>1.00</td>
</tr>
</tbody>
</table>

*Note. n = 30, ** correlation is significant at the 0.01 level (2-tailed).*

Fact error and visual-spatial error has a strong correlation ($r = 0.61, 0.5 < r < 0.8$), and this relationship is statistically significant ($p < 0.01$). Fact error also has a strong correlation with transformation error ($r = 0.52, 0.5 < r < 0.8$), this relationship is statistically significant as well ($p < 0.01$). Most correlations between error types are weak or modest, which suggest that they are quite independent with each other.

**Relationships between eight error types and score of MEPIT**

In the following, eight error types will be used as predictors to predict the score of MEPIT.
Table 15: Using multiple linear regression to look at the relationships between eight error types and score of MEPIT

<table>
<thead>
<tr>
<th>Error Types</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>Significance Test</th>
<th>Collinearity Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Beta</td>
<td>t</td>
<td>Sig.</td>
</tr>
<tr>
<td>Visual-spatial Error</td>
<td>0.22</td>
<td>0.04</td>
<td>0.41</td>
<td>0.67</td>
</tr>
<tr>
<td>Comprehension Error</td>
<td>-1.12</td>
<td>-0.26</td>
<td>-3.59</td>
<td>0.00**</td>
</tr>
<tr>
<td>Transformation Error</td>
<td>-0.61</td>
<td>-0.12</td>
<td>-1.17</td>
<td>0.25</td>
</tr>
<tr>
<td>Relevance Error</td>
<td>-1.70</td>
<td>-0.13</td>
<td>-1.72</td>
<td>0.10</td>
</tr>
<tr>
<td>Fact Error</td>
<td>-1.20</td>
<td>-0.42</td>
<td>-3.77</td>
<td>0.00**</td>
</tr>
<tr>
<td>Procedural Error</td>
<td>-1.40</td>
<td>-0.29</td>
<td>-3.27</td>
<td>0.00**</td>
</tr>
<tr>
<td>Measurement Error</td>
<td>-1.33</td>
<td>-0.31</td>
<td>-4.26</td>
<td>0.00**</td>
</tr>
<tr>
<td>Presentation Error</td>
<td>-0.85</td>
<td>-0.12</td>
<td>-1.65</td>
<td>0.11</td>
</tr>
</tbody>
</table>

Note. n = 30, ** correlation is significant at the 0.01 level (2-tailed).

As shown in Table 15, the results suggest that eight error types are particular good at predicting the score of MEPIT (R² = 0.911, Adjusted R² = 0.877, strong fit) and basically all the relationships are negative. This is not a surprise because it is reasonable to assume that if a child made more errors he/she would end up with a lower score. The unstandardized coefficients, or B values, give the information regarding the value that score of MEPIT will change by if one specific error changes by 1 unit. Since it is unstandardized result, it is preferable to look at the next column – standardized coefficients, or Beta values. Fact error’s Beta is the strongest (-0.42), followed by measurement error (-0.31), procedural error (-0.29), and comprehension error (-0.26). These four error types have stronger effect on score of MEPIT than other error types and these relationships are statistically significant at the 0.01 level (p < 0.01). One condition in regression analysis is that the independent variables shouldn’t be strongly correlated to one another. There are no serious problems in the above analysis as no tolerance values close to 0. Most variables have a tolerance bigger than 0.5; only transformation error and fact error suggest some cause for concern, with a tolerance of around 0.4. One thing to notice is that only visual-spatial error has a positively effect, even though the effect is very little (Beta = 0.04). This could be explained that many visual-spatial
errors were made because children were solving the problems very fast and became careless, and this error had little effect in predicting score of MEPIT with Beta close to 0.

**Test of model fit between the types of errors and the stages of cognition in solving mathematics problems**

Eight error types, including _visual-spatial error_ (ERROR1), _comprehension error_ (ERROR2)…, and _presentation error_ (ERROR8), are manifest variables (which are observed from MEPIT test). Difficulties in four cognitive stages, including _encoding, planning, solving, and responding_, are latent variables (which cannot be observed directly). The researcher assumes that the latent variables cause the observed variables, this is shown in Figure 8 by the single-headed arrows pointing away from the circles (latent variables) to the squares (manifest variables).

![Path diagram](image)

*Figure 8: Path diagram of model fit between the types of errors and the stages of cognition in solving mathematics problems*

The error variances ($R^2$) of ERROR3, ERROR5, and ERROR8 are negative. The $R^2$ of ERROR7 is too small with value of 0.0068. These are bad indicators for a good model fit. The Chi-Square of model fit is 56.89 ($df = 14$), which is large enough to suggest a good fit to the
data. However, the Chi-Square test is widely recognized to be very sensitive to the size of sample and can be problematic sometimes (Albright, 2006). To look at other indices, the goodness of fit index (GFI) is too small with value of 0.741, the comparative fit index (CFI) is also too small with value of 0.446, and the root mean square error of approximation (RMSEA) is too big with value of 0.32. In summary, based on the data collected in this study, the types of errors do not fit in the stages of cognition in solving mathematics problems very well.
5 Discussion and Conclusion

5.1 Discussion

A Mathematics Error Pattern Identification Test (MEPIT) was developed in this study, which aimed to identify mathematical errors among children with MLD. Based on the data analysis of pilot test and official test, MEPIT showed a satisfactory validity and reliability. This test was then used to analyze error patterns of children with MLD in Tibet. According to absolute number, fact error happened most frequently, followed by comprehension error. Relevance error happened least frequently. Fact errors were mainly due to the inadequate mathematics education in primary stage and language barriers in Tibet, which corresponded to other studies (e.g., Wang, 2008; Fang et al., 2008). As for accumulated percentage, children made most comprehension errors and least relevance errors. Comprehension skills is vital to solve mathematics problems, children with MLD are often found to have problems with these skills (e.g., Desoete et al., 2004; Hansen, 2005). In both cases, girls seemed to be more vulnerable to relevance errors and fact errors than boys. Regarding school types, children from rural school tended to make more comprehension errors compared to children from urban school. Most correlations between these eight error types were weak or modest, which indicated that they were quite separate from each other. Among them, comprehension error, fact error, procedure error, and measurement error had a stronger effect on the score of MEPIT. The model fit test suggested that eight error types did not fit into four stages of cognition in solving mathematics problems very well.

This study contributes to ongoing comprehensive diagnosis of the errors and misconceptions shown by children with significant difficulties in learning mathematics. This kind of diagnosis mainly comes in two forms, diagnostic test and diagnostic interviews (including both teachers and students), which has been piloted in several regions but still has not been standardized and await full evaluation (e.g., Kuo et al., 2001; Dowker, 2005; Ketterlin-Geller & Yovanoff, 2009). For a long time, a number of researchers, specialists, and school teachers have emphasized the importance of describing and investigating error patterns manifested in children’s work (e.g., Borasi, 1987; Hansen, 2005; Herholdt & Sapire, 2014). Analysis of error patterns is generally regarded as a prior step to imply better teaching design and to plan intervention programmes for children with MLD (e.g., Mazzocco et al., 2008; McGuire,
The empirical findings of this study provide information concerning significant differences between the groups of children with MLD in making mathematical errors as well as the pattern analysis of these errors from a unique Tibetan perspective. In addition, there is a rich history in cognitive psychology of trying to infer a model to describe the hidden stages in solving mathematics problems (e.g., Sternberg, 1969; Mayer, 2004; Anderson et al., 2016). While such efforts have been successful to some extent, their implications in the field of MLD have been limited. This study constructs a tentative modelling of relationship between error patterns and cognitive stages. Even though it has been suggested to be a poor model fit, this finding could imply improvement for each model and provide possible hypothesis for future study.

5.2 Practical Implication

The MEPIT test developed in this study can be used by teachers in Tibet to identify children’s mathematical errors. The MEPIT test is initially designed to contextualize the mathematics education in Tibet. Teachers in Tibet could take advantage of this test to familiar themselves with the children’s mathematical errors, and then they will be able to plan for and address children’s errors as they arise in mathematics teaching. Given the satisfactory validity and reliability, MEPIT is believed to have a certain extent of generalisability and can be used in other regions. When using this test outside Tibet, it is advised to update the background of word problems as well as other concerning items to their own context. Further, as learning disabilities including MLD are not generally recognized in China, there is still no official diagnosis assessment being widely used (Tong, 2008). The MEPIT test could serve as a pedagogical assessment to identify children with MLD in China.

Based on one of principles of constructivism, it is believed that children construct their own mathematics understanding and knowledge (Piaget, 1953). Therefore, mathematics is not something taught by teachers but rather something learnt by children themselves. In this way, mathematics errors should be regarded as a natural consequences when children’s mathematics development. As children construct their own meanings, they will inevitably make errors. Even mathematicians make errors, mathematics error is not a negative thing (Dowker, 2005). Error patterns should not be interpreted as an instruction to tell teachers to avoid all the errors but to teach absolutely correct mathematics (Hansen, 2005). Children learn from their errors. The value of this study is to support teachers make sense of the errors
that children may make and use the information to plan their lessons appropriately. Error patterns found in this study can also contribute to the preparatory stage of curriculum reform in Tibet. As an autonomous region in China, Tibet is adopting the national curriculum which is mainly designed for Han Chinese students, it is therefore suggested to design a new curriculum to reflect the reality of Tibetan education (Fang et al., 2008). When designing mathematics curriculum, findings of this study would be useful.

5.3 Research Limitation

This study is designed and conducted in a quantitative manner. The data is presented in a precise, quantitative, and numerical manner. The research findings are relatively independent of the researcher (e.g., effect size, statistical significance). However, the researcher's theories that are used may not reflect participants’ true understandings (Johnson & Onwuegbuzie, 2004). For example, the eight types of mathematical errors are concluded mainly according to literature review which may have not completely reflected the actual errors that children with MLD in Tibet would make.

Even though the purpose of this study is to obtain a certain level of variety of participant demographics, the researcher cannot guarantee a randomised and representative sample from the exact target population. The sample size was quite small, so results should be considered as suggestive only. In the process of data analysis, each child with MLD was treated as an independent individual with his/her unique experience in making mathematical errors. In reality, however, children are nested in different classes and schools, it may be that children taught by the same teacher or studying in the same school would make similar errors. The generalisability of findings in this study is thus restricted. This is perhaps the largest methodological limitation of this study.

Regarding research tools, translation and adaptation from previous tests to the new MEPIT test was the crucial part. Items from the MBADT are originally in traditional Chinese, and Tibet is using simplified Chinese. The most obvious variation between these two written systems is the usage of different Chinese characters and a slightly dissimilar vocabulary, especially mathematical terminology. The researcher exchanged emails with the main developer of the MBADT test (Ching-Chih Kuo) and double-checked some key terms (like coordinate) in Taiwan’s textbook. Items from TIMSS Mathematics are originally in English.
There are also Hong Kong and Taiwan’s traditional Chinese versions available, which were used as reference materials in translating. Best efforts had been made to translate and adapt these test items correctly and appropriately to Tibetan context. The draft test was sent to experienced teachers in Tibet for reviewing and their comments had been addressed in the later version. However, due to limited time and financial recourses, the strict double-translation and reconciliation procedure suggested by large international test organisers (e.g., OECD, 2016) was not followed.

When analysing data, such as marking children’s test papers, coding errors manifested in children’s work into error types, it is important to recognize that the researcher has often been working so closely on this research that he might have “established shared and hidden meanings of the coding” (Stemler, 2001, p.5). At any rate, it is the researcher himself who has developed the test and defined the scope of each error type. The potential limitation is that the reliability coefficient reported in this study may be inflated. In addition, many data analysis techniques used in this study are sensitive to the size of sample. Some research findings are not statistically significant may be due to a small sample size. The size of sample could also affect the indices for model fit.

5.4 Future Directions

When answering research questions, very often more new questions will emerge (Malinen, 2013). This is also the case of this study. This study is based on the data collected in Lhasa region. In future studies, it is advisable to investigate the children with MLD in other areas of Tibet and to compare these findings in order to reflect the whole picture. A larger sample size is expected. It is also advisable to investigate how generalizable these findings are to other provinces in China. If MEPIT test can continue to perform satisfactory validity and reliability with large samples, it is worth considering to standardize this test to be a norm-reference test. It would be an effective tool in secondary mathematics education in China. Concerning test marking and error coding process, it is preferable to have at least two people work independently. Developing a set of explicit and detailed marking and coding instructions is therefore necessary.

In terms of identifying error patterns among children with MLD, in prospective studies, a mixed research method is suggested. For example, in addition to test, children could also be
given interviews to determine difficulties in coping with mathematics problems. In this way, children’ knowledge and understandings of mathematics errors could be directly investigated. Teachers’ observations and interviews of teachers are useful data as well. By cross checking results gained through these different methods, data triangulation can be achieved and a better understanding of children’s error patterns is expectable, the effect of researcher’s potential biases could be reduced and the research findings would be stronger (Johnson & Onwuegbuzie, 2004). Lastly, even though the modelling of relationship between error types and cognitive stages was not very successful in this study, it does not mean that this is the way and not worth trying. In fact, it points out the future that both error type’s model and cognitive stages model are in need of more researches.

5.5 Conclusion

Children with MLD in Tibet are especially vulnerable to fact errors and comprehension errors. Tibetan girls seem to be more likely to make mathematical errors compared to boys, especially fact errors and relevance errors. There is no major difference between children from rural school or urban school except comprehension errors.
References


RESULT OF NOTIFICATION TEST: NOT SUBJECT TO NOTIFICATION

You have indicated that neither directly or indirectly identifiable personal data will be registered in the project.

If no personal data is to be registered, the project will not be subject to notification, and you do not have to submit a notification form.

Please note that this is a guidance based on information that you have given in the notification test and not a formal confirmation.

For your information: In order for a project not to be subject to notification, we presuppose that all information processed using electronic equipment in the project remains anonymous.

Anonymous information is defined as information that cannot identify individuals in the data set in any of the following ways:
- directly, through uniquely identifiable characteristic (such as name, social security number, email address, etc.)
- indirectly, through a combination of background variables (such as residence/institution, gender, age, etc.)
- through a list of names referring to an encryption formula or code, or
- through recognizable faces on photographs or video recordings.

Furthermore, we presuppose that names/consent forms are not linked to sensitive personal data.

Kind regards,

NSD Data Protection
### Appendix 2  MEPIT Type I

**数学错误类型鉴定测试——I 型**

测试时间 40 分钟，请合理利用空间，将所有演算步骤写在试卷上面。

姓名:__________  性别:__________  年龄:__________  学校:__________

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>21+43=</td>
</tr>
<tr>
<td>2.</td>
<td>32×26=</td>
</tr>
<tr>
<td>3.</td>
<td>83-37=</td>
</tr>
<tr>
<td>4.</td>
<td>427÷7=</td>
</tr>
<tr>
<td>5.</td>
<td>123÷41=</td>
</tr>
<tr>
<td>6.</td>
<td>4×2-30÷5+8=</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>7.</td>
<td>5.36 + 4.57 =</td>
</tr>
<tr>
<td>8.</td>
<td>9.56 - 7.41 =</td>
</tr>
<tr>
<td>9.</td>
<td>0.4 × 1.2 =</td>
</tr>
<tr>
<td>10.</td>
<td>1.83 ÷ 0.3 =</td>
</tr>
<tr>
<td>11.</td>
<td>( \frac{1}{2} + \frac{2}{3} = )</td>
</tr>
<tr>
<td>12.</td>
<td>( \frac{3}{5} ÷ \frac{6}{7} = )</td>
</tr>
<tr>
<td>13.</td>
<td>15 比 ___ 多 9？</td>
</tr>
</tbody>
</table>
14. 索朗有 1 个姐姐，2 个哥哥，多吉的姐姐比索朗的多 3 个，请问多吉有几个姐姐？

15. 比较 17、17 3/4、17.7 的大小。

16. 文具店里有两种不同价格的铅笔，分别是 3 元及 10 元，丹增如果想每种各买 2 只，怎么写出求总价的算式？

17. 扎西带 100 块钱到书店买文具，买了 2 本 10 元/本的作业本，以及 5 枝 8 元/枝的铅笔，请问扎西还剩多少钱？

18. 请问下面的时钟指的是几点几分？

![时钟图](image)
19. 1 米 = (     ) 厘米

20. 2 千克 = (     ) 克

21. 1.5 小时 = (     ) 分钟

22.

![角的图示]

① 上面哪个角是直角？
② 请将上述四个角由大到小排列。

23. 请在下面的空白处画一个长方形和一个圆。

24. 每边长为 8cm 的正三角形，周长是多少？
25. 长方形的菜园子，长为 3cm，宽为 5cm，请问菜园的面积是多少？

26. 次托想了一个数字，这个数字的两倍再加 4 等于 18，请问这个数字是多少？

27. 拉姆的笔

<table>
<thead>
<tr>
<th>笔的种类</th>
<th>数量</th>
</tr>
</thead>
<tbody>
<tr>
<td>钢笔</td>
<td>3</td>
</tr>
<tr>
<td>铅笔</td>
<td>5</td>
</tr>
<tr>
<td>中性笔</td>
<td>8</td>
</tr>
</tbody>
</table>

①拉姆有几只钢笔？
②拉姆的中性笔比铅笔多几只？
28. 旺堆骑自行车的时速为 30 千米每小时，旺堆每天上学需要骑车半个小时，请问旺堆家离学校有多远？

29. 这是卓玛家所在的乡镇地图，下面标出市场的位置在 F 5，学校的位置在 A 4.

   8
   7
   6
   5
   市场
   4
   学校
   3
   商店
   2
   1
   A B C D E F G H I

   请问商店的位置在哪里？

30. 如果将上图中的线拉直，这条线的长度大概是多少？
数学错误类型鉴定测试——II 型

测试时间 40 分钟，请合理利用空间，将所有演算步骤写在试卷上面。

姓名:__________ 性别:__________ 年龄:__________ 学校:__________

1. $62+25=$

2. $23 \times 45=$

3. $77-29=$

4. $176 \div 8=$

5. $145 \div 29=$

6. $72 \div 6-4 \times 2+9=$
<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>7.</td>
<td>3.72 + 6.17 =</td>
<td>8.</td>
</tr>
<tr>
<td>9.</td>
<td>0.3 \times 1.5 =</td>
<td>10.</td>
</tr>
<tr>
<td>11.</td>
<td>\frac{3}{5} + \frac{1}{4} =</td>
<td>12.</td>
</tr>
<tr>
<td>13.</td>
<td>27 比 ____ 少 3？</td>
<td></td>
</tr>
</tbody>
</table>
14. 央宗家有 10 头牦牛，30 头羊，措姆家的羊比央宗家的少 10 头，请问措姆家有多少头羊？

15. 比较 3, 3.5, $\frac{11}{3}$ 的大小。

16. 扎西和他妹妹同一天过生日，不过妹妹比他小三岁：①扎西的年龄用 $x$ 来表示，妹妹的年龄该用什么算式来表示？②当扎西 12 岁时，妹妹是几岁？

17. 10 元=（ ）个 5 元
   100 元=（ ）个 10 元

18. 拉萨到日喀则坐火车需要 2 小时 40 分，索朗在拉萨做早上 9 点 30 分的火车，到日喀则是几点？
| 19. 1 千米 = (   ) 米 |
| 20. 1 千克 = (   ) 克 |
| 21. 3 分钟 = (   ) 秒 |

22. ① 上面哪个角是锐角？
② 请将上述四个角由小到大排列。

23. 请在下面的空白处画一个三角形和一个正方形。

24. 半径为 5cm 的圆，周长是多少？
25. 底为 10cm，高为 8cm 的三角形，面积是多少?

26. 旺姆想了一个数字，这个数字的三倍再减去 5 等于 16，请问这个数字是多少?

27. 次托家的牦牛

①次托家有多少只杂色牛？
②次托家的黑牦牛比白牦牛多几只？
28. 青藏铁路上有一列火车 5 小时走了 455 千米，请问火车的时速是多少？

29. 这是卓玛家所在的乡镇地图，下面标出学校的位置在 A 5，市场的位置在 F 4。桑旦家的位置在 C 8，请在图中标出桑旦家的位置。

30. 每次将[图片]图形顺时针转动 $\frac{1}{4}$ 圈，请问图形是怎么变化的？