Doctoral Dissertation

The Development of a Theoretical Framework and Tools to Measure Social and Emotional Skills in Mathematics in the Mongolian Lower Secondary Education

KHAJIDMAA OTGONBAATAR

Graduate School for International Development and Cooperation
Hiroshima University

September 2021

The Development of a Theoretical Framework and Tools to Measure Social and Emotional Skills in Mathematics in the Mongolian Lower Secondary Education

D181116

KHAJIDMAA OTGONBAATAR

A Dissertation Submitted to
the Graduate School for International Development and Cooperation
of Hiroshima University in Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy in Education

September 2021

We hereby recommend that the dissertation by Mr. KHAJIDMAA OTGONBAATAR entitled "The Development of a Theoretical Framework and Tools to Measure Social and Emotional Skills in Mathematics in the Mongolian Lower Secondary Education" be accepted in partial fulfillment of the requirements for the degree of DOCTOR OF PHILOSOPHY IN EDUCATION

Committee on Final Examination:

BABA Takuya, Professor, Chairperson

SHIMIZU Kinya, Professor

中矢礼美

NAKAYA Ayami, Associate Professor

渡邊耕二

WATANABE Koji, Professor, School of Education, Miyazaki International College, Japan

ITGEL Miyejav, Associate Professor, School of Arts and Sciences, National University of Mongolia

Date: July 26th. 2021

Approved: 市橋勝 Date: Sep. 3, 2021

ICHIHASHI Masaru, Professor
Dean

Graduate School for International Development and Cooperation
Hiroshima University

DEDICATION

I dedicate this dissertation to four beloved people who mean the world to me.

To my father, Khajidmaa, who started teaching me basic literacy at the tender age of five while battling a life-threatening disease and going through a financial crisis. During the transition period of Mongolia from the socialist system to the free market system, our family had to undergo such challenging times when we were short of food; yet he ensured we were never left short of books. Ever since my childhood, he believed in me and had a firm belief that I would make him proud one day. You are the biggest reason I have come so far.

To my mother, Rentsendulam, who carried the entire family on her back during the rough social transition period. She sacrificed herself to look after her ailing husband, raised eight children, and took every potential job to afford our education. You are the hero who paved the way for our family's hopeful future through your resilience during those challenging times.

To my wife, Munkhzul, who taught me to love unconditionally and always encouraged me to dream big. Her constant support and belief in my potential boosts my confidence and gives me the strength to challenge myself and encourages me to give all it takes to finish what I have started. While I have been tirelessly working on my research, she took a career break and single-handedly took care of the family. You lead me to live my life to the fullest, and I will be indebted to you for the rest of my life.

To my son, Nomundalai, who has given a new meaning to my life, with whom I have discovered the intensity to love that I never knew I had. With your presence, I experience a new level of joy and contentment every day. One day, you will read this, and I hope it will be a source of inspiration for you and help you understand your father more.

ACKNOWLEDGMENT

This dissertation is the outcome of a continuous academic journey that I first started 26 years ago in a humble remote school in the western part of Mongolia. Without a doubt, a large number of unique people have contributed to my accomplishment. Thank you all very much.

First of all, I want to express my sincerest gratitude to my supervisor, Professor Takuya Baba, for his wise guidance, constant support, valuable critique, and warm encouragement during these past three years. His unique expertise and vast professional experience have immensely benefited me. Under his priceless supervision, I learned to be a critical thinker and developed the ability to understand national identity and tackle issues professionally. His wisdom and knowledge are such that a single sentence he utters can be undertaken as a research topic. One of his commonly used terms is 'food for thought,' through which he encourages his students to think, analyze, and question, thus making us critical thinkers. I vividly remember meeting a JICA officer before the orientation meeting at the university. Upon knowing that Professor Takuya Baba was my supervisor, she immediately expressed her belief that I will become a deep thinker. To Professor Takuya Baba, words fall short when I say that I am fortunate enough to be one of your students.

My sincere thanks are extended to my sub-supervisor, Professor Shimizu Kinya, and Professor Nakaya Ayami for the professional assistance, valuable advice, and criticism to improve the quality of my work. I am so thankful.

I also would like to thank the external examiner of my doctoral research, Professor Watanabe Koji from Miyazaki International College, for the valuable advice and professional recommendations. I sincerely appreciate it.

I want to thank Professor Itgel Miyejav from the National University of Mongolia, who encouraged me to apply to the JICA scholarship program, supported me to grow as a researcher, and contributed to my doctoral research as an external examiner with his professional assistance. I am so grateful.

I also would like to express my appreciation to different generations of “Babalab” for helping me to cope with my academic life. I am so grateful for and will remember the experiences and memories we created together.

I warmly thank Japan International Cooperation Agency for the scholarship, which allowed me to perform my research and covered all financial needs during the doctoral program. Additionally, I would like to extend my heartfelt thanks to the Japanese nation for generosity and donation to JICA projects to support developing countries. I will always be grateful.

I also thank the Graduate School of Humanities and Social Sciences staff, Hiroshima University, for the great support and kind cooperation throughout my study.

I wish to thank the Mongolian community in Hiroshima University, especially my friends Bayaraa, Gendee, Zorigt, Unuruu, Ganaa, Tegshee, Munguu, Javzaa, Tsoomoo, Deegii, Ariunaa, Yanjaa, and Altai, and also to my friends from different countries, Samah (Egypt), Nofi (Indonesia), Gaku (Japan), Kusaka (Japan), Putra (Indonesia), Yoshi (Japan), Claude (Rwanda), Sommay (Laos), Soth (Cambodia) for making my life in Japan more fruitful. I cherish the time we spent together.

I sincerely thank Professor Begz Nadmid, who significantly influenced me to start my research career, and Professor Purev Ochir, who mentored me to become a researcher during my time at Mongolian Institute for Educational Research. Thank you for guiding me in the right direction.

Special thanks to Lkhagvasuren Purev, Nyamtseren Rentsenkhand, Tsetsegjargal Khulganaa, Sarantuya Dorj, Enkhtogtokh Amgalan, Tsevegmid Tseden, Tserenkhand Baatarsuuri, Tuvshinjargal Dugarjav, Densmaa Jargalsaikhan, Bayarsaikhan Batmandakh, and many more, from Mongolian Institute for Educational Research for the support and ecouragement to conduct my research.

I also thank my parents, brothers, and sisters for the endless support, care, and encouragement during this academic journey. They have faith in me and always knew that I could complete this dissertation and earn a Ph.D. degree.

My final and most profound gratitude should be given to my wife Enkhbold Munkhzul and my son Otgonbaatar Nomundalai. Thank you for being the reason I go forward, stay strong, and be persistent and confident in any challenging situations during this journey. You have my everlasting gratitude. The gratitude and love I feel for all of you can never be quantified.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AV	Anchoring vignette
AVE	Average variance extraction
CASEL	Collaborative for Academic, Social, and Emotional Learning
CCE	Character and Citizenship Education
CCR	Center for Curriculum Redesign
CFA	Confirmatory Factor Analysis
CFI	Comparative Fit Index
CR	Composite reliability
DF	Degrees of freedom
EFA	Exploratory Factor Analysis
KIPP	Knowledge is Power Program
KMO	Kaiser-Meyer-Olkin
MECSS	Ministry of Education, Culture, Science and Sports
MGI	McKinsey Global Institute
ML	Maximum Likelihood
MOE	Ministry of Education
NFI	Normed-Fit Index

MSV	Maximum Shared Variance
OECD	The Organization for Economic Co-operation and Development
OLS	Ordinary Least Squares
PISA	Programme for International Student Assessment
PCA	Principal Component Analysis
PNFI	Parsimony Normed Fit Index
PSLE	Primary School Leaving Examination
RMSEA	Root Mean Square Error of Approximation
SD	Standard Deviation
SEAL	Social and Emotional Learning Program
SE	Standard Error
TIMSS	Trends in International Mathematics and Science Study
TLI	Tucker Lewis Index
UNESCO	United Nations Educational, Scientific and Cultural Organization
WEF	World Economic Forum

ABSTRACT

In the 21st century, technological advancement, together with the rapidly changing job market and the social and economic climate, requires individuals to learn varying skills to live in an unpredictable world (OECD, 2015; WEF, 2016). Consequently, many organizations and scholars have initiated skills under the term ‘social and emotional skills’ to enable students to face modern challenges. These encapsulate a balanced set of skills, including cognitive (creativity and critical thinking), interpersonal (collaboration and cooperation), and intrapersonal skills (managing emotions, goal-orientation, perseverance). Therefore, many countries have incorporated social and emotional skills into their national education policies and curricula to help future citizens thrive in an advanced society (OECD, 2015; Ontario, 2016).

With the rapid technological advancement, there is a need for change in all aspects of society, including culture and education, and mathematics is no exception. Previously, individuals were required to learn to perform basic arithmetic calculations. With the advent of technology, the need to learn numeric skills has diminished due to the availability of advanced tools that can be used to perform these calculations. Most smartphone users rarely use basic arithmetic calculations due to the rapid advancement in technology (Global Mobile Phone Report, 2016). Findings from a labour market study have revealed that employers in the technology sector look for individuals with sound mathematical knowledge of coding and algorithms and outstanding communication and cooperation skills. Hence, employee interpersonal and intrapersonal skills are equally important, along with their mathematical proficiency. Consequently, it is crucial to rethink mathematical competencies that initially were limited to basic arithmetic operations of addition, subtraction, multiplication, and division of whole numbers, fractions, and decimals (OECD, 2017). Subsequently, the challenges of the 21st century cannot be addressed through the disciplinary

knowledge of only mathematics.

Although some of the social and emotional skills are related to the affective domain in mathematics education, the affective domain concept does not go beyond emotions and motivation. Emotions include enjoyment, anxiety, boredom, and motivation, including individual and situational interest. Due to the limitations involved, the affective domain cannot fully address the challenges of the 21st century. Being an extension of the affective domain, social and emotional skills encapsulate a more comprehensive range of social skills such as communication and cooperation and cognitive skills such as creativity and critical thinking.

Although several countries and organizations have proposed different frameworks to include social and emotional skills in mathematics, most frameworks are based on survey findings from investigations of the most in-demand skills for the 21st-century context. Moreover, the frameworks have some inconsistencies in categorizing the skills into certain conceptual domains. For example, some frameworks identify perseverance (Hasratuddin, 2011; MOE, Singapore, 2012) as an emotion, while others classify it as motivation (Social and Emotional Learning program [SEAL], 2007; Ontario Mathematics Curriculum, 2020). Additionally, some frameworks cover only social skills and emotional management (Collaborative for Academic, Social, and Emotional Learning [CASEL], 2017; SEAL, 2007), while other frameworks include cognitive skills, such as creativity and critical thinking (CCR, 2015; Ontario Mathematics Curriculum, 2020).

Nevertheless, the identification of social and emotional skills in mathematics requires more discussion from a theoretical perspective. Moreover, studies conducted in Asia and Africa have demonstrated that most countries do not have the tools to assess social and emotional skills directly (United Nations Educational, Scientific and Cultural Organization [UNESCO], 2015). Similar to the challenges related to assessing social and emotional skills in a global context, there has been a lack of assessment of pedagogical practice to measure social and emotional skills at the school and system levels in Mongolia. However, the national curriculum includes these skills as assessment objectives. In current assessment practice, educators and researchers have been attempting to measure social and emotional skills using self-ratings and questionnaires. At the same time, these instruments are subject to different types of biases, including socially desirable responses and cultural biases, which can affect the reliability and validity of the instruments. Therefore, further research is required to develop reliable and valid ways to measure these skills. Based on the discussion thus far, the present study aims to answer the following three research questions:

RQ1: What framework can be used to capture social and emotional skills in mathematics?

RQ2: What valid tools can measure social and emotional skills in mathematics reliably?

RQ3: What is the status of social and emotional skills among Mongolian students?

To answer these research questions, first, a systematic literature review was conducted as a tool to 1) identify social and emotional skills related to mathematics in this study, 2) outline conceptual links between the selected skills and existing theories and concepts, and 3) construct the theoretical framework for this study. As an output of the systematic literature review, a theoretically predicted framework was constructed consisting of six components: 'Mathematical creativity', 'Mathematical perseverance', 'Cooperative learning in mathematics', 'Mathematical enjoyment', 'Mathematical self-efficacy' and 'Mathematical anxiety'.

Second, questionnaires followed by anchoring vignettes (AVs) and problem-posing were introduced as a methodological solution against the issues related to measuring social and emotional skills in this study. AVs are short descriptions of hypothetical individuals that illustrate different levels of skills or traits (King et al., 2004). Problem-posing in mathematics is when students use their mathematical knowledge to generate relevant problems from the given information (Yuan & Sriraman, 2012). The questionnaires and problem-posing tasks were adapted from previous works, while the author developed AVs based on the questionnaires. In total, the present study employed three instruments consisting of 34 items, including 1) 15 questionnaires, 2) 15 AVs, and 3) four problem-posing tasks.

Third, the tools were administered to Mongolian ninth-grade students to examine their social and emotional mathematics skills and validate the theoretical framework and tools. In total, 308 ninth-grade students comprising 151 males (49 per cent) and 157 females (51 per cent) from eight public schools in urban, suburban, and rural areas were sampled using the convenience sampling method. The students' ages ranged from 13 to 16, with a mean age of 14.0 (Standard Deviation [SD] = 0.51).

For the first research question, a reflective measurement model was derived from the indicator items using the Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) approaches to validate the theoretically predicted framework. According to the factor structures provided by an EFA, the items were classified into six independent factors without any cross loading, consistent with the structure of the theoretically predicted framework. Next, a CFA was performed to validate the EFA outcome. Various model fit statistics were used for the CFA to assess whether the theoretical model fits the selected data. According to the CFA, all fit indices were within acceptable values, implying that the theoretically predicted model was a good fit for the selected data (Normed Chi-Square [$\chi 2/df$)] = 1.32; Normed Fit Index [NFI] = 0.92; Comparative Fit Index [CFI = 0.98]; Tucker Lewis Index [TLI] = 0.97; Parsimony Normed Fit Index [PNFI] = 0.64; Root Mean Square Error of Approximation [RMSEA] = 0.03). Based on the EFA and CFA results, a theoretically predicted framework on social and emotional skills in mathematics was empirically confirmed for the first time in this study.

For the second research question, to test the reliability and validity of the tools, reliability and validity analyses were conducted in multiple ways. First, a reliability analysis was conducted in two stages using McDonald's ω coefficient and composite reliability (CR). The analysis results indicated that McDonald's ω and the CR value of the tools fulfilled the suggested criteria (> 0.7) in the literature, demonstrating that the items used in this study were internally consistent and reliable. The psychometric results demonstrated that problem-posing tasks and self-rating items adjusted by AVs are valid and reliable tools for measuring social and emotional skills in mathematics among ninth-grade students in Mongolia.

Students' performance on the vignette-corrected new scale and problem-posing tasks were analysed to answer the third research question. According to the findings, Mongolian ninth-grade students tend to be less creative and more anxious about mathematics; however, they demonstrate moderate performance on cooperative learning in mathematics, mathematical perseverance, mathematics enjoyment, and mathematical self-efficacy.

It is hoped that the theoretical framework in this study can provide researchers with a way to conduct empirical studies related to the social and emotional aspects of mathematics. However, the theoretical framework can be extended by adding more social and emotional skills and validated in future studies by applying EFA and CFA approaches in two different sample sizes. Future studies should employ the parametric solution of the AV approach (e.g., compound hierarchical ordered probit model) to handle some of the disadvantages of the non-parametric approach (i.e., ties and order violations). Future studies could also investigate learning activities and other classroom practices to offer insights into curriculum implementation to develop mathematical creativity in Mongolian students.

CHAPTER 1

BACKGROUND AND RESEARCH OBJECTIVES

This chapter presents an introduction and the rationale of this study, underlines the research problems, poses the research objectives, and questions, summarises the research contributions and limitations, and describes the structure of the dissertation.

1.1 Social background and education reform

Accelerated advances in science and technology have caused rapid changes in the labour market and socioeconomic landscape in the 21st century. Because of advanced technologies, there have been changes in the labour market demands and necessary job competencies (UNESCO, 2019). With the creation of new jobs, job displacements increased labour efficiency, and broadening skill gaps, the labour markets are expected to undergo considerable changes (World Economic Forum [WEF], 2016). In the context of job displacements, various studies and reports have predicted that there are high chances of many occupations being automated using artificial intelligence and robotics (Ernest, Merola, & Samaan, 2018; UNESCO, 2019). In the context of job creation, many of the prominent professions today did not exist a decade ago, and this change continues to increase (WEF, 2016). According to the WEF (2016), 65 per cent of today's primary school students will eventually end up in professions that do not exist today. Another challenge that must be dealt with is environmental pressure. The Organisation for Economic Co-operation and Development (OECD) has raised some important issues to preserve the Earth's depleting natural resources. Some of them are rapid climate change, reduction in freshwater availability and biodiversity, and the negative impact of air pollution on human health (OECD, 2012). The conservation of the planet largely depends on the environmental responsibility that the citizens are willing to take. It is certain that technological advancement, together with the rapidly changing job market and the social and economic climate, requires individuals to learn varying skills to live in an unpredictable world (OECD, 2015; WEF, 2016).

This discussion leads to two major questions: 1) Do education systems prepare individuals to deal with the evolving challenges of the 21st century? 2) Do secondary schools provide the necessary skills to their students to help them survive in the 21st century?

According to a recent study conducted by the McKinsey Global Institute (2018), despite the demand for technological skills being expected to exceed the current predictions, social and emotional skills will still be prevalent in workplaces, especially those of advanced economies. Therefore, the role of social and emotional skills becomes even more significant in this diverse and ever-evolving world (OECD, 2015).

Consequently, many organisations and scholars have initiated a set of skills under the term 'social and emotional skills' to enable students to face modern challenges. Particularly, Durlak et al. (2011), the OECD (2015), CASEL (2017), and the Center for Curriculum Redesign (CCR; 2019) have highlighted the importance of social and emotional skills. These encapsulate a balanced set of skills, including cognitive (creativity and critical thinking), interpersonal (collaboration and cooperation), and intrapersonal skills (managing emotions, goal-orientation, and perseverance). Cognitive skills such as numeracy, reading, and scientific literacy provide a better understanding of issues and improve decision-making and problem solving. Perseverance, sociability, and emotional stability are some of the social and emotional skills that play a significant role in an individual's life. These social and emotional skills are as crucial as cognitive skills (Heckman, Stixrud, & Urzua, 2006; Kautz et al., 2014). Therefore, many countries have incorporated social and emotional skills into their national education policies and curricula to help future citizens thrive in an advanced and unpredictable society (OECD, 2015; Ontario, 2016).

1.2 Changes in mathematics education

With the rapid technological advancement, there is a need for change in all aspects of society, including culture and education, and mathematics is no exception. Previously, individuals were required to learn to perform basic arithmetic calculations. With the advent of technology, the need to learn numeric skills has diminished due to the availability of advanced tools that can be used to perform these calculations. For example, most smartphone users rarely use basic arithmetic calculations due to rapid technological advancement (Global Mobile Phone Report, 2016). Findings from a labour market study revealed that around one-third of high school graduates (30 per cent) consider the knowledge of mathematics as 'very important', whereas 70 per cent of respondents from the same category regard communication and collaboration at the workplace as 'very important' (Casner-Lotto and Benner, 2006). According to the study, 81 per cent of college graduates considered creativity as 'very important', and 95 per cent considered communication and collaboration as 'very important', as opposed to 64 per cent who considered mathematics as 'very important'.

Similarly, educators have demonstrated that employers in the technology sector look for individuals with sound mathematical knowledge of coding and algorithms, alongside outstanding communication and cooperation skills. Hence, employees' interpersonal and intrapersonal skills are equally important, along with their mathematical proficiency (Hill, 2019). Furthermore, the Progress of International Students Achievement (PISA) 2021 Mathematics Strategic Advisory Group (2017) states:

> In recent times, the digitization of many aspects of life, the ubiquity of data for making personal decisions involving health and investments, as well as major societal decisions to address areas such as climate change, taxation, governmental debt,

> population growth, the spread of pandemic diseases and the global economy, have reshaped what it means to be mathematically competent and prepared us to be a thoughtful, engaged, and reflective citizen. (p. 3, as cited in Otgonbaatar, 2021, p. 2)

Hence, it is crucial to rethink mathematical competency that was initially limited to basic arithmetic operations of addition, subtraction, multiplication, and division of whole numbers, fractions, and decimals (OECD, 2017). Therefore, it is evident that the challenges of the 21st century cannot be addressed through the disciplinary knowledge of only mathematics.

It is unanimously agreed among educators that with the ever-evolving global economy and society, the secondary education system must enable students to be ready for higher education or the workplace by becoming proficient in advanced mathematics (Partnership for 21st century [P21] skills, 2011). According to the P21 framework, a level of proficiency can be achieved by fusing mathematical knowledge and practices with high-demand skills such as critical thinking, communication, creativity, and cooperation. Merging a core subject with such essential skills makes the process of learning and teaching more interesting, engaging, and rigorous while also improving the ability and understanding of students in advanced mathematics (P21, 2011). In other words, by intentionally infusing social and emotional skills into mathematical education, students can be equipped with the necessary skills to think mathematically and perform mathematical tasks. For instance, a study based on fourth and fifth-grade students in China concluded that high levels of social and emotional skills contribute to high achievement in mathematics, while students with lower social and emotional skills often perform poorly in the subject (Yang, Wang, & Zhang, 2018). Similarly, a study in India found that including social and emotional skills in mathematics education profoundly improves achievements in mathematics (Bhoumick & Saha, 2020). Another study demonstrated that emotional support in the classroom environment explicitly plays a crucial role in developing the mathematical skills of elementary school children coming from low-income urban backgrounds (McCormick et al., 2015).

In contrast, students are given opportunities to learn crucial skills for thriving in modern society through vital mathematical concepts. Mathematics is a lens through which students can view the world; it enables them to contribute to the economy in a meaningful way. It also empowers students to find solutions to problems and devise innovative ways—a skill required to tackle the challenges of the interconnected world (P21, 2011). Moreover, mathematics gives students an understanding of global issues and allows them to collaborate with individuals from different religions, cultures, and lifestyles as a sign of mutual respect; and acquaints them with other cultures and nations. P21 (2011) highlights that the ability to solve mathematical problems relevant at a global level instils empathy and awareness in students, making them more knowledgeable and mindful. It further states that mathematical simulations may be used to deeply analyse environmental issues at a local, national, and global level (P21, 2011). Therefore, mathematics and social and emotional skills are not only interlinked but also help reinforce each other.

Consequently, the aims of mathematical proficiency in the 21st century have evolved, leading to educational reforms in mathematics education worldwide. Many countries in Asia have reformed their mathematics curricula, mainly within the framework of national educational renovation (Catherine & Toh, 2019). Additionally, mathematics curriculum reform has been carried out in Asia and Anglo-Saxon and Northern and Eastern European countries. The top-performing places on large-scale international assessments, particularly Finland, Singapore, and Ontario (Canada), recently revised their curricula by enriching them with social and emotional skills (discussed further in Chapter 2).

Changes apparent in maths education today are consistent with some of the philosophical ideologies presented by Ernest (1991). The industrial trainer, technological pragmatist, progressive educator, and public educator are a few of his prominent educational ideologies.

The ideology of industrial trainer rests on the moral values that propagate authoritarian 'Victorian' values, whose fundamental principles include choice, effort, self-help, work, moral weakness, and the rhetoric of us-good and them-bad. Its theory of learning mathematics is based on hard work, effort, practice, and rote learning. The focus is on individual learning, while activity-based learning is discouraged and deemed ineffective without individual effort.

The ideology of technological pragmatist descends from the industrial trainer and promotes a more modern and pragmatic version of the traditional utilitarian aims. As for the theory of social diversity, the technological pragmatist views it in terms of future occupations. Despite favouring the status quo for education and employment, the skill set is expected to expand with increasing technological progress to meet evolving employment demands. This is in line with the current changes in maths education, emphasizing the need to update the curriculum with the advancements in technology. Moreover, today's labour market also looks for employees with strong interpersonal and intrapersonal skills and mathematical proficiency. Therefore, as the technological pragmatist ideology recognizes the significance of technology in maths education and acknowledges its social role, it has vital relevance with the changes that maths education is experiencing today; however, it does not incorporate it into a comprehensive perspective.

Similarly, the progressive educator ideology is based on the mathematical aims of creativity and self-realization through mathematics, both of which are child-centred approaches. It is also consistent with the changes in maths education today that emphasize inculcating social and emotional skills in students to enable them to think creatively and carry out mathematical tasks.

Finally, the public educator ideology highlights moral values such as social justice, liberty, equality, fraternity, social awareness, engagement, and citizenship. It aims to raise critical awareness and democratic citizenship through mathematics, and its theory of social diversity implies that accommodation of social and cultural diversity is necessary. In today's educational curriculum, various mathematical concepts also equip students with the necessary skills to thrive in a modern world by enabling them to understand different global issues and empowering them to interact respectfully with individuals of other religions and cultures. The empathy and awareness instilled by maths education make one mindful and socially responsible, as implied by the public educator ideology.

1.3 Problem statement

Some of these skills have existed as the affective domain in mathematics for a long time. Despite being widely discussed in mathematics education, the affective domain concept does not go beyond emotions and motivation. Emotions include enjoyment, anxiety, and boredom (Grootenboer and Marshman, 2016), while motivation includes individual interest, preferences, goal-orientation, and effort (Middleton, Jansen, and Goldin, 2016). Due to the limitations involved, the affective domain cannot fully address the challenges related to 21st-century society. Being an extension of the affective domain, social and emotional skills encapsulate a more comprehensive range of social skills such as communication and cooperation and cognitive skills such as creativity and critical thinking. These skills enrich the affective domain with a skill set that can effectively address modern-day challenges. Individuals' personal and social development at an early stage equips them with self-management and self-awareness skills that help them succeed in life. The interpersonal skills also make them more socially aware, enabling them to form positive relationships and make responsible decisions in life. Some other essential social and emotional skills are emotional stability, sociability, and perseverance, which enable individuals to transform their intentions into actions. Skills such as creativity and critical thinking provide individuals with a better understanding of the world and its issues and allow them to solve complicated real-life problems innovatively and creatively. Therefore, social and emotional skills enable individuals to regulate their emotions and provide them with skills to cooperate with others while maintaining healthy behaviours, thus enabling them to pursue long-term goals by thinking creatively and critically.

As the affective domain has limited coverage compared to the wide range of skills covered by social and emotional skills, it is preferable to consider the latter instead of the former. Incorporation of these skills is crucial for raising capable and responsible future citizens. However, it is also vital to develop these skills collectively rather than individually for an individual's holistic development. It is crucial to use various elements in an aligned manner to help contribute to a person's overall development. For instance, mathematics comprises various unique and high-demand skills that equip an individual to thrive in modern society. Skills such as problem solving, creativity, communication, critical thinking, self-regulation, flexibility, adaptability, collaboration, and innovation, when instilled in an individual collectively, empower them to find solutions and tackle modern-day challenges while being innovative, empathetic, mindful, and knowledgeable. The contributions of such an individual to the betterment of the environment and economies can make a significant difference. In contrast, an emphasis on developing each skill one at a time does not enable students to reach their full potential, consequently making them inactive society participants. Hence, the significance of developing these skill sets collectively speaks for itself.

Although several countries and organisations have proposed different frameworks to include social and emotional skills in mathematics, most frameworks are introduced based on survey findings, which investigate skills that are most in-demand for the 21st-century context. Moreover, the frameworks demonstrate some inconsistencies in categorising the skills into certain conceptual domains. For example, some frameworks identify perseverance (Hasratuddin, 2011; MOE, 2012, as cited in Wong, 2016) as an emotion, while others classify it as motivation (SEAL, 2007; Ontario Mathematics Curriculum, 2020). Additionally, some frameworks cover only social skills and emotional management (CASEL, 2017; SEAL, 2007), while other frameworks include cognitive skills—creativity and critical thinking (CCR, 2015; Ontario Mathematics Curriculum, 2020). This indicates a conceptual confusion between the frameworks on social and emotional skills and the initial concept of social and emotional development, which covers only social competence and emotional competence. Nevertheless, the identification of social and emotional skills in mathematics requires more discussion from a theoretical perspective.

Despite the focus on developing social and emotional skills in curricula, teachers cannot identify whether they have been successful in their efforts. A commonly agreed-upon reason for this gap is the difficulty in measuring social and emotional skills. Furthermore, there are limited authentic tools that can help teachers and policymakers to identify the areas that require further improvement (OECD, 2015; Ontario, 2016; OECD, 2018). Studies conducted by UNESCO Bangkok through the Asia Pacific Education Research Institutes Network (ERI-Net) found that most participating countries in Asia and Africa do not have the tools to assess social and emotional skills directly; however, the existing tools can be used to analyse these skills indirectly. Similar challenges can be found in the context of social and emotional skills in Mongolia. The main Mongolian education policies such as the ‘Government Policy on Education (2014–2024)’, the ‘Proper Mongolian Child’ programme, and the new national curriculum have placed great emphasis on developing creativity, teamwork, communication, and life skills and building character traits such as confidence, effort, and curiosity. For example, the mathematics curriculum, especially for the lower and upper secondary levels, focuses on effort, creativity, cooperative learning, aspiration to learn mathematics, and using different criteria to measure these skills (Ministry of Education, Culture, Science and Sports [MECSS], 2015). A domestic and a regional study reported a lack of assessment of pedagogical practice to measure both social and emotional skills at school as well as at the system level in Mongolia; however, the national curriculum has included these skills as assessment objectives (Amarjargal et al., 2016; UNESCO, 2016). In current assessment practice, educators and researchers have been attempting to measure social and emotional skills using self-ratings and questionnaires, while these instruments are subject to different types of biases, including socially desirable responses and cultural biases, which have the potential to affect the reliability and validity of the instruments. The implication is that the non-cognitive or social, and emotional aspects of mathematics need to be carefully developed. Hence, further research is also required to develop reliable and valid ways to measure these skills.

1.4 Research objectives

To address the issues discussed thus far, the present study set the following research objectives (RO).

RO1: To develop a theoretical framework for social and emotional skills in mathematics.

RO2: To develop tools to measure social and emotional skills in mathematics.

RO3: To validate the framework and tools by examining the status of social and emotional skills in mathematics among Mongolian students.

1.5 Research questions

The following research questions are central to achieve the research objectives:

RQ1: What framework can be used to capture social and emotional skills in mathematics?

RQ2: What valid tools can measure social and emotional skills in mathematics reliably?

RQ3: What is the status of social and emotional skills among Mongolian students?

1.6 Definition of the terms

AVs: AVs are short descriptions of hypothetical individuals that illustrate different levels of skills or traits (King et al., 2004).

Bias: '… the tendency of an estimation procedure to produce estimates that deviate in a systematic way from the correct value' (Education Testing Service, 2015, p. 54).

Construct validity: 'Extent to which a set of measured variables actually represents the theoretical latent construct those variables are designed to measure' (Hair et al., 2010, para. 7).

Content validity: 'The aspect of construct validity that emphasizes evidence bearing on the appropriateness of the knowledge, skills, and abilities measured by a test (Education Testing Service, 2015, p. 56).

Convergent validity: 'The items that are indicators of specific construct should converge or share a high proportion of variance in common, known as convergent validity' (Hair et al., 2010, para. 6).

Criterion validity: 'The aspect of construct validity that emphasizes evidence bearing on the statistical relationships between test scores and other variables of interest' (Education Testing Service, 2015, p. 56).

Discriminant validity: 'Extent to which a construct is truly distinct from other constructs both in terms of how much it correlates with other constructs' (Hair et al., 2010, para. 9).

Mathematical problem-posing: Problem-posing in mathematics is when students use their mathematical knowledge to generate relevant problems from the given information (Yuan & Sriraman, 2012).

Nomological validity: 'Test of validity that examines whether the correlations between the constructs in the measurement theory make sense' (Hair et al., 2010, para. 18).

Rater: 'A person or computerized procedure that assigns a score to a constructed response' (Education Testing Service, 2015, p. 60).

Reliability: 'The tendency of test scores to be consistent on two or more occasions of testing, if there is no real change in the test takers' knowledge' (Education Testing Service, 2012, p. 22).

Social and emotional skills: In the context of this study, social and emotional skills can be best defined as the part of an individual's holistic development that cannot be measured through standardised exams or conventional tests and can be improved through formal or informal learning experiences (UNESCO, 2016).

Validity: 'Extent to which the scores on a test are appropriate for a particular purpose' (Education Testing Service, 2012, p. 28).

1.7 Research limitations

This study has the following limitations:

1. First, the current study employed the convenient sampling method for data collection; the sample consisted of students from different geographical regions while considering the balance of the respondents in terms of gender.

2. The study employed a smaller number of questionnaires; however, it was intentionally decided to assess the effect of the AV approach on internal consistency with fewer items.

3. The study excluded some social and emotional skills, such as critical mathematical thinking and mathematical goal-orientation, for the reasons explained in Chapter 4.

4. An EFA and CFA were performed on the same sample in this study. It is desirable to use an EFA and CFA on two groups randomly selected from the same population.

5. The study employed the non-parametric approach to analyse the vignette sets. Researchers (e.g., Vonkova and Hrabak, 2015) have pointed out that the non-parametric approach has a disadvantage in dealing with order violations in the vignette analysis.

1.8 Research contributions

The present study has several important contributions to the existing literature. First, it is a pioneer study that conceptualised, designed, and validated a theoretical framework and instruments to measure students' social and emotional skills in mathematics. Second, the findings of various framework analyses on social and emotional skills revealed that skills under the term 'social and emotional skills' cover not only interpersonal skills and emotional competence but also cognitive skills such as creativity and critical thinking, which are beyond the initial concept of social and emotional development. Third, the study revealed that correcting questionnaires (i.e., self-ratings) with AV increased psychometric properties. Fourth, this study explored the status of social and emotional skills among Mongolian students for the first time using a combination of valid measures. Finally, the study contributes to the development and validation of research instruments.

1.9 Dissertation structure

This dissertation consists of the following six chapters:

Chapter 1 presents an introduction and rationale to this study, underlines the research problems, poses the research objectives and questions, and summarises the research methods used to address the research objectives.

Chapter 2 discusses global practices to develop social and emotional skills by exploring education policies and national curricula in various countries and reviews education policies and curriculum reform in Mongolia, focusing on social and emotional skills in the Mongolian mathematics curriculum.

Chapter 3 explains the sampling, measures, data collection procedures, data analysis, and statistical techniques employed to answer the research questions.

Chapter 4 focuses on designing the theoretical framework of this study by conducting a systematic review of articles from personality research, social and emotional skills, non-cognitive skills in mathematics, affect and attitudes in mathematics, and emotion and motivation in mathematics.

Chapter 5 explains the procedures for developing the tools in detail and provides the rules to correct questionnaire responses using AVs and score students' performance on problem-posing tasks.

Chapter 6 provides the outcomes of the data analysis according to each research question.

Chapter 7 summarises the main findings for each research question, implications, limitations, and recommendations for future research.

CHAPTER 2

SOCIAL AND EMOTIONAL SKILLS IN MATHEMATICS IN SOME COUNTRIES AND MONGOLIA

2.1 Introduction

This chapter discusses various countries' practices to grow social and emotional skills, exploring their national curricula focusing on mathematics curricula. Several countries were selected, including Singapore, Finland, and Canada (Ontario), to be compared to Mongolia for the following reasons. First, the selection of the first three countries is based on their higher performance on large-scale international assessments, which makes them model education systems globally. Second, these countries have also updated their curriculum due to the 21st-century challenges stemming from modern society, as discussed in the previous chapter.

The information and findings were drawn from an analysis of original sources and a literature review. Besides the inclusion of social and emotional skills in national curricula, global issues related to assessing social and emotional skills are also identified at the end of this chapter.

2.2 Education reform and competencies in the selected countries

Only recently, educators and policymakers unanimously decided to include social and emotional skills in education curricula. Various countries have identified the necessary skills for their future citizens and have included them in policy documents and general education. According to Ontario (2015), many of these skills are now being integrated across the education curricula. Many countries around the globe encourage their schools to reform their curricula so that the students' social and emotional skills can be developed across all subjects, including the compulsory subjects of languages and mathematics. This section discusses how different countries included social and emotional skills into their education policies and curricula, focusing on 1) educational reform, 2) curriculum reform, 3) curriculum structure, and 4) mathematics curriculum.

2.2.1 Singapore

Educational reform: Singapore's education system has been working on the vision 'Thinking schools, learning nation' since 1997 to prepare the younger generations to tackle the modern world challenges. This vision aims to create a feasible learning environment for students, teachers, companies, parents, workers, organizations, and the government. Such an environment helps an individual with personality development while preparing him to deal with the world. The policy to teach less so that the students can learn more emphasizes the importance of other skills learned in schools apart from just focusing on grades. To strengthen the character and instil citizenship values in students, the Character and Citizenship Education (CCE) branch was set up in December 2011, and the 'Values in Action' program was started to encourage students to volunteer and become responsible and empathetic citizens (The Ministry of Education, 2012).

Curriculum Reform: Many countries in Asia reformed their mathematics curricula, mainly within the framework of national educational renovation (Catherine & Toh, 2019). The Ministry of Education (MOE) in Singapore constantly revisits and updates its curriculum to ensure that it contains the knowledge and skills required to cater to 21st-century challenges. Recent curriculum reform was initiated in 2012 under the 'Student-Centric and Values-Driven Phase' (MOE, 2014). It aims towards providing students with a standard set of values, knowledge, and competencies while also allowing differentiation to meet the varied needs of students with different abilities and talents.

Curriculum Structure: The MOE is committed to updating the curriculum so that students are better prepared to thrive in modern society nationally and globally. To enable the students to achieve the learning outcomes of each subject, three broad areas are carefully designed: 1) curriculum, 2) teaching strategies, and 3) assessment. The basic education in Singapore comprises ten years, where compulsory primary education is of six years and secondary education for four years, after which students can choose to opt for a post-secondary institute. Students study English, their native language, civics, mathematics, science, social studies, moral education, music, arts and crafts, physical education, and health education for six years at the primary level. By the end of this period, students go through a Primary School Leaving Examination (PSLE), which is meant to examine their ability to enrol in a suitable secondary school according to their talents and inclinations. Upon entering secondary school, students are taught a mix of compulsory and elective subjects. The compulsory subjects are mother tongue language, English language, combined humanities, and a science subject. This is to ensure that students are provided broad-based and balanced education. The elective subjects include a subject of humanities, science, and literature in Chinese. These various subjects are meant to cater to the different interests and abilities of students. Foreign languages like German, French, and Japanese are also offered to students to widen their opportunities.

The Singapore MOE has identified essential skills to prepare students for a globalized and technologically advanced world. The national curriculum of Singapore aims to prepare its young students to grow up to become 1) confident individuals, 2) concerned citizens, 3) self-directed learners, and 4) active contributors. To achieve these objectives, a set of social and emotional skills were included as 21st-century factors in the curriculum, namely 1) critical and inventive thinking; 2) communication, collaboration, and information skills; and 3) civic and global awareness and cross-cultural skills (MOE, 2014). Through these social and emotional skills, a person becomes capable of self-management, relationship management, and responsible decision-making while being more self-aware and socially aware. Under 21st century competencies, these skills are incorporated through languages, mathematics, sciences, humanities, physical education, arts, music, CCE, co-curricular activities, and school-based programmes. Mathematics is essential as it is used to understand real-world problems. Therefore, Singapore gives particular importance to mathematics as it provides essential knowledge and skills to students, enabling them to think critically and rationally to contribute to society.

Singapore mathematics curriculum: Singapore is a top-performing country from East Asia on PISA in mathematics. The Singaporean MOE emphasizes the need to prepare students for a globalized and technologically advanced world. Therefore, the mathematics curriculum primarily focuses on developing mathematical problem-solving competency among students. This main goal is backed by five interrelated components: skills, attitudes, processes, metacognition, and concepts. The mathematical curriculum framework consisting of these components is depicted in Figure 1.

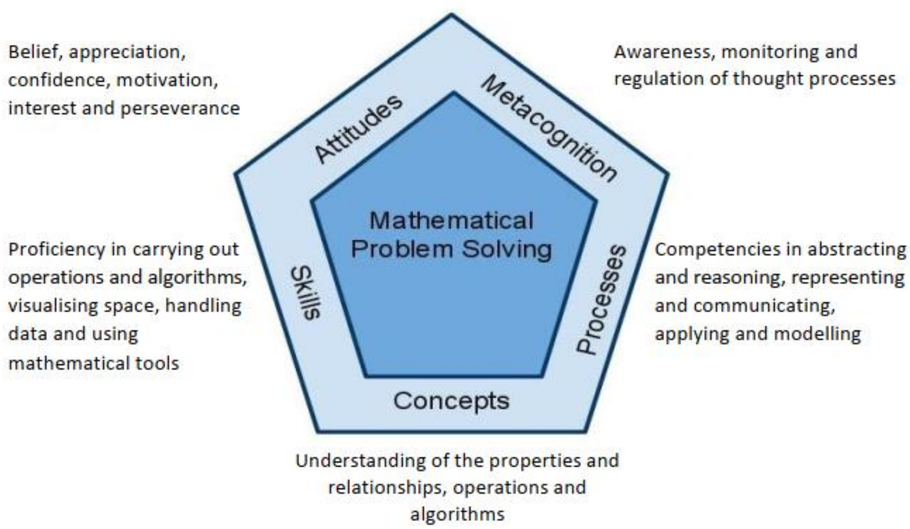


Figure 1. Singaporean Mathematical Curriculum Framework To achieve this aim, the MOE included several skills such as critical thinking, creativity (cognition), learning to learn (metacognition), grit, emotional intelligence (emotions), and communication and collaboration (social intelligence) in

the mathematics curriculum as essential skills for 21st-century society (MOE, 2012, as cited in Wong, 2016). The components of these skills are depicted in Table 1.

Table 1. Social and emotional components in Singapore's mathematics curriculum

Domain	Sub-skills
Cognition	Critical thinking, creativity
Metacognition	Learning to learn
Emotions	Grit, emotional intelligence
Social intelligence	Communication, collaboration

Adopted from Wong, 2016

Singaporean mathematics education provides opportunities for students to develop competencies that are crucial in the 21st century. By engaging in problem-posing, justifications, critiques, and arguments, students learn to reason, think critically, and communicate. By simplifying a complicated real-world problem, they learn to tackle ambiguity. Social and emotional skills are incorporated as a component of the 21st-century competencies in the Singaporean secondary school mathematics curriculum. The content of the curriculum is highly relevant to today's modern society. The pedagogies develop critical thinking, logical reasoning, effective communication, and individual and collaborative work. Problem solving makes them more aware of the global and social issues around them.

2.2.2 Finland

Educational Reform: Finland has one of the best educational systems according to international assessments such as PISA and the Trends in International Mathematics and Science Study (TIMSS) in mathematics, science, and reading. The Constitution of Finland, Government Decrees, the National Core Curriculum, and the Basic Education Act and Decree have made it compulsory for local authorities to educate children in their locality. The Basic Education Act promotes learning according to children's age and abilities to contribute to growth and development. The Pupil and Student Welfare Act ensures the safety of a student in the learning environment. According to the Constitution of Finland and the Non-Discrimination Act, no child can be discriminated against based on age, gender, ethnicity, nationality, religion, language, beliefs, disabilities, or other personal traits. The Act on Equality between Women and Men ensures that all educational institutes provide equal opportunities to men and women. Finland adheres to various international human rights treaties that require the government to ensure the welfare and learning of every child (Finnish National Board of Education, 2016).

Curriculum Reform: Recently, between 2014 and 2017, the Finnish National Board of Education changed its national core curricula at all levels of education—pre-primary, primary, and upper secondary—to ensure that students are prepared to meet the challenges of the modern world and contribute to a sustainable future (Halinen, 2018; Braskén, Hemmi, & Kurten, 2019). The National Core Curriculum is devised in line with the Government Decrees and the Basic Education Act. The national core curriculum is devised considering each child's unique abilities, every child's right to education, development, diverse cultural heritage, and sustainability (Finnish National Board of Education, 2016).

Curriculum Structure: According to the Government Decree, lesson hours are reserved for subjects such as crafts, music, home economics, physical education, and visual arts for Grades 1–9. Moreover, nine annual weekly lessons are also allocated as optional subjects for Grades 1–9. The applied optional subjects also contain various transversal competencies (the regional term for social and emotional skills). One of the core subjects includes mother language and literature in which students study Finnish, Swedish, Roma, Sami, sign language, and other native languages.

Other subjects include English, mathematics, environmental studies, religion, ethics, music, visual arts, crafts, physical education, and guidance counselling (Finnish National Board of Education, 2016).

Finland mathematics curriculum: Mathematics curriculum reform has been carried out in Asia and Northern European countries. The new Finland mathematics curriculum, specifically the basic education mathematics curriculum, emphasizes problem-solving skills, motivation (e.g., perseverance, goal-orientation), positive self-image, and self-confidence among mathematics learners (Finnish National Board of Education, 2016). Apart from the above skills, the new mathematics curriculum contributes to developing transversal competencies (regional terms for social and emotional skills) embedded in all subjects. The competencies include 1) thinking and learning; 2) cultural competence, self-expression, and interaction; 3) managing and taking care of oneself; 4) multiliteracy; 5) ICT competence; 6) entrepreneurship and working life competence; and 7) participation and contribution to a sustainable future (Finnish National Board of Education, 2016).

Thinking and learning (T1) refer to the way individuals see themselves as learners. The way they interact with their surroundings affects their way of thinking and learning. It enables them to observe, analyse, and express their ideas. Cultural competence, interaction, and self-expression (T2) enable individuals to respect cultural diversity and be flexible in different environments. Respectful and constructive interactions are encouraged to make them more empathetic and understanding towards others. The school environment fosters their personal development by developing social skills.

Managing and taking care of oneself in daily life (T3) requires a much more comprehensive range of skills. It encapsulates health, relationships, safety, functioning in a technological world, and personal and financial management. Multiliteracy (T4) is the ability to interpret and judge to understand cultural diversity. It involves critical thinking and is taught in all school subjects. Competence in ICT (T5) is a civic skill and a learning tool. Entrepreneurship and working life competence (T6) equip individuals with collaboration and communication skills to take risks and be self-employed. By learning to work independently and in teams, students gain the necessary skills to be an entrepreneur. Participation and contribution to a sustainable future (T7) are the competencies offered in a safe school environment. They enable students to become responsible and active future citizens. By being guided about the consequences of their choices, actions, and lives, they are taught to be responsible for themselves and their environment, contributing to a sustainable future.

2.2.3 Canada (Ontario)

Educational reform: Canada is one of the top ten performing countries in mathematics on PISA 2018. Even before the educational reforms, Ontario's educational performance met international standards. However, literacy and numeracy were two aspects where students needed improvement. To cater to this need, the Literacy and Numeracy Strategy was introduced in 2004 to assist students in reading, writing, and mathematics (Boyd, 2021). These reforms proved to be successful in improving the students' performance in literacy and numeracy. Despite meeting the target set by the education reforms, students' performances are often lagging. For example, while the reading scores improved, the maths scores remained the same (O'Grady et al., 2019). Thus, a critical challenge for Ontario lies in changing the trajectory of its math scores. Therefore, improvement of maths scores has been a challenging task for Ontario over the last decade (Boyd, 2021).

Moreover, an analysis conducted in 2013 revealed that Ontario policy documents have no formal discussion on the inclusion of 21st-century-related skills. To prepare the students to be a part of the workforce in the future, the policy framework emphasized 'hard skills' without including any skills related to character development (Action Canada Task Force, 2013). Consequently, a learning movement was initiated in 2015 in Ontario and encapsulated skills such as citizenship, communication, character, collaboration, critical thinking, and creativity as the learning outcomes of the new movement (Boyd, 2021).

Curriculum reform: Following the movement, Ontario revised its primary and secondary education curriculum in 2020. The new curriculum aims to develop social and emotional skills across the curriculum.

Curriculum structure: As for Ontario, the subjects that inculcate social and emotional skills in basic education curriculum are arts, business studies, classical studies and international languages, computer studies, cooperative education, English, French as a second language, guidance and career education, health and physical education, interdisciplinary studies, mathematics, native languages, science, social sciences and humanities, and technological education.

Ontario mathematics curriculum: The revisions to the new basic education curriculum (Grade 1 to Grade 8) specifically concerning social and emotional skills in the mathematics curriculum. The revised mathematics curriculum aims to equip students with social and emotional skills through specific mathematical processes. The social and emotional skills in the Ontario Mathematics Curriculum are as follows: 1) recognizing and regulating emotions; 2) identifying causes of stress and coping with challenges; 3) being motivated and persistent; 4) building relationships and communicating effectively; 5) promoting self-awareness and a sense of identity; and 6) thinking critically and creatively (The Ontario Math Curriculum, 2020, pp. 79–84). The students must learn these skills to develop understanding, resilience, optimism, decision-making, problem solving, and a sense of identity. The social and emotional skills are learned through the mathematical processes, as illustrated in Figure 2.

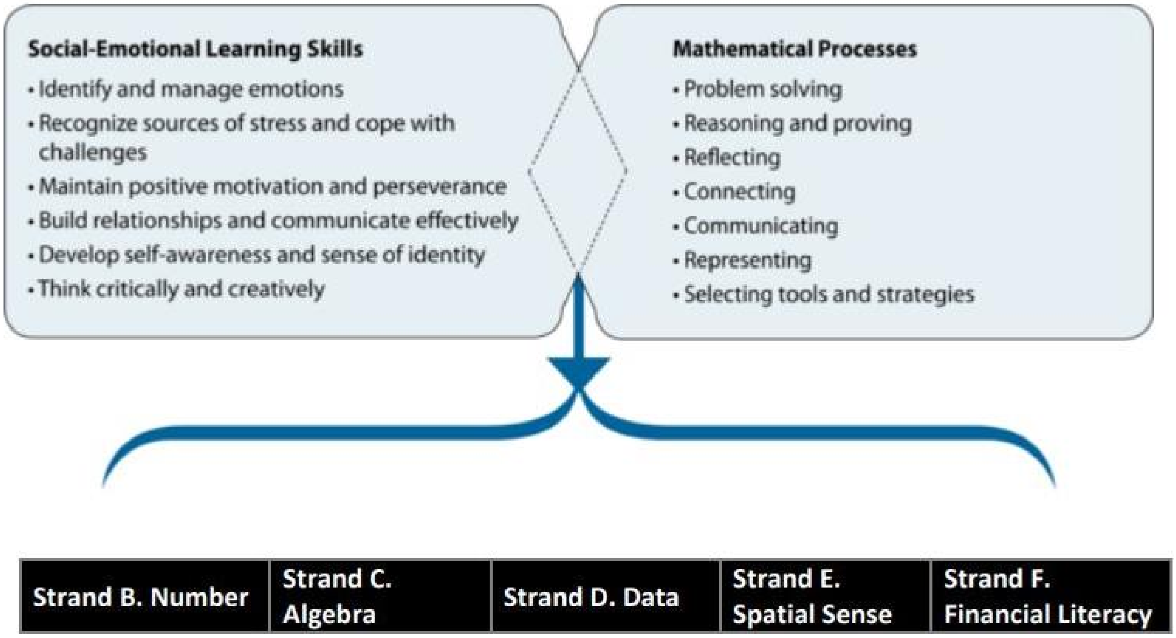


Figure 2. Social and emotional skills and mathematical processes in the Ontario mathematics curriculum

2.2.4 Mongolia

Educational reform: This section clarifies education policies and curriculum reform in Mongolia, focusing on the mathematics curriculum. The main goal of the Government Education Policy in Mongolia (2014–2024) is to educate individuals by developing their skills to work efficiently and live content and ethical lives while aspiring to learn for a lifetime. According to the Policy, primary education aims to make Mongol children competent in their mother tongue and promote creativity in the learning process. In contrast, lower secondary education intends to inculcate life skills, basic science knowledge, and independent and creative learning in students. Similar to the countries mentioned above, the Government of Mongolia also launched education reforms in 2013 to address the challenges stemming from 21st-century-society. Within the framework of the educational reform, the 'Proper Mongolian Child' programme (2013–2016) was introduced to raise confident, creative, and decisive citizens who are eager to learn and live together while adopting a national culture. Hence, there was a need to devise policies to implement an updated national curriculum for primary and secondary school students.

Curriculum reform: Because of the 'Proper Mongolian Child' programme, a new curriculum for the primary level was implemented across the country in 2014–2015, the lower secondary level in 2015–2016, and the upper secondary level in 2016–2017. A new curriculum for the primary level was introduced in 2014, for the lower secondary level in 2015, and the upper secondary level in 2016. The new national curriculum has a concept note that states:

> Nowadays, many countries have been developing educational policies aiming to prepare citizens who are capable of being flexible and adaptable to science and technology advancement and to live in an open society in the future. Therefore, the lower secondary curriculum will be designed within the goal of developing 'patriotic' Mongolian citizens who are creative, confident, and proficient in decision-making, cooperating, and lifelong learning. (MECSS, 2015, p. 18)

The primary education national curriculum sets out eight general skills in three domains, as illustrated in Table 2.

Table 2. Primary education learning goals on skills

Domain	General skills
Cognitive skills	Creative thinking Problem solving
Linguistic and social skills	Communication skills Collaborative learning
Life skills	Ability to learn independently Information and technology skills Healthy lifestyle Environmental awareness

Source: Mongolian Primary Education Curriculum, 2014, Article 5.1.2, p. 5.

After implementing an updated primary level curriculum in the academic year 2014–2015, preparations were undertaken to implement the lower secondary curriculum from the academic year 2015–2016 at a national level.

The three main goals of assessment are:

1. To develop a learning strategy
2. To encourage scientific knowledge and understanding
3. To develop willingness, interest, and motivation to learn and inquire.

In the context of the lower secondary education national curriculum, inquiry-based learning is the keyword.

Alongside subjects, extra-curricular activities are also meant to promote social and emotional skills in students. The school-based curriculum, integrated study hours, civic education, and extra-curricular activities were designed as a package called ‘Learning support activities’ under the 2013 curriculum reform. The national curriculum handbook states that learning life skills also involves more complicated skills such as teamwork and problem solving. Similarly, the assessments focus more on learning techniques and interpersonal skills such as communication and cooperation. Civic and moral education is meant to develop student habits and instil respect for national customs and traditions. As for daily habits, the primary curriculum aims to work on student behaviours such as respect, sociability, communication, and participation within the class and family environment. Although social and emotional skills are incorporated in learning support activities, the standard assessments do not suffice when assessing these skills. As opposed to the primary and lower secondary curricula, the upper secondary curriculum comprises two main parts: 1) mandatory curriculum and 2) selective curriculum. The general objective of the mandatory curriculum is to raise individuals with fundamental knowledge of science and research methods and instil creativity and civic and social values in them while encouraging lifelong learning behaviour. However, the selective curriculum is designed to provide an opportunity for students to gain a deeper understanding of science and be able to identify their areas of interest and develop their strengths in those areas, which can then help them decide on suitable careers.

Curriculum structure: Mongolia's primary curriculum devised for Grades 1-5 mainly focuses on providing knowledge to students by improving their thinking and learning skills. Every Mongolian child must learn Mongolian language; however, English is also taught as a language. Other subjects must learn to include mathematics, human and environment, human and nature, human and society, arts, physical education, and health education. These subjects also cover some aspects of social and emotional skills.

The lower secondary education curriculum is designed for students of Grade 6-9. It focuses on scientific knowledge and understanding and developing a sense of motivation, curiosity, and willingness to learn among students. To achieve these goals, it offers compulsory subjects such as Mongolian language, literature, Mongolian script, mathematics, information technology, physics, chemistry, biology, geography, history, social studies, fine arts, music, graphics, technology, physical and health education, English, and Russian. Moreover, civic education is also taught as a subject and aims to raise dutiful students who respect their customs and traditions. Likewise, the co-curricular activities are also an integral part of the lower secondary curriculum and are meant to develop social and emotional skills in students. Like the primary education curriculum, the lower secondary education curriculum also focuses on developing social and emotional skills in students; however, the assessment criteria of both curricula differ.

Similarly, the upper secondary education curriculum is designed explicitly for Grade 10-12 students to develop their scientific knowledge, research skills, creativity, and learning. It includes a mix of compulsory and elective subjects such as Mongolian language, literature, Mongolian script, civic and ethical education, mathematics, information technology, physics, chemistry, biology, geography, history, social studies, English, Russian, design, graphics, technology, and physical and health education. The curriculum structure is designed to enable students to identify their interests and choose a suitable career path.

Mongolian mathematics curriculum: The general objective of the Primary mathematics curriculum is to provide fundamental mathematical knowledge to students and improve their thinking skills while developing their interest in mathematics so that they can solve daily life problems involving calculations. The primary mathematics curriculum has five major components: 1) Objectives (what is the need to learn mathematics?); 2) Content (what should be learned?); 3) Teaching and learning strategy (how should the learning take place?); 4) Teaching aids and materials (what teaching aids should be provided?); and 5) Assessment (what should be assessed and how?).

In the context of social and emotional skills, the 'teaching and learning strategy' includes inculcating creativity and confidence in students through the learning process to express their ideas freely and learn mathematics individually and cooperatively.

The assessment part states that students' mathematical knowledge and skills must be assessed through summative and formative assessments. Hence, the mathematics curriculum for primary education places greater emphasis on assessing skills and knowledge. Moreover, the curriculum objectives and teaching and learning strategies also cover some social and emotional aspects.

Lower secondary mathematics curriculum aims to enable students to work with numbers, shapes, variables, and dimensions and help them to become proficient in problem solving so that they can apply their mathematical knowledge in their daily lives while also aspire to learn mathematics and understanding its usefulness (Lower Secondary Education Curriculum 7.1). The curriculum document also includes the assessment criteria. However, the association between primary and lower secondary still needs to be explained to teachers so that they are aware of what and how to assess. This is also because the assessment criterion of lower secondary education is very different from that of the primary level.

In the lower secondary mathematics curriculum, creativity, cooperative learning, aspiration, and effort are considered assessment goals. The curriculum also endeavours to provide a criterion for measurement; however, the criterion remains unclear as it contains only a few keywords (Table 3).

Table 3. Assessment objectives and criteria: Lower secondary

#	Assessment objectives	Assessment criteria
1	Mathematical knowledge and understanding	To know and understand: – mathematics terminology – mathematics laws and rules – mathematics principles and relationships
2	Application of mathematical knowledge and understanding and strategies for mathematical problem solving	To be able to: – use mathematical knowledge and understanding in learning and their lives – choose and plan their strategies to solve mathematical problems – use mathematical knowledge in implementing their plan to solve mathematical problems – draw conclusions and provide well-reasoned results.
3	Elaboration and thinking skills	To be able to: – express their ideas using mathematical language and symbols – identify mathematical reasons and relationships in

		objects and phenomenon – look at the problems from various angles and solve them in multiple ways – categorise, compare and order, and identify special and particular cases
4	Effort, interest, and attitude	To be able to: – learn cooperatively and independently – be interested in and curious about mathematics

Source: Mongolian Lower Secondary Mathematics Curriculum, 2015, 7.6.2, p. 45

Furthermore, the upper secondary mathematics curriculum aims to enable students to process scientific information in their daily lives to help them in problem solving and decision-making by using their mathematical knowledge and creativity (Upper Secondary Mathematics Curriculum 1.1).

Table 4. Assessment objectives and criteria: Upper secondary mathematics curriculum

#	Assessment objectives	Assessment criteria
1	Mathematical knowledge	To able to: – convert numbers into their equivalent form and approximate (with given precision) – do calculations – do geometry in the plane and space and their relationship – know mathematical terminology – organise, interpret, and explain information and data (in written, table, graph, and diagrammatic forms)
2	Application of mathematical knowledge in problem solving	To be able to: – calculate following the rules – use materials (teaching aids) to measure, draw, and construct – use standardisation and measurement in problem solving – use mathematics in daily life

		– detect structures, principles, and general forms of situations – solve problems in unfamiliar situations based on familiar situations – use appropriate strategies for better outcomes – use mathematical knowledge and skills in problem solving
3	Thinking skills	To be able to: – design the problem mathematically, analyse, synthesise, and generalise – draw conclusions based on mathematical evidence – propose mathematical proofs and conclusions – propose and prove hypotheses using mathematics
4	Attitude	– Interest – Effort and engagement – Creativity

Source: Mongolian Upper Secondary Mathematics Curriculum, 2016, 4.2, pp. 47–48

According to the policy review, creativity is an essential skill out of all social and emotional skills in the education system of Mongolia. Therefore, significant policies such as the Government Education Policy (2014–2024), Primary Curriculum (2014), 'Proper Mongolian Child' program (2013–2016), Lower Secondary Curriculum (2015), and Upper Secondary Curriculum (2016) consider creativity as a crucial skill that every Mongolian child must possess.

Moreover, cooperative learning and teamwork have been explicitly mentioned in the policy documents of all three levels of curricula. The updated lower secondary education curriculum focuses on developing learning skills in students, such as cooperative learning and inquiry-based learning, which allow them to be curious and explore.

It is also interesting to note that critical thinking, which is one of the most widely discussed social and emotional skills, has no explicit mention in the education policies of Mongolia.

In general, these policy documents in Mongolia aim to raise self-regulated and self-motivated learners who are creative, confident, and cooperative and exhibit curiosity in problem solving and decision-making while also aspiring to be lifelong learners with a positive attitude.

In the context of social and emotional skills, Mongolia's revised mathematical curriculum, especially for the lower and upper secondary levels, focuses on effort, creativity, cooperative learning, aspiration to learn mathematics, and using different criteria to measure these skills (MECSS, 2015). It is evident that the Mongolian educational policies, especially its mathematics curriculum, have laid massive emphasis on developing creativity, teamwork, cooperation, and building character traits such as confidence, effort, and curiosity. However, the primary mathematics curriculum categorizes creativity as a thinking skill in its general objectives, whereas the upper secondary mathematics curriculum categorizes it as 'attitude' under its assessment criteria. This is in line with the argument that there is a weak correlation between the objectives, content, and assessment of all three curricula.

Unlike other countries, such as Singapore, Finland, and Canada, Mongolia's mathematics curriculum has no explicit inclusion of critical thinking in all three levels of education. As evident from the curriculum structure, all four countries have included social and emotional skills in their curricula using different terms. These skills are included as '21st-century skills' in Singapore's educational curriculum, 'social and emotional learning skills' in Ontario's curriculum, and 'transversal competencies' in Finland's curriculum. As for Mongolia's curricula, similar coverage of these skills is included despite using no term for these skills.

Despite all the focus on developing social and emotional skills in education policies across different education systems, educators and teachers can still be unable to identify whether they have been successful in their efforts. A commonly agreed-upon reason for this gap is the difficulty in measuring social and emotional skills. Furthermore, there are limited authentic tools that can help teachers and policymakers to identify the areas that require further improvement (OECD, 2015; Ontario, 2016; OECD, 2018). Studies conducted in Asia and Africa have demonstrated that most countries do not have the tools to assess social and emotional skills directly (UNESCO, 2015).

Like the challenges related to assessing social and emotional skills in a global context, there has been a lack of assessment of pedagogical practice to measure social and emotional skills at the school level and the system level in Mongolia. However, the national curriculum has included these skills as assessment objectives.

Assessment of social and emotional skills is not easy for practical teaching, despite the teachers being aware of policy documents (Munkhjargal et al., 2016). The only guides available are primary, lower secondary, and upper secondary education national curricula, some of which include a very short introduction of assessment. Apart from that, no other handbooks or guidelines explicitly describe the assessment of social and emotional skills (Amarjargal et al., 2016).

CHAPTER 3

METHODS FOR DEVELOPING AND VALIDATING THE TOOLS AND FRAMEWORK AND EVALUATING STUDENTS' PERFORMANCE

3.1 Introduction

This chapter explains the methods and techniques used in this study. The chapter begins with methodological solutions against the biases related to self-rating items, followed by an introduction of AVs and the problem-posing approach. The chapter explains the sampling, measures, and data collection procedures. Finally, statistical techniques such as the EFA and CFA, which were employed to answer the research questions, are explained at the end of the chapter.

3.2 Methodological issues and the solutions related to measuring social and emotional skills

Previous studies have demonstrated that the most common method of assessing students' social and emotional skills is questionnaires and self-rating (Otgonbaatar, 2021). Many mathematics researchers also make use of questionnaire surveys to evaluate these skills. For instance, creativity researchers use this approach to analyse creative thinking and willingness towards new experiences (Lubart & Guignard, 2004). In the context of mathematics, measures of emotions such as enjoyment, anxiety, motivation, and perseverance are determined through self-rating (Perels et al., 2005; Schillinger et al., 2018; Gundersan et al., 2017). Questionnaire is also employed to measure mathematical self-efficacy, cooperation skills, and learning strategies in mathematics (Lavasani & Khandan, 2011; Hossain et al., 2013; Zakaria et al., 2010).

3.2.1. Biases

Questionnaires and self-rating are considered a more practical, efficient, and cost-effective method of collecting information related to an individual. However, when individuals self-rate their traits and behaviours, it leads to systematic biases that affect the utility of this method and data validity (Paulhus, 1991; Kyllonen & Bertling, 2014; Weiss & Roberts, 2018). Cultural bias is one of the most frequent response biases that cause differences in responses due to varying communication styles among different cultures (Van de Vijver & Leung, 1997; Fischer, 2004). Another frequent response bias is social desirability responding, which includes midpoint responding, extreme responding, and acquiescent responding (which refers to the recurring use of the upper half of the response option scale due to the tendency of agreeing with items; Paulhus, 1991).

Cultural biases: Studies employing large-scale international databases have revealed that students in the US exhibit higher levels of mathematical self-efficacy than students in East Asia do; however, their mathematical achievement remains lower than the OECD average levels (OECD, 2004; Lee, 2009). Nevertheless, students in East Asian countries, particularly Korea and Japan, have demonstrated lower levels of mathematical self-efficacy than the OECD average, despite their high achievement in mathematics (Lee, 2009; Scholz et al., 2002). This might be due to the influence of the Chinese heritage culture, which makes them more cautious of opting for end-point items on the Likert-type response scale (Leung 2001, 2002), and because Chinese and Japanese students are more inclined towards choosing the midpoint of a Likert-type response scale (Chen et al., 1995; Lee et al., 2002). Similarly, individuals in East Asia are more likely to rate themselves much lower in the conscientiousness category than people in any other region (Schmitt et al., 2007).

Social desirability bias: Socially desirable responding (SDR) refers to a respondent's tendency to choose options that make them look good (Paulhus, 1991). For instance, if a questionnaire includes an item like 'I am hardworking', a child, teacher, or parent might be inclined to opt for a highly rated option to appear good to either the observer or themselves. Although social desirability bias is uniform within a study population, it can affect the absolute level of individual responses without altering the rank order. Some individuals can be affected more by social pressure than others; hence, it affects their relative placement in the overall distribution of responses. The issue of SDR on the validity of questionnaires was raised by psychometricians many years ago (e.g., Bernreuter, 1933; Vernon, 1934), and researchers have devised various methods of data collection through self-rating while also minimising biases (Paulhus, 1991)

3.2.2 Methodological considerations

Forced-choice questions: Instead of asking the students to rate their liking of mathematics on a five-point Likert scale, this technique asks them to choose between mathematics and science. The objective of the forced-choice format is to address the faking problem. Based on a ranking or preference format, the forced-choice format requires respondents to choose a statement that describes them better. Successful experiments for forced-choice methods were carried out in PISA 2012 field trials. The style problems common in cross-cultural and international surveys are reduced to a minimum level in forced-choice methods.

Situational judgment tests: These contain items with different situations, and the respondents are asked to opt for their typical response to a situation in the form of text or multimedia. The responses can be in the form of a multiple-choice (i.e., choosing the best option), constructed response (i.e., providing a response to a particular situation), or rating (i.e., rating each response according to its effectiveness on a Likert-type scale; e.g., McDaniel, Morgesen, Finnegan, Campion, & Braverman, 2001). This technique allows the measurement of various attributes such as social competence, critical thinking, leadership, and communication skills (e.g., Oswald et al., 2004; Waugh & Russell, 2003). With their ability to measure these constructs, situational judgment tests are more likely to resolve the validity issue of conventional tests for university admissions and personnel selection. The internal consistency reliabilities for situational judgment tests are lower (0.40) than other tests

(Catano, Brochu, & Lamerson, 2012).

Anchoring vignettes: AVs are another useful method to improve data quality and the cross-cultural comparability of social and emotional skill assessments while reducing response biases (Kyllonen & Bertling, 2014). It is a relatively new method that aims to improve the validity of ratings that describe hypothetical situations or persons. AVs require the respondents to rate various hypothetical persons or situations on the construct of interest and later uses those ratings as anchors to place a respondent's actual ratings (Hopkins & King, 2010). For instance, a respondent might be asked to rate three individuals based on a skill, such as teamwork, and then rate themselves. The rating by the respondent is completed using a rating scale that is also used for self-rating (e.g., a five-point agreement scale to represent choices from 'strongly disagree' to 'strongly agree'). The self-rating is then recoded into a new score by a non-parametric scoring procedure that demonstrates if the respondent rated themselves above the highest-rated vignette, at the level of the rated vignette, or below the lowest-rated vignette. The new score generated is used for further analysis. PISA 2012 employed this method to analyse questionnaires, and the results demonstrated a substantial improvement in the validity of the questionnaire response. This resolves the rating scale issues stated above.

The AVs approach was found to have a positive and significant impact on the validity and reliability of survey instruments in various aspects of social sciences and provide evidence about the positive effect of AVs, the method is still new in the education field, particularly in mathematics education.

In this regard, the current study attempted to use this method as a practical solution for measuring social and emotional skills in mathematics education by employing information from AVs to correct self-ratings.

Problem-posing: According to Kyllonen (2012), it is not easy to measure some cognitive skills such as creativity. On the contrary, a study by the National Research Council (2012) demonstrated that certain authentic and well-established testing techniques are used to measure cognitive competencies.

An instrument to measure mathematical thinking abilities has been developed by creative researchers such as Torrance and Balka (Silver, 1997): the Torrance Tests of Creative Thinking (TTCT) were developed by Torrance, while Balka developed the Creative Mathematical Ability Test (CAMT). However, these tests rely solely on multiple-choice questions, which have the potential to assess only a few aspects of a person's creativity (Sternberg, 2006).

Jensen (1973) reported that the ability of students to pose mathematical problems is dependent on their creativity. Problem-posing in mathematics is when students use their mathematical knowledge to generate relevant problems from the given information (Sriraman and Lee, 2011). For instance, students are given a bar chart or graph and are instructed to formulate questions from the given information (e.g., Prouse, 1964; Jensen, 1973; Balka, 1974). Posing the problem is considered the most crucial part of creative problem solving (Kim, 2009).

A mathematical example of problem-posing presented by Silver (1997) demonstrates different ways to assess mathematical creativity in students. It includes measures of mathematical flexibility (posing problems with various solutions), mathematical fluency (generation of maximum problems), and mathematical originality (posing a unique problem).

Silver (1997) claims that the practice of problem-posing helps students with creativity in mathematics. Hence, while assisting researchers with mathematical creativity, it also develops basic elements of creativity such as flexibility, fluency, and originality in students (Torrance, 1988;

Silver, 1997). Therefore, the present study endeavoured to measure mathematical creativity, emphasising flexibility, fluency, and originality while applying the problem-posing approach.

In summary, AVs and problem-posing are introduced as a methodological solution against the issues related to the measurement of social and emotional skills in this study.

3.3 Research sample

The present study had several concerns in terms of deciding on the sample. First, the students' experiences related to the new curriculum were considered, as the development of specific social and emotional skills (e.g., creativity) is emphasised in the curriculum objectives. In this regard, ninth-grade students were selected according to their experience in relation to the new curriculum because ninth-grade students had been taught the new curriculum from primary school.

A total of 308 ninth-grade students comprising 151 males (49 per cent) and 157 females (51 per cent) from eight public schools located in urban and rural areas were sampled for this study using the convenience sampling method (Table 5). The ages of students ranged from 13 to 16, with a mean age of 14.0 (SD = 0.51). The school information is provided in Table 6.

Table 5. Social demographic characteristics of the study sample

	N	Per cent
Individual characteristics		
Male	151	49.0
Female	157	51.0
Age groups		
13	36	11.7
14	237	76.9
15	32	10.4
16	3	1.0

Ethnicity		
Mongolian	271	88.0
Kazakh	37	12.0
Regional characteristics		
Urban	116	38.7
Rural	192	61.3

Table 6. School information

School	School size	Type	Regional characteristics	Selected sample
School 1	1,100	Public	Urban	36
School 2	1,356	Public	Urban	28
School 3	1,760	Public	Urban	52
School 4	2,150	Public	Rural	39
School 5	2,410	Public	Rural	58
School 6	2,560	Public	Rural	40
School 7	832	Public	Rural	28
School 8	1502	Public	Rural	27

According to National Statistics Committee Mongolia, Mongolia is categorised into five main geographical regions: Central, Western, Eastern, Forest, and Desert.

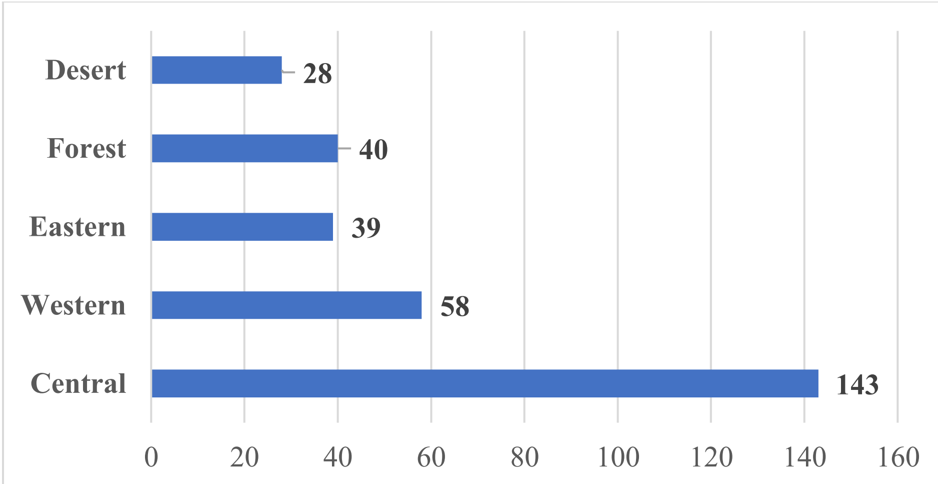


Figure 3. Distribution of the participants by geographical region N = 308 Some international and domestic studies have found statistically significant differences in students' mathematics achievement between rural and urban areas in Mongolia (Otgonbaatar, 2013; Education Evaluation Centre, 2017, 2018). It has also been demonstrated that there were statistically significant differences in Grade 9 students' mathematics achievements among the five geographical regions (Education Evaluation Centre, 2017, 2018). It was also revealed that ethnic minority (Kazakh) students performed worse in mathematics than their Mongolian peers did (Education Evaluation Centre, 2016, 2017, 2018). In this regard, regional and geographical characteristics were considered, and research participants were distributed across the five regions (Figure 3). It is hoped that the diverse distribution of the sample supports the research findings as it represents social and emotional aspects in mathematics across the country.

3.4 Data collection procedure and measures

The data were gathered during regular class time (45 minute) in September and October of academic year 2019–2020. To collect the data, visits were made to schools, and the research objective was explained to the principals and mathematics teachers. The data collection process continued for two consecutive days at each school. On the first day, 15-self-rating items were administered to the participants to state their degree of agreement or disagreement regarding how accurately each self-rating item described them using a five-point Likert rating scale (1 = disagree strongly, 2 = disagree, 3 = neither agree nor disagree, 4 = agree, 5 = agree strongly). After completing the self-rating items, the participants were then asked to rate imaginary students in the vignette sets using the same scale as with the self-ratings. On the second day, four problem-posing tasks were administered to the participants during a 45-minute class period. The procedures of selecting the self-rating items and problem-posing tasks and designing the vignette sets will be explained in detail in Chapter 5.

The data were collected using a paper-and-pencil version of the items, and the responses were entered into Microsoft Excel for analysis using SPSS 20.0, R studio 1.2.1335, JAMOVI 1.6.23, and AMOS 26. The statistical techniques that were used for each research question are explained below.

3.5 Research method to address Research Question 1

RQ1: What framework can be used to capture social and emotional skills in mathematics?

A systematic literature review was conducted to 1) identify social and emotional skills related to mathematics in this study, and 2) outline conceptual links between the selected skills and existing theories and concepts, and 3) construct the theoretical framework in this study (further discussed in Chapter 4).

3.6 Research methods to address Research Question 2

RQ2: What valid tools can measure social and emotional skills in mathematics reliably?

The current study used EFA and CFA approaches. The main goal of an EFA is to define a few meaningful structures from all the variables, while a CFA is used to determine the extent to which a theoretical pattern of factor loadings on pre-specified constructs represents the real data. The EFA and CFA statistics depict the level to which theoretical specifications are in line with reality. In other words, a CFA is used to test the validity of a literature review or a theoretical model.

The items of a construct do not all need to be of the same scale type, nor is it required for different values to be normalised before conducting an EFA and CFA (Hair et al., 2010). Each construct can be measured using a scale with different point values (e.g., five points, seven points, ten points, or 100 points). It is possible to transform these scale points to a common scale before estimation to normalise them (for instance, all with seven points). However, it is not essential because an EFA and CFA can analyse multiple variables using different scale points. In this regard, AV-corrected scales and problem-posing tasks can be analysed together using EFA and CFA techniques without normalisation. Figure 4 presents the statistical techniques for validating the theoretically predicted framework.

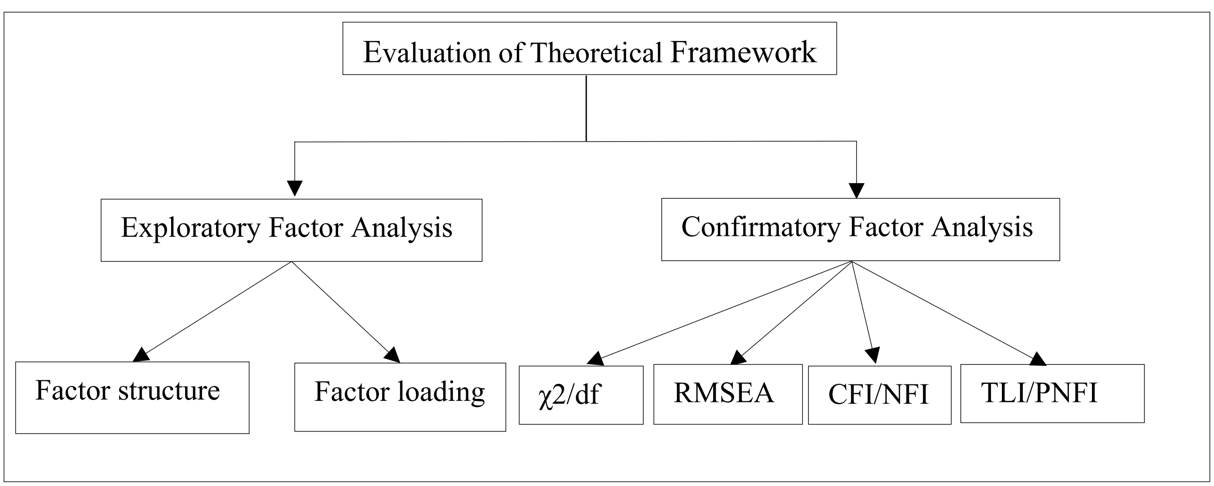


Figure 4. Statistical techniques for evaluating the theoretically predicted model

To validate the survey instruments, a psychometric analysis should be conducted. Several techniques need to be carried out to achieve the reliability and validity of the instruments, as indicated in Figure 5.

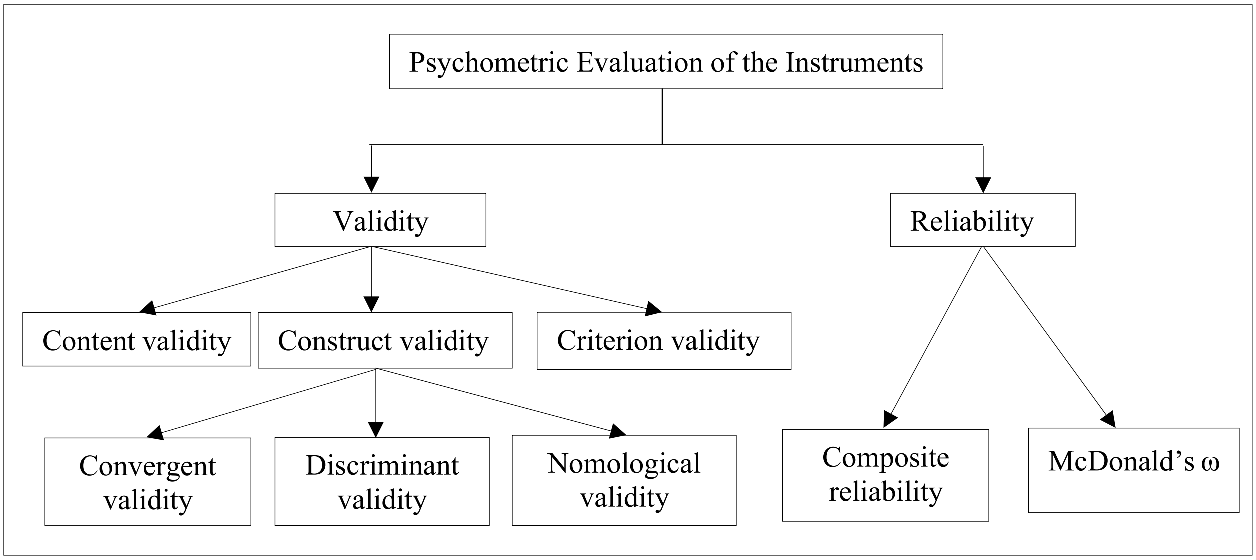


Figure 5. Statistical techniques for evaluating the instruments

3.7 Reliability analysis

Reliability refers to the extent to which an item or a set of items demonstrate consistency in their measurements. When there are multiple measurements, the reliable measures will be those that have consistent values. Reliability is different from validity because it focuses on how to measure instead of what to measure (Hair et al., 2010). Cronbach's alpha coefficient (Cronbach, 1951) and McDonald's omega coefficient (McDonald, 1970) are commonly used measures to analyse the consistency of the entire scale. However, McDonald's omega is considered a preferable indicator of reliability instead of Cronbach's α (Zinbarg et al., 2005).

Another measure of reliability is composite reliability (CR) introduced by Jöreskog (1971). Researchers recommend using CR instead of Cronbach's alpha for measurement models as the latter is not a precise measure of reliability due to items being unweighted (Hair et al., 2018). By contrast, items under CR are weighted based on the construct indicators' loadings, making it more reliable than Cronbach's alpha (Hair et al., 2018). Based on the above discussion, this study employed McDonald's omega and CR as a measure of reliability. The commonly agreed-upon lowest value for McDonald's omega is 0.7. A satisfactory CR value is between 0.7 and 0.9. Any value equal to or larger than 0.95 is problematic as it indicates redundant items, which then reduces validity (Hair et al., 2018).

3.8 Validity analysis

Validity refers to the extent to which a measure or set of measures accurately represents the concept of interest. Unlike reliability, it demonstrates how well a measure defines a concept (Hair et al., 2010). Various forms of validity, such as content, construct, convergent, discriminant, nomological, and criterion validity, are crucial for evaluating survey instruments.

Content validity is the extent to which the content of the items is consistent with the definition of the construct (Hair et al., 2010). This type of validity subjectively analyses the association between a concept and the individual items based on the judgment given by experts. Construct validity refers to the extent to which a set of measured variables represent the latent theoretical construct that those variables are meant to measure (Hair et al., 2010). Convergent validity examines the extent of correlation between two measures of the same concept. A higher convergent validity is obtained when the CR value is between 0.7 and 0.9, and all factor loadings are above 0.5 (Hair et al., 2010).

Discriminant validity analyses the extent to which a particular construct is different from others in terms of correlation and representation by distinctly measured variables (Hair et al., 2010). Therefore, a high discriminant validity proves that a construct is unique, as it contains some phenomenon not possessed by any other construct. Discriminant validity is identified by comparing the value of the average variance extracted (AVE) with the maximum shared variance

(MSV). According to this approach, discriminant validity is established when the AVE is higher than the MSV for each construct (Hair et al., 2014). Discriminant validity is also achieved when the square root of the AVE is higher than the correlation of a specific construct with other constructs (Fornell and Larcker, 1981). The above two approaches were used to analyse the discriminant validity.

Nomological validity refers to the form of test validity that assesses the meaningfulness of the construct correlations in the measurement theory (Hair et al., 2010). The correlation between constructs is useful in the assessment. In this sense, a correlation analysis was conducted among the constructs.

Criterion validity can be defined as the part of construct validity that focuses on the statistical evidence showing the relationships between test scores and other relevant variables (Education Testing Service, 2015). In this regard, regression analysis was conducted to test if there is consistency between responses in the questionnaire (old measure) and the vignettes (new measure).

3.9 Research methods to address Research Question 3

RQ3: What is the status of social and emotional skills among Mongolian students?

The students' general performance on the vignette-corrected new scale and the problem-posing tasks was analysed using descriptive statistics. The significant mean difference in mathematical creativity, cooperative learning in mathematics, mathematical perseverance, mathematics anxiety, mathematics enjoyment, and mathematical self-efficacy was examined by students' gender, ethnicity, and regions, using an independent samples t-test (Figure 6).

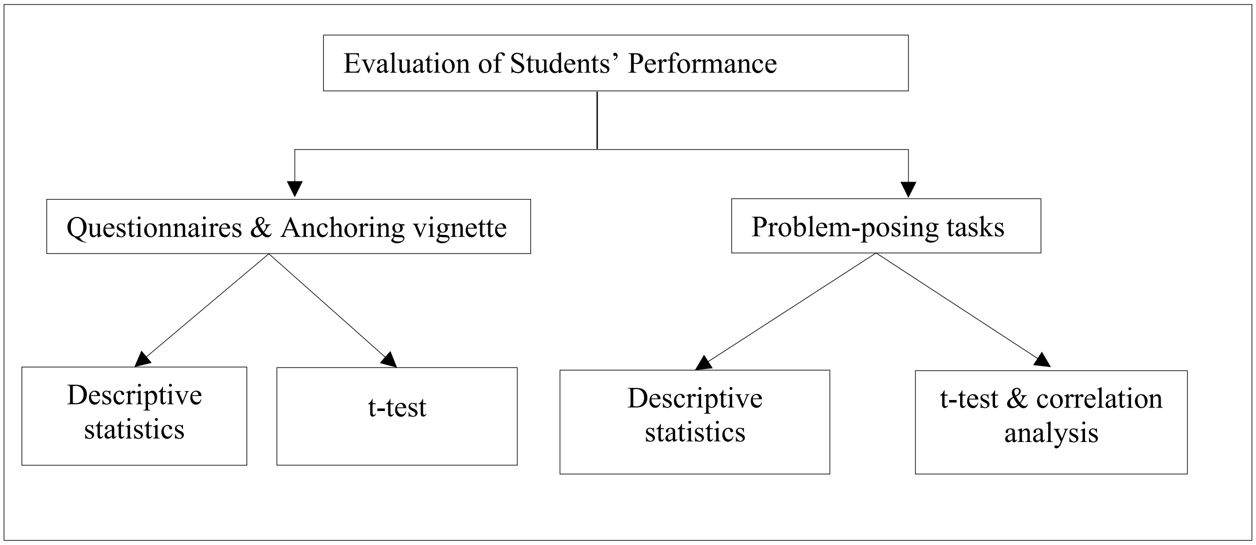


Figure 6. Statistical techniques for evaluating students' performance

CHAPTER 4

BIG FIVE PERSONALITY TRAITS AND THEORETICAL FRAMEWORK OF SOCIAL AND EMOTIONAL SKILLS IN MATHEMATICS EDUCATION

4.1 Introduction

This chapter aims to present a theoretical framework to capture social and emotional skills in mathematics. Literature was reviewed to 1) discuss the relationship between social and emotional skills and the Big Five personality traits, 2) identify social and emotional skills related to mathematics in this study, and 3) outline conceptual links between the selected skills and the Big Five personality traits in detail. Based on the gap in the literature review and conceptual relationship between the Big Five personality traits and mathematics, the theoretically predicted framework in this study was constructed.

4.2 Correspondence between Big Five personality traits and social and emotional skills in mathematics education

Discussions pertaining to social and emotional skills have been an integral part of education systems for over a decade. However, researchers such as Rotherham and Willingham (2010) and Mishra and Kereluik (2011) have claimed that these skills were in demand centuries ago. In this context, John and De Fruyt (2015) analysed various frameworks of social and emotional skills to conclude that these skills are in line with the Big Five personality traits (McCrae & John, 1992). Some other relevant studies have also asserted that the Big Five personality traits can serve as a theoretical framework to explore and analyse social and emotional skills (De Fruyt et al., 2006; Anglim & O'Connor, 2018).

Extensive research on the Big Five personality traits has revealed important evidence. All the five personality traits have identical composition; hence, this consistency in results leads to the wide acceptance of this model. The personality traits that are the components of the Big Five personality traits are openness and intellect, conscientiousness, extraversion, agreeableness, and neuroticism (emotional stability).

Openness and intellect describe an individual's variability in creativity, originality, and fantasy (Soto & John, 2017; Primi et al., 2016). Conscientiousness refers to individual differences in achievement orientation, self-discipline, and performance (e.g., Costa & MacCrae, 1992; Primi et al., 2016; Soto & John, 2017). Extraversion refers to positive emotion, excitement seeking, and warmth (Costa & MacCrae, 1992). Agreeableness represents an individual's differences in social interaction. Neuroticism describes an individual's variation in anxiety, vulnerability to stress, and emotional regulation (Costa & MacCrae, 1992; Primi et al., 2016; Soto & John, 2017). Apart from providing a strong empirical foundation, the Big Five framework also provides a comprehensive summary of individuals' unique social and emotional skills.

Table 7. Mapping social and emotional skills into Big Five personality traits

Collaboration (related to the agreeableness domain)
Compassion, care, cooperation Trust, appreciation of others, empathy
Task performance (related to the conscientiousness domain)
Perseverance, goal-orientation, motivation Self-discipline, productivity, self-control
Emotion regulation (related to low levels of negative emotionality)
Self-confidence, self-esteem, tackling problems Optimism, self-kindness, self-compassion (being positive and understanding towards yourself

when you fail)
Engagement with others (related to extraversion) Enthusiasm, zest, social awareness Teamwork, passion, inspiration
Open-mindedness: The inquiring mind (related to openness) Creativity, innovation, willingness to try new ideas The excitement of creating something new, curiosity

Derived from John & Fruyt (2015)

The CASEL (2017) highlights five competency aspects that must be included in SEAL programmes: self-awareness, social awareness, relationship skills, responsible decision-making, and self-management. The categories of the framework are in line with the Big Five personality traits in terms of content. Social awareness (empathy for others while taking their perspective) and relationship skills (positive relationships, conflict resolution, and teamwork) both belong to the agreeableness domain. Self-awareness (acknowledging one's strengths and limitations) is associated with the emotional stability domain. Self-management (the ability to regulate one's emotions with perseverance to achieve goals) is mainly related to emotional stability, but it is also related to conscientiousness. Lastly, responsible decision-making (making conscious decisions after evaluating risks and respecting others) is related mainly to conscientiousness and agreeableness.

The CCR (2015) has also proposed an integrative set of skills and constructs. These skill sets can also be classified within the Big Five personality traits. Mindfulness (e.g., tranquillity and self-awareness) is related to emotional stability, while courage (e.g., energy, bravery) and leadership (e.g., responsibility, charisma, and assertiveness) refer to 'getting ahead' and are, thus, related to extraversion. Ethics (e.g., honesty, kindness) is related to agreeableness, and resilience encapsulates self-discipline, perseverance, and grit and is, therefore, associated with conscientiousness. Curiosity is related to the domain of openness and intellect. As for the 'Four Cs' in the CCR framework, creativity and critical thinking fall within the openness and intellect domain, whereas communication and collaboration belong to the agreeableness domain. The OECD (2015) highlights that the CCR framework, when analysed from the perspective of social and emotional skills, appears to be more comprehensive than the CASEL framework because of its better coverage of the openness and intellect domain.

The SEAL (2010) aims to develop social and emotional skills through effective learning. Goleman (1995) devised a model of emotional intelligence that identifies five skills: self-awareness, managing emotions, empathy, social skills, and motivation. Self-awareness can be defined as knowing one's worth. Being aware of one's strengths and limitations makes one feel confident about oneself, thus enabling one to engage in productive interactions. This category is related to the conscientiousness domain. Managing one's emotions refers to dealing with unpleasant feelings and boosting positive feelings. Having the skills to cope with difficult emotions enables one to concentrate better and engage in positive interactions with those around them, and managing emotion belongs to the emotional stability domain. Motivation involves having the resilience and optimism to work towards goals. Setting goals, devising effective strategies to reach those goals, and reacting positively to setbacks maximises an individual's potential to achieve those goals. Indeed, motivation in this framework belongs to the conscientiousness domain. Empathy can be defined as understanding other people's thoughts and feelings by stepping into their shoes and demonstrating support. Understanding other people's feelings and beliefs enables one to engage with people with diverse backgrounds and build healthy relationships. Social skills include problem-solving skills and skills to form and maintain relationships. These skills improve one's learning experiences and social interactions. In this sense, empathy and social skills in the SEAL framework belong to the agreeableness domain.

The Knowledge is Power Program (KIPP) is a charter school network that has played an important role in improving students' academic performance in underserved communities. Schools that are in the KIPP have put forward seven constructs in relation to social and emotional skills. These include self-control, optimism, zest, gratitude, social intelligence, grit, and curiosity. Similar to the above frameworks, these skills are conceptually incorporated into the Big Five personality traits. For instance, self-control and grit are related to conscientiousness, optimism is categorised as emotional stability, gratitude and social intelligence belong to agreeableness, and curiosity is associated with openness and intellect (John and Fruyt, 2015).

According to the discussion about the various frameworks, social and emotional skills are not limited to the initial concept of social and emotional development but cover broader concepts such as social skills, communication and collaboration, emotional regulation, managing emotions, optimism, and cognitive skills such as creativity and critical thinking. Moreover, some frameworks cover only social skills and emotional management (CASEL, 2017; SEAL, 2007), while other frameworks include cognitive skills such as creativity and critical thinking (CCR, 2015; Ontario Mathematics Curriculum, 2020). This indicates a conceptual confusion between the frameworks on social and emotional skills and the initial concept of social and emotional development, which covers only social competence and emotional competence.

As illustrated in Table 8, numerous social and emotional skills have been consciously included in different frameworks, using either similar or different terms.

Table 8. Different frameworks on social and emotional skills

No.	Framework	Skills or constructs
1	OECD (2015) 'social and emotional skills'	Self-esteem Optimism Confidence Sociability Caring Respect Perseverance Self-control Passion for goals
2	CASEL (2017) 'social and emotional learning competencies'	Self-awareness Relationship skills Social awareness Responsible decision-making Self-management
3	CCR (2015) 'social and emotional learning skills'	Mindfulness Courage Leadership Collaboration Communication Ethics Resilience Creativity Curiosity Critical thinking
4	KIPP Schools 'social and emotional learning'	Self-control Zest/optimism Gratitude Social intelligence Grit Curiosity
5	SEAL (2010) 'social and emotional skills'	Self-awareness Managing feelings Empathy Social skills Motivation

The OECD (Organization for Economic Co-operation and Development; 2015) uses the term 'social and emotional skills' to cover skills like self-esteem, optimism, confidence, sociability, care, respect, perseverance, self-control, and passion for achieving goals. CASEL (2017) used the term 'social and emotional learning competencies' to emphasize skills such as self-awareness, self-management, social awareness, responsible decision making, and relationship skills. CCR (2015) uses the term 'social and emotional learning skills in the domain character' to foster mindfulness, courage, leadership, collaboration, communication, ethics, resilience, creativity, curiosity, and critical thinking. KIPP School uses the simple term 'social and emotional learning' to incorporate self-control, zest/optimism, gratitude, social intelligence, grit, and curiosity.

Like the OECD, SEAL (2010) uses the common and the most simplistic term of 'social and emotional skills' to instil skills including self-awareness, emotion management, empathy, motivation, and social skills in individuals. The coverage of skills in each framework is varied despite the use of similar terms. Ranging from personal development to an individual's social and emotional development, many of these skills aim to prepare individuals to thrive in a technologically advanced and globalized modern society. Despite similar terms in these frameworks, social and emotional skills are not restricted to interpersonal and intrapersonal skills; they also encapsulate cognitive skills such as creativity and critical thinking, as mentioned in the CCR framework. Each of these skills is unique and contributes to the holistic development of an individual. A combination of these skills enables an individual to develop the high-demand competencies of the 21st century.

Chernyshenko et al. (2018) suggest that there are other social and emotional skills outside of the Big Five framework as well. These skills are sometimes referred to as 'compound skills' (Chernyshenko et al., 2018, p. 19), as they combine many individual skills. Some examples are metacognition and self-efficacy. Hence, as an extension to the framework introduced by John and Fryut (2015), Chernyshenko et al. (2018) have devised a conceptual framework related to social and emotional skills by including a 'compound skill domain' (Table 9).

Table 9. Conceptual framework designed by Chernyshenko et al. (2018)

Big Five Domains	Sub-skills
Conscientiousness	Persistence, achievement orientation
Neuroticism	Stress resistance, emotional control
Agreeableness	Cooperation, empathy
Openness and intellect	Creativity, curiosity
Extraversion	Energy, assertiveness
Compound skills	Self-efficacy, self-reflection

To narrow down the discussions on social and emotional skills in the previous chapters, the frameworks introduced by John and Fruyt (2015) and Chernyshenko et al. (2018) were employed as a lens to determine social and emotional skills in mathematics and devise a theoretical framework for this study. Thornton (2016) suggests that the new skills demanded in the 21st century tend to undermine competencies that are otherwise appreciated in the traditional education systems. Hence, previously valued dispositions must be re-discovered in the 21st-century context instead of proposing entirely new skill sets (Thornton, 2016). Researchers have also argued that skills or competencies, particularly for mathematical education, should not be limited to memory-based or factual knowledge; rather, they should encapsulate the development of cognitive and creative abilities and the motivation for individuals to handle difficult tasks (Toh & Kaur, 2016).

The above discussion is significant as it reviews the relevant social and emotional skills in mathematics. As the term ‘social and emotional skills’ is relatively new in mathematics education, the present study focused on the skills and constructs highlighted in the framework devised by Chernyshenko et al. (2018). For mathematics education, most of these skills were examined under the term ‘affective domain’ (e.g., mathematics enjoyment and mathematical self-efficacy), and the umbrella terms ‘non-cognitive predictors’ or ‘non-cognitive constructs’ (e.g., mathematical perseverance and cooperative learning in mathematics), ‘attitudes towards mathematics’ (e.g., mathematics anxiety and mathematics enjoyment), ‘motivation and emotion’ (e.g., mathematical perseverance, mathematics anxiety; mathematical self-efficacy), and ‘dispositional factors’ (e.g., mathematical perseverance, and mathematical self-efficacy). These terms were used to search relevant articles from the Education Resources and Information Center databases (Figure 7). Additionally, the term ‘mathematical creativity’ was added to the search as it corresponds to the framework on social and emotional skills.

4.3 Identification of social and emotional skills related to mathematics

A thorough initial search on the ERIC database using different combinations of keywords generated over 350 documents ranging from 2012 to 2020. Many of these articles were then eliminated due to recurring overlaps, thus reducing the number of unique articles to 50. Each article was then analysed to ascertain its relevance to the present study. Finally, by identifying relevant social and emotional skills in mathematics, specific skills were selected for this study.

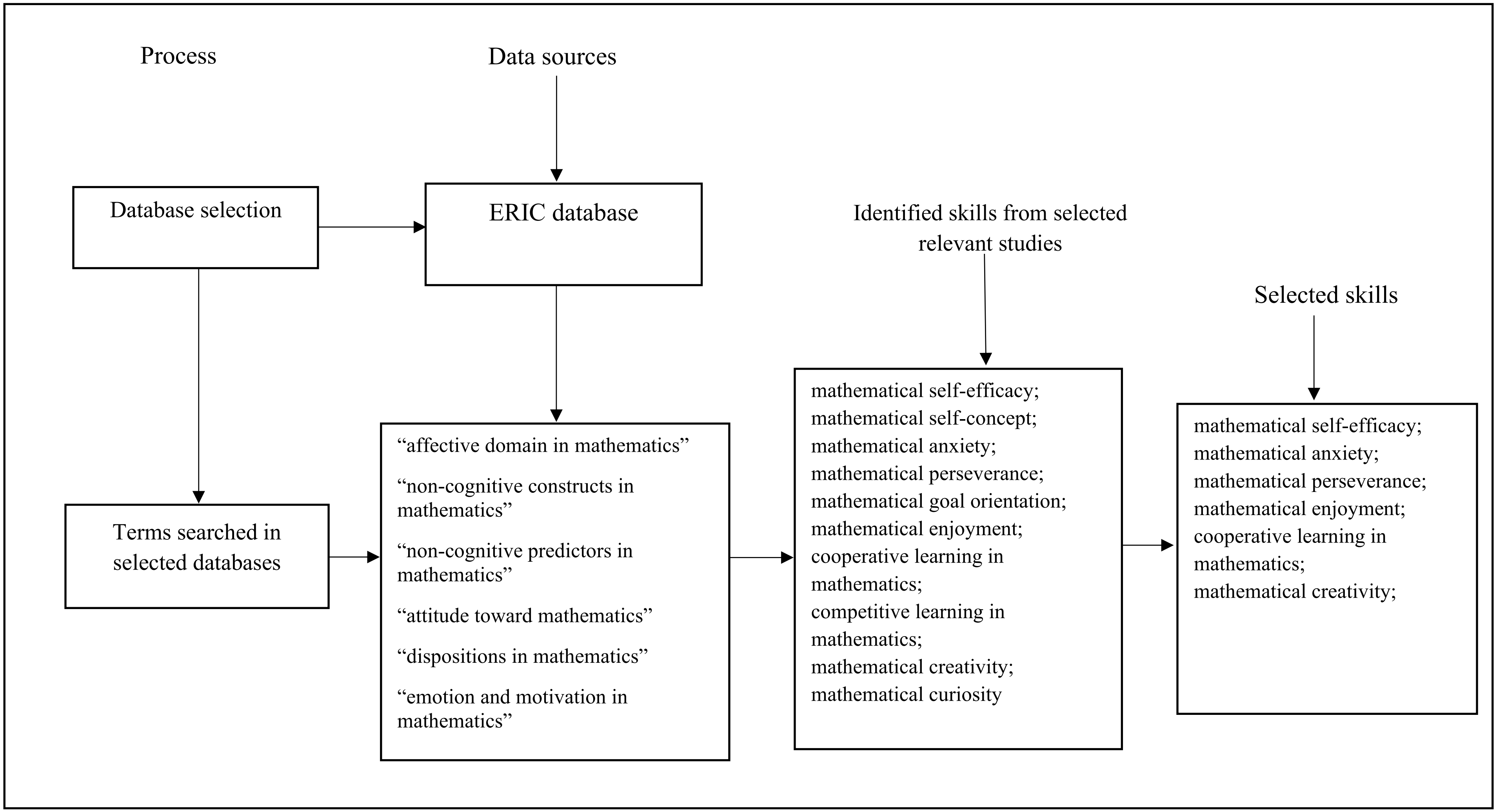


Figure 7. Flowchart of the literature search process

The selection criterion of the six social and emotional skills to be included in this study was first based on empirical evidence provided by John and Fruyt (2015), Kankaras (2017), and Chernyshenko et al. (2018). In other words, the relevance of these skills to the frameworks of social and emotional skills was considered. Second, the significance of these skills for mathematics education was considered. Third, the emphasis on each skill in the Mongolian education policies, particularly in the mathematics curriculum, was also considered. Finally, it is also important to note that in case the research participants were children aged 14 and below, a limited number of skills were analysed so as not to burden them.

Considering the criteria mentioned above, the present study selected mathematical creativity instead of mathematical curiosity under the cognitive aspects of social and emotional skills, mathematical self-efficacy, mathematical perseverance instead of mathematical self-concept and mathematical goal-orientation under intrapersonal aspects, and cooperative instead of competitive learning in mathematics under interpersonal aspects. Further details on the significance of the selected social and emotional skills and their relevance to the frameworks mentioned above will be provided in the following section.

Consequently, the following seven skills were selected from the existing literature as social and emotional skills relevant to mathematics (Table 10).

Table 10. Selected social and emotional skills related to mathematics in this study

	Selected social and emotional skills	Some sources
1	Mathematical self-efficacy	Lee, 2008; Lee & Stankov, 2013; Tuohilampi et al., 2013; Pipere & Mierina, 2017;
2	Mathematical creativity	Grégoire, 2016; Kattou et al., 2012; Simonton, 2000; Sriraman, 2009; Haylock, 1987; Sriraman, 2005; Mann, 2006
3	Mathematics anxiety	Lee & Stankov, 2013; Pipere & Mierina, 2017; Gunderson et al., 2017; Imam Kusmaryono et al., 2018
4	Mathematical perseverance	Hannula et al., 2016; Zakaria & Yusoff, 2009; Imam Kusmaryono et al., 2018
5	Cooperative learning in mathematics	PISA, 2000, 2003, Lee & Stankov, 2013
6	Mathematics enjoyment	Tuohilampi et al., 2013; Imam Kusmaryono et al., 2018; Tuohilampi et al., 2013; Pekrun et al., 2017

4.4 Conducting a literature review between the selected skills in mathematics and Big Five personality traits

A literature review for each skill was conducted separately to assess the association that each selected skill has with the Big Five personality traits and mathematical achievement.

4.4.1 Conceptual link between mathematical creativity and openness and intellect

Various frameworks related to modern 21st-century skills have identified creativity as a major skill that can help children face society's current and future challenges (see, e.g., National Research Council, 2012; UNESCO, 2016; Partnership for 21st-century skills, 2018). With rapidly evolving technology, dealing with new challenges requires the citizens to think creatively (Lubart & Guignard, 2004). Hence, education systems around the world tend to encourage creativity in students. An OECD study conducted in 2015 reported that countries like Sweden, Ireland, Norway, Israel, South Korea, Japan, and Turkey actively update their national education curricula while focusing on inculcating creativity in students.

A great amount of research and discussion on creativity has taken place in psychology over the last 60 years. Researchers suggest that creativity or mathematical creativity has no unanimously agreed-upon definition (Haylock, 1987; Sriraman, 2005; Mann, 2006). However, one study has also presented more than 100 of the latest definitions of creativity (Treffinger, Young, Selby, & Shepardson, 2002). According to Guilford (1960), a diverse thought process defines creativity. However, mathematical creativity is defined by Poincaré (1948) as the art of identifying useful combinations from useless ones. According to some other researchers, creativity refers to originality (Simonton, 2000; Sriraman, 2009). Chamberlin and Moon (2005) describe mathematical creativity as using mathematical modelling to derive solutions to assumed or real problems. Laycock (1970) implies that mathematical creativity is the ability to come up with multiple solutions for a particular problem by observing patterns. Mann (2005, p. 29) describes mathematical creativity as ‘an essential aspect in the development of mathematical talent’.

Despite being diverse, many of these definitions align with the concept of divergent thinking proposed by Guilford (1960), who presented a model of the structure of intellect in his pioneering work (Wechsler et al., 2018). This model was introduced as the basis for creativity. The study implied that factors associated with creativity such as originality, flexibility, and fluency must be considered when the concept of divergent thinking is being discussed. Fluency is the ability to come up with multiple solutions for a particular problem. Flexibility is the ability to present varied solutions to a problem, while originality is the ability to present a unique solution to a problem (Guilford, 1960; Torrance, 1966, 1974). According to some mathematical researchers like Chamberlin and Moon (2005), one of the core concepts of mathematical creativity is divergent thinking. Various studies in mathematics education have used this concept by emphasising flexibility, fluency, and originality to analyse mathematical creativity (e.g., Sriraman and Lee, 2011, 2012; Katou et al., 2015).

Creativity has a vital role in advanced mathematical thinking as it helps draw logical inferences, which then prove to be useful in developing mathematical theories and advanced mathematical knowledge (Ervynck, 1991). Research has also demonstrated the correlation between mathematical creativity and mathematical ability. According to Meissner (2000), having sound mathematical knowledge is a prerequisite for mathematical creativity, and a person's ability to present various solutions to a specific mathematical problem depicts their competence in mathematical reasoning. Moreover, mathematical creativity in students can also be encouraged by solving open-ended and challenging problems (Grégoire, 2016).

Empirical research on creativity demonstrates a strong relationship between mathematical creativity and mathematical ability. According to Kattou et al. (2012), the level of mathematical ability in students is determined by their mathematical creativity.

Neumann (2007) states that an interactive environment is necessary for the development of mathematical creativity. Sawyer (2007) and Sriraman (2009) also report that social interaction and cooperation play a crucial role in stimulating creativity in mathematical activities. Working in groups in a classroom environment provides an opportunity for students to flourish in their creative thinking skills (Molad, Levenson, & Levy, 2020). In other words, an interactive classroom environment promotes creativity in students.

For researchers like Balka (1974) and Silver (1997), mathematical creativity is a part of general creativity. In particular, openness is associated with creativity and intellectual aspects required for mathematical testing (Saucier & Goldberg, 1996). For many researchers, creativity is the major element of openness and intellect, as they have devised a model to measure personality traits (DeYoung et al., 2007; Ashton, Lee, et al., 2014; Chernyshenko, Kankaras, & Drasgow, 2018). Considering the discussion, this study analysed mathematical creativity as a factor of openness and intellect.

4.4.2 Conceptual link between mathematics anxiety and neuroticism

A plethora of literature highlights the important role that emotions and the regulatory system play in the learning process of a child. Blair (2002) takes on a developmental neurobiological perspective to state that emotionality in young children develops at a very early stage, largely affected by a biologically based temperament.

How children regulate their emotions directly affects their performance in school, regardless of their cognitive abilities (Blair, 2002). Evidence suggests that negative emotionality negatively affects success (Gumora & Arsenio, 2002), and a child's ability to regulate emotions can help them learn fast, thus increasing their chances of success.

Researchers have also investigated certain types of negative emotionality apart from general negative emotionality. Mathematics anxiety is one of them, and it has been long established that anxiety is a basic human emotion (Bodas & Ollendick, 2005). Mathematics anxiety is generally defined as 'feelings of tension and anxiety that interfere with manipulating numbers and solving mathematical problems in a wide variety of ordinary life and academic situations' (Richardson & Suinn, 1972, p. 551).

Research has also demonstrated that students at all educational levels suffer from mathematics anxiety worldwide (Dowker, Sarkar, & Looi, 2016; Foley et al., 2017). According to the PISA report in 2012, 59 per cent of ninth-grade students expressed a fear of mathematics (OECD, 2013).

Empirical research has demonstrated how mathematics anxiety leads to students performing poorly in mathematics (Foley et al., 2017) because it causes individuals to worry constantly, thus disrupting the thinking skills required to solve mathematical problems (Ashcraft & Kirk, 2001; Park, Ramirez, & Beilock, 2014).

When students suffering from mathematics anxiety are asked about their perception of mathematics, their responses often indicate a level of enjoyment, motivation, and self-confidence that is much lower than those with less or no mathematics anxiety have. These maths-anxious students try to avoid the subject, often by not opting for it during high school (Hembree, 1990). An unenthusiastic approach towards mathematics, coupled with attempts to avoid the subject, are the major reasons behind maths-anxious students performing poorly in mathematics (Carey, Hill, Devine, & Szücs, 2016).

In summary, relevant literature has demonstrated that mathematics anxiety is a major determinant of students' poor performance in mathematics. Considering the evidence, there is a dire need to upgrade the tools used to evaluate mathematics anxiety so that the association between mathematics anxiety and maths performance can be explored further.

Mathematics anxiety has also been associated with other forms of anxiety, such as trait anxiety (general anxiety not necessarily triggered by a particular situation) or test anxiety (which usually occurs before one takes a test; Dowker et al., 2016). Hembree (1990) reported that students with mathematics anxiety tend to have much higher test anxiety and trait anxiety levels than students with less or no mathematics anxiety do. The findings also demonstrated that the correlations between mathematics anxiety and other types of anxiety are either very small or moderate (Dew, Galassi, & Galassi, 1983; Hembree, 1990). This demonstrates the independent nature of the underlying constructs. Therefore, despite its association with trait anxiety and test anxiety, many researchers consider mathematics anxiety a unique psychological construct (Dowker et al., 2016).

From the Big Five perspective, neuroticism or negative emotionality characterises individual personality traits such as anxiety, fearfulness, and emotional volatility (Soto & John, 2017; Costa & MacCrae, 1992). In the case of social and emotional skills at the individual level, negative emotionality typically differentiates between three types of negative traits: a) fear and anxiety, b) sadness and depression, and c) irritation and anger. The most significant trait among these is the tendency to experience fear and anxiety. Therefore, based on the previous discussions pertaining to the characteristics of mathematics anxiety, its association with the neuroticism domain traits such as depression, vulnerability to stress, and emotional volatility become clear. Therefore, mathematics anxiety was considered a component of the neuroticism domain in this study.

4.4.3 Conceptual link between mathematical perseverance and conscientiousness

The term ‘perseverance’ has been used interchangeably with ‘grit’ and ‘effort’. For example, some studies have examined perseverance using the term ‘grit’ (Duckworth et al., 2007), while others used the term ‘effort’ (Tuohilampi et al., 2013). Despite the different terms, the common concept emphasises an individual’s persistent attitude to accomplish a task and achieve a goal regardless of feeling challenged or unsuccessful (Child Trends, 2014). Different frameworks regarding social and emotional skills emphasise perseverance more (see, e.g., Durlak et al., 2011; CCR, 2019; OECD, 2015).

In addition to the significance of perseverance in the general domain, its importance in mathematics education has long been highly appreciated. Persistence in solving mathematical problems is one of the seven main components of mathematical disposition (National Council of Teachers of Mathematics [NCTM], 1989, as cited in Kusmaryono et al., 2018). The more the students persist, the less anxiety they feel and the more they attempt to solve difficult mathematical tasks compared to those who are less persistent would do (Lufi & Cohen, 1987). Similarly, a 2015 OECD study highlighted that '… a child who is more persistent is likely to increase his or her math skills more than a child with equal levels of math skills but lower levels of persistence' (p. 39).

Furthermore, the persistent effort that a student puts into learning plays a crucial role in solving mathematical problems (Kusmaryono et al., 2018; Rahayu & Kartono, 2014). Tuohilampi et al. (2013) found a positive relationship between mathematical perseverance and mathematical performance when comparing Finnish and Chilean students. Moreover, PISA 2021 added a score for perseverance as an extension to mathematical literacy (OECD, 2017). From the perspective of the Big Five traits, many studies (e.g., John & Fruyt, 2015; OECD, 2015, 2018) classified perseverance into the conscientiousness domain. According to the above discussion, mathematical perseverance is related to the conscientiousness domain.

4.4.4 Conceptual link between cooperative learning in mathematics and agreeableness

Similar to creativity and perseverance, various social and emotional skills frameworks have focused more on social skills such as teamwork, cooperation, and collaboration (e.g., CASEL, 2017; CCR, 2019; OECD, 2015). However, these skills have been introduced interchangeably to represent sociability or the interpersonal domain. Casner-Lotto and Benner (2006) examined which skills are most in-demand among high school, two-year college, and four-year college graduates in the 21st-century labour market. They reported that teamwork/collaboration was rated 'very important' by 74.7 per cent of high school graduates, 82.7 per cent of two-year college graduates, and 94.4 per cent of four-year college graduates. Ultimately, teamwork/collaboration ranked as the second most important skill after oral communication. Interpersonal skills such as communication, cooperation, and teamwork skills help students work effectively with peers and teachers and manage their feelings and emotions towards school while indirectly increasing academic performance (Buhs et al., 2006; Ladd et al., 2006; Valiente et al., 2008).

Cooperative work refers to a group of individuals with varying abilities working together to find a common solution to a specific problem by using everyone's abilities (Karali and Aydemir, 2018). Cooperative learning is also a contemporary way to improve the social and emotional aspects of students along with the cognitive aspects (Erden, 1988). In the context of mathematics, cooperative learning refers to students working with their peers to learn together and solve mathematical tasks while also enjoying themselves (Lee & Stankov, 2013). Cooperative learning has been reported to be efficient in helping students excel in mathematics. Zakaria et al. (2010) found cooperative learning helpful in improving students' mathematics performance. However, the PISA studies conducted in 2000 and 2003 found no significant relationship between cooperative learning and academic achievement (Lee & Stankov, 2013).

Research has also found the cooperative approach to positively affect the students' affect towards mathematics (Akman & Koçoğlu, 2016). Teachers who applied this learning method in their classes expressed their belief in the educational, psychological, and social benefits of cooperative learning, such as developing confidence, motivation, and positive attitude in students (Macit, 2013, as cited in Karali and Aydemir, 2018).

Perels et al. (2005) observed learning groups of lower secondary students for a significant period to conclude that the students demonstrated increased levels of motivation and self-efficacy while solving mathematical problems. Schukajlow et al. (2011) also confirmed that students at the lower secondary level demonstrated higher interest levels and motivation in solving mathematical problems under the cooperative learning approach than the students who studied under teachers' instructions. According to Arısoy (2011, as cited in Karali and Aydemir, 2018), cooperative learning in mathematics also improves the level of mathematical perseverance in students.

From the Big Five traits perspective, individuals who focus more on their interpersonal relationships tend to be more cooperative and are more popular among their fellows (John & Fruyt, 2015; Soto & John, 2017; OECD, 2018). Such agreeable individuals study easily and can live in harmony with others while being appreciative of close relationships. Moreover, John and Fruyt (2015) analysed various frameworks related to social and emotional skills from the perspective of the Big Five traits and suggested cooperation and teamwork skills as a sub-construct of the agreeableness domain. Similarly, other studies (OECD, 2015; Chernyshenko et al., 2018) have classified cooperation within the agreeableness domain. Therefore, interpersonal skills such as working in teams, setting positive relationships, and handling conflicts are certainly associated with the agreeableness domain. In this regard, the present study explored cooperative learning in mathematics as a component of the agreeableness domain.

4.4.5 Conceptual link between mathematics enjoyment and extraversion

Previous studies have revealed that emotions in mathematics can significantly affect the process of learning, and hence, achievement. Recent literature has also emphasised positive emotions apart from anxiety. Pekrun et al. (2017) conducted a longitudinal study to analyse the emotions and performance of secondary school students and concluded that positive emotions such as enjoyment have a positive impact on the grades and test scores attained by students in mathematics. As a positive emotion, enjoyment is considered a crucial factor in the learning process (Pekrun et al., 2002), and mathematics enjoyment is when one finds pleasure in solving mathematical problems.

Pekrun and Linnenbrink-Garcia (2012) state that a student's academic performance in mathematics mostly depends on emotions such as enjoyment, enthusiasm, frustration, anxiety, and boredom. Both enjoyment and enthusiasm, being the positive emotions, induce pleasurable feelings; however, what highly activates the peripheral nervous system is enthusiasm. More findings affirm that positively activating emotions such as pride and enjoyment exert a positive effect on achievement in mathematics (Villavicencio & Bernardo, 2015 as cited in Gómez et al., 2020).

The Big Five traits perspective also brings into light positive emotions as one of the behavioural examples of extraversion. As the name implies, the extraversion domain includes behavioural traits such as excitement-seeking, positive emotions, cheerfulness, and enthusiasm (Costa and MacCrae, 1992; DeYoung et al., 2007; OECD, 2018). In the context of this study, mathematics enjoyment refers to the positive emotions or pleasure that a student experiences while solving mathematical problems. As per the discussions above, traits such as positive emotions, excitement seeking, enthusiasm, and energy level can be found in mathematics enjoyment; hence, it has been considered a component of the extraversion domain in this study.

4.4.6 Mathematical self-efficacy as a compound skill

Self-efficacy is one of the most frequently examined motivational constructs in education (Schukajlow et al., 2017). In the contrary, it has been disregarded in mathematics education for many years (Zan et al. 2006). According to Öztürk (2017, as cited in Görgün & Tican, 2020), self-efficacy has more to do with the feeling of competence experienced during the successful completion of a task than the actual competence itself. In other words, a person's perception of his ability to complete a task successfully is based on their mathematical self-efficacy. This means that mathematical self-efficacy refers to believing in one's ability to achieve in mathematics. In other words, it is the self-confidence towards mathematics that eventually determines their ability to solve mathematical problems (Görgün & Tican, 2020).

When it comes to a school environment, self-efficacy has a considerable effect on students' choices, their willingness to make an effort, and their determination to keep going, which then has a combined impact on their academic performance (Pajares & Miller, 1994). A considerable amount of available literature implies that students' mathematical self-efficacy is positively related to their achievement in mathematics (Pajares & Miller, 1994; Skaalvik, Federici, & Klassen, 2015). It has also been established that students' achievement in mathematics can also affect their mathematical self-efficacy. This demonstrates that mathematical self-efficacy and mathematical achievement reinforce each other positively (Bandura, 1993; Pajares & Miller, 1994).

Students with low mathematical self-efficacy usually tend to be less confident about their performance and achievement in mathematics. By contrast, students with high mathematical self-efficacy can organise mathematical activities while relying on their own knowledge and skills (Görgün and Tican, 2020).

Researchers have different views about self-efficacy in relation to the Big Five personality traits. Some suggest that self-efficacy is a component of the conscientiousness domain (Education Testing Service, 2012), while others argue that self-efficacy can be found in the conscientiousness, neuroticism, and, to a lesser degree, extraversion domains (OECD, 2018). Admittedly, according to Bandura's social cognitive theory (1986), self-efficacy has characteristics of cognitive (intellectual), affective (emotional), and conative (motivational) dimensions. Consequently, one of the most influential studies on developing a framework on social and emotional skills, conducted by Chernyshenko et al. (2018), examined self-efficacy as a 'compound skill' due to its combination of characteristics from different categories of the Big Five personality traits. In this regard, the present study examined mathematical self-efficacy as a compound skill.

4.5 Theoretical framework

Based on the theoretical and empirical discussion thus far, the following six-domain framework was constructed as a theoretically hypothesised framework in this study (Table 11).

Table 11. A theoretically predicted framework in this study

Big Five Domains	Sub-skills
Conscientiousness	Mathematical perseverance
Neuroticism	Mathematics anxiety
Agreeableness	Cooperative learning in mathematics
Openness and intellect	Mathematical creativity
Extraversion	Mathematics enjoyment
Compound skills	Mathematical self-efficacy

In summary, the literature review was conducted to identify social and emotional skills from a theoretical perspective. It indicates that 1) social and emotional skills in mathematics are consistent with the Big Five personality traits, and 2) these skills are conceptually different constructs. Moreover, this framework contributes to enhancing social and emotional skills in the general domain and plays a pivotal role in mathematics education. Hence, the theoretical framework can be meaningful for mathematics education and the development of social and emotional skills. Moreover, the skills in this theoretical framework are consistent with some social and emotional skills in mathematics curricula of the selected countries either expressly or implicitly (Table 12). Considering the challenges of 21st-century society, the Singaporean mathematics curriculum focuses on incorporating social and emotional skills in the maths curriculum as a component of 21st-century competencies. Particularly, skills such as mathematical creativity, mathematical perseverance, cooperative learning in maths, maths enjoyment, and mathematical self-efficacy are explicitly incorporated in the maths curriculum; however, mathematical anxiety is implicitly stated (MOE, 2014; Wong, 2016, p. 35).

Similarly, Finland’s basic education mathematics curriculum incorporates the social and emotional skills under the regional term of transversal competencies. The new maths curriculum at all levels of education is based on the demands of the modern world. The basic education mathematics curriculum particularly emphasizes the development of mathematical creativity, cooperative learning in mathematics, mathematical self-efficacy (confidence as a learner of mathematics), and mathematical perseverance (Finnish National Board of Education, 2014, pp. 481-482) and implicitly include mathematical enjoyment under enthusiasm for mathematics and mathematical anxiety under positive attitude towards mathematics (Finnish National Board of Education, 2014, p. 176).

As for Canada, the recent revision of Ontario's national mathematical curriculum for Grade 1 to 8 ensures the inclusion of these skills to prepare the students for an advanced world through mathematical processes. The curriculum has explicitly endorsed the skills such as mathematical creativity, mathematical perseverance, cooperative learning in maths, mathematical enjoyment, mathematical self-efficacy, and mathematical anxiety, within the framework of social and emotional learning skills along with their supporting mathematical processes (The Ontario Math Curriculum, 2020, p. 176-178).

Mongolia's mathematics curriculum has also been updated to provide mathematical knowledge to its students to improve their thinking skills and solve daily life problems pertaining to calculations. The learning process focuses on communication, creativity, and cooperation for the students to develop social and emotional skills. Hence, the updated mathematics curriculum in the education policies of Mongolia explicitly mentions mathematical creativity, mathematical perseverance, cooperative learning in maths, and mathematical self-efficacy; however, mathematical anxiety and mathematical enjoyment are implicitly included under a positive attitude towards mathematics (MECSS, 2014, p. 5; MECSS, 2016, p. 47-48).

Table 12. Consistency between the framework and social and emotional skills in mathematics curricula in different countries

	The selected social and emotional skills	Countries			
		Canada	Singapore	Finland	Mongolia
1	Mathematical creativity	●	●	●	●
2	Mathematical perseverance	●	●	●	●
3	Cooperative learning in mathematics	●	●	●	●
4	Mathematical enjoyment	●	●	○	○
5	Mathematical anxiety	●	○	○	○
6	Mathematical self-efficacy	●	●	●	●

Note: ● : Specifically stated; ○ : Implicitly stated;

CHAPTER 5

DEVELOPMENT OF THE TOOLS AND SCORING RULES

5.1 Introduction

This chapter aims to explain the procedures of selecting self-rating items and problem-posing tasks and designing the AV sets in detail. The chapter also explains the rules to correct self-rating items using the vignette sets and to score students' responses on the problem-posing tasks regarding the main components of creativity, such as fluency, flexibility, and originality.

5.2 Questionnaires

The present study employed questionnaires to measure each skill except mathematical creativity. According to Hair et al. (2010), the decision of how many items need to be added per construct is a matter of predicament. It is possible to use several questionnaires or items to represent a construct to the fullest and maximise reliability. However, a small number of items can also be used for adequate representation of a construct. It is important to note that having more items does not necessarily make it better. Although more items generally increase reliability estimates, they also require huge sample sizes, making it difficult to generate unidimensional factors.

Consequently, three items per construct are considered acceptable (Hair et al., 2010). In this regard, the present study employed three questionnaires to measure skills adapted from previous works. There are two reasons for adopting the measures from previous works. First, theorists have suggested using existing measures to validate a theoretical framework for better psychometric properties and less measurement error (John & Fruyt, 2015). Second, this study was intended to improve existing tools by applying a novel approach called AVs.

The questionnaires were adapted from previous works while also considering several important criteria such as 1) its age and grade appropriateness for the study group; 2) its content appropriateness for the Mongolian lower secondary mathematics curriculum and textbook; and 3) its psychometric properties and cross-cultural validation. The majority of the self-rating items, for example, mathematical anxiety, cooperative learning in mathematics, and mathematical self-efficacy were adapted from PISA 2003 as the items were validated across diverse cultures within the same age (15) and grade groups (Grade 7 and above).

Mathematical anxiety

Mathematical anxiety was measured using items such as, 'I feel worried about solving problems like finding the size of angle *x* in a given figure' and 'I get very tense when I have to solve mathematics problems like finding the area of a parallelogram', which were adapted from a previous study (OECD, 2003). Some adjustments were made to the items based on the Mongolian lower secondary mathematics curriculum and textbooks (Lower Secondary Mathematics Textbook, pp. 86-87). For example, the phrase 'doing mathematics problems' was replaced with 'finding the area of parallelogram' and 'doing mathematics homework' was changed to 'finding the size of angle *x* in a given figure'. In other words, mathematics anxiety items became more content specific instead of being general (See Item 1, 2, & 3 in the Appendix).

Cooperative learning in mathematics

Cooperative learning in mathematics was measured using items such as 'I think that mathematics is about working together with others to solve problems' and 'I enjoy helping others to work well in a group in mathematics' (OECD, 2003).

Mathematical self-efficacy

Mathematical self-efficacy was assessed through items such as 'I feel confident in finding the height of a pyramid' and 'I feel confident in finding the unknown side of a rectangle if its perimeter is given', which were developed based on existing items (OECD, 2003). Similar to mathematics anxiety items, some adjustments were also made based on the Mongolian lower secondary mathematics curriculum and textbooks (Lower Secondary Mathematics Textbook, p. 88). For instance, the phrase 'calculating the number of square meters of tile you need to cover a floor' was changed to 'finding the unknown side of a rectangle, if its perimeter is given' (See Item 10, 11, & 12 in the Appendix). The reason behind designing the items for mathematical anxiety and mathematical self-efficacy as content specific is related to Mongolian students' poor performance on geometry content area (Otgonbaatar, 2013; Education Evaluation Center, 2016; 2018)

Mathematical enjoyment

Items for mathematical enjoyment (e.g., 'I feel mathematics is fun'; 'I feel happy when dealing with mathematics') were adapted from Grootenboer and Marshman (2017, pp. 131, 134). The items were validated in a middle school sample in Australia and New Zealand among Grades 5 to 8 students (11–14 years of age).

Mathematical perseverance

Mathematical perseverance was measured using items such as 'I usually keep trying a difficult problem until I have solved it' (Grootenboer & Marshman, 2017, p. 25) and 'I feel challenged to work hard to find a solution when I get a difficult mathematics problem' (Kusmaryono et al., 2017, p. 3). The items were validated in a middle school sample in Australia, New Zealand, and Indonesia among Grades 5 to 8 students (11–14 years of age). Considerations of the item selection process are presented in Table 13.

Table 13. Considerations for selecting questionnaires (items)

Construct	Previous works		The present study	
	Source	Grade & Age	Grade & Age	Content relevance
Mathematical perseverance	Kusmaryono et al. (2018) (validated in a junior secondary sample in Indonesia)	Grade 8 Age: 12–14	Grade 9 13–16	Content general (no specific content was applied)
Cooperative learning in mathematics	PISA 2003 (validated in Lee and Stankov, 2013)	Grade 7 or above Age: 15		Content general (no specific content was applied)
Mathematics enjoyment	Grootenboer and Marshman (2017) (validated in a middle school sample in Australia and New Zealand)	Grades 5 to 8 Age: 11–14		Content general (no specific content was applied)
Mathematics anxiety	PISA 2003 (validated in a study by Lee and Stankov, 2013)	Grade 7 or above Age: 15		Geometry content (Mongolian Lower Secondary Mathematics Textbook, pp. 86-87)
Mathematical self-efficacy	PISA 2003 (validated in Lee and Stankov, 2013)	Grade 7 or above Age: 15		Geometry content (Mongolian Lower Secondary Mathematics Textbook, p. 88)

5.3 Problem-posing tasks for mathematical creativity

Mathematical creativity was measured using mathematical problem-posing tasks. The tasks were adapted from previous works while considering several important criteria such as 1) their content appropriateness to Mongolian lower secondary mathematics curriculum and textbook; 2) their age appropriateness for the study group; and 3) their psychometric properties and cross-cultural comparison (Table 14).

Task 1. Make as many groups of numbers as you can, using the numbers given below. Label each group with its characteristics.

2, 3, 7, 9, 13, 15, 17, 25, 36, 39, 49, 51, 60, 64, 91, 119, 121, 125, 136, 143, 150

This task was adapted from Kattou and colleagues (2013) and is related to Mongolian Grade 9 mathematics content (see, Lower secondary mathematics curriculum 7.4, Grade 9 content, Numeracy & Algebra, p. 39). Moreover, the task was validated using a CFA, a promising statistical technique, in Cyprus's primary and lower secondary samples.

Task 2. Last night there was a party at your cousin's home, and the door opened ten times. The first time the door opened, only one guest arrived. Each time the door opened, three more guests arrived than had arrived on the previous opening. Ask as many questions as you can that are in some way related to this problem.

This task was adapted from Sriraman and Lee (2011) and was originally developed by Stoyanova (1997). The task was tested in different cultural contexts such as Australia, America, and China among students of different ages from different lower secondary and upper secondary schools. In this regard, the task was deemed suitable for the present study as it demonstrates consistency in cross-cultural comparability. However, the task was adjusted by considering social and cultural factors. For instance, the phrase ‘cousin’s house’ was replaced with ‘cousin’s home’, and ‘doorbell rang ten times;’ was changed to ‘door opened ten times’. Two main concerns led to the adjustments. First, the sample consisted of students from urban, suburban, and rural schools. Many students lived in a traditional yurt called a ‘ger’, which is very different from a house in the Western context. Second, the rural students are not very familiar with the term ‘doorbell’. It is recommended to ensure that the items apply to the social and cultural context of the sample size whenever the previously used items are adapted.

Task 3. In the picture below, there is a triangle and an inscribed circle. Make up as many problems as you can that are in some way related to this picture.

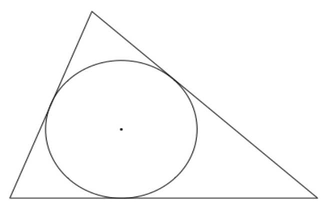


Figure 8. Figure for the semi-structured problem-posing situation example

Task 3 was adapted from Stoyanova’s (1997) dissertation and was also tested in a study by Sriraman and Lee (2011). As the task was adapted from the same source, the same reasons were considered to select the task for this study. In addition to its consistency in cross-cultural comparability and age appropriateness, the task content is relevant to Mongolian Grade 9 mathematics content (see, Lower secondary mathematics curriculum 7.4, Grade 9, Geometry & Measurement, p. 41, Grade 9 Mathematics textbook, p. 37).

Task 4. Example problem: Naran is three times older than Gerel. The sum of their two ages is 48. What are their ages?

Make up questions similar to the example, which require you to find two unknowns. Make sure the information provided in your question is adequate to find the desired unknowns.

This task was adapted from Akgul and Kahveci (2016). The task content is related to Mongolian Grade 9 mathematics content (Lower secondary mathematics curriculum 7.4, Grade 9, Numeracy & Algebra, p. 39, Grade 9 Mathematics textbook, p. 12). Moreover, psychometric properties were tested using promising statistical techniques such as the EFA and the test-retest approach with a study group consisting of students aged 10–15 in Turkey. Children's names in the task were adjusted to the Mongolian context by changing 'Ali' to 'Naran' and 'Ahmed' to 'Gerel'.

Table 14. Considerations for selecting problem-posing tasks

Task	Previous works			The present work	
	Source	Grade & Age	Psychometric analysis	Grade & Age	Content relevance
1	Kattou et al. (2013) developed and tested the item in a Cyprus primary and lower secondary sample.	Grade 6 (Age 10)	CFA	Grade 9 (Age 13–16)	Lower secondary mathematics curriculum 7.4, Grade 9 content, Numeracy & Algebra, p. 39
2	Stoyanova (1997) originally developed and tested the item in an Australian middle school sample. Sriraman and Lee (2011) validated the item in Chinese and American secondary school samples.	Grades 7 & 8 (Ages 13–14) Grades 10–12 (Ages 15–17)	Cross-cultural comparison analysis (Linguistic and cultural validity)		The task is not necessarily related to a particular content rather it enables students to create problems in particular mathematical area that came to mind.
3	Stoyanova (1997) originally developed and tested in an Australian middle school sample. Sriraman and Lee (2011) validated the item in Chinese and American secondary school samples.	Grades 7 & 8 (Ages 13–14) Grade 10 (Ages 15–17)	Cross-cultural comparison analysis (Linguistic and cultural validity)		Lower secondary mathematics curriculum 7.4, Grade 9, Geometry & Measurement, p. 41, Grade 9 Mathematics textbook, p. 37
4	Akgul & Kahveci (2016)	Grades 5–8 (Ages 10–15)	EFA, Test-rest analysis		Lower secondary mathematics curriculum 7.4, Grade 9, Numeracy & Algebra, p. 39, Grade 9 Mathematics textbook, p. 12

It should be noted that content validity was performed to understand whether the selected items accurately reflected the corresponding skills. Content validity is the extent to which the content of the items is consistent with the definition of the construct (Hair et al., 2010). This type of validity subjectively analyses the association between a concept and the individual items using ratings given by judges. In this regard, four experts were chosen based on their expertise in the subject matter. Two experts were specialised in mathematics and were responsible for designing the mathematics curriculum and mathematics textbooks at the Mongolian Institute for Educational Research. One expert was a specialist from the Education Evaluation Centre who was responsible for the Assessment Development Section. Another expert was a language expert from an International High School who was bilingual in Mongolian and English to ensure linguistic and cultural validation. Initially, there were 19 self-rating items and 5 problem-posing tasks. Based on the experts' feedback, four self-rating items and one problem-posing task were deleted, and 19 items, including 4 problem-posing tasks, were retained as appropriate to use.

5.4 Vignette sets

The vignette development process comprises the following three stages based on Zigler et al. (2020).

Stage 1: Vignette development

The vignette scenarios were designed based on the target questionnaires (item) adapted from the previous studies. Three vignettes were generated for each target item to reflect the response options at the five-point Likert scale, which resulted in 15 separate AVs. The AVs included the commonly used names in ethnic groups such as Kazakh and Mongol to ensure the actual representation of participants. The selected names were then given to the vignettes. The vignettes for mathematical perseverance, cooperative learning in mathematics, and mathematical enjoyment were carefully designed to be content-general.

In contrast, the vignettes for mathematical anxiety and mathematical self-efficacy were content-specific, including geometry content for Mongolia's ninth grade. Studies showed that the Mongolian students exhibited poor performance in geometry content compared to other aspects (Otgonbaatar, 2013; Education Evaluation Center, 2016; 2018). The level of the vignette sets except mathematical self-efficacy were decided by aligning with the vignette sets for Weiss and Roberts' Big Five personality traits (2018). The level of the vignette set for mathematical self-efficacy were decided based on the guide for constructing self-efficacy scales developed by Bandura (2006). The vignette sets had a random order in the administration context, but they were kept constant across participants. The respondents were given the vignette sets after responding to the questionnaires.

Stage 2: Assessment of content validity

First, the vignette sets were written in English and then translated into Mongolian. A bilingual language expert checked the translated versions of the vignette sets in Mongolian and English to ensure linguistic and cultural validation. In contrast, the math expert (math curriculum developer) checked the appropriateness of the math content in the vignette scenarios.

Stage 3. Quantitative analysis

The final stage was the quantitative analysis which tests vignette equivalence (if all respondents perceived the vignettes in the same way) and psychometric analysis of vignette corrected scores (details are discussed in Chapter 6). The processes of the three steps and the considerations undertaken are listed in Table 15.

Table 15. Procedures for the development of Anchoring vignettes

Stage	Process	Considerations
Stage 1: Vignette development	Develop the content. Decide the number of vignettes. Determine names and the implied gender of vignette characters.	Content general vs. content specific Variation of attributes (Levels) Order and number of vignettes Length of vignettes
Stage 2: Assessment of content validity	Revise vignettes considering the feedback	Readability and understanding Cultural validity
Stage 3: Quantitative analysis	Administer AVs alongside questionnaires in the target sample	Vignette equivalence (if all respondents perceived the vignettes in the same way) Psychometric analysis of vignette corrected scores
Note: The bold stage was not discussed in this Chapter.		

As an output of Stage 1 & 2, the vignette sets were developed. Sample vignette sets for mathematical perseverance, mathematical self-efficacy, and cooperative learning in mathematics are presented in Table 16.

Table 16. Sample vignette sets with their target items

Target characteristics	Being cooperative in mathematics class	Being persistent in solving mathematics problems	Being confident in solving mathematics problems
Target items	Thinks that mathematics is about working together with others to solve problems' (See Item 7, 8, & 9 in the Appendix)	Usually keeps trying a difficult problem until I have solved it' (See Item 13, 14, &15 in the Appendix)	Feels confident in finding the height of a pyramid' (See Item 10, 11, & 12 in the Appendix)
Target level	**Vignettes**		
Low	Bataa tends to disagree with others and, as a result, often starts quarrels. Therefore, he prefers to work on his own in mathematics even if he is stuck with a problem. He thinks that working with others in mathematics class does not help to perform better in mathematics.	Tuya easily feels desperate and gives up quickly if she faces difficulties in solving mathematics problems. She is unaware of the resources.	Zaya feels that he cannot do it at all when he is assigned to solve problems like finding the size of angle x in a given figure. Generally, he often thinks, 'I can't do it', when he is assigned to solve geometry problems.
Medium	Solongo does not really like working in a group during maths class. However, sometimes she thinks it is helpful to discuss with others when she is stuck when solving a mathematics problem.	Tulgaa tries to complete mathematical tasks when the answers or solutions are not readily available but gives up when the task is too difficult. He gets off the task easily. He draws on a limited range of resources. He does not put in enough effort to solve mathematics problems.	Ganaa feels moderately certain that he can solve geometry problems like finding the area of a triangle, but he feels somewhat unsure whether he can continue with some of the procedures to find the answer.
High	argal finds it easy to cooperate with others. He thinks that it is a good idea to combine all the students' ideas in a group when they work on a mathematics project. He also enjoys helping others to work well in a group on mathematics and listening to how others solve mathematics problems. Therefore, Jargal thinks he could do better in mathematics when he works with other students.	Oyunaa stays on a mathematical task no matter how difficult it is to find the answers. She searches for more information to clarify if she faces difficulties in solving the mathematics problem. Oyunaa always keeps trying a difficult mathematics problem until she has solved it.	Delger feels highly certain that she can accomplish the task when she is assigned to solve a geometry problem like finding the height of a pyramid. Most of the time, she feels geometry problems too easy and unchallenging.

In sum, the present study employed three instruments consisting of 34 items, including 1) 15 self-rating items, 2) 15 AVs, and 3) four problem-posing tasks (Table 17). For data analysis, the items were coded as it is presented in Table 18.

Table 17. Summary of the instruments used in this study

Instruments	Number of items	Comments
Questionnaires	15	Adapted from previous studies
AVs	15	Developed by the author
Problem-posing tasks	4	Adapted from previous studies
Total	34	

Table 18. Item coding

Domain	Sub-skill	Item	Coding
Openness and intellect	Mathematical creativity	1	MCQ1
		2	MCQ2
		3	MCQ3
		4	MCQ4
Conscientiousness	Mathematical perseverance	1	MPQ1
		2	MPQ2
		3	MPQ3
Agreeableness	Cooperative learning in mathematics	1	MCLQ1
		2	MCLQ2
		3	MCLQ3
Extraversion	Mathematics enjoyment	1	MEQ1
		2	MEQ2
		3	MEQ3
	Mathematics anxiety	1	MAQ1
		2	MAQ2

Neuroticism (Emotional stability)		3	MAQ3
Compound skill	Mathematical self-efficacy	1	MSEQ1
		2	MSEQ2
		3	MSEQ3

5.5 Scoring rules

5.5.1 Adjusting self-rated responses using AVs

To transform self-rated responses into vignette-corrected new responses, the present study employed a simple non-parametric approach introduced by King and Wand (2007). The five-point self-rating scales were converted into the vignette adjusted on seven-point scales as presented in Table 19.

Table 19. Possible values of self-rated and AV-adjusted item responses

Responses to self-rating	Strongly disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly agree 5		
AV-adjusted responses	Lower than the low vignette 1	Same as the low vignette 2	Between low and medium vignettes 3	Same as the medium vignette 4	Between the low and medium vignettes 5	Same as the high vignette 6	Higher than the high vignette 7

Adapted from Kyllonen and Bertling (2013)

The scoring rule for AVs is applied only when the vignettes are rated in the expected order (i.e., low vignette < medium vignette < high vignette). In some cases, responses to the vignettes can be tied (i.e., low vignette = medium vignette < high vignette) or violated (i.e., medium vignette < low vignette < high vignette). Kyllonen and Bertling (2014) suggest choosing the lowest possible score among the range of possible scores to modify the original scale into an AV-adjusted new scale to address this issue. For example, if a low vignette is tied with a medium vignette, the range of possible scores will be {2,3,4}. Then, 2 is assigned as an AV-adjusted new score.

Additionally, if there is order violation in the vignettes' responses and the low vignette is rated higher than the medium vignette, then the two vignettes are considered tied. Note that the tie should be created at the value that is given to the higher vignette. Then the ties are analysed as explained earlier. The same procedure was applied in this study. The procedure for correcting responses to a self-rating item for mathematics anxiety using the corresponding vignette set is presented in Table 20, including expected order, tied order, and violated order.

Table 20. Sample procedure for correcting responses to a mathematical anxiety item

Type of response	Response to an item for mathematics anxiety	Response to the **low vignette**	Response to the **medium vignette**	Response to the **high vignette**	Possible values	Vignette-corrected new scale
Expected order	1 (2) 3 4 5	1 2 (3) 4 5	1 2 3 (4) 5	1 2 3 4 (5)	1	1
Tie	1 (2) 3 4 5	1 2 (3) 4 5	1 2 (3) 4 5	1 2 3 (4) 5	{**2**,3,4}	2
Order violation	1 2 3 4 (5)	1 2 (3) 4 5	1 (2) 3 4 5	1 2 3 (4) 5	7	7

5.5.2 Scoring rule for problem-posing tasks

For the mathematical creativity test, fluency was scored by giving one point for every idea produced by the student. Flexibility was scored by categorising ideas produced and awarding one point for every category, and the originality score was determined by giving the highest mark to the rarest mark (Table 21).

Table 21. Scoring rubric for originality

Frequency	Score
1	9
2	8
3–4	7
5–8	6
9–16	5
17–32	4
33–64	3
65–128	2
129–256	1
257–445	0

Source: Adapted from Akgul and Kahveci (2016)

Therefore, three different scores were produced for each student in each task. The test's total score was derived by adding the fluency, flexibility, and originality scores on the four tasks. The scores for each component were carefully determined based on the agreement between the two experts. Some non-mathematical responses were eliminated from the analysis (e.g., responses such as 'my favourite numbers' for Task 1; 'Please find whose birthday it was. How can we find it?' and 'Were all the guests' classmates?' for Task 2; and 'Why is this circle inside the triangle?' and 'The triangle is stronger than the circle as it is bigger than the circle. Is it true?' for Task 3). The responses were entered into Microsoft Word and were coded into Microsoft Excel (Figure 9). The data was analysed using SPSS 20.0, JAMOVI 1.6.23, R 4.1.0, and AMOS 26.

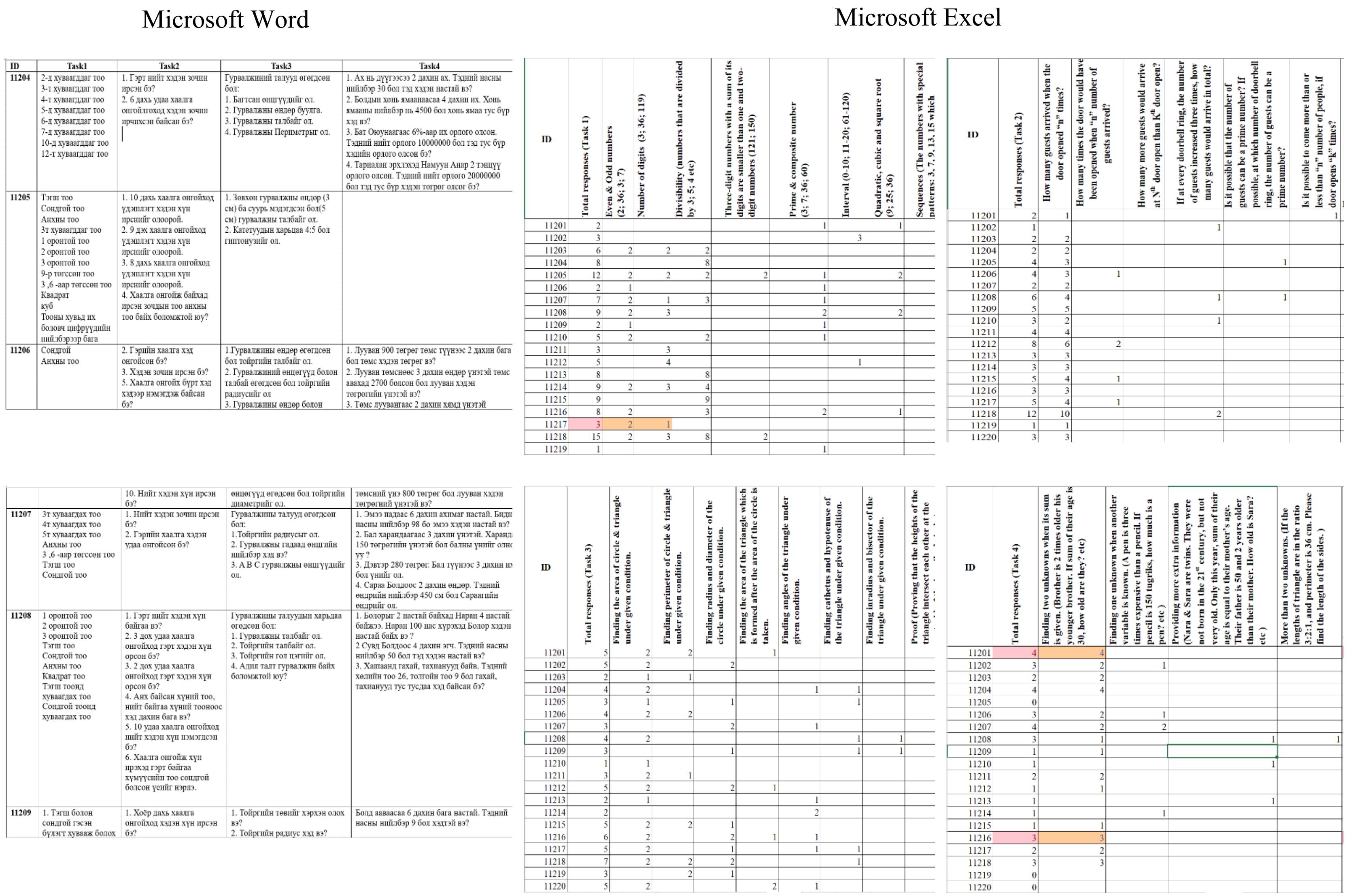


Figure 9. Process of organizing students' responses to the problem – posing tasks

CHAPTER 6

RESULTS AND DISCUSSIONS

6.1 Introduction

This chapter presents the outcomes of the analysis, including the vignette analysis and descriptive statistics of the general performance on the vignette-corrected scales, and the results of the reliability analysis and sampling adequacy for the factor analysis. The chapter further presents the results of the data analysis according to each research question outlined in Chapter 1. The results of the EFA and CFA for confirming the theoretically predicted framework and the results of different types of validity analysis for evaluating the tools are presented. Finally, the results of the descriptive statistics regarding students' performances on the vignette-corrected new scales and problem-posing tasks are presented, and the results of the independent samples t-test are presented with respect to the mean differences in students' mathematical creativity, cooperative learning in mathematics, mathematical perseverance, mathematical anxiety, mathematical enjoyment, and mathematical self-efficacy.

6.2 Analysis

6.2.1 Vignette analysis

Before interpreting the AV-corrected scores, the current study analysed the consistency of student responses with the vignettes. Descriptive statistics were examined for response patterns such as ties (i.e., same rating of two adjacent vignettes), correct ordering (i.e., conventional ordering of three hypothetical persons), and violations (a contradiction in the proposed anchored order of vignettes). Each scale's vignette ordering was evaluated using the 'anchor' package (Wand, King, & Lau, 2016) and R 4.1.0.

Table 22. illustrates the vignette orderings for mathematical perseverance with 308 total respondents rating vignettes. As evident from the first row of the table, the most common ordering for mathematical perseverance is '1,2,3' as rated by 188 respondents (61 per cent). The second row illustrates that the most frequent ordering is '{1,2},3', with 60 respondents (19 per cent) tying vignettes 1 and 2. The vignette ordering violations for mathematical perseverance can be seen in rows 4, 5, 7, 8, 9, and 10. However, these violations existed in fewer than 10 per cent of the sample (8.1 per cent for mathematical perseverance), which indicates that there are no arguable vignettes; hence, the slight violation can be regarded as a measurement error (Weiss & Roberts, 2018, p. 10).

Table 22. Vignette orderings for mathematical perseverance (N = 308)

Order	Frequency	Proportion	Violation
1,2,3	188	0.61	0
{1,2},3	60	0.19	0
1,{2,3}	31	0.10	0
2,{1,3}	8	0.03	1
2,1,3	8	0.03	1
{1,2,3}	4	0.01	0
1,3,2	4	0.01	1
{1,3},2	2	0.006	1
3,{1,2}	2	0.006	2
{2,3},1	1	0.003	2

Table 23. represents the vignette ordering for cooperative learning in mathematics with 308 respondents. The first row illustrates that the most common ordering for cooperative learning in mathematics is '1,2,3', with 185 respondents (60 per cent) rating vignettes. The second most common ordering is displayed in the second row as '1,{2,3}', with 59 respondents (19 per cent)

tying vignettes 2 and 3. Any violations in vignette ordering are depicted in rows 4, 5, 6, 8, 9, and 10. However, this violation existed in fewer than 10 per cent of the sample (8.4 per cent for cooperative learning in mathematics). This suggests the absence of any ambiguous vignettes, pointing towards the possibility of a measurement error (Weiss & Roberts, 2018, p. 10).

Table 23. Vignette orderings for cooperative learning in mathematics (N = 308)

Order	Frequency	Proportion	Violation
1,2,3	185	0.60	0
1,{2,3}	59	0.19	0
{1,2},3	35	0.11	0
1,3,2	7	0.02	1
2,1,3	7	0.02	1
2,{1,3}	6	0.01	1
{1,2,3}	3	0.01	0
3,1,2	3	0.010	2
{2,3},1	2	0.006	2
{1,3},2	1	0.003	1

Table 24. illustrates the vignette ordering for mathematics enjoyment, with the first row demonstrating ‘1,2,3’ as the most frequent ordering as rated by 207 respondents (67 per cent). The second most common ordering as depicted by the second row is ‘{1,2},3’, with 65 respondents (21 per cent) tying vignettes 2 and 3. The violations in vignette ordering occur in rows 3, 5, 6, 7, 9, and 10; however, they exist in less than 10 per cent of the sample (8.4 per cent for mathematical enjoyment). Considering the possibility of a measurement error, it is safe to assume that no problematic vignette exists (Weiss & Roberts, 2018, p. 10).

Table 24. Vignette orderings for mathematics enjoyment

Order	Frequency	Proportion	Violation
1,2,3	207	0.67	0
{1,2},3	65	0.21	0
2,1,3	14	0.05	1
1,{2,3}	9	0.03	0
2,{1,3}	5	0.02	1
2,3,1	3	0.01	2
3,1,2	2	0.01	2
{1,2,3}	1	0.003	0
{2,3},1	1	0.003	2
1,3,2	1	0.003	1

As for mathematics anxiety, the first row of Table 25. illustrates '3,2,1' as the most frequent ordering as rated by 171 respondents (56 per cent) rating vignettes. Similarly, the second row illustrates '3,{1,2}' as the second most frequent ordering as rated by 66 respondents (21 per cent) tying vignettes 1 and 2. Moving further down the table, it is evident that violations occur in rows 2, 3, 4, 5, 6, 7, 8, 9, and 10. In this particular case, violations occur in more than 10 per cent of the sample (i.e., 44 per cent of the mathematics anxiety sample), which indicates problematic vignettes. One possible factor might be the negative statement of vignettes, which often tends to confuse students about the rating process of hypothetical individuals.

Table 25. Vignette orderings for mathematics anxiety (N = 308)

Order	Frequency	Proportion	Violation
3, 2, 1	171	0.56	0
3,{1,2}	66	0.21	1
{2,3},1	28	0.09	1
3,1,2	15	0.05	1
2,{1,3}	9	0.03	2
{1,3},2	6	0.02	2
2,3,1	5	0.02	1
{1,2},3	4	0.01	2
{1,2,3}	2	0.007	2
2,1,3	2	0.007	2

Table 26. illustrates the results for mathematical self-efficacy, with its first row depicting '1,2,3' as the prevalent ordering as rated by 184 respondents (60 per cent) rating vignettes. The second row illustrates '{1,2},3' as the second most prevalent ordering as rated by 86 respondents (28 per cent) tying vignettes 1 and 2. Proceeding further, the violations in vignette ordering can be seen in rows 3, 5, 6, 7, 9, and 10. As the order violation is less than 10 per cent of the sample (9.4 per cent for the mathematical self-efficacy sample), it can conveniently be regarded as a measurement error. Hence, it is safe to say that there is no questionable vignette (Weiss & Roberts, 2018, p. 10).

Table 26. Vignette orderings for mathematical self-efficacy (N = 308)

Order	Frequency	Proportion	Violation
1,2,3	184	0.60	0
{1,2},3	86	0.28	0
2,1,3	16	0.05	1
1,{2,3}	8	0.03	0
{1,3},2	5	0.02	1
2,{1,3}	4	0.01	1
3,1,2	2	0.007	2
{1,2,3}	1	0.003	0
{2,3},1	1	0.003	2
1,3,2	1	0.003	1

The study also evaluated the vignette equivalence assumption, which implies a similar understanding of the variable represented in the vignette among all respondents involved (King & Wand, 2007). The mean and standard deviation (SD) values of self-rating for all three vignettes are presented in Table 27.

Table 27. Descriptive statistics of vignettes

Vignette	Mean	SD	n
Mathematical perseverance			
Vignette 3 (High)	4.74	0.60	308
Vignette 2 (Medium)	2.83	1.18	308
Vignette 1 (Low)	1.52	0.95	308
Cooperative learning in mathematics			
Vignette 3 (High)	4.74	0.73	308
Vignette 2 (Medium)	3.16	1.24	308
Vignette 1 (Low)	1.34	0.82	308
Mathematics enjoyment			
Vignette 3 (High)	4.76	0.60	308
Vignette 2 (Medium)	2.53	1.08	308
Vignette 1 (Low)	1.41	0.82	308
Mathematics anxiety*			
Vignette 1 (High)	4.35	0.85	308
Vignette 2 (Medium)	3.24	1.19	308
Vignette 3 (Low)	1.60	1.01	308
Mathematical self-efficacy			
Vignette 3 (High)	4.69	0.87	308
Vignette 2 (Medium)	2.43	1.07	308
Vignette 1 (Low)	1.49	0.90	308

Notes: *negative statements

The mean values indicate the characterisation of the vignettes and demonstrate consistency among vignette orderings. In other words, the high vignette has been rated above the medium vignette on average, which is in a higher ranking than the low vignette. Hence, the vignette equivalence assumption is backed.

The overall results demonstrate that the percentages of correct order vignettes range from 55 per cent to 66 per cent. Any violations of 10 per cent or less can be considered measurement errors (Weiss & Roberts, 2018). Vignettes such as mathematical perseverance, mathematics enjoyment, cooperative learning in mathematics, and mathematical self-efficacy demonstrated less than 10 per cent order violations, which indicate unproblematic vignettes. As for mathematics anxiety, the violations were quite high. However, the present study continued the analysis despite the violations in mathematics anxiety since the vignette equivalence assumption is acceptable in the case of partial violations (see Weiss & Roberts, 2018). Overall, the vignettes were well functioning and applicable to correct self-rated responses.

6.2.2 Descriptive statistics of vignette-corrected new scales

As the vignettes were concluded to be applicable for further analysis, self-rated responses were corrected using the vignette evaluation. Descriptive statistics for each item were computed after vignette correction, and the score distribution was tested for normality. The skewness and the distribution's kurtosis ranged from -1.082 to 1.214 and from -0.955 to 1.932, respectively (Table 28). These fulfil the criteria for the absolute values of skewness (<3) and kurtosis (<8) (Kline, 2010).

Table 28. Descriptive statistics of vignette-corrected new scales and problem-posing tasks

	Item	Min	Max	Mean	SD	Skewness	Kurtosis
1	MCLQ1	1	7	4.35	1.76	-0.334	-0.955
2	MCLQ2	1	7	4.52	1.74	-0.492	-0.845
3	MCLQ3	1	7	4.55	1.71	-0.512	-0.789
4	MPQ1	1	7	4.42	1.76	-0.546	-0.917
5	MPQ2	1	7	4.48	1.67	-0.555	-772
6	MPQ3	1	7	4.41	1.69	-0.534	-0.796
7	MAQ1	1	7	4.83	1.67	-0.523	-0.606
8	MAQ2	1	7	4.88	1.73	-0.672	-0.502
9	MAQ3	1	7	5.26	1.65	-1.082	0.294
10	MEQ1	1	7	4.80	1.71	-0.873	-0.391
11	MEQ2	1	7	4.74	1.67	-0.834	-0.394
12	MEQ3	1	7	4.51	1.72	-0.698	-0.631
13	MSEQ1	1	7	4.56	1.73	-0.784	-0.461
14	MSEQ2	1	7	4.58	1.73	-0.728	-0.589
15	MSEQ3	1	7	4.53	1.69	-0.706	-0.507
16	MCQ1	2	27	7.36	4.20	1.177	1.932
17	MCQ2	2	23	6.43	4.27	0.768	-0.038
18	MCQ3	3	17	7.35	3.26	0.622	-0.365
19	MCQ4	2	15	4.48	2.87	1.214	0.759

6.2.3 Reliability analysis

To test the effect of AVs on reliability, the current study compared McDonald's ω to measure the internal consistency of the original and vignette-corrected scales. The commonly agreed-upon lowest value of McDonald's ω is 0.7. McDonald's ω coefficient was computed using JAMOVI 1.6.23.

Table 29. Internal consistency of original and vignette-corrected scales

Scale	Original scale	Vignette-corrected scale
	McDonald's ω	McDonald's ω
Mathematical perseverance	0.699	0.833
Mathematical self-efficacy	0.651	0.802
Mathematics anxiety	0.640	0.718
Mathematics enjoyment	0.848	0.893
Cooperative learning in mathematics	0.667	0.823
Mathematical creativity	0.707	
Total scale	0.738	0.743

As presented in Table 29, the values of McDonald's ω for most of the original scales failed to satisfy the threshold value except the mathematics enjoyment scale. However, after the original scales were corrected by the vignettes, the values for internal consistency increased, for example, from 0.640 to 0.718 for the mathematics anxiety scale and from 0.699 to 0.833 for the mathematical perseverance scale. According to the reliability analysis, the vignette-corrected scales demonstrated higher reliability than the self-rating scales. This finding suggests that the AV approach has the potential to improve internal consistency. Based on the vignette analysis, the vignette-corrected new scales were used for further analysis.

6.2.4 Sampling adequacy

Before proceeding to the extraction process of constructs, certain tests are carried out to verify the adequacy of the data and sample for a factor analysis (Burton and Mazerolle, 2011). The Kaiser-Meyer-Olkin (KMO) test can be used to authenticate sampling adequacy (Kaiser, 1970), and a KMO score more than 0.50 is regarded as adequate for an EFA output analysis. Bartlett's Test of Sphericity (Bartlett, 1950) generates a chi-square output that must be significant. This demonstrates that the matrix is not an identity matrix; hence, it must be significant for a factor analysis to be deemed appropriate (Hair et al., 1995a; Tabachnick & Fidell, 2001). The Kaiser Meyer-Olkin's sample adequacy measure (KMO = 0.716) and Bartlett's Test of Sphericity ($\chi 2$ = 2136/df = 171; $p<.0001$) results demonstrate that the data selected were suitable for factor analysis (Table 30).

Table 30. KMO & Bartlett's Test of Sphericity (Sample size = 308)

Kaiser Meyer-Olkin		0.716
Bartlett's Test of Sphericity	$\chi 2$	2136
	df	171
	Sig.	$p < .0001$

A guide to sampling sizes was presented by Comrey (1973), who categorised the sample into 100 (poor), 200 (fair), 300 (good), 500 (very good), and 1,000 or above (excellent) for a factor analysis. Although suggestions were also made for the sample size to be larger than 100 (Hair, Anderson et al., 1995a), the present study employed a sample size of 308, which fulfils the minimum criteria for a factor analysis.

6.3 Results

This section presents the results of the data analysis according to each research question outlined in Chapter 1.

6.3.1 Evaluation of the measurement model (RQ1)

RQ1: What framework can be used to capture social and emotional skills in mathematics?

The measurement model often serves as a framework that builds an association between data and theory. It can either be reflective (where items are determined by constructs) or formative (where constructs are determined by the items). In the case of reflective models, a similar theme exists among items as they emerge from the same domain. On the contrary, the items in the formative models do not have a similar theme, as each item represents different aspects of the construct. Therefore, the items in the reflective model are more likely to be correlated as they measure the same underlying construct, while the possibility of correlation is negligible between the formative construct items (Quoquab & Mohammad, 2020). Hence, the study presents a six-factor reflective model (Figure 10) based on the theoretically predicted framework for social and emotional skills in mathematics (Table 31). To validate the framework, the data were analysed using SPSS 20.0 for EFA, and AMOS 26.0 for CFA.

Table 31. The theoretically predicted framework used in this study

	Domain (Big Five)	Sub-skills
1	Openness and intellect	Mathematical creativity
2	Conscientiousness	Mathematical perseverance
3	Extraversion	Mathematical enjoyment
4	Agreeableness	Cooperative learning in mathematics
5	Neuroticism	Mathematical anxiety
6	Compound skill	Mathematical self-efficacy

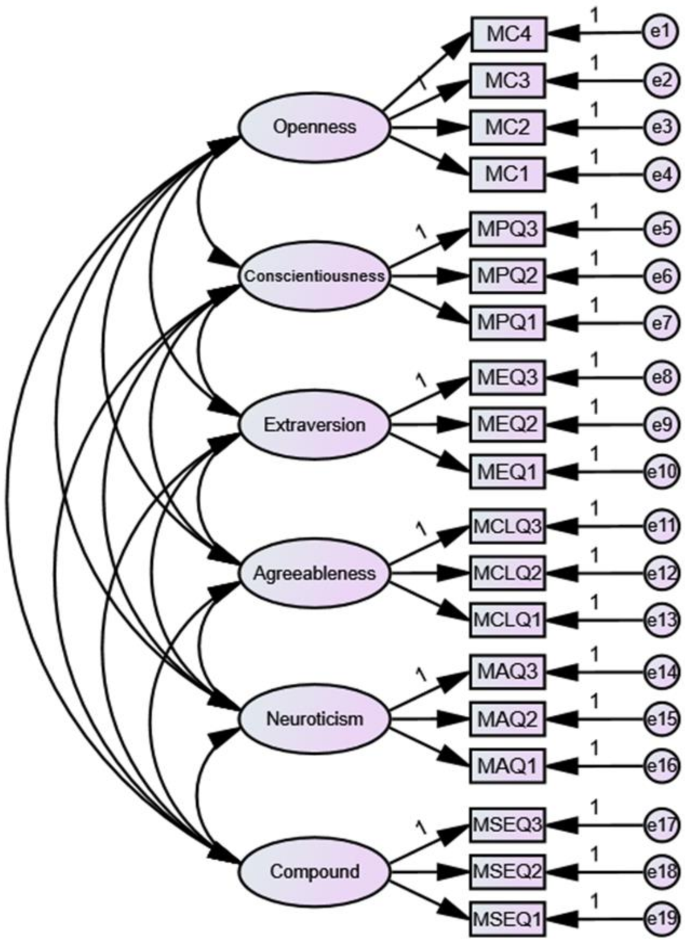


Figure 10. Six-factor hypothetical model for social and emotional skills in mathematics

6.3.1.1 Exploratory factor analysis results

Leandre et al. (2012) stressed that EFA is frequently converged with Principal Component Analysis (PCA) despite both being different in several ways. EFA identifies the probable underlying constructs for a set of items. It is assumed that for EFA, the shared variance among items depicts the construct, while the non-shared variance shows the measurement errors. On the contrary, PCA is a technique for data reduction without assuming an underlying construct. Under PCA, the number of observed variables is reduced to minor components, which contain most of the variance (Knekta, Runyon, and Eddy, 2019). Different EFA methods are employed for different purposes. If the purpose of the analysis is to only deal with a sample for further analysis, it is suggested to use Principal Axis Factoring (Samuel, 2017). If it aims to develop an instrument to be employed in the future, it is recommended to use Maximum Likelihood (ML). Thus, in the context of this study, ML was applied to the data, and factor loadings were analysed after the Promax rotation technique was employed to deduce whether it was possible to divide the scale into independent factors. Items with factor loading values of less than 0.30 and those distributed under multiple factors with a difference of 0.10 must be taken off the scale (Field, 2013). In the context of the analysis conducted in this study, the eigenvalues of items should be at least 1.00, with their factor load being at least 0.40 according to Kaiser's method (Kaiser, 1974).

Before the factor extraction, an item correlation analysis was conducted if there is any redundant item in the scale. A correlation matrix is presented in Table 32. For a correlation matrix analysis, it is important to examine two assumptions to avoid redundant items in a scale. First, any correlation greater than 0.8 should be removed. Second, the determinant value of the correlation matrix should be greater than 0.00001 (Field, 2013). As evident from the correlation matrix, there is no correlation coefficient greater than 0.8, and the determinant of the correlation matrix is 0.001, which is greater than the threshold value (0.00001). As the critical values are in the safe range, it can be concluded that there are no problematic items for a factor analysis.

Table 32. Correlation matrix

		MCLQ1	MCLQ2	MCLQ3	MPQ1	MPQ2	MPQ3	MSEQ1	MSEQ2	MSEQ3	MEQ1	MEQ2	MEQ3	MAQ1	MAQ2	MAQ3	MCQ1	MCQ2	MCQ3	MCQ4
Correlation	MCLQ1	1.000	.713	.548	.128	.111	.146	.131	.077	.186	.161	.157	.134	-.055	.037	.000	-.007	.035	.007	.099
	MCLQ2	.713	1.000	.527	.192	.134	.116	.184	.166	.171	.223	.243	.171	-.033	-.058	.034	-.093	-.005	-.014	.036
	MCLQ3	.548	.527	1.000	.062	.069	.075	.130	.207	.125	.066	.105	.034	-.025	.004	.000	-.016	-.016	.120	.147
	MPQ1	.128	.192	.062	1.000	.647	.593	.139	.155	.177	.217	.197	.260	.017	.001	.030	-.047	-.012	.025	.034
	MPQ2	.111	.134	.069	.647	1.000	.633	.125	.059	.109	.179	.191	.289	-.016	.051	.043	-.079	-.027	-.005	.020
	MPQ3	.146	.116	.075	.593	.633	1.000	.111	.132	.108	.218	.185	.252	-.039	.024	-.002	.016	.010	.040	.085
	MSEQ1	.131	.184	.130	.139	.125	.111	1.000	.563	.633	.090	.073	.095	.119	.101	.146	-.043	-.025	-.044	-.018
	MSEQ2	.077	.166	.207	.155	.059	.132	.563	1.000	.514	.171	.183	.094	.056	.088	.108	.022	.073	.114	.060
	MSEQ3	.186	.171	.125	.177	.109	.108	.633	.514	1.000	.184	.178	.158	.112	.163	.192	-.059	-.015	-.045	-.015
	MEQ1	.161	.223	.066	.217	.179	.218	.090	.171	.184	1.000	.797	.655	.024	-.006	.113	.059	.066	.035	.053
	MEQ2	.157	.243	.105	.197	.191	.185	.073	.183	.178	.797	1.000	.738	.027	.018	.087	.003	.014	-.012	.008
	MEQ3	.134	.171	.034	.260	.289	.252	.095	.094	.158	.655	.738	1.000	.034	.104	.082	.019	.025	.001	.000
	MAQ1	-.055	-.033	-.025	.017	-.016	-.039	.119	.056	.112	.024	.027	.034	1.000	.528	.439	-.010	.022	-.051	-.035
	MAQ2	.037	-.058	.004	.001	.051	.024	.101	.088	.163	-.006	.018	.104	.528	1.000	.391	.017	.120	-.039	-.015
	MAQ3	.000	.034	.000	.030	.043	-.002	.146	.108	.192	.113	.087	.082	.439	.391	1.000	.078	.100	.043	.003
	MCQ1	-.007	-.093	-.016	-.047	-.079	.016	-.043	.022	-.059	.059	.003	.019	-.010	.017	.078	1.000	.476	.370	.300
	MCQ2	.035	-.005	-.016	-.012	-.027	.010	-.025	.073	-.015	.066	.014	.025	.022	.120	.100	.476	1.000	.359	.363
	MCQ3	.007	-.014	.120	.025	-.005	.040	-.044	.114	-.045	.035	-.012	.001	-.051	-.039	.043	.370	.359	1.000	.381
	MCQ4	.099	.036	.147	.034	.020	.085	-.018	.060	-.015	.053	.008	.000	-.035	-.015	.003	.300	.363	.381	1.000
Sig. (1-tailed)	MCLQ1		.000	.000	.012	.026	.005	.011	.089	.001	.002	.003	.009	.167	.258	.494	.455	.268	.450	.042
	MCLQ2	.000		.000	.000	.009	.021	.001	.002	.001	.000	.000	.001	.280	.156	.276	.052	.462	.404	.266
	MCLQ3	.000	.000		.137	.113	.094	.011	.000	.014	.126	.032	.276	.328	.474	.499	.391	.389	.018	.005
	MPQ1	.012	.000	.137		.000	.000	.007	.003	.001	.000	.000	.000	.380	.496	.303	.204	.419	.331	.276
	MPQ2	.026	.009	.113	.000		.000	.014	.150	.028	.001	.000	.000	.387	.185	.226	.083	.320	.465	.361
	MPQ3	.005	.021	.094	.000	.000		.026	.010	.030	.000	.001	.000	.250	.340	.484	.391	.428	.244	.069
	MSEQ1	.011	.001	.011	.007	.014	.026		.000	.000	.057	.101	.047	.018	.039	.005	.224	.331	.219	.378
	MSEQ2	.089	.002	.000	.003	.150	.010	.000		.000	.001	.001	.051	.163	.061	.029	.351	.102	.022	.145
	MSEQ3	.001	.001	.014	.001	.028	.030	.000	.000		.001	.001	.003	.025	.002	.000	.149	.395	.218	.397
	MEQ1	.002	.000	.126	.000	.001	.000	.057	.001	.001		.000	.000	.335	.458	.024	.150	.126	.270	.177
	MEQ2	.003	.000	.032	.000	.000	.001	.101	.001	.001	.000		.000	.318	.375	.064	.480	.404	.416	.443
	MEQ3	.009	.001	.276	.000	.000	.000	.047	.051	.003	.000	.000		.278	.034	.075	.368	.331	.494	.498
	MAQ1	.167	.280	.328	.380	.387	.250	.018	.163	.025	.335	.318	.278		.000	.000	.430	.351	.185	.268
	MAQ2	.258	.156	.474	.496	.185	.340	.039	.061	.002	.458	.375	.034	.000		.000	.383	.017	.248	.394
	MAQ3	.494	.276	.499	.303	.226	.484	.005	.029	.000	.024	.064	.075	.000	.000		.086	.039	.226	.476
	MCQ1	.455	.052	.391	.204	.083	.391	.224	.351	.149	.150	.480	.368	.430	.383	.086		.000	.000	.000
	MCQ2	.268	.462	.389	.419	.320	.428	.331	.102	.395	.126	.404	.331	.351	.017	.039	.000		.000	.000
	MCQ3	.450	.404	.018	.331	.465	.244	.219	.022	.218	.270	.416	.494	.185	.248	.226	.000	.000		.000
	MCQ4	.042	.266	.005	.276	.361	.069	.378	.145	.397	.177	.443	.498	.268	.394	.476	.000	.000	.000	

a. Determinant = .001

As evident from Table 34, 69.604 per cent of the total variance can be explained by the factors and the items in the scale. According to Hensen and Roberts (2006), the minimum ratio can be 52 per cent; hence, the attained value is sufficient. Factor names were determined based on the content of the items placed in all six factors. Factors such as 'Mathematical enjoyment', 'Mathematical perseverance', 'Cooperative learning in mathematics', 'Mathematical self-efficacy', 'Mathematical anxiety', and all contained three items, whereas the factor named 'Mathematical creativity' contained four items. As evident from Table 34, the 'Mathematical enjoyment' factor contains items with a factor load varying between 0.782 and 0.952. Its eigenvalue in the general scale was 3.699, while it contributed 19.469 per cent to the general variance. The 'Mathematical perseverance' factor contained items with factor load varying between 0.760 and 0.840. Its eigenvalue was found to be 2.179, while it contributed 11.469 per cent to the general variance. The 'Cooperative learning' factor contained items with factor load varying between 0.639 and 0.862. Its eigenvalue was found to be 2.138, while it contributed 11.251 per cent to the general variance. The 'Mathematical self-efficacy' factor contained items with factor values varying between 0.699 and 0.824. Its eigenvalue was found to be 2.014, while it contributed 10.602 per cent to the general variance. The 'Mathematical creativity' factor contained items with factor loads varying between 0.547 and 0.669. Its eigenvalue was found to be 1.713, while it contributed 9.013 per cent to the general variance. The 'Mathematical anxiety' factor contained items with factor loads varying between 0.572 and 0.729. Its eigenvalue was found to be 1.482, while it contributed 7.80 per cent to the general variance. Moreover, acceptable commonalities should be at least above 0.2 (Samuels, 2017). Communalities for the items varied between 0.312 and 0.911, which satisfies the criteria, and so all the items can be retained.

Table 33. Factor loadings and factor structures retrieved from Exploratory Factor Analysis

Items	Component						Communalities
	Factor 1 Mathematical enjoyment	Factor 2 Mathematical perseverance	Factor 3 Cooperative learning in mathematics	Factor 4 Mathematical self-efficacy	Factor 5 Mathematical creativity	Factor 6 Mathematical anxiety	
MCLQ1	.170	.165	**.862**	.171	.042	-.003	.747
MCLQ2	.252	.186	**.831**	.228	-.040	-.041	.700
MCLQ3	.098	.089	**.639**	.190	.077	-.024	.417
MPQ1	.255	**.773**	.171	.198	-.010	.020	.600
MPQ2	.247	**.840**	.135	.131	-.048	.038	.710
MPQ3	.244	**.760**	.149	.146	.059	-.007	.583
MSEQ1	.091	.162	.193	**.824**	-.061	.176	.689
MSEQ2	.190	.132	.168	**.699**	.106	.110	.505
MSEQ3	.196	.164	.217	**.759**	-.056	.222	.585
MEQ1	**.839**	.248	.202	.192	.084	.051	.708
MEQ2	**.952**	.235	.215	.183	.003	.057	.911
MEQ3	**.782**	.337	.156	.154	.011	.113	.626
MAQ1	.037	-.020	-.055	.128	-.036	**.726**	.532

MAQ2	.036	.037	-.008	.139	.035	**.729**	.534
MAQ3	.108	.030	.012	.193	.082	**.572**	.339
MCQ1	.028	-.053	-.047	-.030	**.647**	.038	.426
MCQ2	.041	-.013	.017	.012	**.669**	.119	.457
MCQ3	.010	.021	.024	.007	**.597**	-.041	.361
MCQ4	.028	.054	.105	.014	**.547**	-.027	.312
Eigenvalue	3.699	2.179	2.138	2.014	1.713	1.482	
Explained variance	19.469 per cent	11.469 per cent	11.251 per cent	10.602 per cent	9.013 per cent	7.800 per cent	
Total variance	69.604 per cent						
Extraction Method: Maximum Likelihood. Rotation Method: Promax with Kaiser Normalization.							

Cattell's Scree test (Cattell, 1966) is another widely used method to determine the number of factors that need to be retained. It visually explores the graphical representation of the eigenvalues to look for possible disruptions. The number of values above the break (excluding the point at which the break occurs) is the number of factors that need to be retained. The rationale behind this method is that the point of break segregates the significant factors from minor factors (Ledesma & Valero-Mora, 2007). The results show that six factors had eigenvalues higher than 1, which indicates that there were six independent factors (Figure 15).

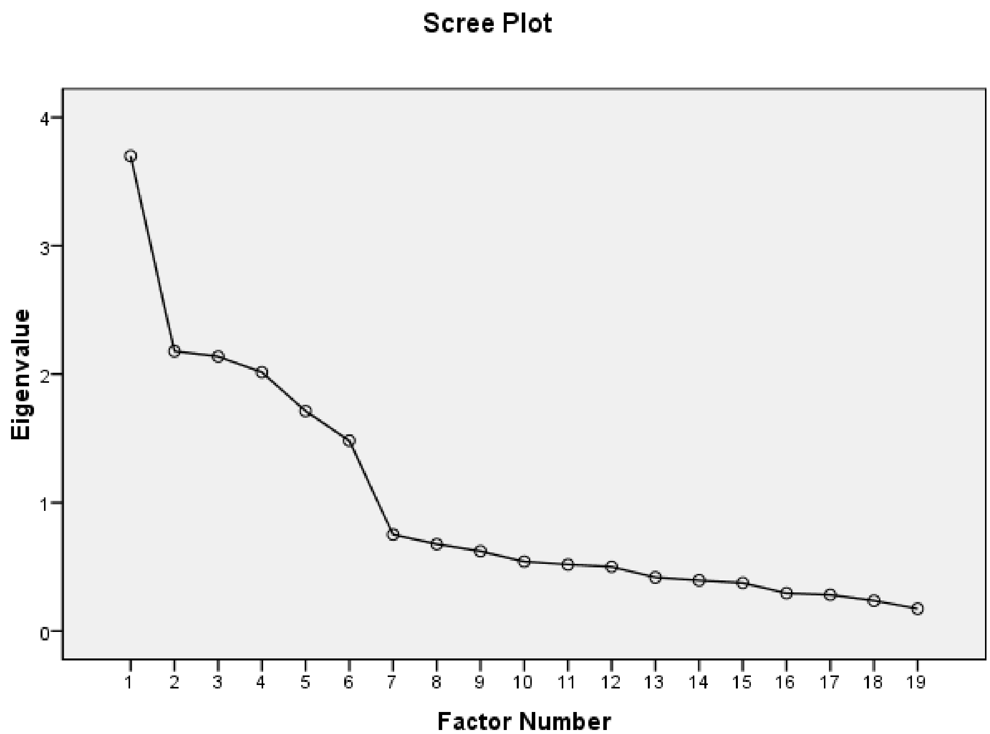


Figure 11. *Scatter diagram for the eigenvalues of the factors* Based on the factor structures provided by the EFA, the items were classified into their theoretically corresponding domains without any cross-loading and satisfying suggested criteria for the necessary statistical assumptions.

6.3.1.2 Confirmatory factor analysis results

The EFA implied that six factors determine the dimensions of social and emotional skills in mathematics. A CFA was conducted on the same sample to confirm these factors. This was because one of the fundamental objectives of a CFA is to examine whether the proposed measurement theory is valid.

For the CFA analysis, several model fit statistics are recommended to assess whether the theoretical model fits the selected data (Kline, 2005; Parry, 2017). These include Normed Chi-Square ($\chi2$/df), Root Mean Square Error of Approximation (RMSEA), Comparative Fit Index (CFI), Normed Fit Index (NFI), Tucker Lewis Index (TLI), and Parsimony Normed Fit Index (PNFI). These fit indices are grouped into three categories: *absolute*, *incremental*, and *parsimony fir indices* (Hair et al., 2010). Absolute fit indices evaluate the theoretical model considering the observed data. This category includes $\chi2$/df and RMSEA. Incremental fit indices assess the hypothesized model while comparing it to baseline models. This category includes NFI, CFI, and TLI. In contrast, Parsimony fit indices are devised to identify the best-suited model among a set of varying models. PNFI is an example.

$\chi2$/df: A measure of goodness of fit for a model with a ratio of $\chi2$ to degrees of freedom. The ratios of $\chi2$:df in the order of 3:1 or less are generally considered better-fitting models, except when the sample size is larger than 750 (Hair et al., 2010).

RMSEA: A measure of goodness of fit for statistical models that require the population to have a close fit with the model rather than a complete fit, which is not practical for large population sizes (Kaplan, 2000).

CFI: The CFI examines the model fit by observing the discrepancy between the hypothesised model and the data while resolving the sample size issues in the chi-squared test of model fit (Gatignon, 2010).

NFI: The NFI is one of the original incremental goodness of fit measures for a statistical model that remains unaffected by the number of variables/parameters. The goodness of fit is measured by comparing the model of interest with a model containing entirely uncorrelated variables (Ullman, 1996).

TLI: An incremental fit index devised against the NFI, which is disadvantaged by the sample size. The key advantage of TLI lies in its ability to remain unaffected by changes in the sample size (Cangur & Ercan, 2015).

PNFI: The PNFI can adjust the NFI by multiplying it with the parsimony ratio. It also has more characteristics of incremental fit indices than absolute fit indices to favour less complicated models (Hair et al., 2010).

Table 34 illustrates the calculated values and the range of values that are acceptable for these indices. In the context of the CFI, values greater than 0.90 demonstrate an acceptable fit, and values greater than 0.95 demonstrate an excellent fit. For RMSEA, any value less than 0.06 illustrates a good-fitting model, and values below 0.08 represent a good fit. For $\chi 2/df$, values are obtained based on the ratio of chi-square and the degree of freedom. A value less than 2.0 is regarded as good, whereas any value between 2.0 and 5.0 is acceptable. The table illustrates that all fit indices are within acceptable value ranges, implying that the theoretically predicted model is a good fit for the selected data. Details of the fit indices can be found in literature: for chi-square statistics ($\chi 2/df$) and the PNFI, see Wu, Yang, and Koo (2017); and for the RMSEA, CFI, and NFI see Hu and Bentler (1999); for the TLI, see Schumacker and Lomax (2010).

Table 34. Fit indices

	Fit statistics	Values	Acceptance level
Absolute Fit Indices			
1	$\chi 2$/df	1.32 (p-value <0.05)	<1 → Very good 1–2 → Good 2–5 → Acceptable >5 → Unacceptable
2	RMSEA	0.03	<0.05 → Very good 0.05–0.08 → Good >0.08 → Unacceptable
Incremental Fit Indices			
3	NFI	0.922	>0.90 → Very good 0.80–0.90 → Acceptable <0.80 → Unacceptable
4	CFI	0.980	>0.95 → Very good 0.90–0.95 Acceptable <0.90 → Unacceptable
5	TLI	0.974	>0.95 → Very good
Parsimony Fit Indices			
6	PNFI	0.647	≥ 0.50 → Acceptable

Besides the fit statistics, several findings were observed from the analysis of the CFA model. First, the existence of the six domains found during the EFA was supported by the CFA. Second, it demonstrated moderate (0.54) to higher (0.94) factor loadings for all six factors without cross-loadings, which means all the factors corresponded to their theoretically relevant domains (see Figure 12).

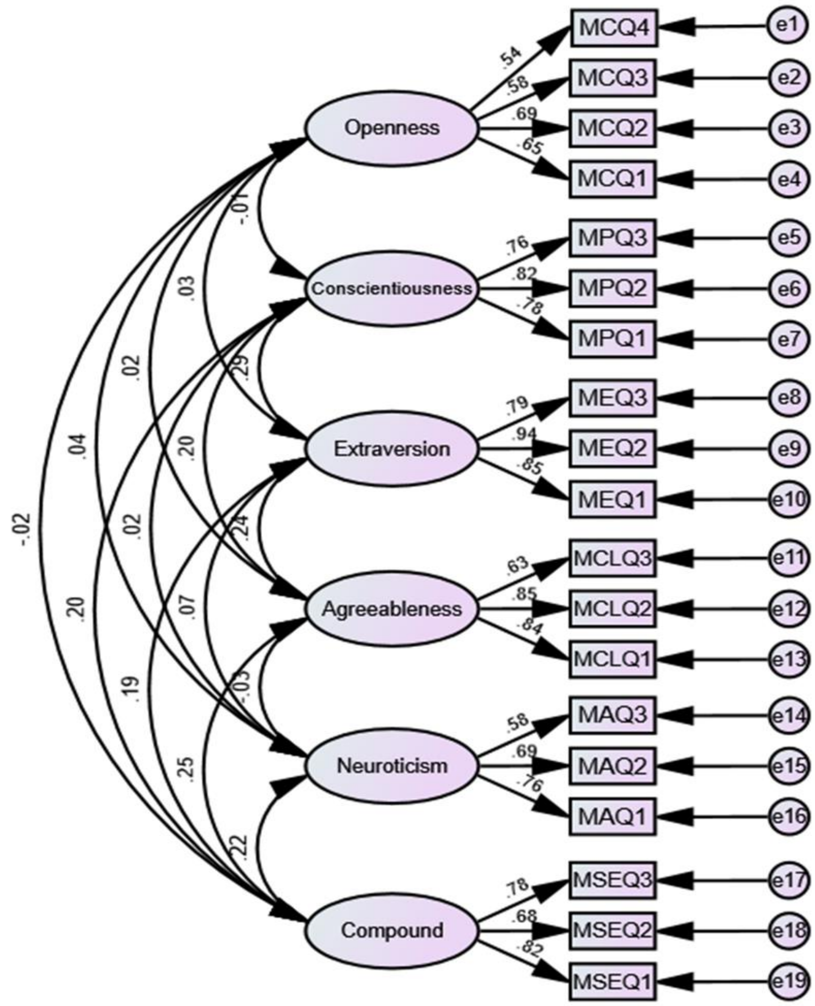


Figure 12. Six-factor hypothetical model for social and emotional skills in mathematics

Overall, the theoretically predicted framework for identifying social and emotional skills in mathematics was empirically confirmed by the findings from the EFA and CFA.

6.3.2 Validation of the tools (RQ2)

RQ2: What valid tools can measure social and emotional skills in mathematics reliably?

After the EFA and the CFA were conducted to confirm the theoretically predicted framework, several techniques, such as content, convergent, discriminant, and nomological validity, were carried out to test the reliability and validity of the survey instruments.

Content validity is the extent to which the content of the items is consistent with the definition of the construct. This type of validity is tested in developing tools and subjectively analyses the association between a concept and the individual items using judges' ratings (Hair et al., 2010). The establishment of content validity was discussed in the Methodology section above.

6.3.2.1 Results of the convergent validity analysis

Convergent validity examines the extent of correlation between two measures of the same concept (Hair et al., 2010). The first step is to examine the factor loadings, as it is important to consider the size of the factor loadings for establishing convergent validity. When the convergent validity is high, all factor loadings must be above 0.5 (Hair et al., 2010). As evident in Table 34, the factor loadings for all constructs are above the cut-off value, meaning that convergent validity was achieved.

The next step is to assess reliability, which is another indicator of convergent validity. A higher convergent validity is obtained when the CR value is between 0.7 and 0.9. Any value equal to or larger than 0.95 is problematic since it indicates redundant items, which then reduces validity. Table 38 illustrates that CR exceeds the minimum acceptable value for all constructs ranging from 0.707 to 0.894, establishing convergent validity. Additionally, for higher convergent validity, the CR value must be greater than the average variance extracted (AVE). The AVE is a summary measure of convergence of a set of items that represent a latent construct. It refers to the average percentage of variation that can be explained by the items included in a construct (Hair et al., 2010). According to Table 39, the CR value is greater than the AVE for all constructs, indicating convergent validity.

Table 35. Indices of composite reliability and average variance extracted for convergent validity

Construct	CR	AVE
Mathematical creativity	0.707	0.378
Mathematical perseverance	0.834	0.626
Cooperative learning in mathematics	0.822	0.610
Mathematics enjoyment	0.894	0.739
Mathematics anxiety	0.718	0.462
Mathematical self-efficacy	0.802	0.576

6.3.2.2 Results of the discriminant validity analysis

Discriminant validity analyses the extent to which a particular construct is different from other constructs in terms of correlation and representation by distinctly measured variables. Therefore, a high discriminant validity proves that a construct is unique, as it contains some phenomenon not possessed by any other construct (Hair et al., 2010). There are two standard methods of analysing discriminant validity under CFA. The first one is the criteria put forward by Fornell and Larcker (1981). According to the Fornell-Larcker criteria, the square root of the AVE must be much higher than the correlation of a specific construct with other constructs. The minimum value of the square root of AVE for each construct must be 0.50. The second method is to compare the value of the AVE with that of the maximum-shared variance (MSV). According to the MSV, discriminant validity is established when AVE is higher than MSV for each construct (Hair et al., 2010). Evidence of significant cross loading also demonstrates a lack of discriminant validity.

The first method requires the AVE values to be calculated and compared with each construct's correlation coefficients. For this purpose, a matrix that demonstrates each construct's correlation with other constructs must be obtained. Later, the AVE value is inserted in the diagonal compared with the other correlation coefficient (Table 40). The AMOS 26 software has been used to compute

AVE values and correlation coefficients for all constructs. The results of the AVE analysis are depicted in Table 40. It is evident from the results that the AVE values were above 0.5 and higher than the correlation coefficients for each construct. The results demonstrate that discriminant validity was achieved for all constructs.

Table 36. Fornell-Larker criteria for discriminant validity

Construct	1	2	3	4	5	6
Mathematical creativity **(1)**	**0.615**					
Mathematical perseverance **(2)**	-0.007	**0.791**				
Cooperative learning in mathematics **(3)**	0.017	0.196**	**0.781**			
Mathematics enjoyment **(4)**	0.031	0.291**	0.241**	**0.859**		
Mathematics anxiety **(5)**	0.044	0.017	-0.032	0.067	**0.680**	
Mathematical self-efficacy **(6)**	-0.024	0.202	0.248	0.193	0.223	**0.759**

For the second method, AMOS 26 software was used to compute both MSV and AVE values. The results are illustrated in Table 41. All constructs with an AVE higher than MSV have indices with good validity. Hence, both convergent validity and discriminant validity were achieved for all the constructs.

Table 37. Comparison of AVE and MSV for discriminant validity

Construct	AVE	MSV
Mathematical creativity	0.378	0.002
Mathematical perseverance	0.626	0.085
Cooperative learning in mathematics	0.610	0.061
Mathematics enjoyment	0.739	0.085
Mathematics anxiety	0.462	0.050
Mathematical self-efficacy	0.576	0.061

6.3.2.3 Results of the nomological validity analysis

Nomological validity refers to the form of test validity that assesses the meaningfulness of the construct correlations in the measurement theory. The correlation between constructs is useful in the assessment. For instance, social interaction and cooperation play a crucial role in stimulating creativity in mathematical activities (Sawyer, 2007; Sriraman, 2004). Moreover, working in groups in a classroom environment provides an opportunity for students to improve their creative thinking skills (Molad, Levenson, & Levy, 2020). Similarly, it was confirmed that the more the students cooperate in solving mathematical problems, the more they enjoy the mathematics class (Schukajlow et al., 2012). Moreover, according to Arısoy (2011, as cited in Karali and Aydemir, 2018), cooperative learning in mathematics improves the level of mathematical perseverance in students. Furthermore, when lower secondary students are taught in cooperative learning for a significant period, the students demonstrate increased levels of self-efficacy in solving mathematical problems (Perels et al., 2005). Finally, the more students are cooperative in mathematical problem solving, the less they feel anxiety (Lavasani & Khandan, 2011). In sum, cooperative learning in mathematics is negatively correlated with mathematics anxiety and positively correlated with other constructs. Theoretically, most of these skills are related to cooperative learning. A correlation analysis among the constructs was conducted to test if the above correlational structure is supported in the measurement model. The correlations among the constructs are presented in Table 42.

Table 38. Correlation coefficients among the constructs for nomological validity

	Construct	1	2	3	4	5	6
1	Mathematical creativity (1)	-					
2	Mathematical perseverance (2)	-0.007	-				
3	Cooperative learning in mathematics (3)	0.017	0.196**	-			
4	Mathematics enjoyment (4)	0.031	0.291**	0.241**	-		
5	Mathematics anxiety (5)	0.044	0.017	-0.032	0.067	-	
6	Mathematical self-efficacy (6)	-0.024	0.202	0.248	0.193	0.223	-

According to Table 42, cooperative learning in mathematics was positively correlated with mathematical creativity and mathematical self-efficacy, mathematical perseverance, and mathematics enjoyment and was negatively correlated with mathematics anxiety. As the correlations among the constructs are consistent with the theoretical foundation, this can be regarded as evidence for the establishment of nomological validity.

6.3.2.4 Results of the criterion validity analysis

One way to determine criterion validity is to pose the same question in varying ways or repeat the question at the next phase in the questionnaire to assess consistency in the responses (Harvey, 2020). This approach is suitable for the current study to test criterion validity, as the vignette sets were developed based on the existing questionnaires and posed differently. Pacheco (2019) applied a similar technique using Ordinary Least Squares regression (OLS) to test response consistency in the vignette analysis. In case a set of skills is to be measured on the same scale, it is optional for researchers to fuse them to be made into a single measure. While it is possible to consider each rating as a separate dependent variable, another option is to converge them into one dependent variable by calculating their mean. However, it is crucial to ensure correlation between individual dependent variables by calculating an internal consistency measure. In case of no correlation, combining them into a single measure is neither required nor plausible. Internal consistency for each construct was well above the minimum acceptable value ($\omega > 0.7$). Therefore, like Pacheco (2019), the dependent variable was created by combing all responses to the vignettes. In contrast, the independent variable was created by combining all responses to the questionnaires. For each scale, a higher value indicates higher social and emotional skills. The regression analysis was performed and revealed that vignette ratings are positively and significantly associated with the questionnaire responses (Table 39).

Table 39. OLS regression analysis summary for questionnaire score predicting the vignette score

Variable	B	95% CI	β	t	p
(Constant)	2.788	[2.593 – 2.983]	0.014	28.141	0.000***
Mean score for questionnaire	0.056	[0.001 – 0.112]		2.007	0.046*

Note: R^2 adjusted = 0.010. *** indicates $p < 0.001$, * indicates $p < 0.05$. CI = confidence interval for B.

In another way, criterion validity examines if the new measure is parallel with what is previously known (Harvey, 2020). In this regard, two relevant papers (Primi et al., 2016; Weiss and Roberts, 2018) that applied AVs to explore social and emotional skills were compared. These studies found that AVs increase internal consistency to measure social and emotional skills in the general domain. Particularly, Primi et al. (2016) reported that AVs positively affect reliability in exploring Big Five Openness and Big Five Conscientiousness. Similarly, Weiss and Roberts (2018) found that AVs increased reliability to adjust self-reported personality measures. The present study found that internal consistency increased after AVs adjusted the questionnaires which is consistent with what the previous studies have already addressed (see Table 40).

Table 40. The effect of AVs on reliability in three different studies

	Primi et al. (2016)	Weiss and Roberts (2018)	The present study
Sample	Brasil	Rwanda	Mongolia
A measure	Cronbach's alpha	McDonald's omega	McDonald's omega
Effect	(Openness) From 0.83 to 0.91 (Conscientiousness) From 0.87 to 0.95	(Openness) From 0.66 to 0.91 (Conscientiousness) From 0.74 to 0.92	(Mathematical perseverance) From 0.69 to 0.83 (Cooperative learning in math) From 0.65 to 0.81

Besides AVs effect on reliability, a previous study also explored their effect on validity. Primi et al. (2016) found that when vignettes corrected the questionnaire responses, the correlation between the constructs (Openness and Conscientiousness) dropped from 0.36 to 0.18. Therefore, a substantial reduction in collinearity leads to improvement in discriminant validity (Table). Similarly, the present study found that the vignette corrected scale had the higher square root of AVE, which is well above the correlation between the constructs, showing higher discriminant validity (Table 41).

Table 41. The correlation coefficient between two constructs (Primi et al., 2016)

Constructs	Original scale		AVs adjusted scale	
	A	**B**	**A**	**B**
Openness **(A)**	**--**	0.36	**--**	0.18
Conscientiousness **(B)**	0.36		0.18	

Table 42. Correlation coefficient and AVE matrix of the two constructs (Present study)

Constructs	Original scale		AVs adjusted scale	
	A	**B**	**A**	**B**
Mathematical self-efficacy **(A)**	**0.606**	**--**	**0.758**	**--**
Mathematical enjoyment **(B)**	0.645	**0.784**	0.214	**0.858**

In sum, the results of OLS regression analysis and literature review provided evidence for establishing criterion validity of the new measures in this study.

6.3.3 Results of the analysis of vignette-corrected new scales and problem-posing tasks (RQ3)

RQ3: What is the status of social and emotional skills among Mongolian students?

This section illustrates the analysis of students' performance on the vignette-corrected new scale and problem-posing tasks. General performance and significant mean differences in mathematical creativity, cooperative learning in mathematics, mathematical perseverance, mathematics anxiety, mathematics enjoyment, and mathematical self-efficacy were examined by students' gender, ethnicity, and region, using descriptive statistics and independent samples t-test.

6.3.3.1 Results of the analysis of vignette corrected scales

First, the vignette-corrected new scales were analysed for all scales. Figure 12 illustrates the distribution of students' performance in cooperative learning in mathematics on three vignettes. As illustrated in the graph, the distribution tendency is the highest between the medium and high vignettes. This implies that most student performance in cooperative learning in mathematics lies between the medium and high vignettes. The aggregate mean performance for cooperative learning is 4.47, which is above the medium vignette (Figure 13).

Mean	4.47
SD	1.49
Min	1
Max	7

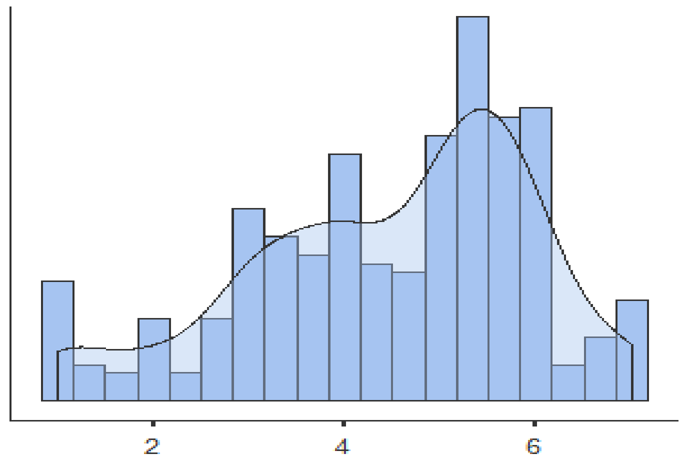


Figure 13. Distribution of the vignettes for cooperative learning in mathematics

Notes: 2 = low vignette; 4 = medium vignette; 6 = high vignette

Figure 14 illustrates the distribution of students' performance in mathematical perseverance on three vignettes. As illustrated in the graph, the distribution tendency is the highest on the high vignette and mainly lies between the medium and high vignettes. This implies that most of the students had high mathematical perseverance. The aggregate mean performance for mathematical perseverance was 4.44, which is above the medium vignette.

Mean	4.44
SD	1.48
Min	1
Max	7

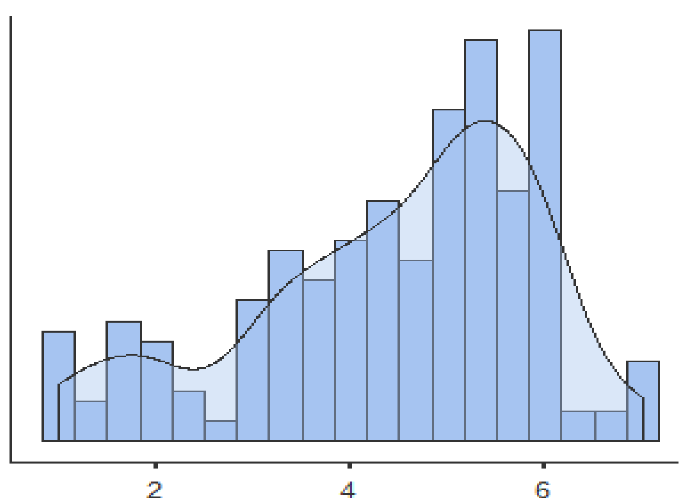


Figure 14. Distribution of the vignettes for mathematical perseverance

Notes: 2 = low vignette; 4 = medium vignette; 6 = high vignette

Figure 15 illustrates students' mathematics anxiety on three vignettes. Unlike other skills, the vignette set for mathematics anxiety had negative statements. For factor analysis suitability, the students' responses were reverse-coded. As illustrated in the graph, the distribution tendency was the highest between the medium and high vignettes. The aggregate mean performance for mathematics anxiety was 5, which was above the medium vignette. This implies that most students had higher mathematics anxiety than the hypothetical individuals described in the medium vignette did.

Mean	5
SD	1.34
Min	1
Max	7

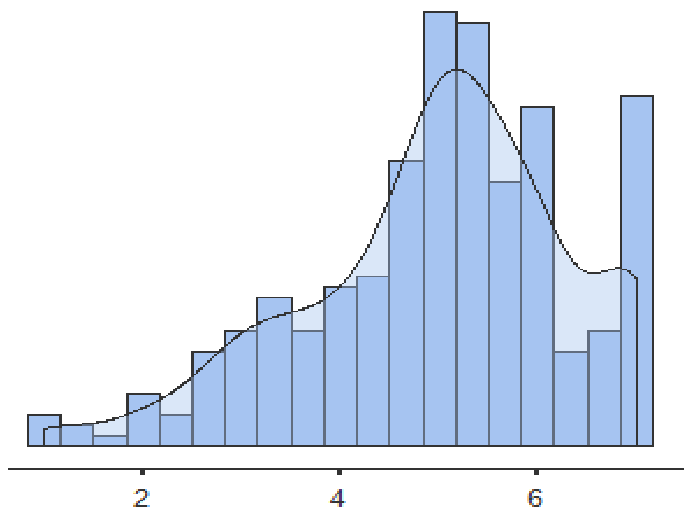


Figure 15. Distribution of the vignettes for mathematical anxiety

Notes: 2 = low vignette; 4 = medium vignette; 6 = high vignette.

Figure 16 illustrates students' performance in mathematics enjoyment on three vignettes. As illustrated in the graph, the distribution tendency was the highest on the high vignette. The aggregate mean performance for mathematics enjoyment was 4.68, which was above the medium vignette. This implies that most of the students had higher mathematics enjoyment than the hypothetical individuals described in the medium vignette.

Mean	4.68
SD	1.34
Min	1
Max	7

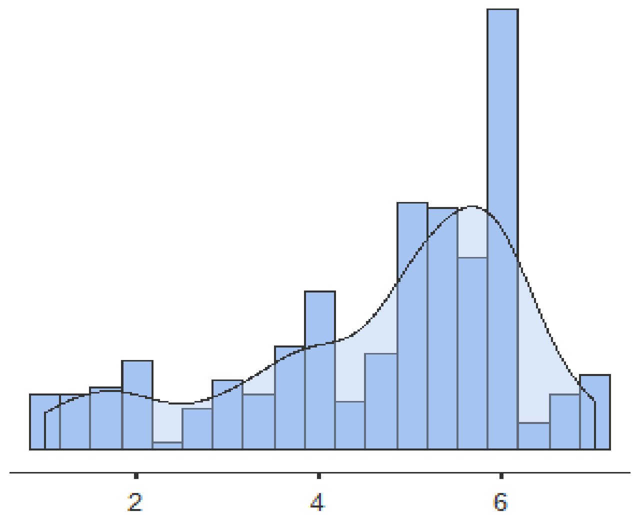


Figure 16. Distribution of the vignettes for mathematical enjoyment

Notes: 2 = low vignette; 4 = medium vignette; 6 = high vignette

Figure 17 illustrates the distribution of students' performance in mathematical self-efficacy on three vignettes. As illustrated in the graph, the distribution tendency was the highest between the medium and high vignettes. The aggregate mean performance for mathematical self-efficacy was 4.55, which was above the medium vignette. The finding implies that most students had better mathematical perseverance than the hypothetical individuals described in the medium vignette did.

Mean	4.55
SD	1.45
Min	1
Max	7

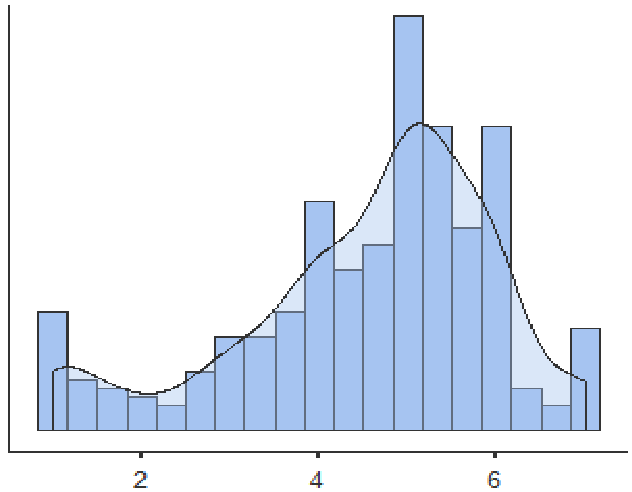


Figure 17. Distribution of the vignettes for mathematical self-efficacy

Notes: 2 = low vignette; 4 = medium vignette; 6 = high vignette

Next, students' performance based on gender, ethnicity, and region skills was analysed. Following the criteria outlined by Cohen (1988), effect sizes[1] in t-test were computed: $d < 0.20$ = insignificant effect; $d \geq 0.20$ and $d < 0.50$ = minor effect; $d \geq 0.50$ and $d < 0.80$ = moderate effect; and $d \geq 0.80$ = considerable effect. Cooperative learning in mathematics was the first skill to be evaluated. The analysis indicated a statistically significant difference between males and females, with females having a higher mean value (4.68) than males (4.26), which demonstrates that the females had higher cooperative learning in mathematics (Table 43).

---

[1] The formula to calculate the effect size is $\text{Cohen's d} = \frac{m_A - m_B}{SD_{pooled}}$

where,

m_A and m_B depict the mean value of the group A and B, respectively.

SD_{pooled} estimates the pooled standard deviation of both groups.

Table 43. Independent sample t-test results of Cooperative learning in mathematics by gender, ethnicity, and region

Group	n	Mean	SD	Standard Error (SE)	df	t	p-value	Cohen's d
Male	151	4.26	1.61	0.13	306	-2.46	0.01*	0.28
Female	157	4.68	1.34	0.10				
Mongol	271	4.45	1.51	0.09	306	-0.76	0.44	0.14
Kazakh	37	4.65	1.30	0.26				
Urban	116	4.42	1.47	0.13	306	-0.47	0.63	0.06
Rural	192	4.51	1.50	0.10				

There was no statistically significant difference between the students' performance in cooperative learning in mathematics based on ethnicity; however, Kazakh students had a slightly higher mean value (4.65) than Mongol students did (4.45).

Regarding regions, again, there was no statistically significant difference between the cooperative learning of urban and rural students; however, students in rural areas had a higher mean value (4.51) than that of the students in urban areas (4.42).

In terms of mathematical perseverance, the analysis indicated a statistically significant difference between males and females, with females having a higher mean value (4.61) than males did (4.26), which demonstrates that the females had higher mathematical perseverance than males did (Table 44).

Table 44. Independent sample t-test results of Mathematical perseverance by gender, ethnicity, and region

Group	n	Mean	SD	Standard Error (SE)	df	t	p-value	Cohen's d
Male	151	4.26	1.53	0.12	306	-2.05	0.04*	0.23
Female	157	4.61	1.41	0.11				
Mongol	271	4.44	1.51	0.09	306	0.01	0.44	0.00
Kazakh	37	4.43	1.31	0.21				
Urban	116	4.22	1.58	0.14	306	-0.19	0.04*	0.23
Rural	192	4.57	1.41	0.10				

There was no statistically significant difference between mathematical perseverance based on ethnicity; however, Mongolian students had a slightly higher mean value (4.44) than Kazakh students did (4.43).

About regions, there was a statistically significant difference between the mathematical perseverance of urban and rural students, with students in rural areas having a higher mean value (4.57) than students in urban areas did (4.22).

The mathematical enjoyment analysis indicated a statistically insignificant difference between males and females, with both genders having the same mean value (4.68) (Table 45).

Table 45. Independent sample t-test results of Mathematical enjoyment by gender, ethnicity, and region

Group	n	Mean	SD	Standard Error (SE)	df	t	p-value	Cohen's d
Male	151	4.68	1.57	0.12	306	0.01	0.98	0.00
Female	157	4.68	1.52	0.12				
Mongol	271	4.67	1.56	0.09	306	-0.31	0.75	0.05

Kazakh	37	4.76	1.44	0.23				
Urban	116	4.76	1.54	0.14	306	0.71	0.47	0.08
Rural	192	4.63	1.55	0.11				

Regarding mathematical enjoyment based on ethnicity, there was no statistically significant difference between Mongolian and Kazakh students; however, Kazakh students had a slightly higher mean value (4.76) than Mongolian students did (4.67).

Mathematical enjoyment between students in rural areas and those in urban areas was also statistically insignificant; however, students in urban areas had a higher mean value (4.76) than those in rural areas (4.63).

The mathematical anxiety analysis indicated a statistically insignificant difference between males and females, with males having a higher mean value (5.11) than females did (4.87) (Table 46).

Table 46. Independent sample t-test results of Mathematical anxiety by gender, ethnicity, and region

Group	n	Mean	SD	Standard Error (SE)	df	t	p-value	Cohen's d
Male	151	5.11	1.30	0.10	306	1.57	0.11	0.17
Female	157	4.87	1.38	0.11				
Mongol	271	5.03	1.36	0.08	306	1.26	0.20	0.23
Kazakh	37	4.73	1.22	0.20				
Urban	116	5.07	1.47	0.13	306	0.76	0.44	0.09
Rural	192	4.94	1.26	0.09				

Regarding ethnicity, there was no statistically significant difference between the mathematics anxiety of Mongolian and Kazakh students; however, Mongolian students had a slightly higher mean value (5.03) than Kazakh students did (4.73). Mathematical anxiety between students in rural areas and those in urban areas was also statistically insignificant; however, students in urban areas had a higher mean value (5.07) than those in rural areas did (4.94).

The mathematical self-efficacy analysis indicated a statistically insignificant difference between males and females; however, females had a slightly higher mean value (4.67) than males did (4.43) (Table 47).

Table 47. Independent sample t-test results of Mathematical self-efficacy by gender, ethnicity, and region

Group	n	Mean	SD	Standard Error (SE)	df	t	p-value	Cohen's d
Male	151	4.43	1.55	0.12	306	-1.44	0.15	0.16
Female	157	4.67	1.34	0.10				
Mongol	271	4.59	1.47	0.08	306	1.15	0.25	0.20
Kazakh	37	4.30	1.34	0.22				
Urban	116	4.52	1.55	0.14	306	-0.31	0.75	0.23
Rural	192	4.57	1.39	0.10				

Regarding ethnicity, there was no statistically significant difference between the mathematical self-efficacy of Mongolian and Kazakh students; however, Mongolian students had a slightly higher mean value (4.59) than Kazakh students did (4.30). Mathematical self-efficacy between students in rural and urban areas was also statistically insignificant; however, students in rural areas had a higher mean value (4.57) than those in urban areas did (4.52).

6.3.3.2 Results of analysis of problem-posing tasks

In this section, students' performance on problem-posing tasks was analysed. First, the total number of responses and different responses provided by the students on the four tasks were analysed (Table 48).

Table 48. Total number of responses and different types of responses

Task	Total number of responses	Number of different types of responses
1	1,020	8
2	646	6
3	632	8
4	396	4

Table 48 illustrates the total number of responses for each of the four mathematical tasks and the number of different types of responses. For Task 1, students provided 1020 responses, which can then be categorized into eight different types. For Task 2, students give 646 responses that can be categorized into six different types. Task 3 is a geometrical task with 632 responses that can be categorized into eight different types. Similarly, Task 4 is an algebraic task with a total of 396 responses and 4 different types. The types of responses for each task have been further elaborated below In Table 49, Table 50, Table 51, and Table 52.

Table 49. Types of responses for Task 1

Task	Types of response	Per cent
Task 1	Even & odd numbers (2; 36; 3; 7)	30.91
	Number of digits (3; 36; 119)	29.93
	Divisibility (numbers that are divided by 3; 5; 4; etc.)	15.31
	Three-digit numbers with a sum of its digits are smaller than one and two-digit numbers (121; 150)	4.51
	Prime & composite number (3; 7; 36; 60)	8.54
	Interval (0-10; 11-20; 61-120)	7.46
	Squared, cubed, and square root (9; 25; 36)	3.14
	Sequences (The numbers with special patterns: 3, 7, 9, 13, 15 which increases in following pattern +4 +2 + 4+2)	0.20

Table 50. Types of responses for Task 2

Task	Types of response	Per cent
Task 2	How many guests arrived when the door opened ‘n’ times?	82.4
	How many times would the door have been opened when ‘n’ number of guests arrived?	7.0
	If the number of guests increased three times at every doorbell ring, how many guests would arrive in total?	1.9
	How many more guests would arrive at the nth door open than the kth door open?	5.1
	Is it possible to come more than or less than ‘n’ number of people, if the door opens ‘k’ times?	1.9
	Is it possible that the number of guests can be a prime number? If possible, at which number of doorbell rings can the number of guests be a prime number?	1.7

Table 51. Types of responses for Task 3

Task	Types of response	Per cent
Task 3	Finding the perimeter of circles and triangles under given conditions.	28.16
	Finding the area of circles and triangles under given conditions.	15.82
	Finding the radius and diameter of a circle under given conditions.	22.47
	Finding the area of the triangle, which is formed after the area of the circle is taken.	4.75
	Finding angles of the triangle under given conditions.	11.87
	Finding cathetus and hypotenuse of the triangle under given conditions.	14.24
	Finding inradius and bisector of the triangle under given conditions.	1.74
	Proof (proving that the heights of the triangle intersect each other at the centroid of the inscribed circle.)	0.95

Table 52. Types of responses for Task 4

Task	Types of response	Per cent
Task 4	Finding two unknowns when its sum is given. (A man is twice as old as his younger brother is. If the sum of their age is 30, how old are they? etc.)	82.07
	Finding one unknown when another variable is known. (e.g., A pen is three times expensive than a pencil. If a pencil is 150 tugriks, how much is a pen?)	11.62
	Providing more information (E.g., Nara & Sara are twins. They were not born in the 21st century but are not very old. This year, the sum of their age is equal to their mother's age. Their father is 50, and he is two years older than their mother is. How old is Sara?)	4.04
	More than two unknowns. (If the lengths of a triangle are in the ratio 3:2:1, and perimeter is 36 cm. Please find the length of the sides.)	2.27

As mentioned earlier, the problem-posing tasks have been adapted from previous studies. The direct comparison of students' performance is not entirely possible due to various considerations such as age, grade, level of mathematical knowledge, and different raters. However, it is still interesting to see students' performance from different countries on the same mathematical problem-posing tasks. For Task 1 and Task 4, it was impossible to extract the required information about student performance from previous studies as the original study used the tasks to devise a scale for measuring mathematical creativity without providing any in-depth information about a student's performance on individual performance tasks. However, information on total responses and different responses for Task 2 and Task 3 was obtained from a previous study (Sriraman and Lee, 2011). Based on the data obtained and the data analysis of the present study, Table 53 presents the total responses and the different types of student responses from the US, China, and Mongolia.

Table 53. Comparison between the US, China and Mongolia for Task 2 and Task 3

Task	US		China		Mongolia	
	Total responses	Response type	Total responses	Response type	Total responses	Response type
Task 2	91	12	207	13	646	6
Task 3	106	8	200	10	632	8
Grade	Grade 11 & Grade 12		Grade 12		Grade 9	
Sample	30		55		308	

Participants from China were Grade 12 students from one of the top five high schools in Jiaozhou, while participants from the US were from a laboratory high school of Illinois State University, both considered at an advanced level in mathematics (Sriraman and Lee, 2011). In contrast, participants from Mongolia were Grade 9 students at traditional public schools in rural and urban areas. Table 53. illustrates that the sample size for the US was 30, 55 for China, and 308 for Mongolia. For Task 2, students from the US gave 91 responses, with the responses being categorized into 12 types. Similarly, students from China provided 207 responses to Task 2, with their responses being categorized into 13 types. As for the Mongolian students, the total number of responses was 646, with six different types. For Task 3, US students provided 106 responses from eight different types, Chinese students provided 200 responses from ten different types. In comparison, Mongolian students provided 632 responses from eight different types.

As evident from Table 53., the differences between sample size, age, and mathematical knowledge have been considered. Notably, students from Mongolia posed 646 responses for Task 2 and 632 responses for Task 3, which is the highest number of responses. However, the underlying reason is the large sample size of Mongolian students. The results would have favoured the Chinese and US students if the ratio between the number of students and total responses was considered. This is because the students included in the sample for the US and China are high school students with advanced knowledge of mathematics. Considering the different nature of the three samples, it was impossible to discuss the initial responses; however, the analysis provided an interesting result. According to Sriraman and Lee (2011), for Task 3, i.e., the geometry content task, Chinese students provided a proof type of response (5 per cent) while US students did not provide this type of response. Interestingly, a proof type of response was found in the Mongolian sample; however, the percentage was very small (0.95 per cent). It is also important to highlight that the US, China, and Mongolia analyses are meant to provide a rough picture of students' performance on different tasks rather than draw a comparison between them.

The next analysis focused on fluency, flexibility, and originality for each problem-posing task. Fluency score refers to the number of correct mathematical responses given by a student. A flexibility score refers to the number of different types of responses given by a student. Similarly, originality refers to the unique responses given by a student in the sample. The study analysed the ability of students to present correct problems to given tasks, that is, fluency. Figure 18 illustrates the distribution of fluency scores achieved by students for Task 1. As evident from the graph, 25.3 per cent of the students were able to produce up to three correct responses, 20.9 per cent provided two correct responses, 16.5 per cent gave five correct responses, 11.2 per cent gave four correct responses, 6.4 per cent gave one correct response, 5.6 per cent gave seven correct responses, and 5.2 gave six correct responses. The number of correct responses between 8 and 15 was given by fewer students, with only 0.4 per cent giving 10, 14, and 15 correct responses. The total number of responses given by students for Task 1 was 1,020 (as illustrated in Table 48). The mean value for fluency was 4.09, which means that the students provided four responses to this task on average. Hence, only a few students could give a higher number of correct responses.

Mean	4.09
SD	2.43
Mode	3
Min	1
Max	15

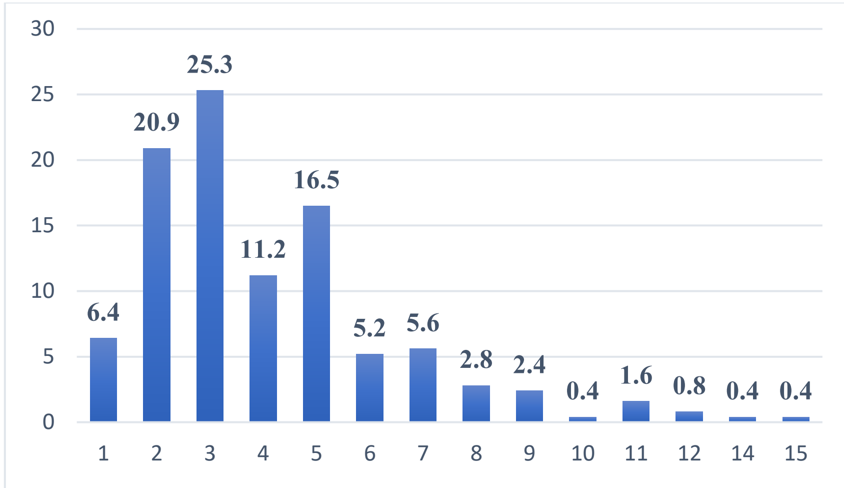


Figure 18. Distribution of fluency score for Task 1

Next, the study examined the different types of problems identified by students, that is, flexibility. Figure 19 illustrates the distribution of the flexibility scores attained by students for Task 1. The graph illustrates that most students (41.1 per cent) presented only one type of problem, 33.9 per cent of students gave two types of problems, 16.5 per cent presented three types of problems, 6.9 per cent presented four types of problems, and only 1.2 per cent and 0.4 per cent of students presented five and seven types of problems. The responses given by the students for Task 1 can be classified into eight types (Table 49). The mean value for flexibility in Task 1 was 1.94, which means that the students gave two different responses to the task. Hence, the lack of students' ability to produce different responses is evident from the findings.

Mean	1.94
SD	1.03
Mode	1
Min	1
Max	7

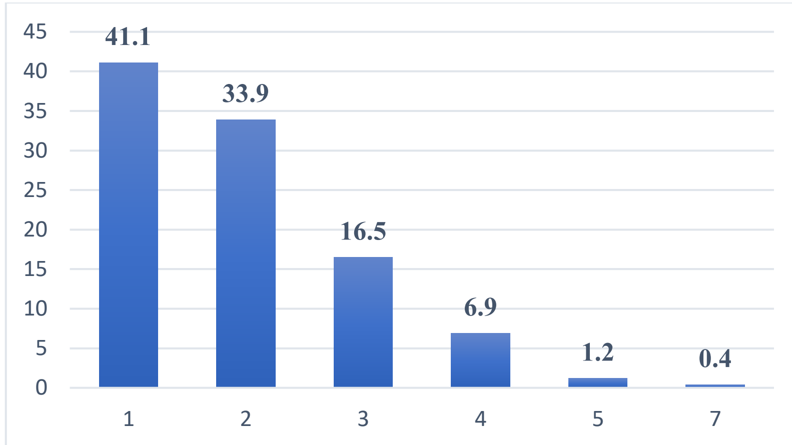


Figure 19. Distribution of flexibility score for Task 1

Finally, the uniqueness of responses given by the students was examined, that is, originality. Figure 20 illustrates the distribution of the originality scores attained by the students for Task 1, demonstrated that most students (54.2 per cent) were not able to present any unique response, attaining a score of 0, which means that 257–445 students gave the same response. Similarly, 8.4 per cent of the students gave one unique response (meaning that 129–256 students gave the same response); 24 per cent of the students gave two unique responses, demonstrating that 65–128 students gave the same response; 4.9 per cent of the students gave three unique responses, demonstrating that 33–64 students gave the same response; 7.8 per cent of the students gave four unique responses, demonstrating that 17–32 students gave the same response; and only 0.6 per cent of the students gave eight unique responses, demonstrating the rarest response in the sample (only two students gave the same response). The mean value for originality in Task 1 was 1.07, which illustrates that most students (ranging between 129 and 246) in the sample gave a similar response. Therefore, the findings indicate that most students gave similar responses, which demonstrates a lack of originality.

Mean	1.07
SD	1.41
Mode	0
Min	0
Max	8

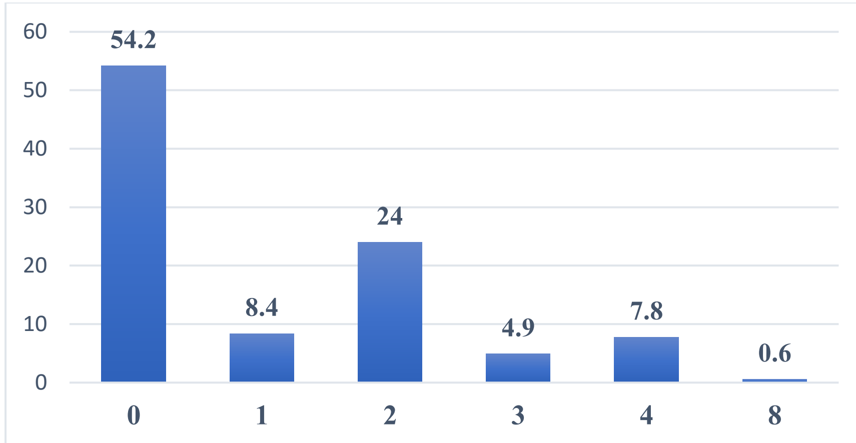


Figure 20. Distribution of originality score for Task 1

Task 2, again, analysed the ability of students to respond with fluency, flexibility, and originality. Figure 21 illustrates the distribution of fluency scores achieved by the students for Task 2. The total number of responses given by students for Task 2 was 646 (as illustrated in Table 48). The mean value for fluency was 3.3, which means that the students provided an average of three responses to this task. Hence, only a few students could give a high number of correct responses.

Mean	3.3
SD	2.66
Mode	2
Min	1
Max	17

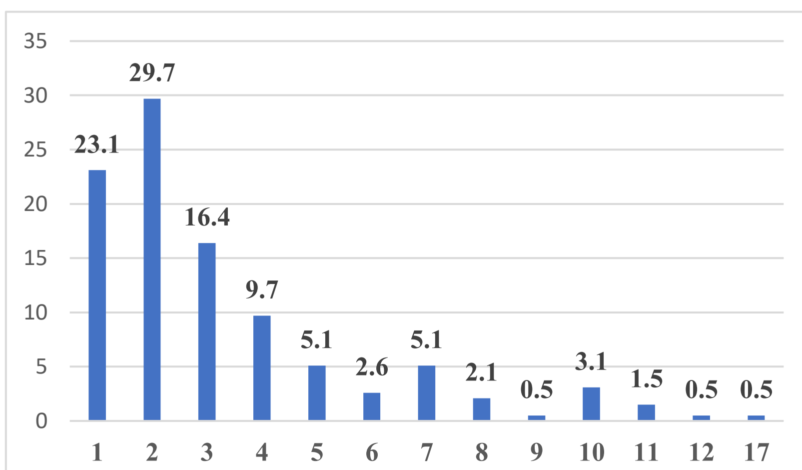


Figure 21. Distribution of fluency score for Task 2

Figure 22 illustrates the distribution of the flexibility score attained by students for Task 2. The graph illustrates that most students (60 per cent) presented only one type of problem, 35.4 per cent presented two types of problems, and only 4.6 per cent presented three types of problems. The responses given by students for Task 2 can be classified into six types (as illustrated in Table 50). The mean value for flexibility in Task 2 was 1.43, which means that students gave one different response to the task. Hence, the lack of students' ability to produce different responses is evident from the findings.

Mean	1.43
SD	0.59
Mode	1
Min	1
Max	3

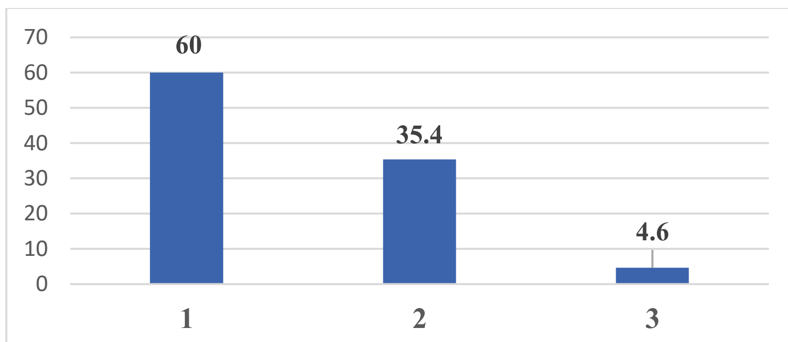


Figure 22. Distribution of flexibility score for Task 2

Figure 23 illustrates the distribution of the originality score attained by the students for Task 2. The graph illustrates that most students (73.7 per cent) could not present any unique response, attaining a score of 0, which means that 257–445 students in the sample responded similarly. Moreover, 9.1 per cent of the students attained a score of 3, meaning that 33–64 students gave the same response; 7.5 per cent of the students attained a score of 4, demonstrating that 17–34 students gave the same response; and 9.7 per cent of the students attained a score of 5, demonstrating that 9–16 students gave the same response. The mean value for originality in Task 2 was 1.05, demonstrating that most students (ranging between 129 and 246) in the sample gave a similar response. Hence, the findings indicate a lack of originality in the students.

Mean	1.05
SD	1.82
Mode	0
Min	0
Max	5

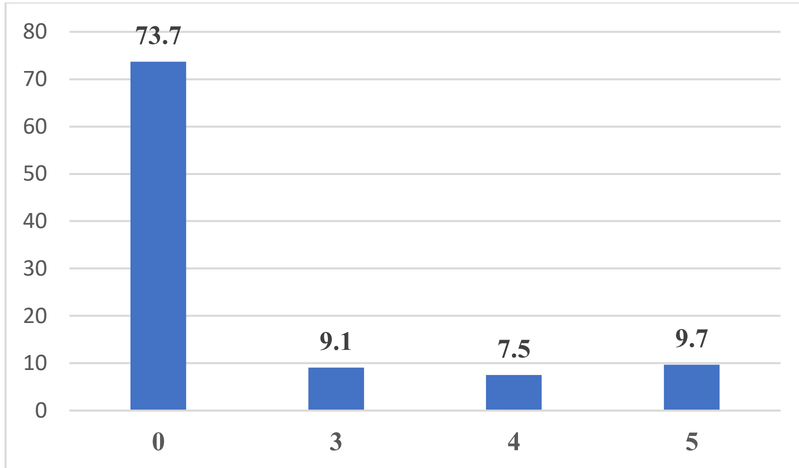


Figure 23. Distribution of originality score for Task 2

Figure 24 illustrates the distribution of fluency scores achieved by students for Task 3. The total number of responses given by the students for Task 3 was 632 (as illustrated in Table 48). The mean value for fluency was 2.8, which means that the students provided an average of three responses to this task. Hence, only a few students were able to give a high number of correct responses.

Mean	2.8
SD	1.59
Mode	2
Min	1
Max	8

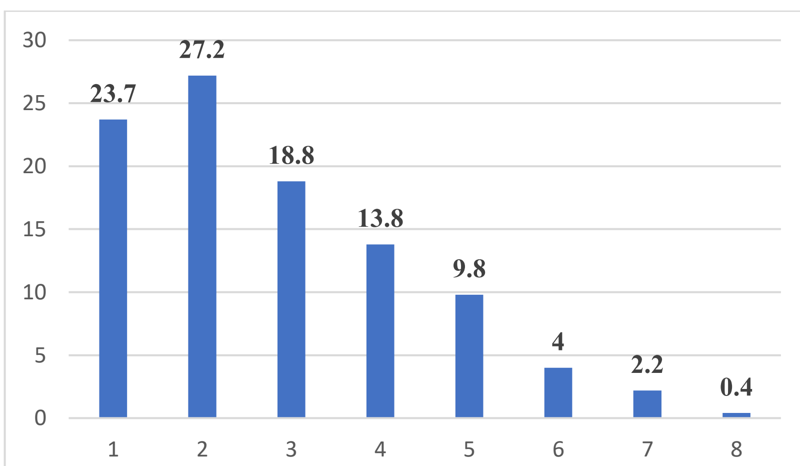


Figure 24. Distribution of fluency score for Task 3

Figure 25 illustrates the distribution of the flexibility score attained by the students for Task 3. The responses given by the students for Task 1 can be classified into eight types (Table 51). The mean value for flexibility in Task 3 was 2.25, which means that students gave two different responses to the task. Hence, the lack of students' ability to produce different responses is evident from the findings.

Mean	2.25
SD	1.07
Mode	2
Min	1
Max	6

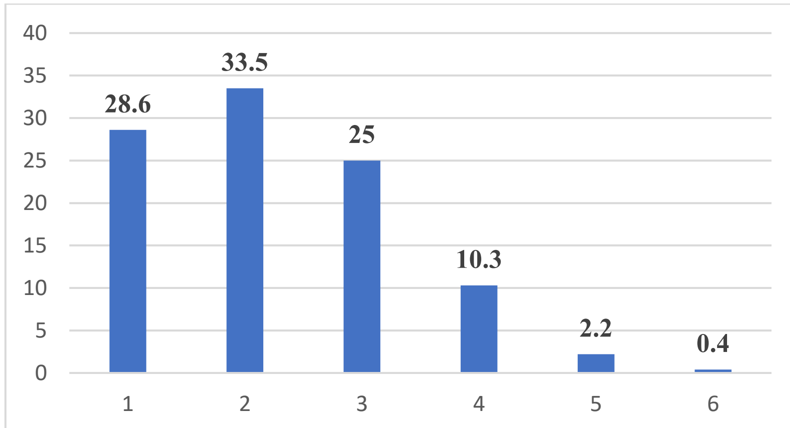


Figure 25. Distribution of flexibility score for Task 3

Figure 26 illustrates the distribution of the originality score attained by the students for Task 3. The graph illustrates that most students (45.1 per cent) attained a score of 2, demonstrating that 65–128 students in the sample gave the same response. Moreover, 27.3 per cent of the students could not present any unique response, attaining a score of 0, which means that 257–445 students responded similarly. Additionally, 14 per cent of the students achieved a score of 1, implying that 129–445 students gave the same response; 8.8 per cent of the students attained a score of 4, demonstrating that 17–32 or more students gave the same response; 2.9 per cent of the students attained a score of 5, implying that 9–16 or more students gave the same response; and 1.9 per cent students attained a score of 6, implying that it was the rarest response in the sample as only 5–8 students gave a similar response. The mean value for originality in Task 3 was 1.65, which demonstrates that most students (ranging between 129 and 246) in the sample gave a similar response. Hence, the findings indicate a lack of originality.

Mean	1.65
SD	1.42
Mode	2
Min	0
Max	6

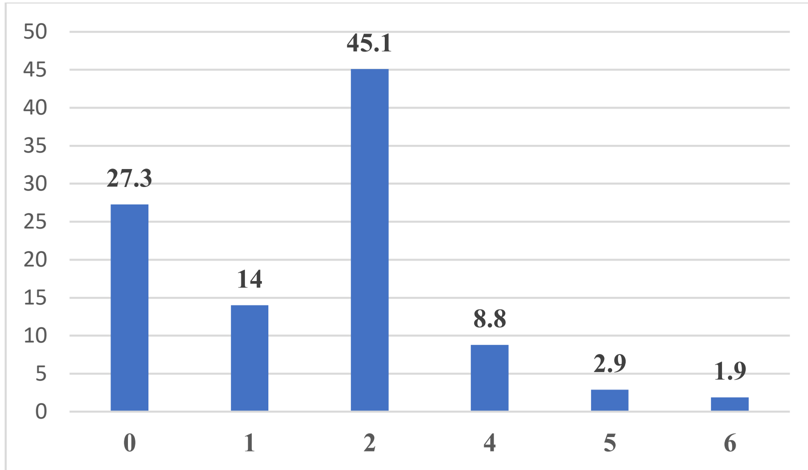


Figure 26. Distribution of originality score for Task 3

Figure 27 illustrates the distribution of the fluency score achieved by the students for Task 4. As evident from the graph, 40.5 per cent of the students gave only one correct response, 28.9 per cent provided two correct responses, 21.1 per cent gave three correct responses, 4.7 per cent gave four correct responses, 2.6 per cent gave five correct responses, 1.6 per cent gave six correct responses, while only 0.5 per cent were able to give ten correct responses. The total number of responses given by students for Task 4 was 396 (as illustrated in Table 48). The mean value for fluency was 2.08, which means that the students provided an average of two responses to this task. Hence, only a few students were able to give a high number of correct responses.

Mean	2.08
SD	1.27
Mode	2
Min	1
Max	10

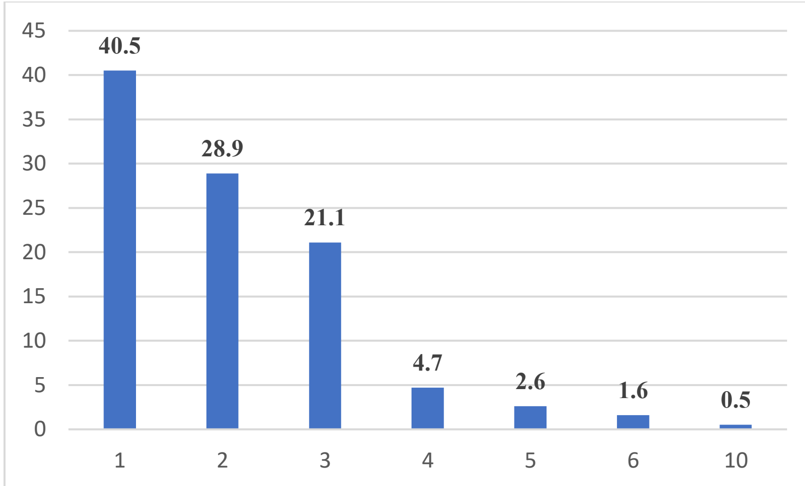


Figure 27. Distribution of fluency score for Task 4

Figure 28 illustrates the distribution of the flexibility score attained by the students for Task 4. The graph illustrates that most students (78.4 per cent) were able to present only one type of problem, 17.4 per cent of the students presented two types of problems, while 3.7 per cent and 0.5 per cent of the students presented three and four types of problems, respectively. The responses given by the students for Task 4 can be classified into four types (Table 52). The mean value for flexibility in Task 3 was 1.25, which means that students gave one type of response to the task. Hence, the lack of students' ability to produce different responses is evident from the findings.

Mean	1.25
SD	0.55
Mode	1
Min	1
Max	4

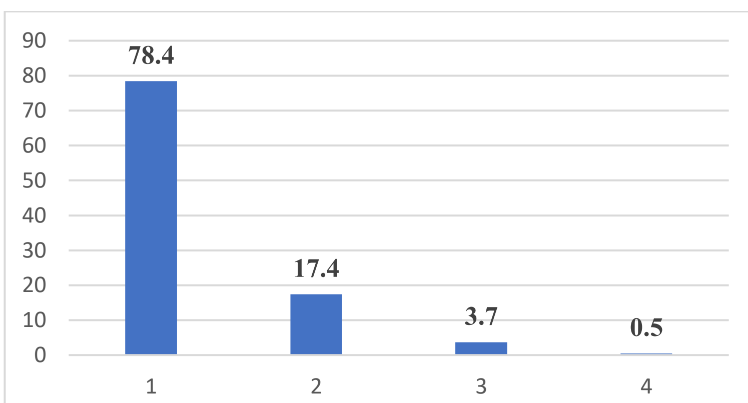


Figure 28. Distribution of flexibility score for Task 4

Figure 29 illustrates the distribution of the originality score attained by students for Task 4. The graph illustrates that most students (80.8 per cent) attained a score of 1, demonstrating that 129–445 students gave the same response. Similarly, 12.7 per cent of the students attained a score of 2, implying that 65–128 students gave the same response, and 6.5 per cent of the students attained a score of 3, demonstrating that 33–64 students gave the same response. The mean value for originality in Task 4 was 0.70, demonstrating that most students (ranging between 129 and 246) in the sample gave a similar response. Hence, the findings demonstrate students' lack of originality.

Mean	0.70
SD	1.50
Mode	0
Min	0
Max	5

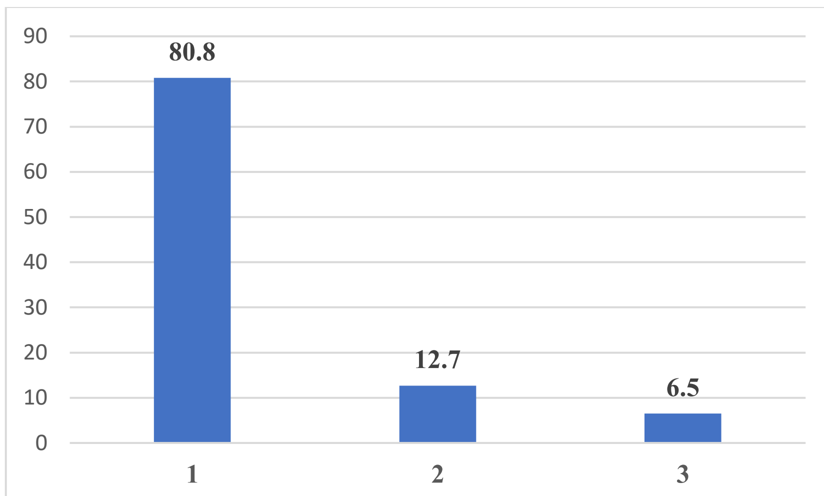


Figure 29. Distribution of originality score for Task 4

The overall performance is presented in Table 54., and as depicted, the mean value for fluency is 3.36, which means that the students provided three responses to the tasks on average. Hence, it is evident that only a few students could provide a high number of correct responses. As for flexibility, the mean value was 1.86, which means that students provided one different response to the task on average. This demonstrates the students' lack of ability to come up with different types of responses. As for originality, the mean value was 1.12, which means that most of the students (i.e., between 129 and 246) in the sample gave a similar response. This demonstrates a lack of originality.

Table 54. General performance on problem-posing tasks

	Fluency		Flexibility		Originality	
N = 308	Mean	SD	Mean	SD	Mean	SD
Performance	3.36	1.37	1.86	0.52	1.12	0.99

Therefore, the overall results demonstrate that the students' performance lacked fluency, flexibility, and originality while doing problem-posing tasks. Thus, it can be concluded that the ninth-grade Mongolian students lacked mathematical creativity.

Moreover, in terms of mathematical creativity, there exists a statistically significant difference between males and females, with females having a higher mean value (5.29) than males do (4.49), which demonstrates that females are more creative in mathematics (Table 55). This finding parallels some previous studies (Evans, 1964; Jensen, 1973; Prouse, 1967).

Table 55. Independent sample t-test results of Mathematical creativity by gender, ethnicity, and region

Group	n	Mean	SD	SE	df	t	p-value	Cohen's d
Male	151	4.49	2.79	0.23	306	2.33	0.02*	0.27
Female	157	5.29	2.97	0.24				
Mongol	271	5.09	2.95	0.30	306	-3.59	0.00*	0.89
Kazakh	37	2.98	1.55	0.18				
Urban	116	5.69	2.89	0.27	306	3.69	0.00*	0.44
Rural	192	4.41	2.82	0.21				

The ethnicity of student performance in mathematical creativity was also statistically significant. Mongol students had a higher mean value (5.09) than Kazakh students (2.98), implying that Mongol students are more creative in mathematics than Kazakh students are (Table 55).

Regarding the distribution across regions, it is evident that students in urban areas had higher mathematical creativity, with a mean value of 5.69 than those in the rural areas, whose mean value was 4.41 (Table 55).

As the values were highly statistically significant, it can be concluded that urban students are more mathematically creative than rural students are. The underlying reason might be that the curriculum and the teaching strategies for mathematical creativity in classrooms in rural areas lag behind. Finally, correlation among the total scores for fluency, flexibility, originality, and total mathematical creativity was analysed. Table 56 illustrates the correlation between the four fluency scores: MC1, MC2, MC3, and MC4. A statistically significant correlation can be observed between the fluency scores of each task, which implies that the students' ability to give correct responses is not affected by the content of the problem-posing task. In other words, students who posed more problems in the geometry task (Task 3) also posed more problems on the algebra task (Task 4).

Table 56. Correlation among fluency scores

	MC1 Fluency	MC2 Fluency	MC3 Fluency	MC4 Fluency
MC1_Fluency	1	.430**	.353**	.233**
MC2_Fluency	.430**	1	.350**	.293**
MC3_Fluency	.353**	.350**	1	.414**
MC4_Fluency	.233**	.293**	.414**	1
**. Correlation is significant at the 0.01 level (2-tailed).				

Table 57. depicts the correlation between the four flexibility scores. As evident from the Pearson correlation results, there is a statistically significant correlation between the flexibility scores for each task, indicating that the content of the problem-posing tasks does not influence the students' ability to provide different responses. This demonstrates that students who provided several different types of problems for the geometry task (Task 3) also provided different problems for the algebra task (Task 4).

Table 57. Correlation among flexibility scores

	MC1 Flexibility	MC2 Flexibility	MC3 Flexibility	MC4 Flexibility
MC1_Flexibility	1	.425**	.322**	.308**
MC2_Flexibility	.425**	1	.336**	.327**
MC3_Flexibility	.322**	.336**	1	.387**
MC4_Flexibility	.308**	.327**	.387**	1
**. Correlation is significant at the 0.01 level (2-tailed).				

Table 58. presents the correlation between the originality scores. The results demonstrate a statistically significant correlation between the originality scores for each task, indicating that the students' ability to provide unique problems remains unaffected by the content of the problem-posing tasks. This signifies that those students who gave more unique problems on the geometry task also provided many unique problems for the algebra task (Task 4).

Table 58. Correlation among originality scores

	MC1 Originality	MC2 Originality	MC3 Originality	MC4 Originality
MC1_Originality	1	.264**	.272**	.188**
MC2_Originality	.264**	1	.223**	.199**
MC3_Originality	.272**	.223**	1	.173**
MC4_Originality	.188**	.199**	.173**	1
**. Correlation is significant at the 0.01 level (2-tailed).				

Table 59. depicts a strong and significant correlation between total fluency, total flexibility, total originality, and total mathematical creativity. As evident from the results, the mathematical creativity score is highly correlated with the students' ability to pose correct (fluency) and different responses (flexibility).

Table 59. Correlation between fluency, flexibility, originality, and mathematical creativity

	Total fluency	Total flexibility	Total originality	Total Creativity
Total fluency	1	.853**	.717**	.948**
Total flexibility	.853**	1	.830**	.946**
Total originality	.717**	.830**	1	.890**
Total Creativity	.948**	.946**	.890**	1
**. Correlation is significant at the 0.01 level (2-tailed).				

CHAPTER 7

SUMMARY, CONTRIBUTIONS, AND RECOMMENDATIONS

This research emphasised the need for a theoretical framework and reliable and valid tool to identify and measure social and emotional skills in mathematics as a research problem. The study has three research objectives: 1) develop a theoretical framework to capture social and emotional skills in mathematics; 2) develop tools to measure social and emotional skills in mathematics; and 3) validate the framework and tools by examining the status of social and emotional skills in mathematics among Mongolian students. To achieve these research objectives, this study answered three research questions. This chapter summarises the main findings for each research question, the implications and limitations of the study, and recommendations for future research.

7.1 Summary of the study findings

RQ1: What framework can be used to capture social and emotional skills in mathematics?

First, a theoretically predicted framework was constructed based on a systematic review of existing theories and literature. To validate the theoretically predicted framework, a reflective measurement model was derived from indicator items using the EFA and CFA approaches. According to the factor structures provided by EFA, the items were classified into six independent factors without any cross loading, which is consistent with the structure of the theoretically predicted framework. Moreover, the factor structure was supported by Kaiser's method (Kaiser, 1974), and factor loading criteria (above 0.5) suggested by Hair et al. (2010). Next, a CFA was performed to validate the EFA outcome. Various model fit statistics were used for the CFA to assess whether the theoretical model fits the selected data. According to the CFA, all fit indices were within acceptable values, which implies that the theoretically predicted model was a good fit for the selected data ($\chi 2/df = 1.32$; NFI = 0.92; CFI = 0.98; TLI = 0.97; PNFI = 0.64; RMSEA = 0.03).

The findings demonstrated that cooperative learning in mathematics was negatively correlated with mathematics anxiety and positively correlated with other skills, as theoretically predicted in mathematics research. Additionally, this finding was consistent with the theoretical patterns of the Big Five personality model in the general domain, which demonstrated that the agreeableness domain was negatively correlated with the neuroticism (emotional stability) domain and positively correlated with other domains such as openness and intellect, conscientiousness, and extraversion. Hence, this demonstrates that the selected skills in this study represent their theoretically corresponding domain in relation to the Big Five personality traits.

Based on the EFA and CFA results, a theoretically predicted framework on social and emotional skills in mathematics was empirically confirmed for the first time in this study.

RQ2: What valid tools can measure social and emotional skills in mathematics reliably?

To test the reliability and validity of the tools, reliability and validity analyses were conducted in multiple ways. First, a reliability analysis was conducted in two stages. In the first stage, internal consistency was computed for the raw scales and the vignette-corrected scales to examine the effect of AVs on internal consistency. In this stage, internal consistency was measured using McDonald's ω coefficient. The analysis results indicated that McDonald's ω coefficient was below the threshold (0.7) for the raw scales and were above the threshold after vignette correction, which suggests that the AV approach has the potential to increase reliability. In the second stage, the internal consistency of the vignette-corrected new scales was assessed using CR in the measurement model after the factors were loaded on the items. With this approach, internal consistency is measured more precisely. The CR value for all scales fulfilled the suggested cut-off value (0.7–0.9). In sum, McDonald's ω, and the CR value of the instruments fulfilled the suggested criteria in the literature, demonstrating that the items used in this study were internally consistent and reliable.

Second, a validity analysis was conducted in various forms, including convergent, discriminant, nomological validity, and criterion validity. Convergent validity was assessed through factor loadings and CR values for each construct. The factor loadings were substantially significant and above the suggested criteria (>0.5), and the CR values were well above the threshold (>0.7), indicating that the items that measured the same latent construct were positively correlated and internally consistent. Next, discriminant validity was assessed in two ways: 1) using Fornell and Larcker's (1981) criteria and 2) comparing the value of the AVE with that of the MSV. According to the results, discriminant validity was established for all constructs, demonstrating that each construct was unique and different from the other constructs. Nomological validity was assessed by testing if existing theories and literature supported the correlation among the constructs. The results demonstrated that the correlations among the constructs are consistent with the theoretical foundation, indicating that nomological validity was achieved. Finally, criterion validity was tested by examining response consistency between questionnaires (old measure) and vignettes (new measure). OLS regression analysis was performed for the dependent (vignette responses) and independent (questionnaires responses) variables. According to the regression analysis result, vignette ratings are positively and significantly associated with the questionnaire responses. Next, it was tested that if the new measure is parallel with what is previously known. Previous studies found that AVs had a positive effect on reliability (Weiss and Roberts, 2018; Primi et al., 2016) and discriminant validity (Primi et al., 2016). Similarly, the present study also found that AVs increases reliability and discriminant validity. Therefore, the results of OLS regression analysis and literature review provided evidence for establishing criterion validity of the new measures in this study. In sum, based on the psychometric analysis results, problem-posing tasks and questionnaires adjusted by AVs are valid and reliable tools to measure social and emotional skills in mathematics among ninth-grade students in Mongolia.

RQ3: What is the status of social and emotional skills among Mongolian students?

To answer this research question, students' performance on the vignette-corrected new scale and problem-posing tasks were analysed. As explained in Chapter 5, three hypothetical individuals were proposed, representing low, medium, and high levels of each skill under investigation (e.g., cooperation in mathematics and mathematical perseverance). The students' vignette-corrected responses were distributed from 1 to 7 (1 < low vignette; 2 = low vignette; 3 > low vignette; 4 = medium vignette; 5 > medium vignette; 6 = high vignette; 7 > high vignette). According to the vignette-corrected scales analysis results, the aggregate mean performance for cooperative learning in mathematics is 4.47, which indicates that Mongolian students tend to be more cooperative in mathematics than the hypothetical student described in the medium vignette. Similar findings were obtained for mathematical perseverance (Mean = 4.44), mathematics enjoyment (Mean = 4.68), and mathematical self-efficacy (Mean = 4.55), which implies that the students possess better respective skills than the fictitious individuals described in the corresponding medium vignettes. Moreover, for mathematics anxiety, the students are higher anxious in mathematics than the hypothetical student in the medium vignette is (Mean = 5), which is much closer to the high vignette (higher anxiety).

For mathematical creativity, students' performance on problem-posing tasks was analysed based on three major components of mathematical creativity, namely fluency, flexibility, and originality. The findings of the fluency, flexibility, and originality analysis on problem-posing tasks indicate that Mongolian students lack mathematical creativity (the details are discussed in Chapter 6). Moreover, there is a statistically significant difference between male and female students, and female students have a higher mean value (5.29) than male students do (4.49), which demonstrates that female students are more creative in mathematics. The results also demonstrated that the difference in students' performance in mathematical creativity was statistically significant according to their ethnicity. Mongol students had a higher mean value (5.09) than Kazakh students (2.98), implying that Mongol students are more creative in mathematics than Kazakh students are. However, it should be noted that the Kazakh sample was considerably smaller than the Mongol sample.

Regarding the distribution across regions, the students in urban areas had higher mathematical creativity with a mean value of 5.69 than the rural areas did, whose mean value was 4.41. As the values are highly statistically significant, it can be concluded that urban students are more mathematically creative than rural students are. Overall, Mongolian students tend to be less creative and more anxious about mathematics; however, they demonstrate moderate performance on cooperative learning in mathematics, mathematical perseverance, mathematics enjoyment, and mathematical self-efficacy.

7.2 Implications

The present study has both theoretical and practical implications. This pioneering study conceptualises, designs, and validates a theoretical framework and instruments to measure social and emotional skills in mathematics. Moreover, the analysis of various frameworks on social and emotional skills reveals that the skills under the term 'social and emotional skills' go beyond the initial concept of social and emotional development, which covers social skills and emotional intelligence. Furthermore, the results of nomological validity indicate that cooperative learning in mathematics is negatively correlated with mathematics anxiety (neuroticism) and is positively correlated with other constructs such as mathematical creativity (openness and intellect), mathematical perseverance (conscientiousness), and mathematics enjoyment (extraversion). The finding supports the theoretical pattern in the Big Five personality traits where agreeableness is negatively associated with neuroticism and positively correlated with openness and intellect, conscientiousness, and extraversion. Additionally, as it was confirmed that the AV approach increases discriminant validity in this study, it may help to distinguish conceptual overlap between motivation and emotion in mathematics education. There has been a serious discussion about the conceptual overlap between these two constructs in mathematics education for many years (Pekrun, 2006; Schukajlow et al., 2017).

Moreover, the theoretical framework in this study will enable researchers to conduct empirical studies related to the social and emotional aspects of mathematics. Additionally, the systematic procedures carried out in this study can be adapted to construct and validate a conceptual or theoretical framework in various areas of social sciences. Furthermore, the current study demonstrates that students who are more cooperative in learning mathematics tend to be more creative, confident, and less anxious and have higher enjoyment and higher persistence in mathematics. This finding has a practical implication that teachers should encourage students' cooperative learning in the mathematics classroom for better mathematical attainment. Finally, in measurement practice, internal consistency is heavily dependent on the number of items. The fewer the items used, the lower the internal consistency. In this study, AVs positively affected internal consistency, with fewer items and increased validity. In this regard, the vignette sets used in this study can be applied along with the questionnaires to measure social and emotional skills in mathematics for better psychometric properties. Besides the vignette sets and questionnaires, problem-posing can be a reliable measure to assess mathematical creativity while also contributing to the development of mathematical creativity among students.

7.3 Limitations and future research

This study has some limitations that could be explored in future research. First, the study employed the convenient sampling method for data collection; however, the sample size consisted of students from different geographical regions while considering a balance of the respondents' gender. Second, the study employed a smaller number of self-rating items; however, the decision to examine the effect of the AV approach on internal consistency using fewer items was intentional because internal consistency has a positive association with the number of items. Future studies could validate the framework by employing a satisfactory number of items.

Third, items for mathematics anxiety and mathematical self-efficacy were content-specific, which is related to geometry content. In contrast, items for cooperative learning in mathematics, mathematical perseverance, and mathematics enjoyment were content-general. This may have caused a weak correlation among the constructs. In this regard, further studies may examine nomological validity while considering the content of the items.

Fourth, this study excluded some social and emotional skills such as critical mathematical thinking and mathematical goal-orientation for the reasons explained in Chapter 4. In this regard, the theoretical framework can be extended by adding more social and emotional skills and validated in future studies. Fifth, the EFA and CFA approaches were performed on the same sample in this study. It is desirable to use the EFA and CFA approaches on two different groups that are randomly selected from the same population. For the present study, it was not possible to conduct the second study due to specific circumstances.

Sixth, this study did not test criterion validity due to the confined time allotted for data collection from schools and the constricted size of the survey instruments.

Seventh, the present study employed a non-parametric approach to analyse the vignette sets. Researchers (Vonkova & Hrabak, 2015) have pointed out that the non-parametric approach has a disadvantage regarding dealing with order violations in the vignette analysis. This study used a method introduced by Kyllonen and Bertling (2014), which suggests analysing order violation as a tie. However, this might lead to the loss of information in the data. Considering this limitation, future studies could employ the parametric solution of the AV approach (e.g., compound hierarchical ordered probit model) to handle the limitation of the non-parametric approach. Furthermore, as discussed in Chapter 6, larger measurement errors occurred in the vignette set for mathematics anxiety, indicating order violations in more than 20 per cent of the sample size. Future research should consider ways to minimise the size of the order violation while revising and adjusting the negative statements of the vignette set to the scale options.

Finally, according to the findings in Chapter 6, Mongolian ninth-grade students indicated lower levels of mathematical creativity. However, the proposed curriculum and policies place more emphasis on developing mathematical creativity. Future studies should investigate learning activities and other classroom practices to offer insights into curriculum implementation to develop mathematical creativity in Mongolian students. Particularly, future studies may explore if the textbook tasks, teaching strategies, and mathematics classroom environment in Mongolian secondary schools contribute to increasing students' mathematical creativity by enabling them to think in different and original ways.

REFERENCES

Action Canada Task Force. (2013). Future tense: Adapting Canadian education systems for the 21st century. Retrieved from https://actioncanada.ca/publications/future-tense-adapting-canadian-education-systems-for-the-21st-century/

Akgül, S., & Kahveci, N. (2016). A Study on the Development of a Mathematics Creativity Scale. *Eurasian Journal of Educational Research*, *16*(62). https://doi.org/10.14689/ejer.2016.62.5

Akman, O., & Koçoğlu, E. (2016). Investigation 8th Grade Students Secondary School Cognitive Structure about Principles of Ataturk through Word Association Test. *Journal of Education and Training Studies*, *4*(11). https://doi.org/10.11114/jets.v4i11.1696

Albin, E. J., & Vice, G. (2019). Power to Change: Math as a Social-Emotional Language in a Classroom of 4 and 5 Year Olds. Occasional Paper Series, 2019 (41). https://educate.bankstreet.edu/occasional-paper-series/vol2019/iss41/7/

Amarjargal, A., Munkhjargal, D., Khishigbayar, B., & Otgonbaatar, Kh. (2016). "Policy and implementation of assessment of new competencies", *Proceedings of International Conference on Educational Evaluation* (pp. 100-105). Ulaanbaatar, Mongolia: Soyombo Printing. [In Mongolian]

Anglim, J., & O'connor, P. (2019). Measurement and research using the Big Five, HEXACO, and narrow traits: A primer for researchers and practitioners. Australian Journal of Psychology, 71(1), 16-25. https://doi.org/10.1111/ajpy.12202

Ashcraft, M., & Kirk, E. (2001). The relationships among working memory, math anxiety, and performance. *Journal of Experimental Psychology: General*, *130*(2), 224-237. https://doi.org/10.1037/0096-3445.130.2.224

Ashton, M., Lee, K., & de Vries, R. (2014). The HEXACO Honesty-Humility, Agreeableness, and Emotionality Factors. *Personality and Social Psychology Review*, *18*(2), 139-152. https://doi.org/10.1177/1088868314523838

Balka, D. S. (1974). "Creative ability in mathematics", *The Arithmetic Teacher, 21*, 633-363.

Ballarino G., Meschi E., Scervini F. (2013). The expansion of education in Europe in the 20th Century. AIAS, GINI Discussion Paper 83. http://archive.uva-aias.net/uploaded_files/publications/83-3-3-6.pdf

Bandura, A. (1986). The Explanatory and Predictive Scope of Self-Efficacy Theory. *Journal of Social and Clinical Psychology*, *4*(3), 359-373. https://doi.org/10.1521/jscp.1986.4.3.359

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. Educational Psychologist, 28(2), 117-149. http://dx.doi.org/10.1207/s15326985ep2802_3

Bartlett, M. (1950). Tests of Significance in Factor Analysis. *British Journal of Statistical Psychology*, *3*(2), 77-85. https://doi.org/10.1111/j.2044-8317.1950.tb00285.x

Bhoumick, P., & Saha, R. (2020). Social and Emotional Learning (SEL) in Mathematics Classroom and Its Impact on Mathematics Achievement. Research and Reflections On Education, 18(2), 34-38.

Bernreuter, R. G. (1933). Validity of the personality inventory. Personnel Journal, 11, 383–386.

Blair, C. (2002). School readiness: Integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. *American Psychologist*, *57*(2), 111-127. https://doi.org/10.1037/0003-066x.57.2.111

Bodas, J., & Ollendick, T. (2005). Test Anxiety: A Cross-Cultural Perspective. *Clinical Child And Family Psychology Review*, *8*(1), 65-88. https://doi.org/10.1007/s10567-005-2342-x

Boyd T. (2021) Education Reform in Ontario: Building Capacity Through Collaboration. In: Reimers F.M. (eds) Implementing Deeper Learning and 21st Century Education Reforms. Springer, Cham. https://doi.org/10.1007/978-3-030-57039-2_2

Braskén, M., Hemmi, K., & Kurtén, B. (2019). Implementing a Multidisciplinary Curriculum in a Finnish Lower Secondary School – The Perspective of Science and Mathematics. *Scandinavian Journal of Educational Research*, *64*(6), 852-868. https://doi.org/10.1080/00313831.2019.1623311

Buhs, E., Ladd, G., & Herald, S. (2006). Peer exclusion and victimization: Processes that mediate the relation between peer group rejection and children's classroom engagement and achievement?. Journal of Educational Psychology, 98(1), 1-13. https://doi.org/10.1037/0022-0663.98.1.1

Burton, L., & Mazerolle, S. (2011). Survey Instrument Validity Part I: Principles of Survey Instrument Development and Validation in Athletic Training Education Research. Athletic Training Education Journal, 6(1), 27-35. https://doi.org/10.4085/1947-380x-6.1.27

Cai, J. (2000). "Mathematical thinking involved in U.S. and Chinese students' solving of process-constrained and process-open problems", *Mathematical Thinking and Learning, 2*(4), 309-340. https://doi.org/10.1207/s15327833mtl0204_4

Cangur, S., & Ercan, I. (2015). Comparison of Model Fit Indices Used in Structural Equation Modeling Under Multivariate Normality. Journal Of Modern Applied Statistical Methods, 14(1), 152-167. https://doi.org/10.22237/jmasm/1430453580

Carey, E., Hill, F., Devine, A., & Szücs, D. (2016). The Chicken or the Egg? The Direction of the Relationship Between Mathematics Anxiety and Mathematics Performance. *Frontiers In Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.01987

Casner-Lotto, L. (2020). Are They Really Ready to Work? Employers' Perspectives on the Basic Knowledge and Applied Skills of New Entrants to the 21st Century U.S. Workforce. Retrieved 23 March 2020, from https://eric.ed.gov/?id=ED519465

Catano, V., Brochu, A., & Lamerson, C. (2012). Assessing the Reliability of Situational Judgment Tests Used in High-Stakes Situations. *International Journal of Selection And Assessment*, *20*(3), 333-346. https://doi.org/10.1111/j.1468-2389.2012.00604.x

Cattell, R. (1966). The Scree Test for The Number of Factors. *Multivariate Behavioral Research*, *1*(2), 245-276. https://doi.org/10.1207/s15327906mbr0102_10

Center for Curriculum Redesign. (2019). *Framework of Competencies/Sub-competencies*. Curriculumredesign.org. Retrieved 15 April 2020, from https://curriculumredesign.org/framework/

Chamberlin, S., & Moon, S. (2005). Model-Eliciting Activities as a Tool to Develop and Identify Creatively Gifted Mathematicians. Journal Of Secondary Gifted Education, 17(1), 37-47. https://doi.org/10.4219/jsge-2005-393

Chen, C., Lee, S., & Stevenson, H. (1995). Response Style and Cross-Cultural Comparisons of Rating Scales Among East Asian and North American Students. *Psychological Science*, *6*(3), 170-175. https://doi.org/10.1111/j.1467-9280.1995.tb00327.x

Chernyshenko, O., Kankaraš, M., & Drasgow, F. (2018). *"Social and emotional skills for student success and wellbeing: Conceptual framework for the OECD study on social and emotional skills*. Paris: OECD Publishing. Retrieved from http://dx.doi.org/10.1787/db1d8e59-en

Child, D. (2006). The Essentials of Factor Analysis. 3rd edition, New York: Continuum.

Child Trends. (2014). Measuring Elementary School Students' Social and Emotional Skills: Providing Educators with Tools to Measure and Monitor Social and Emotional Skills that Lead to Academic Success. Retrieved 12 April 2020, from https://www.childtrends.org/wp-content/uploads/2014/08/2014-37CombinedMeasuresApproachandTablepdf1.pdf

Cohen, J. (1988). *Statistical Power Analysis for the Behavioural Sciences.* New Jersey: Lawrence Erlbaum.

Collaborative for Academic, Social, and Emotional Learning. (2017). *Core SEL Competencies*. Casel.org. Retrieved 15 April 2020, from https://casel.org/core-competencies/

Comrey, A., & Lee, H. (1992). *A First Course in Factor Analysis* (2nd ed.). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Costa, P. T., & MacCrae, R. R. (1992). Revised NEO personality inventory (NEO PI-R) and NEO five-factor inventory (NEO FFI): Professional manual: Psychological Assessment Resources.

Cronbach, L. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297-334. https://doi.org/10.1007/bf02310555

De Fruyt, F., et al. (2006). Five types of personality continuity in childhood and adolescence. Journal of Personality and Social Psychology, 91(3), 538-552.

Dew, K., & Galassi, J. (1983). Mathematics anxiety: Some basic issues. *Journal of Counselling Psychology*, *30*(3), 443-446. https://doi.org/10.1037/0022-0167.30.3.443

DeYoung, C., Quilty, L., & Peterson, J. (2007). Between facets and domains: 10 aspects of the Big Five. *Journal of Personality and Social Psychology*, *93*(5), 880-896. https://doi.org/10.1037/0022-3514.93.5.880

Dowker, A., Sarkar, A., & Looi, C. (2016). Mathematics Anxiety: What Have We Learned in 60 Years?. *Frontiers In Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00508

Duckworth, A., Peterson, C., Matthews, M., & Kelly, D. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, *92*(6), 1087-1101. https://10.1037/0022-3514.92.6.1087

Durlak, J., Weissberg, R., Dymnicki, A., Taylor, R., & Schellinger, K. (2011). The Impact of Enhancing Students' Social and Emotional Learning: A Meta-Analysis of School-Based Universal Interventions. *Child Development*, *82*(1), 405-432. https://doi.org/10.1111/j.1467-8624.2010.01564.x

Education Evaluation Center, Mongolia. (2016). National Assessment for Education Quality Report. Ulaanbaatar. [Mongolian]

Education Evaluation Center, Mongolia. (2017). National Assessment for Education Quality Report. Ulaanbaatar. [Mongolian]

Education Evaluation Center, Mongolia. (2018). National Assessment for Education Quality Report. Ulaanbaatar. [Mongolian]

Education Testing Service (2012). Facets of the Big Five. Princeton, USA.

Educational Testing Service. (2012). *The "Plain English" Pocket Guide to Testing and Psychometric Terms* (2nd ed.). Princeton, USA.

Educational Testing Service. (2015). ETS Standards for Quality and Fairness. Princeton, USA.

Erden M (1988). Teaching Techniques Based on Collaboration Among Students. Education and Science pp. 57-60.

Ernest, P. (1991). *The Philosophy of Mathematics Education* (1st ed., pp. 137-139). Basingstoke, UK: Burgess Science Press.

Ernst, E., Merola, R., & Samaan, D. (2021). The Economics of Artificial Intelligence: Implications of for the Future of Work. Geneva, Switzerland: International Labor Organization. https://www.researchgate.net/publication/328353684_The_economics_of_artificial_intelligence_Implications_for_the_future_of_work/link/5bc81e60299bf17a1c5b28d2/download

Ervynck, G. (1991). Mathematical creativity. In D. Tall, Advanced mathematical thinking (pp. 42-52). Kluwer Academic Publishers New York.

Field, A. (2013). Discovering Statistics using SPSS, 4th edition, London: SAGE

Finnish National Board of Education. (2016). National Core Curriculum for Basic Education. Helsinki.https://kosovo.finnish.school/wp-content/uploads/2020/11/Basic-Education-Grades-1-9.pdf

Fischer, R. (2004). Standardization to Account for Cross-Cultural Response Bias. Journal Of Cross-Cultural Psychology, 35(3), 263-282. https://doi.org/10.1177/0022022104264122

Foley, A., Herts, J., Borgonovi, F., Guerriero, S., Levine, S., & Beilock, S. (2017). The Math Anxiety-Performance Link. Current Directions in Psychological Science, 26(1), 52-58. https://doi.org/10.1177/0963721416672463

Fornell, C., & Larcker, D. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research*, *18*(1), 39. https://doi.org/10.2307/3151312

Gatignon, H. (2010). Statistical analysis of management data. Boston, MA: Kluwer Academic Publishers.

Goldin, G., Hannula, M., Heyd-Metzuyanim, E., Kaasila, R., Lutovac, S., & Di Martino, P. et al. (2016). Motivation. In J. Middleton, A. Jansen & G. Goldin, *Attitudes, Beliefs, Motivation, and Identity in Mathematics Education* (1st ed., pp. 17-21). Switzerland: Springer International Publishing.

Goleman, D. (1995). Emotional intelligence. New York, NY: Bantam Books.

Gómez, O., García-Cabrero, B., Hoover, M., Castañeda-Figueiras, S., & Benítez, Y. (2020). Achievement Emotions in Mathematics: Design and Evidence of Validity of a Self-Report Scale. Journal Of Education and Learning, 9(5), 233. https://doi.org/10.5539/jel.v9n5p233

Görgün, S., & Tican, C. (2020). Investigation of Middle School Students' Math Self-Efficacy Perceptions and Math Problem Posing Attitudes. *International Education Studies*, *13*(11), 86. https://doi.org/10.5539/ies.v13n11p86

Grégoire, J. (2016). "Understanding creativity in mathematics for improving mathematical education", *Journal of Cognitive Education and Psychology, 15*(1), 24-36. https://doi.org/10.1891/1945-8959.15.1.24

Grootenboer, P., & Marshman, M. (2016). *Mathematics, Affect and Learning* (1st ed., pp. 131-135). Springer Science + Business Media Singapore. https://doi.org/10.1007/978-981-287-679-9

Guilford, J. P. (1960). "The structure of the intellect model: Its use and implications", New York: McGraw-Hill.

Gumora, G., & Arsenio, W. (2002). Emotionality, Emotion Regulation, and School Performance in Middle School Children. *Journal of School Psychology*, *40*(5), 395-413. https://doi.org/10.1016/s0022-4405(02)00108-5

Gunderson, E., Park, D., Maloney, E., Beilock, S., & Levine, S. (2017). Reciprocal relations among motivational frameworks, math anxiety, and math achievement in early elementary school. Journal Of Cognition and Development, 19(1), 21-46. https://doi.org/10.1080/15248372.2017.1421538

Gutman, L., & Schoon, I. (2013). *The impact of non-cognitive skills on outcomes for young people*. Available at https://pdfs.semanticscholar.org/f4a5/2db3001fb6fb22eef5dc20267b5b807fd8ff.pdf.

Hair, J., R. Anderson, et al. (1995). Multivariate data analysis. New Jersey, Prentice-Hall Inc

Hair, J.F., Black, W. C., Babin, B. J. and Anderson, R. E. (2010). "Multivariate Data Analysis". (7th Ed.). Prentice Hall, Englewood Cliffs.

Hair, J. F., Hult, G.T.M., Ringle, C.M. and Sarstedt, M.M. (2014). "A Premier on Partial Least Squares Structural Equation Modeling (PLS-SEM)". SAGE Publications, California, U.S.A.

Hair, J., Risher, J., Sarstedt, M., & Ringle, C. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, *31*(1), 2-24. https://doi.org/10.1108/ebr-11-2018-0203

Halinen, I. (2018). The New Educational Curriculum in Finland. In M. Matthes, L. Pulkkinen, C. Clouder & B. Heys, Improving the Quality of Childhood in Europe (1st ed., pp. pp. 75-89). Brussels, Belgium: Alliance for Childhood European Network Foundation. https://www.allianceforchildhood.eu/publications

Harvey, L. (2012–2020). Researching the Real World. http://www.qualityresearchinternational.com/methodology/

Hasratuddin, H. (2011). Improving Student's Emotional Intelligence by Mathematics Learning. In International Seminar and the Fourth National Conference on Mathematics Education (pp. 539-546). Yogyakarta: Yogyakarta State University. https://core.ac.uk/download/pdf/11059747.pdf

Haylock, D. (1987). "A framework for assessing mathematical creativity in schoolchildren", *Educational Studies in Mathematics, 18*(1), 59-74. https://doi.org/10.1007/bf00367914

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. Journal of Labor economics, 24(3), 411-482.

Hembree, R. (1990). The Nature, Effects, and Relief of Mathematics Anxiety. Journal For Research in Mathematics Education, 21(1), 33-46. https://doi.org/10.5951/jresematheduc.21.1.0033

Henson, R., & Roberts, J. (2006). Use of Exploratory Factor Analysis in Published Research. Educational And Psychological Measurement, 66(3), 393-416. doi: https://doi.org/10.1177/0013164405282485

Hess, F. M. (2008). Book Review: Charter Schools: Hope or Hype? by Jack Buckley and Mark Schneider. Princeton, NJ: Princeton University Press, 2007. 376 pp. Urban Affairs Review, 43(6), 887–889. https://doi.org/10.1177/1078087407309225

Hinz, A., Häuser, W., Glaesmer, H., & Brähler, E. (2016). The relationship between perceived own health state and health assessments of anchoring vignettes. *International Journal of Clinical and Health Psychology*, *16*(2), 128-136. https://doi.org/10.1016/j.ijchp.2016.01.001

Hopkins, D., & King, G. (2010). Improving Anchoring Vignettes: Designing Surveys to Correct Interpersonal Incomparability. *Public Opinion Quarterly*, *74*(2), 201-222. https://doi.org/10.1093/poq/nfq011

Humphrey, N., Lendrum, A., & Wigelsworth, M. (2010). Social and emotional aspects of learning (SEAL) programme in secondary schools: national evaluation. University of Manchester. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/181718/DFE-RR049.pdf

Jensen, L. R. (1973). "The relationships among mathematical creativity, numerical aptitude, and mathematical achievement", Unpublished dissertation. The University of Texas at Austin, Austin.

John, O., & Fruyt, F. (2015). Education and Social Progress: Framework for the Longitudinal Study of Social and Emotional Skills in Cities. OECD. Retrieved from https://one.oecd.org/document/EDU/CERI/CD(2015)13/en/pdf

Jöreskog, K. (1970). A General Method for Estimating a Linear Structural Equation System. ETS Research Bulletin Series, 1970(2), i-41. https://doi.org/10.1002/j.2333-8504.1970.tb00783.x

Jürges, H., & Winter, J. (2011). Are anchoring vignettes ratings sensitive to vignette age and sex?. *Health Economics*, *22*(1), 1-13. https://doi.org/10.1002/hec.1806

Kaiser, H. (1974). An index of factorial simplicity. *Psychometrika*, *39*(1), 31-36. https://doi.org/10.1007/bf02291575

Kankaraš, M. (2017). Personality matters: Relevance and Assessment of Personality Characteristics. Paris, France: OECD. Retrieved from https://doi.org/10.1787/19939019

Kaplan, D. (2000). Structural equation modeling: Foundation and extensions. Thousand Oaks, CA: Sage Publications.

Kattou, M., Christou, C., & Pitta-Pantazi, D. (2015). Mathematical creativity or general creativity?. In *Ninth Congress of the European Society for Research in Mathematics Education* (pp. pp.1016-1023). Prague, Czech Republic. Retrieved from https://hal.archives-ouvertes.fr/hal-01287305/document

Kattou, M., Kontoyianni, K., Pitta-Pantazi, D., & Christou, C. (2012). "Connecting mathematical creativity to mathematical ability", ZDM, 45(2), 167-181. https://doi.org/10.1007/s11858-012-0467-1

Kautz, T., Heckman, J., Diris, R., Weel, B., & Borghans, L. (2014). Fostering and Measuring Skills Improving Cognitive and Non-cognitive Skills to Promote Lifetime Success. Paris, France: OECD Publishing. Retrieved from https://doi.org/10.1787/19939019

Kim, K. H. (2009). Creative Problem Solving. In B. Kerr (Ed). Encyclopedia of Giftedness, Creativity and Talent. Sage Publications. pp. 188-191.

King, G., Murray, C., Salomon, J., & Tandon, A. (2004). Enhancing the Validity and Cross-Cultural Comparability of Measurement in Survey Research. *American Political Science Review*, *98*(1), 191-207. https://doi.org/10.1017/s000305540400108x

King, G., & Wand, J. (2007). Comparing Incomparable Survey Responses: Evaluating and Selecting Anchoring Vignettes. *Political Analysis*, *15*(1), 46-66. https://doi.org/10.1093/pan/mpl011

Kline, R. (2005). Principles and practice of structural equation modelling. New York: Guilford Press.

Knekta, E., Runyon, C., & Eddy, S. (2019). One Size Doesn't Fit All: Using Factor Analysis to Gather Validity Evidence When Using Surveys in Your Research. CBE—Life Sciences Education, 18(1), rm1. https://doi.org/10.1187/cbe.18-04-0064

Krapp, A. (2005). Basic needs and the development of interest and intrinsic motivational orientations. *Learning and Instruction*, *15*(5), 381-395. https://doi.org/10.1016/j.learninstruc.2005.07.007

Kristensen, N., & Johansson, E. (2008). New evidence on cross-country differences in job satisfaction using anchoring vignettes. *Labour Economics*, *15*(1), 96-117. https://doi.org/10.1016/j.labeco.2006.11.001

Kusmaryono, I., Hardi, S., & Nur, K. (2018). Evaluation on Dispositional Mental Functions of Cognitive, Affective, and Conative in Mathematical Power Problems-Solving Activity. *Journal of Mathematics Education*, *11*(1), 81-102. https://doi.org/10.26711/007577152790022

Kyllonen, P. C. (2012). *Measurement of 21st Century Skills Within the Common Core State Standards*. Retrieved 12 April 2020, from https://www.ets.org/Media/Research/pdf/session5-kyllonen-paper-tea2012.pdf.

Kyllonen, P. C., and Bertling, J. P. (2014). *Draft report: Anchoring Vignettes reduce Bias in Noncognitive Rating Scale Responses.* Princeton, NJ: ETS/OECD

Ladd, G., Herald, S., & Kochel, K. (2006). School Readiness: Are There Social Prerequisites?. Early Education & Development, 17(1), 115-150. https://doi.org/10.1207/s15566935eed1701_6

Lavasani, M., & Khandan, F. (2011). The effect of cooperative learning on mathematics anxiety and help seeking behavior. Procedia - Social and Behavioral Sciences, 15, 271-276. https://doi.org/10.1016/j.sbspro.2011.03.085

Laycock, M. (1970). “Creative mathematics at Nueva”, *The Arithmetic Teacher, 17*, 325-328. Retrieved from https://www.jstor.org/stable/41186201?seq=1#metadata_info_tab_contents

Leandre, R., Fabrigar, L. R., & Wegener, D. T. (2012). Exploratory factor analysis. Oxford, UK: Oxford University Press.

Ledesma, R., Valero-Mora, P., & Macbeth, G. (2015). The Scree Test and the Number of Factors: A Dynamic Graphics Approach. *The Spanish Journal of Psychology*, *18*. https://doi.org/10.1017/sjp.2015.13

Leung, F. (2001). Educational Studies in Mathematics, 47(1), 35-51. https://doi.org/10.1023/a:1017936429620

Leung, F. (2002). Behind the High Achievement of East Asian Students. Educational Research and Evaluation, 8(1), 87-108. https://doi.org/10.1076/edre.8.1.87.6920

Lee, J. (2009). Universals and specifics of math self-concept, math self-efficacy, and math anxiety across 41 PISA 2003 participating countries. Learning and Individual Differences, 19(3), 355-365. https://doi.org/10.1016/j.lindif.2008.10.009

Lee, J., & Stankov, L. (2013). Higher-order structure of noncognitive constructs and prediction of PISA 2003 mathematics achievement. *Learning and Individual Differences*, *26*, 119-130. https://doi.org/10.1016/j.lindif.2013.05.004

Lee, J., Jones, P., Mineyama, Y., & Zhang, X. (2002). Cultural differences in responses to a likert scale. *Research In Nursing & Health*, *25*(4), 295-306. https://doi.org/10.1002/nur.10041

Lipnevich, A., MacCann, C., & Roberts, R. (2013). Assessing Non-Cognitive Constructs in Education: A Review of Traditional and Innovative Approaches. Oxford Handbooks Online. https://doi.org/10.1093/oxfordhb/9780199796304.013.0033

Lubart, T. & Guignard, J. H. (2004). “The generality-specificity of creativity: A multivariate approach”, In R. J. Sternberg, E. L. Grigorenko, & J. L. Singer (Eds.), Creativity: From potential to realization (pp. 43-56). Washington, D.C.: American Psychological Association.

Lufi, D., & Cohen, A. (1987). A Scale for Measuring Persistence in Children. *Journal Of Personality Assessment*, *51*(2), 178-=185. https://doi.org/10.1207/s15327752jpa5102_2

Macit E (2013). Teacher Opinions on the Use of Cooperative Learning Method in Elementary Education Second Level Mathematics Lessons. Malatya: İnönü University Institute of Educational Sciences.

Mann, E. (2005). "Mathematical creativity and school mathematics: Indicators of mathematical creativity in middle school students", Doctoral Dissertations. AAI3205573. https://opencommons.uconn.edu/dissertations/AAI3205573

Mann, E. (2006). “Creativity: The essence of mathematics”, *Journal for The Education of The Gifted, 30*(2), 236-260. https://doi.org/10.4219/jeg-2006-264

McDaniel, M., Morgeson, F., Finnegan, E., Campion, M., & Braverman, E. (2001). Use of situational judgment tests to predict job performance: A clarification of the literature. *Journal Of Applied Psychology*, *86*(4), 730-740. https://doi.org/10.1037/0021-9010.86.4.730

McDonald, R. (1970). The Theoretical Foundations of Principal Factor Analysis, Canonical Factor Analysis, and Alpha Factor Analysis. British Journal of Mathematical and Statistical Psychology, 23(1), 1-21. https://doi.org/10.1111/j.2044-8317.1970.tb00432.x

McKinsey Global Institute. (2018). *Skill shift. Automation and the future of the workforce*. Washington DC.

McLeod, D. (1994). Research on Affect and Mathematics Learning in the JRME: 1970 to the Present. *Journal for Research in Mathematics Education*, *25*(6), 637-647. https://doi.org/10.5951/jresematheduc.25.6.0637

Messiner, H. (2000, July-August). Creativity in mathematics education. Paper presented at the meeting of the International Congress on Mathematical Education, Tokyo.

Ministry of Education, Culture, Science and Sports. (2014). "Primary education core curriculum", Retrieved from https://mecss.gov.mn/news/95/ [In Mongolian]

Ministry of Education, Culture, Science and Sports. (2015). "Lower secondary education core curriculum", Retrieved from https://mecss.gov.mn/news/94/ [In Mongolian]

Ministry of Education, Culture, Science and Sports. (2016). "Upper secondary education core curriculum", Retrieved from http://www.mier.mn/wp-content/uploads/2018/11/Math-2.pdf [In Mongolian]

Ministry of Education, Culture, Science and Sports. (2019). "Grade 9 Mathematics Textbook", Retrieved from https://econtent.edu.mn/book/9rangi [In Mongolian]

Ministry of Education, Ontario. (2020). The Ontario Curriculum: Grades 1–8, Mathematics. Ontario. https://www.dcp.edu.gov.on.ca/en/curriculum/elementary-mathematics

Ministry of Education, Singapore. (2019). Mathematics Syllabuses. Singapore: Curriculum Planning and Development Division. https://www.moe.gov.sg/-/media/files/secondary/syllabuses/maths/2020-express_na-maths_syllabuses

Mishra, P., & Kereluik, K. (2011). What 21st Century Learning? A Review and a Synthesis. In Society for Information Technology & Teacher Education International Conference.

Nashville: Association for the Advancement of Computing in Education. Retrieved from https://www.learntechlib.org/primary/p/36828/

Molad, O., Levenson, E., & Levy, S. (2020). Individual and group mathematical creativity among post–high school students. Educational Studies in Mathematics, 104(2), 201-220. https://doi.org/10.1007/s10649-020-09952-5

Munkhjargal, D., Khishigbayar, B., Otgonbaatar, K., & Amarjargal, A. (2016) Network on Education Quality Monitoring in the Asia-Pacific Comparative Study on Assessment of Transversal Competencies.

National Research Council. (2012). "Education for life and work: Developing transferable knowledge and skills in the 21st century", Washington, DC: National Academies Press. Retrieved 15 June 2020, from https://www.nationalacademies.org/our-work/defining-deeper-learning-and-21st-century-skills

Neumann, C. J. (2007). Fostering creativity—A model for developing a culture of collective creativity in science. EMBO Reports, 8(3), 202–206

Newzoo. (2016). *Global Mobile Market Report*. Retrieved from https://newzoo.com/insights/trend-reports/2016-global-mobile-market-report/

OECD (2013). PISA 2012 results: Ready to learn – Students' engagement, drive and self-beliefs. Vol. III. PISA: OECD Publishing.

OECD. (2017). PISA 2021 Mathematics: A Broadened Perspective. Retrieved from https://www.oecd.org/pisa/pisaproducts/PISA2021_Mathematics_StrategicDirectionPaper.pdf.

OECD (2015), Skills for Social Progress: The Power of Social and Emotional Skills, OECD Skills Studies, OECD Publishing, Paris, https://doi.org/10.1787/9789264226159-en.

OECD. (2014). *PISA 2012 Technical Report*. Paris: OECD Publishing.

OECD. (2005). *PISA 2003 Technical Report*. Paris: OECD Publishing.

OECD. (2004). Learning for tomorrow's world: First results from PISA 2003. Paris Author. http://dx.doi.org/10.1787/9789264006416-en

Oswald, F. L., Schmitt, N., Kim, B. H., Ramsay, L. J., & Gillespie, M. A. (2004). Developing a Biodata Measure and Situational Judgment Inventory as Predictors of College Student Performance. Journal of Applied Psychology, 89(2), 187–207. https://doi.org/10.1037/0021-9010.89.2.187

Otgonbaatar, K. (2013). Gender and Location Difference in Mathematics Achievement in Mongolia: An Item Response Theory Approach (Master). National Taichung University of Education.

Otgonbaatar, K. (2020). Examining Mathematical Creativity Among Mongolian Ninth-Grade Students Using Problem-Posing Approach. *Journal of Education and Practice*, *11*(27), 69-75. https://doi.org/10.7176/jep/11-27-08

Otgonbaatar, K. (2021). Effectiveness of anchoring vignettes in re-evaluating self-rated social and emotional skills in mathematics. *International Journal of Evaluation and Research In Education (IJERE)*, *10*(1), 237. https://doi.org/10.11591/ijere.v10i1.20716

Pacheco, J. (2019). Using Anchoring Vignettes to Re-evaluate the Link between Self-Rated Health Status and Political Behavior. *Journal of Health Politics, Policy and Law*, *44*(3), 533-558. https://doi.org/10.1215/03616878-7367060

Pajares, F., & Miller, M. (1994). Role of self-efficacy and self-concept beliefs in mathematical problem solving: A path analysis. *Journal of Educational Psychology*, *86*(2), 193-203. https://doi.org/10.1037/0022-0663.86.2.193

Park, D., Ramirez, G., & Beilock, S. (2014). The role of expressive writing in math anxiety. *Journal of Experimental Psychology: Applied*, *20*(2), 103-111. https://doi.org/10.1037/xap0000013

Partnership for 21st Century Skills. (2009). *P21 Framework Definitions*. Retrieved 17 April 2020, from https://eric.ed.gov/?id=ED519462.

Partnership for 21st century skills. (2011). 21st Century Skills Map (p. 4). Washington, DC. Retrieved from https://files.eric.ed.gov/fulltext/ED543032.pdf

Parry, S. (2017). Fit Statistics commonly reported for CFA and SEM. Cornell Statistical Consulting Unit: Cornell University, 2.

Paulhus, D. (1991). *Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), Measures of personality and social psychological attitudes* (pp. 17-59). San Diego: Academic Press.

Pekrun, R., Goetz, T., Titz, W., & Perry, R. (2002). Academic Emotions in Students' Self-Regulated Learning and Achievement: A Program of Qualitative and Quantitative Research. Educational Psychologist, 37(2), 91-105. https://doi.org/10.1207/s15326985ep3702_4

Pekrun, R., Goetz, T., Frenzel, A., Barchfeld, P., & Perry, R. (2011). Measuring emotions in students' learning and performance: The Achievement Emotions Questionnaire (AEQ). Contemporary Educational Psychology, 36(1), 36-48. https://doi.org/10.1016/j.cedpsych.2010.10.002

Pekrun, R., & Linnenbrink-Garcia, L. (2012). Academic Emotions and Student Engagement. In S. Christenson, A. Reschly & C. Wylie (Eds.), Handbook of research on student engagement (pp. 259‒282). Springer. https://doi.org/10.1007/978-1-4614-2018-7_12

Perels, F., Gürtler, T., & Schmitz, B. (2005). Training of self-regulatory and problem-solving competence. Learning and Instruction, 15(2), 123-139. https://doi.org/10.1016/j.learninstruc.2005.04.010

Poksinska, B., & Cronemyr, P. (2017). Measuring quality in elderly care: possibilities and limitations of the vignette method. *Total Quality Management & Business Excellence*, *28*(9-10), 1194-1207. doi: 10.1080/14783363.2017.1303875

Poincaré, H. (1948). "Science and method", New York: Dover.

Primi, R., Zanon, C., Santos, D., De Fruyt, F., & John, O. (2016). Anchoring Vignettes. European Journal of Psychological Assessment, 32(1), 39-51. https://doi.org/10.1027/1015-5759/a000336

Prouse, H. L. (1964). "The construction and use of a test for the measurement of certain aspects of creativity in seventh-grade mathematics", Dissertation Abstracts, 26 (01), 394.

O'Grady, K., Deussing, M., Scerbina, T., Tao, Y., Fung, K., Elez, V., & Monk, J. (2019). Measuring up: Canadian results of the OECD PISA 2018 study: The performance of Canadian 15-yearolds. Canada: Council of Ministers of Education.

Quoquab, F., & Mohammad, J. (2020). Cognitive, Affective and Conative Domains of Sustainable Consumption: Scale Development and Validation Using Confirmatory Composite Analysis. *Sustainability*, *12*(18), 7784. https://doi.org/10.3390/su12187784

Rahayu, R., & Kartono. (2014). The Effect of Mathematical Disposition toward Problem Solving Ability Based on IDEAL Problem Solver. International Journal Of Science And Research, 3(10), 1315-1318. https://www.ijsr.net/archive/v3i10/MjAxMDE0MDI=.pdf

Richardson, F., & Suinn, R. (1972). The Mathematics Anxiety Rating Scale: Psychometric data. *Journal of Counseling Psychology*, *19*(6), 551-554. https://doi.org/10.1037/h0033456

Rotherham, A., & Willingham, D. (2010). 21st-Century Skills: Not New, but a Worthy Challenge. American Educator, (Spring), 17-20. Retrieved from https://www.aft.org//sites/default/files/periodicals/RotherhamWillingham.pdf

Samuels, P. (2017). Advice on Exploratory Factor Analysis (pp. 2-6). Birmingham: Centre for Academic Success.

Sawyer, K. (2007). Group genius: The creative power of collaboration. New York: Basic Books.

Schillinger, F., Vogel, S., Diedrich, J., & Grabner, R. (2018). Math anxiety, intelligence, and performance in mathematics: Insights from the German adaptation of the Abbreviated Math Anxiety Scale (AMAS-G). *Learning And Individual Differences*, *61*, 109-119. https://doi.org/10.1016/j.lindif.2017.11.014

Schmitt, D., Allik, J., McCrae, R., & Benet-Martínez, V. (2007). The Geographic Distribution of Big Five Personality Traits. Journal Of Cross-Cultural Psychology, 38(2), 173-212. https://doi.org/10.1177/0022022106297299

Scholz, U., Doña, B. G., Sud, S., & Schwarzer, R. (2002). Is general self-efficacy a universal construct? Psychometric findings from 25 countries. European Journal of Psychological Assessment, 18(3), 242–251. https://doi.org/10.1027/1015-5759.18.3.242

Schukajlow, S., Leiss, D., Pekrun, R., Blum, W., Müller, M., & Messner, R. (2011). Teaching methods for modelling problems and students' task-specific enjoyment, value, interest, and self-efficacy expectations. *Educational Studies in Mathematics*, *79*(2), 215-237. https://doi.org/10.1007/s10649-011-9341-2

Schukajlow, S., Rakoczy, K., & Pekrun, R. (2017). Emotions and motivation in mathematics education: theoretical considerations and empirical contributions. *ZDM*, *49*(3), 307-322. https://doi.org/10.1007/s11858-017-0864-6

Schumacker, R., & Lomax, R. (2010). A beginner's guide to structural equation modeling (3rd ed.). New York: Taylor and Francis Group, LLC.

Silver, E. (1997). "Fostering creativity through instruction rich in mathematical problem solving and problem posing", *Zentralblatt Für Didaktik Der Mathematik, 29*(3), 75-80. https://doi.org/10.1007/s11858-997-0003-x

Simonton, D. (2000). "Creativity: Cognitive, personal, developmental, and social aspects", *American Psychologist, 55*(1), 151-158. https://doi.org/10.1037/0003-066x.55.1.151

Skaalvik, E., Federici, R., & Klassen, R. (2015). Mathematics achievement and self-efficacy: Relations with motivation for mathematics. *International Journal of Educational Research*, *72*, 129-136. https://doi.org/10.1016/j.ijer.2015.06.008

Soto, C., & John, O. (2017). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology*, *113*(1), 117-143. https://doi.org/10.1037/pspp0000096

Sriraman, B., & Lee, K. (2011). “The elements of giftedness and creativity in mathematics”, Rotterdam, The Netherlands: Sense Publishers.

Sriraman, B. (2005). “Are giftedness and creativity synonyms in mathematics?”, *Journal of Secondary Gifted Education, 17*(1), 20-36. https://doi.org/10.4219/jsge-2005-389

Sriraman, B. (2009). “The characteristics of mathematical creativity”, *ZDM Mathematics Education 41*, 13. https://doi.org/10.1007/s11858-008-0114-z

Sternberg, R. (Ed.). (1998). “Handbook of Creativity”, Cambridge: Cambridge University Press. https://doi:10.1017/CBO9780511807916

Sternberg, R. J. (2006). The nature of creativity. Creativity Research Journal, 18(1), 87-98.

Stoyanova, E. (1997). “Extending and exploring students’ problem solving via problem posing: A study of years 8 and 9 students involved in mathematics challenge and enrichment stages of Euler enrichment program for young Australians”, Unpublished doctoral dissertation submitted to Edith Cowan University, Perth, Australia.

Tabachnick, B. G. and L. S. Fidell (2001). Using multivariate statistics. Needham Heights, MA, Allyn & Bacon.

The Government of Mongolia. (2013). "Proper Mongolian Child Program”, Retrieved 12 May 2020, from https://www.legalinfo.mn/annex/details/5962?lawid=9349

The Parliament of Mongolia. (2015). “The Government Policy in Education”, Retrieved 12 May 2020, from https://www.legalinfo.mn/annex/details/6661?lawid=10935

The Ontario Public Service. (2016). 21st Century Competencies. http://edugains.ca/resources21CL/About21stCentury/21CL_21stCenturyCompetencies.pdf

Thornton, S. (2021). Mathematics Education, Virtues and 21st Century Competencies. In P. Toh & B. Kaur, Developing 21st Century Competencies in the Mathematics Classroom (1st ed.). Singapore: World Scientific Publishing.

Toh, P., & Kaur, B. (2016). 21st Century Competencies in Mathematics Classrooms. In P. Toh & B. Kaur, Developing 21st Century Competencies in the Mathematics Classroom (1st ed.). Singapore: Word Scientific Publishing.

Toh, P., & Kaur, B. (2016). Developing 21st century competencies in the mathematics classroom. Singapore: World Scientific Publishing Co. Pte. Ltd.

Toh, T., & Vistro-Yu, C. (2019). School Mathematics Curricula: Asian Perspectives and Glimpses of Reform (1st ed.). Springer, Singapore.

Torrance, E. P. (1966). "Torrance Tests of Creative Thinking: Norms technical manual", Princeton, NJ: Personnel Press.

Torrance, E. P. (1974). "Norms-technical manual: Torrance Tests of Creative Thinking", Lexington, MA: Ginn and Company.

Torrance, E. P. (1988). "The nature of creativity as manifest in its testing", In R. J. Sternberg (Ed.), The nature of creativity: Contemporary psychological perspectives (pp. 43–75). New York: Cambridge University Press.

Treffinger, D. J., Young, G. C., Selby, E.C., & Shepardson, C. (2002). "Assessing creativity: A guide for educators (RM02170)", Storrs, CT: The National Research Center on the Gifted and Talented, University of Connecticut.

Tuohilampi, L., Hannula, M., Varas, L., Giaconi, V., Laine, A., Näveri, L., & i Nevado, L. (2014). Challenging the Western Approach to Cultural Comparisons: Young Pupils' Affective Structures Regarding Mathematics in Finland and Chile. *International Journal of Science and Mathematics Education*, *13*(6), 1625-1648. https://doi.org/10.1007/s10763-014-9562-9

Ullman JB. "Structural equation modeling." In: Tabachnick BG, Fidell LS (ed). Using Multivariate Statistics. 5th ed. Boston, MA: Pearson Education Inc. 1996. 676-780.

UNESCO. (2016). "Assessment of transversal competencies: Current tools in the Asian region", Retrieved from https://bangkok.unesco.org/content/assessment-transversal-competencies-current-tools-asian-region

United Nations Educational, Scientific and Cultural Organization. (2019). Final Report: Planning Education in the AI era: Lead the leap. Paris, France. Retrieved from https://unesdoc.unesco.org/ark:/48223/pf0000370967

Valiente, C., Lemery-Chalfant, K., Swanson, J., & Reiser, M. (2008). Prediction of children's academic competence from their effortful control, relationships, and classroom participation. *Journal of Educational Psychology*, *100*(1), 67-77. https://doi.org/10.1037/0022-0663.100.1.67

Van de Vijver, F., & Leung, K. (1997). Methods and data analysis of comparative research. In J. W. Berry, Y. H. Poortinga, & J. Pandey (Eds.), Handbook of cross-cultural psychology: Theory and method (pp. 257–300). Allyn & Bacon.

Van Harpen, X., & Sriraman, B. (2012). "Creativity and mathematical problem posing: An analysis of high school students' mathematical problem posing in China and the USA", *Educational Studies in Mathematics, 82*(2), 201-221. https://doi.org/10.1007/s10649-012-9419-5

Vernon, P. E. (1934). The attitude of the subject in personality testing. Journal of Applied Psychology, 18, 165-177.

Vonkova, H., & Hrabak, J. (2015). The (in) comparability of ICT knowledge and skill self-assessments among upper secondary school students: The use of the anchoring vignette method. *Computers & Education*, *85*, 191-202. https://doi.org/10.1016/j.compedu.2015.03.003

Wadlington, E., & Wadlington, P. (2008). Helping Students With Mathematical Disabilities to Succeed. *Preventing School Failure: Alternative Education For Children And Youth*, *53*(1), 2-7. https://doi.org/10.3200/psfl.53.1.2-7

Wand, J., King, G., & Lau, O. (2011). anchors: Software for Anchoring Vignette Data. *Journal of Statistical Software*, *42*(3). https://doi.org/10.18637/jss.v042.i03

Wand, J., King, G., & Lau, O. (2016). Statistical analysis of surveys with anchoring vignettes (Version 3.0-8) [Windows]. USA: GPL.

Wardani, S.B., Sumarmo, U., & Nishitani, I. (2011). "Mathematical creativity and disposition: experiment with grade-10 students using silver inquiry approach", https://gair.media.gunma-u.ac.jp

Waugh, G., & Russell, T. (2003). Scoring both Judgment and Personality in a Situational Judgment Test. In 45th Annual Conference of the International Military Testing Association (pp. 540-548). Pensacola, Florida: International Military Testing Association. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.115.772&rep=rep1&type=pdf#page=556

Wechsler, S., Saiz, C., Rivas, S., Vendramini, C., Almeida, L., Mundim, M., & Franco, A. (2018). "Creative and critical thinking: Independent or overlapping components?", *Thinking Skills and Creativity*, *27*, 114-122. https://doi.org/10.1016/j.tsc.2017.12.003

Weiss, S., & Roberts, R. (2018). Using Anchoring Vignettes to Adjust Self-Reported Personality: A Comparison Between Countries. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.00325

West, M., Kraft, M., Finn, A., Martin, R., Duckworth, A., Gabrieli, C., & Gabrieli, J. (2016). Promise and Paradox: Measuring Students' Non-Cognitive Skills and the Impact of Schooling. *Educational Evaluation And Policy Analysis*, *38*(1), 148-170. https://doi.org/10.3102/0162373715597298

Wong, K. (2016). Enriching Secondary Mathematics Education with 21st Century Competencies. In P. Toh & B. Kaur, Developing 21st Century Competencies in the Mathematics Classroom (1st ed.). Singapore: World Scientific Publishing.

World Economic Forum. (2016). *The Future of Jobs Employment, Skills and Workforce Strategy for the Fourth Industrial Revolution*. Cologny/Geneva Switzerland. Retrieved from http://www3.weforum.org/docs/WEF_Future_of_Jobs.pdf

Wu, L., Yang, S., & Koo, M. (2017). Psychometric properties of the Chinese version of Spiritual Index of Well-Being in elderly Taiwanese. BMC Geriatrics, 17(1). https://doi.org/10.1186/s12877-016-0392-1

Yalçin, K., & Hasan, A. (2018). The effect of cooperative learning on the academic achievement and attitude of students in Mathematics class. *Educational Research and Reviews*, *13*(21), 712-722. https://doi.org/10.5897/err2018.3636

Yang, Z., Wang, Y., & Zhang, Y. (2019). Students' Social-emotional competency and mathematics academic development: A Clustering analysis on China. In 9th ICEEPSY 2018 The International Conference on Education and Education Psychology. UK: Future Academy. https://doi.org/10.15405/epsbs.2019.01.1

Zakaria, E., Chin, L., & Daud, Y. (2010). The Effects of Cooperative Learning on Students' Mathematics Achievement and Attitude towards Mathematics. *Journal of Social Sciences*, *6*(2), 272-275. https://doi.org/10.3844/jssp.2010.272.275

Zan, R., Brown, L., Evans, J., & Hannula, M. (2006). Affect in Mathematics Education: An Introduction. *Educational Studies in Mathematics*, *63*(2), 113-121. https://doi.org/10.1007/s10649-006-9028-2

Zigler, C., Jacobe, H., Ardalan, K., Coles, T., Lane, S., Schollaert, K., & Torok, K. (2020). The importance of development standards for anchoring vignettes: an illustrative example from pediatric localized scleroderma quality of life. Quality Of Life Research, 29(12), 3263-3272. https://doi.org/10.1007/s11136-020-02575-6

Zimmermann, M., Bescherer, C., & Spannagel, C. (2010). A Questionnaire for Surveying Mathematics Self-Efficacy Expectations of Future Teachers.

Zinbarg, R., Revelle, W., Yovel, I., & Li, W. (2005). Cronbach's α, Revelle's β, and Mcdonald's ωH: their relations with each other and two alternative conceptualizations of reliability. *Psychometrika*, *70*(1), 123-133. https://doi.org/10.1007/s11336-003-0974-7

APPENDICES

GENERAL INSTRUCTIONS

This survey contains questions about your general information, general characteristics, views regarding others, and mathematical problem-posing tasks. The questions and tasks in this questionnaire have no ‘right’ or ‘wrong’ answers. Your answer should reflect what is ‘right’ for you. Everyone is different; hence, each answer can be different. This survey is by no means a test, and you will not be judged or graded based on your responses. In case you do not feel comfortable participating, you do not have to take this survey. It will not affect your class attendance.

Please go through all questions carefully and respond as accurately as you can. After answering a question, check the corresponding box. A few questions require you to write an answer in words.

Paper instructions: If you check an incorrect box by mistake, cross it out and check the box you think is correct. If you make a mistake while writing your answer, cross it out and write the correct answer next to it. Your responses will be used along with those of other respondents to compute aggregate and average values. Confidentiality of everyone’s identity and responses will be maintained.

Are you willing to take this survey? Yes No

GENERAL INFORMATION ABOUT YOU

Age: ……………………	Ethnicity: ……………………
Gender: ………………….	School: ………………………

PART I: YOUR GENERAL CHARACTERISTICS

Instructions: This section presents various characteristics that you may or may not find relatable. As a response to each statement, please specify the extent to which you agree or disagree with it.

#	I am someone who …	Disagree Strongly	Disagree	Neither agree nor disagree	Agree	Agree strongly
MAQ1	Feels worried to solve problems like finding the size of angle x in a given figure					
MAQ2	Gets very tense when I have to solve mathematics problems like finding the area of a parallelogram					
MAQ3	Feels nervous when doing a mathematics problem like finding the measure of the smallest angle of a triangle					
MEQ1	Feels happy when dealing with mathematics					
MEQ2	Feels mathematics is fun					
MEQ3	Thinks mathematics is an enjoyable and cool subject					
MCLQ1	Does my best work in mathematics when working					

	with other students					
MCLQ2	Enjoys helping others to work well in a group in mathematics					
MCLQ3	Thinks that mathematics is about working together with others to solve problems					
MSEQ1	Feels confident in finding the height of a pyramid					
MSEQ2	Feel confident in finding the unknown side of a rectangle if its perimeter is given					
MSEQ3	Feel confident in finding the area of a parallelogram					
MPQ1	Searches for more information to clarify the problem if I face difficulties in solving mathematics problems					
MPQ2	Feels challenged to work hard to find a solution when I get a difficult mathematics problem					
MPQ3	Usually keeps trying a difficult problem until I have solved it					

PART II: HOW YOU RATE OTHER PEOPLE

Instruction: This section describes the characteristics of some individuals. Please indicate how much you agree or disagree with each statement by marking an appropriate rating for each statement.

	How much do you agree with this statement?	Disagree strongly	Disagree	Neither agree nor disagree	Agree	Agree strongly
1	One day, during a mathematics class, Nasaa was asked to answer the teacher's question related to geometry. He gave a wrong answer upon which the teacher made a negative remark, and his peers laughed at him. Since then, he is scared of being called to the board and even dreads going to mathematics class. Now, Nasaa gets very tense, feels helpless, and starts sweating when he has to solve problems like finding the area of a parallelogram. Based on the information provided, to what extent do you agree with the statement, 'Nasaa feels worried about mathematics'?					
2	Nara feels that sometimes mathematics is hard and sometimes it is not. She feels that problems in mathematics are not always complicated. However, Nara tends to worry quite a lot when she has to solve problems like finding the measure of the smallest angle of a triangle. In this sense, when she can avoid doing mathematics, she does so. The only reason she takes mathematics is that she has to. Based on the information provided, to what extent do you agree with the statement, 'Nara feels worried about mathematics'?					

3	Sara always appears relaxed and feels at ease in a mathematics class. She never comes across as upset, even if she feels like it is challenging to solve mathematics problems like finding the length of a segment. Sara has good feelings towards mathematics. Sometimes she works on more mathematics problems than are assigned in class. Based on the information provided, to what extent do you agree with the statement, 'Sara feels worried about mathematics'?					
4	Zoloo feels mathematics is a boring and dull subject. He does not enjoy the mathematics class. Even upon hearing the word 'mathematics', he has feelings of dislike. He would like to spend less time in school doing mathematics. Based on the information provided, to what extent do you agree with the statement, 'Zoloo enjoys doing mathematics'?					
5	Sometimes Gerel feels what she learns in her mathematics class is uninteresting and is mostly about numbers. She thinks mathematics is not very enjoyable. In the mathematics class, she does learn about the things that interest her. Based on the information provided, to what extent do you agree with the statement, 'Gerel enjoys doing mathematics'?					
6	Tsogoo looks forward to his mathematics class because he is interested in the things he learns in mathematics. He enjoys attempting to solve mathematics problems, no matter how hard or easy they are. He also enjoys talking to other people about mathematics. Based on the information provided, to what extent do you agree with the statement, 'Tsogoo enjoys doing mathematics'?					
7	Tuya easily feels desperate and gives up quickly if she					

	faces difficulties in solving mathematics problems. She is unaware of resources. Based on this information, to what extent do you agree with the statement, 'Tuya is persistent in solving mathematics problems'?					
8	Tulgaa tries to complete mathematical tasks when the answers or solutions are not readily available but gives up when the task is too difficult. He gets off the task easily. He draws on a limited range of resources. He does not put in enough effort to solve mathematics problems. Based on this information, to what extent do you agree with the statement, 'Tulgaa is persistent in solving mathematics problems'?					
9	Oyunaa stays on a mathematical task no matter how difficult it is to find the answers. She searches for more information to clarify if she faces difficulties in solving the mathematics problem. Oyunaa always keeps trying a difficult mathematics problem until she has solved it. Based on this information, to what extent do you agree with the statement, 'Oyunaa is persistent in solving mathematics problems'?					
10	Bataa tends to disagree with others and, as a result, often starts quarrels. Therefore, he prefers to work on his own in mathematics even if he is stuck with a problem. He thinks that working with others in mathematics class does not help to perform better in mathematics. Based on this information, to what extent do you agree with the statement, 'Bataa is a cooperative learner in solving mathematics problems'?					
11	Solongo does not really like working in a group during maths class. However, sometimes she thinks it is helpful to discuss with others when she is stuck when solving a					

	mathematics problem. Based on this information, to what extent do you agree with the statement, 'Solongo is a cooperative learner in solving mathematics problems'?					
12	Jargal finds it easy to cooperate with others. He thinks that it is a good idea to combine all the students' ideas in a group when they work on a mathematics project. He also enjoys helping others to work well in a group on mathematics and listening to how others solve mathematics problems. Therefore, Jargal thinks he could do better in mathematics when he works with other students. Based on this information, to what extent do you agree with the statement, 'Jargal is a cooperative learner in solving mathematics problems'?					
13	Zaya feels that he cannot do it at all when he is assigned to solve problems like finding the size of angle x in a given figure. Generally, he often thinks, 'I can't do it', when he is assigned to solve geometry problems. Based on this information, to what extent do you agree with the statement, 'Zaya is confident in solving geometry problems'?					
14	Ganaa feels moderately certain that he can solve geometry problems like finding the area of a triangle, but he feels somewhat unsure whether he can continue with some of the procedures to find the answer. Based on this information, to what extent do you agree with the statement, 'Ganaa is confident in solving geometry problems'?					
15	Delger feels highly certain that she can accomplish the task when she is assigned to solve a geometry problem like finding the height of a pyramid. Most of the time,					

	she feels geometry problems too easy and unchallenging. Based on this information, to what extent do you agree with the statement, 'Delger is confident in solving geometry problems'?					

PART III: PROBLEM-POSING TASKS

Task 1: Make as many groups of numbers as you can, using the numbers given below. Label each group with its characteristics.

2, 3, 7, 9, 13, 15, 17, 25, 36, 39, 49, 51, 60, 64, 91, 119, 121, 125, 136, 143, 150

Instructions: You can use each number in more than one group. Each group should contain more than two numbers.

Task 2 Last night, there was a party at your cousin's home, and the door opened ten times. The first time the door opened, only one guest arrived. Each time the door opened, three more guests arrived than had arrived on the previous opening. Ask as many questions as you can that are in some way related to this problem.

Task 3 In the picture below, there is a triangle and an inscribed circle. Make up as many problems as you can that are in some way related to this picture.

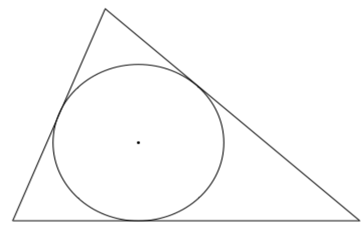
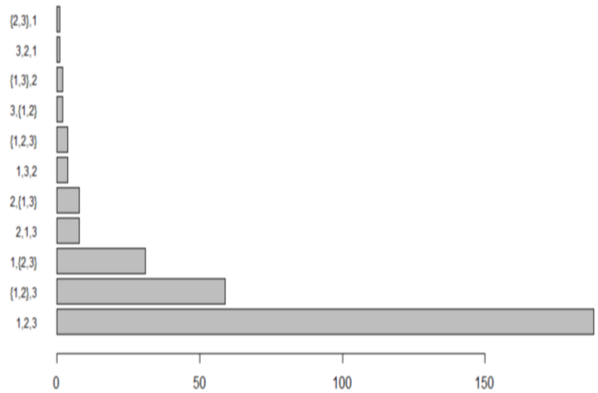


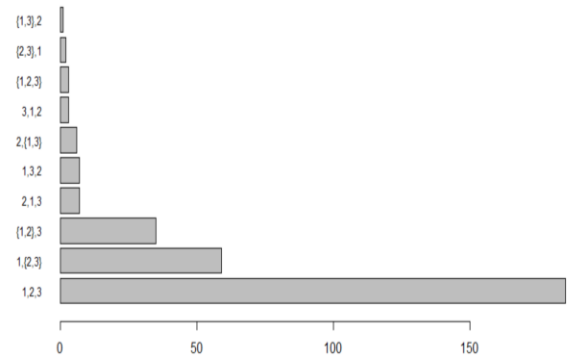
Figure 8. A semi-structured problem-posing situation example.

Task 4 Naran is three times older than Gerel. The sum of their two ages is 48. What are their ages? Make up questions similar to the example, which require you to find two unknowns. Make sure the information provided in your question is adequate to find the desired unknowns. Make sure the problem is correct. You do not have to write the answer to the questions.

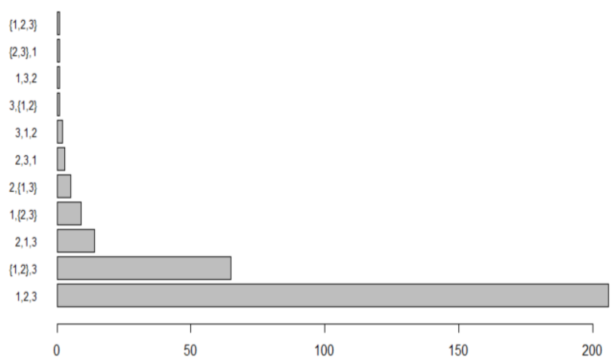
Distribution of the vignette responses



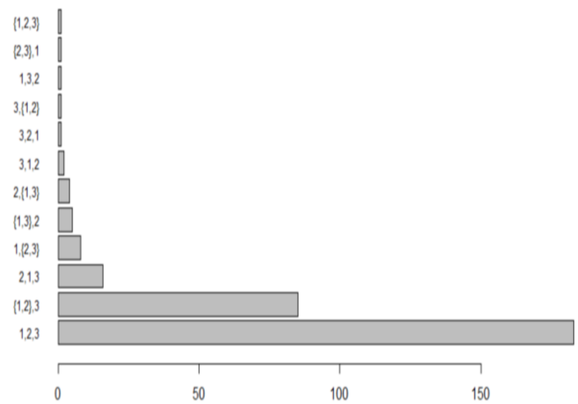
Vignette set for Mathematical perseverance



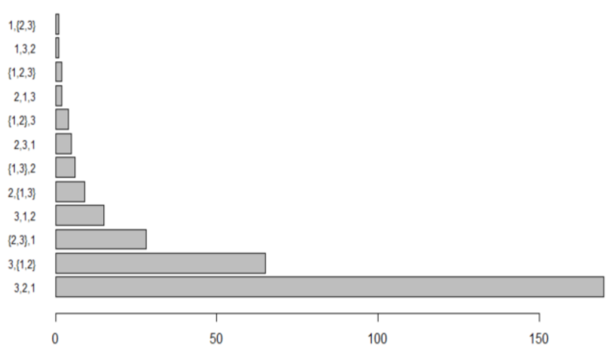
Vignette set for Cooperative learning in mathematics



Vignette set for Mathematical enjoyment



Vignette set for Mathematical self-efficacy



Vignette set for Mathematical anxiety