JIRSEA

JOURNAL OF INSTITUTIONAL RESEARCH SOUTH EAST ASIA

JIRSEA Issue: Vol. 22 No. 1

JIRSEA/UPM Special Issue

April 2024

ISSN 1675-6061

http://www.seaairweb.info/journal/index.aspx

Submission Timeline

Received: 30/10/2023

Revised as per Preliminary Review: 15/01/2024

Final Revision & Acceptance: 15/03/2024

Validity of STEM-Based Modelling Instrument for Pre-Service Teachers of Mathematics Education

Riyan Hidayat

CITE ARTICLE: Hidayat, R. (2024). Validity of STEM-Based Modelling Instrument for Pre-Service Teachers of Mathematics Education, *Journal of Institutional Research South East Asia*, 22(1), 154-178

SEAAIR
SOUTH EAST ASIAN ASSOCIATION FOR INSTITUTIONAL RESEARCH

UPM
UNIVERSITI PUTRA MALAYSIA
BERILMU BERBAKTI

Publisher: SEAAIR Secretariat

C/O Suan Dusit Rajaphat University
295 Nakhon Ratchasima Rd, Dusit, Dusit District, Bangkok 10300, THAILAND
email:seaair.info@gmail.com
http://www.seaairweb.info/

VALIDITY OF STEM-BASED MODELLING INSTRUMENT FOR PRE-SERVICE TEACHERS OF MATHEMATICS EDUCATION

Riyan Hidayat

Faculty of Educational Studies, Universiti Putra Malaysia, Malaysia

Institute for Mathematical Research, Universiti Putra Malaysia

riyan@upm.edu.my

ABSTRACT

Mathematical modelling involves a set of interconnected skills that allow students to translate real-world situations into mathematical representations and vice versa. This research aimed to develop the instrument consisting of sub-competency; simplifying, mathematising, computing, interpreting and validating. The study utilized a cross sectional survey research design. A total of 135 pre-service teachers of mathematics voluntarily selected using convenience sampling methods, participated in this study. The data was analysed by item-CVI (I-CVI), S-CVI and exploratory factor analysis (EFA). This tool's preliminary versions demonstrated good content validity for both the individual items and overall, for STEM-based material (S-CVI/UA = 0.88; S-CVI/Ave = 0.93). The Kappa value was K = 0.88, indicating an excellent value of validity. The development of STEM-Based modelling instrument revealed great item-content validity and scale-validity for assessing sub competency in pre-service teachers. At the same time, EFA revealed that STEM-based modelling instrument had five sub-components; simplifying, mathematising, computing, interpreting and validating. The results showed that the STEM-based modelling instrument's reliability was good. Creating and verifying the STEM-focused modelling tool designed for pre-service teachers is essential in mathematics education and research. It offers a valuable means to evaluate and improve the essential skillset related to mathematical modelling, benefiting educators and STEM students alike. Moreover, the research's approach and discoveries have the potential to influence forthcoming investigations and educational methodologies across diverse fields.

Keywords: factor analysis, modelling, STEM, validation

1. Introduction

Science, Technology, Engineering and Mathematics (STEM) education plays a significant role in influencing cultural and economic growth, embracing innovation and caring about creativity and problem-solving (Cooper & Hearverlo, 2013). One of the most important tools for transition in STEM education is mathematical modelling. As pointed out by Minarni and Napitupulu (2020), students can apply modelling abilities to describe context problems mathematically, organize tools, discover relationships, transfer between real-world and mathematical problems, and visualize problems in various ways. Mathematical modelling encompasses a range of interconnected abilities that enable students to convert real-life scenarios into mathematical forms and vice versa. Mathematical modelling applications are composed of concepts related to different disciplines by their nature. Anhalt et al., (2018) indicate that mathematical modelling is a method in which students think about and make sense of a real-life issue which will be analysed using mathematics in order to comprehend, explain or predict something. Modelling exercises help students improve their conceptual knowledge and the procedures they establish while attempting to address a real-world situation. Thus, teachers cannot rely solely on textbooks, since the majority of the exercises are designed to engage modelers in the use of models but not to construct their own models for specific scenarios. An authentic problem is also known as a project that students perform which are relevant to them, as opposed to assignments which are unrelated to any type of work that would be done outside of the classroom. Kaiser et al. (2011) revealed that students did not see the necessary type of mathematics which can be used in real situations. They did not want to perform well in mathematics because their focus is not apparent. In addition, students always imagine that mathematics is a strenuous subject. Therefore, the teacher or pre-service teacher should find the way to assists student to perform well in all subject by doing modelling questions. To utilize mathematical modelling in the classroom, pre-service teachers must first comprehend the significance of mathematical modelling in STEM. Pre-service teachers must plan their teaching skills to develop STEM based mathematical modelling that requires sub competency in simplifying, mathematising, computing, interpreting and validating the solution in order to achieve the students with six key skills.

Mathematical modelling consists of holistic and atomistic approach. Holistic approach is an overall concept of mathematical competence where the subject is taken as a whole instead of through the individual parts that make it up. Alternatively, an atomistic approach focuses on specific stages of the modelling process, particularly the mathematizing and analysis of models. It can be said that the atomistic approach is a detailed approach, while holistic approach is an overall approach. Cevikbas et al. (2021) show that the holistic approach accommodates eight mathematical competencies. The holistic definition refers specifically to the term ‘modelling competency’ as a ‘distinct, recognised, and more or less well-defined thing’. Brand’s (2014) study revealed the effect of a holistic versus an atomistic modelling approach on students' mathematical modelling competencies. The result showed that the holistic approach seemed superior to the atomistic approach because students of the comprehensive schools acquired higher performance in the holistic group. A study by Hankeln et al. (2019) developed a new test instrument and evaluated sub-competencies in mathematical modelling. They employed atomistic test questions designed to evaluate each sub-skill of mathematical modelling individually. The findings revealed that test instruments which evaluated mathematical modelling sub-competencies incorporated several sub-competencies rather than treating them as distinct dimensions of a larger general modelling competency. In this study, we used an atomistic test consisting of five sub-competencies which were simplifying, mathematising, computing, interpreting and validating. This is because the atomistic test was more suitable to test the sub-competency in a more detailed manner. The pre-service teacher can be tested by referring to five sub-competencies about STEM-based modelling which include STEM subjects such as biology, physic, chemistry, probability, mathematical reasoning. Moreover, in a systematic literature review conducted by Hidayat et al., (2022), the holistic approach was used by the majority of scholars in the examined study to evaluate the modelling proficiency of pre-service mathematics teachers.

To date, various tools have been employed in recent studies to assess mathematical modeling competence (Haines & Crouch, 2001; Hankeln et al., 2019; Zöttl et al.,2011). Despite the availability of several modern instruments designed for evaluating mathematical modeling competence, these tools exhibit diversity in application contexts and lack a specific emphasis on pre-service mathematics education teachers. The current instruments lack an emphasis on STEM content, processes, and contexts, implying a significant oversight, as the intricate interplay of science, technology, engineering, and mathematics is not adequately addressed. There is a need for assessment tools that go beyond conventional approaches, ensuring a comprehensive understanding and measurement of competencies within the specific domains of STEM education. This article reported the findings from a study involving pre-service teachers in Malaysia on the development of STEM-based modelling proficiency. The main objectives were to develop and validate STEM-based modelling instrument for pre-service teachers of mathematics education. To accomplish this, we set out to discover how pre-service teachers' understanding of STEM-based modelling was manifested in their work through the STEM-based modelling instruments, which included five sub-competencies: simplifying, mathematising, computing, interpreting and validating. As far as the researcher's understanding is concerned, there are limited past studies which develop and validate STEM-based modelling instruments for pre-service mathematics education teachers.

This study contributes to the literature on the development and validation of STEM-based modelling instruments for pre-service mathematics teachers. As such, it can also be used as a benchmark process for the assessment process of mathematical modelling for pre-service mathematics education teachers. Therefore, this study aimed to develop and validate a STEM-based modelling instrument for pre-service mathematics education teachers.

2. Theoretical Perspectives

2.1. Mathematical Modelling

The utilization of mathematical modeling is pivotal in nurturing students' mathematical skills as it enables them to apply theoretical concepts to practical scenarios. This method fosters meaningful engagement with mathematical principles, thereby enhancing comprehension and problem-solving capabilities (Albarracín & Gorgorió, 2020). Through mathematical modeling, students deepen their understanding of the world, thus fueling their learning motivation and proficiency in tackling real-world issues (Kurniadi et al., 2022). Additionally, integrating mathematical modeling into education positively impacts the development of students' creativity and problem-solving prowess (Salingkat & Bilalu, 2021).

Although model and modelling have different connotations, they both serve as important tools for problem solving, forecasting, decision-making, and communication which have been researched and examined in engineering science as well as in the history, philosophy, and sociology of science and technology (Muller, 2009). Hallstrom and Ankiewicz (2019) indicated that models can be anything from simple conceptual sketches and crude prototypes to advanced mathematical models that indicate something about reality. Thus, prediction requires correlation but not causal connection. The capacity to construct, utilize, apply, assess, and revise models is a vital skill for gaining a thorough grasp of technology development processes and scientific practice, as well as a key component of pursuing real learning in technology, math, and science classrooms. Besides, mathematics modeling includes the process, teacher preparation, and theoretical framework that captures mathematical modeling development through a series of sub-competencies which were utilized to monitor modeling activity and cumulatively build modeling competency (Maaβ, 2006). National Council of Teachers of Mathematics (NCTM) indicated that mathematical practices can be used to observe pre-service teacher's preparation as characterized in the Common Core (Common Core State Standard Initiative, 2010).

Students are required to interpret, describe, explain, justify, reject, or revise model in modeling activities (English, 2003). They must reduce an actual scenario by making justified assumptions and identifying those factors which they believe are important, resulting in an idealized representation of reality (Kaiser & Stender, 2013). Real model is also known as simplified reality (Borromeo, 2006). A real model can be represented mathematically with equations, numeric tables, diagrams, or other relevant representations to answer a mathematical question. One is required to solve the model, which must be interpreted considering the original circumstance. In this situation, the original selections must be altered in order to develop an improved model which leads to a better conclusion using a similar method. Referring to Common Core State Standard in Mathematics (CCSSM), the cycle diagram has been modified by the author in a particular way. We did not utilize the situation or problem as sub-competencies.

2.2. STEM-based Modelling

A positive advantage of mathematical modeling activities is its ability to challenge problem solvers and help them learn in general in the STEM field (Chamberlin et al., 2020). For example, technology and engineering have a modeling sub-domain (Arikan et al., 2020) where modeling is very important for predictive analysis and design level testing in an engineering context (Fan et al., 2020). Furthermore, English (2017) proposed STEM-based modeling with the aim of discussing the competence of mathematical modeling in the perspective of STEM education, STEM integration approach, representation of STEM disciplines and equity in access to STEM education. Blomhøj (2009) states that discussions about models, modeling, modeling processes, modeling competencies and applications are important aspects of the study under the perspective of the educational modeling framework. Therefore, one example of a good study in an educational perspective about mathematical modeling competence is the framework proposed by Stillman et al. (2007). The modeling process in this framework is different from other modeling processes, because it includes the metacognition process of each transition of the mathematical modeling process.

English (2017) proposed STEM-based modeling as a cyclical generic (generative) learning activity where the modeling and engineering processes share the same characteristics and facilitate authentic problem solving involving the content, process and context of STEM. This generative concept is in line with the concept of emergent modeling proposed by Gravemeijer (2008). It refers to the characteristics of the problem where the learning content or process is acquired by the student, rather than provided by the teacher. For France (2018), models and modelling techniques can, through genuine experience, bridge the gap between STEM fields. Subsequently, models and modeling should be used as tools to promote STEM literacy and the transfer of knowledge and skills between contexts, both within and outside the STEM discipline (Hallström, & Schönborn, 2019). Therefore, it must be seen as a basic component of STEM literacy (Williams, 2017).

STEM professionals, from a broader political viewpoint, are those who possess the requisite skills in science, technology, engineering, and mathematics, which may appear to be straightforward to acquire (Hidayat & Wardat, 2023). The Malaysia Education Blueprint 2013-2025 had invested in innovation for future generations of STEM professionals by inaugurating three stages to reinforce STEM based education in schools. However, from an educational standpoint, we must define what STEM education entails, how it should be thought, and how it may be implemented (Kertil & Gurel, 2016). In other words, learning information, attitudes, and skills to spot real-world problems through an awareness of the characteristics of the STEM courses should be connected to both national economic growth goals and individual student development (Hallström & Schönborn, 2019). Sanders (2009) stated that STEM literacy aims for broad educational objectives, but these objectives must solve real-world concerns by combining two or more STEM fields. Creating authentic learning scenarios is perhaps one of the most difficult aspects of STEM literacy education initiatives. The fundamental characteristics of authenticity, according to Herrington and Parker (2013), include a genuine setting, an authentic task, the presence of expert performances, different viewpoints, cooperation, reflection, articulation, metacognitive assistance, and authentic assessment. Mathematical modeling is a technique involved in all STEM-related applications. All STEM activities are not modeling activities, but many of them allow students to gain expertise with the mathematical modeling process. By using STEM-based mathematical modeling context, pre-service teachers can optimize the knowledge to implement STEM based in class using mathematical modeling instruments. The teachers can implement the mathematical modeling in STEM subject such as biology, physics, mathematics, chemistry, and other related subjects.

Mathematical modeling competency within the realms of STEM content, processes, and contexts is systematically developed by leveraging established theories and insights gleaned from previous research endeavors. Earlier investigations have predominantly employed the Realistic Mathematics Education (RME) theory and the Model and Modeling Perspective (MMP) as foundational frameworks to enhance mathematical modeling competence. However, insights from Carreira and Baioa (2011) suggest a convergence between these theoretical perspectives, emphasizing shared similarities while delineating a limited number of distinctions. In the context of this study, we draw inspiration from the educational modeling paradigm proposed by Stillman et al. (2007), offering a cohesive integration of STEM content, processes, and contextual considerations. A thorough examination of relevant theories and preceding studies forms the basis of this research, culminating in the identification of a singular dimension within the study's scope: the STEM-based modeling instrument. This instrument encompasses a nuanced perspective with five distinct sub-dimensions, namely simplifying, mathematizing, computing, interpreting, and validating. Each of these sub-dimensions contributes uniquely to the multifaceted landscape of STEM-based mathematical modeling, reflecting the intricate interplay between theoretical underpinnings and practical applications in educational settings. In this study, we have developed a research framework which draws upon theoretical underpinnings and prior research findings (see Figure 1). This theoretical framework not only guides our current research endeavors but also serves as a foundation for future explorations in the field.

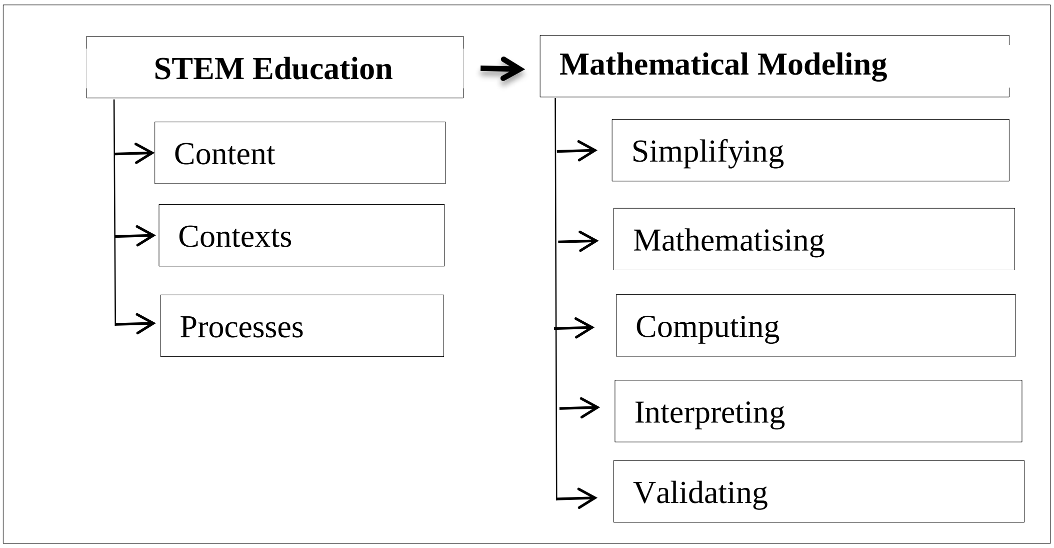

Figure 1: Research framework

2.3. Assessment in Mathematical Modelling

Mathematical modeling is defined as the process of using mathematics to depict, analyze, predict, or otherwise provide insight into real-world occurrences. Mathematical modeling offers far more ‘potent and successful strategies to help students become (a) better problem solvers and (b) better equipped to use mathematics in real-life circumstances outside of school. In mathematical modeling, one identifies a scenario in the actual world, makes certain assumptions, and then utilizes a mathematical model to produce a mathematical formulation to obtain conclusions that can be translated back into the real world to validate the practicality of the result. Furthermore, using a real-world issue, students construct assumptions, use a model to obtain a mathematical formula, and apply mathematical tools to that formula to obtain a reasonable result. Blum and Leiss (2007) also stated that mathematical modeling is a process in which an issue in the real world is identified, certain assumptions and choices are made, and then a mathematical model is used to generate a solution that can be translated back into the real world. Modeling is an iterative process that consists of the following stages: (a) comprehending the phenomenon, (b) developing a physical representation or model, (c) mathematizing the phenomenon and doing computations, (d) interpreting results, (e) validating them in the context of the real world, and (f) spreading them through debates and writing. Cevikbas at al. (2021) indicated that before developing mathematical modeling, the researcher must conceptualize modeling competencies using particular questions: Is the empirical research reflecting the theoretical views defined and expressed in the theoretical frameworks? Modeling skills are understood as an atomistic construct, or are they differentiated as analytic constructions using distinct sub-competencies? Next, which measures for measuring modeling competences have researchers used to study the modeling competencies of (pre-service) teachers or school students? What equipment and data collection procedures were utilized, which groups were targeted, and how large were the sample sizes? Lastly, which strategies for creating and monitoring modeling competencies have researchers utilized to support the modeling competencies of (pre-service) teachers or school students? In this section, we have to consider the assessment of mathematical modeling before developing a mathematical model to answer research questions.

Recent studies (Haines & Crouch, 2001; Hankeln et al., 2019; Xu et al., 2023; Zöttl et al., 2011) have utilized a range of tools to assess mathematical modeling competence. For example, Xu et al. (2023) present a cognitive diagnostic analysis of students' mathematical competency, which can offer valuable insights into developing a framework for assessing preservice teachers' mathematical modeling competency. However, even with the advancements in assessment tools, there persists a conspicuous void in resources customized explicitly to assess the proficiency of pre-service mathematics education teachers. The current array of evaluation instruments demonstrates a wide spectrum of application contexts, yet frequently overlooks a concentrated emphasis on the integration of STEM content, processes, and contexts. This gap is particularly consequential due to its failure to comprehensively acknowledge the intricate interconnectedness inherent in science, technology, engineering, and mathematics. The absence of emphasis on STEM domains in current assessment tools highlights the need for instruments that surpass conventional approaches. It is crucial to develop assessment tools that offer a more comprehensive understanding and measurement of competencies within the specific realms of STEM education. Such tools should account for the intricate interplay between these disciplines, ensuring that future educators are adequately prepared to teach and integrate STEM concepts effectively. Expanding assessment frameworks to incorporate STEM content, processes, and contexts will better equip pre-service mathematics education teachers to meet the evolving demands of modern education. Therefore, bridging this gap becomes imperative for fostering well-rounded educators equipped to meet the evolving needs of STEM education. The proposed study seeks to address this deficiency by creating and employing an innovative evaluation framework customized for assessing the proficiency of pre-service mathematics education teachers within the STEM environments.

3. Method

3.1. Participant and Design

The research employed a quantitative research model within the framework of a cross-sectional survey research design (Creswell, 2012). In the realm of research, a quantitative research model serves as a practical and systematic method for delving into and comprehending various phenomena. This approach involves the meticulous collection and interpretation of numerical data, providing a robust foundation for analysis. This approach involved collecting data at a single point in time to provide a snapshot of the research variables under investigation. The cross-sectional survey research design facilitated the systematic gathering of quantitative data, allowing for the analysis of relationships and patterns within a specified timeframe. Given the nature of the research and its aim for inclusivity, we employed simple random sampling techniques to meticulously identify and invite prospective mathematics teachers to partake in the online survey. Prospective mathematics teachers in this study pertains to college students enrolled in mathematics education programs who possess comparable modeling experiences. These individuals represent aspiring mathematics educators who are being equipped to teach mathematics at the secondary school level.

A commendable total of 135 prospective mathematics teachers willingly participated in this project, contributing to the richness of our dataset with their diverse perspectives and insights. It is worth noting that the decision to work with a sample size of 135 was purposeful, considering the unique context of developing a novel measurement scale. This size was deemed appropriate to ensure a robust exploration of the self-efficacy dimensions within the specific realm of STEM-based mathematical modeling. The voluntary engagement of this sizable cohort allowed for a comprehensive and well-rounded understanding of the subject matter. Within this participant pool, it is noteworthy that the majority, precisely 135 pre-service teachers, were women. This demographic composition introduces an additional layer of insight, potentially shedding light on gender-specific perspectives and experiences related to self-efficacy in STEM-based mathematical modeling. The diversity within the sample not only enhances the external validity of the findings but also opens avenues for nuanced analyses and interpretations based on gender dynamics within the field of prospective math educators.

3.2. Instruments

The generated items were derived from a thorough examination of pertinent studies and literature focusing on mathematical modeling. To ensure a comprehensive understanding, a systematic review was undertaken to identify the definitions and existing descriptions of STEM-based mathematical modeling. This involved delving into relevant literature to extract valuable insights, ultimately leading to the identification of four primary dimensions integral to STEM-based mathematical modeling. In the process of scale development, we seamlessly integrated insights from previous literature and theories pertaining to mathematical modeling and self-efficacy contexts. This holistic approach not only informed the creation of the scale but also enriched its foundations with a nuanced understanding gleaned from the broader academic landscape.

The meticulous examination of the item set was initiated with the explicit goal of addressing concerns related to content validity. The overarching objective was to ensure a comprehensive representation of both the theoretical foundations and empirical aspects inherent in STEM-based mathematical modeling. To achieve this, an initial collection comprising 25 items was carefully crafted, strategically employing the principles of mathematical modeling to delve into the nuances of STEM-based mathematical modeling. The conceptual framework for the inaugural version of the STEM-based Mathematical Modeling found its roots in pertinent literature, particularly delving into mathematical modeling (Maaß, 2006) and the models and modeling perspective (MMP). This foundational knowledge served as the bedrock for creating a robust scale that encapsulated the multidimensional nature of self-efficacy in the context of STEM-based mathematical modeling. In the process of generating items for STEM-based mathematical modeling, researchers assumed the responsibility of adhering to the definition and dimensions inherent in STEM-based mathematical modeling. This involved utilizing the established framework as a guide for item development. The formulation of these items was a nuanced process, drawing insights from both qualitative data and existing instruments found in the expansive body of literature. This comprehensive approach aimed to capture the intricacies of self-efficacy within the dynamic realm of STEM-based mathematical modeling.

The STEM-based mathematical modeling was multidimensional with five sub-competencies including simplifying, mathematising, computing, interpreting and validating. The 25 questions on the five scale were changed using item analysis and factor analysis, as well as cognitive evaluations with subject matter experts. Each item was constructed using 5 multiple choice question with score 0 to score 4. Score 0 represents an incorrect answer followed by each score and the correct answer is represented by score 4. Examples of items on the scale included the following: (1) pre-service teacher learn to make assumptions for the problem and simplify the problem; (2) pre-service teacher learn to mathematize relevant quantities and their relations; (3) pre-service teacher learn to use mathematical knowledge to solve the problem; (4) pre-service teacher learn to interpret mathematical results in a real situation; and (5) pre-service teacher learn to reflect other way to solve the problem if solution can be developed differently. The content experts contributed feedback on the items' applicability, sufficiency, accuracy, and language in order to establish the scale's content validity. The selection of the subject matter specialists was made using purposeful sampling strategies according to their availability, accessibility, and expertise in mathematical modelling. Four mathematics professors with a variety of interests in mathematical modelling and operations research were among the subject matter experts. Three full professors with PhDs in mathematics education and expertise in mathematical modelling were also included.

3.3. Data Collection and Analysis

Potential respondents received invitations to participate in the online poll using Google Form. This approach involved distributing invitations to individuals identified as potential participants in the study. These invitations typically contained a concise description of the survey's objectives, guidance on accessing and completing the questionnaire via Google Form, and details regarding confidentiality and data handling practices. Upon receiving the invitation, participants could access the survey link provided and respond to the questionnaire at their convenience. For statistical analysis, their responses were downloaded and coded. The data analysis for this study's primary objective was to produce solid evidence to support this new assessment scale for evaluating pre-service instructors. The survey data was only acceptable to the degree to which they are determined valid and reliable. Before beginning the survey, the participants were informed that the study was optional and anonymous and they were also given informed consent. In the process of gathering data, the Human Research Ethics approved this study as ethical. The respondents' answers to a STEM-based modelling instrument and demographic data (such as gender and age) were collected. Each participant's data gathering process took about 30 minutes to complete.

The data analysis for this research was done in stages. Firstly, in this study, the index of content validity (CVI) was determined empirically. Item-CVI (I-CVI) can be used to calculate an instrument's content validity (Zamamzadeh et al., 2015). I-CVI is computed as number of experts providing a rating of 'strongly agree' for each item divided by the total number of experts (Rodrigues et al., 2017). The item is considered acceptable when I-CVI > 0.79 when the values range from 0 to 1. Meanwhile, when I-CVI is between 0.70 and 0.79, the item needs to be revised, and if the value of I-CVI is below 0.70, the item is eliminated. Similarly, S-CVI is calculated using the number of items in a tool which have achieved a rating of 'very relevant'. There are two methods to calculate S-CVI. Firstly, we can utilise Universal Agreement (UA) among experts (S-CVI/ UA). Secondly, the Average CVI (S-CVI/Ave) (Zamamzadeh, et. al, 2015). S-CVI/UA is calculated by adding all items with I-CVI equal to 1 divided by the total number of items, while S-CVI/Ave is calculated by taking the sum of the I-CVIs divided by the total number of items. The output of S-CVI for S-CVI/UA $\geq$ 0.8 and a S-CVI/Ave $\geq$ 0.9 have excellent content validity (Shi et al., 2012). In this study, we utilised S-CVI/Ave as it was easier to calculate. In this study, the result for I-CVI was that 95% acquired the value of 1. Only four questions acquired 0.75 in I-CVI but they were still accepted. Meanwhile, the value of S-CVI was 0.96 which was a higher result for the overall items.

The second phase of our analytical approach involved the utilization of SPSS version 23.0 to conduct exploratory factor analysis (EFA). This statistical technique was employed to delve into the intricate structure of our data, aiming to discern the underlying factors that contribute to the complexity of the STEM-based modelling instrument. EFA served as a powerful tool to scrutinize the interrelationships among variables, unveiling the latent factors that emerged from the set of items designed to measure STEM-based modeling competency. To determine the number of factors, several key metrics were scrutinized. The Kaiser-Meyer-Olkin (KMO) value was assessed to gauge the adequacy of the sample for factor analysis. Factor loading, Bartlett's test of sphericity, scree plot, and eigenvalues were also pivotal elements in this analytical process. These indicators collectively contributed to a comprehensive understanding of the underlying structure, enabling us to identify and interpret the primary dimensions shaping the STEM-Based Modelling instrument. In a complementary approach, we employed EFA with varimax rotation. This rotation method was chosen to enhance the interpretability of the factors and simplify the structure, facilitating a clearer representation of the theoretical underpinnings of the STEM-Based Modelling instrument. Through this multifaceted analysis, we aimed to not only uncover the inherent factors within our measurement scale but also to gain deeper insights into the theoretical foundations that govern STEM-based modeling competency.

4. Results

The instrument was developed consisting of 25 questions with 5 questions in each sub-competency (simplifying, mathematising, computing, interpreting and validating). Each sub-competency shows the level of difficulty in solving the question. For example, the question in the simplified sub-competency was much easier than the mathematical model sub-competency, the question in the mathematical model sub competency was easier than the question in the mathematical question sub-competency and its follow-on result of question in the validation sub-competency was more strenuous.

4.1. Validity of STEM-Based Modelling instrument

Table 1: The I-CVI Result for Each Item

Questions	I-CVI
1	$\text{I-CVI} = \frac{\text{Total of experts give score 3 and 4}}{\text{Total experts}} = \frac{4}{4} = 1.00$
2	$\text{I-CVI} = \frac{\text{Total of experts give score 3 and 4}}{\text{Total experts}} = \frac{4}{4} = 1.00$
3	$\text{I-CVI} = \frac{\text{Total of experts give score 3 and 4}}{\text{Total experts}} = \frac{4}{4} = 1.00$
4	$\text{I-CVI} = \frac{\text{Total of experts give score 3 and 4}}{\text{Total experts}} = \frac{4}{4} = 1.00$
5	$\text{I-CVI} = \frac{\text{Total of experts give score 3 and 4}}{\text{Total experts}} = \frac{4}{4} = 1.00$
6	$\text{I-CVI} = \frac{\text{Total of experts give score 3 and 4}}{\text{Total experts}} = \frac{3}{4} = 0.75$
7	$\text{I-CVI} = \frac{\text{Total of experts give score 3 and 4}}{\text{Total experts}} = \frac{4}{4} = 1.00$
8	$\text{I-CVI} = \frac{\text{Total of experts give score 3 and 4}}{\text{Total experts}} = \frac{3}{4} = 0.75$
9	$\text{I-CVI} = \frac{\text{Total of experts give score 3 and 4}}{\text{Total experts}} = \frac{4}{4} = 1.00$
10	$\text{I-CVI} = \frac{\text{Total of experts give score 3 and 4}}{\text{Total experts}} = \frac{4}{4} = 1.00$
11	$\text{I-CVI} = \frac{\text{Total of experts give score 3 and 4}}{\text{Total experts}} = \frac{4}{4} = 1.00$
12	$\text{I-CVI} = \frac{\text{Total of experts give score 3 and 4}}{\text{Total experts}} = \frac{3}{4} = 0.75$
13	$\text{I-CVI} = \frac{\text{Total of experts give score 3 and 4}}{\text{Total experts}} = \frac{4}{4} = 1.00$
14	$\text{I-CVI} = \frac{\text{Total of experts give score 3 and 4}}{\text{Total experts}} = \frac{4}{4} = 1.00$
15	$\text{I-CVI} = \frac{\text{Total of experts give score 3 and 4}}{\text{Total experts}} = \frac{4}{4} = 1.00$
16	$\text{I-CVI} = \frac{\text{Total of experts give score 3 and 4}}{\text{Total experts}} = \frac{4}{4} = 1.00$
17	$\text{I-CVI} = \frac{\text{Total of experts give score 3 and 4}}{\text{Total experts}} = \frac{4}{4} = 1.00$
18	$\text{I-CVI} = \frac{\text{Total of experts give score 3 and 4}}{\text{Total experts}} = \frac{4}{4} = 1.00$
19	$\text{I-CVI} = \frac{\text{Total of experts give score 3 and 4}}{\text{Total experts}} = \frac{4}{4} = 1.00$
20	$\text{I-CVI} = \frac{\text{Total of experts give score 3 and 4}}{\text{Total experts}} = \frac{4}{4} = 1.00$
21	$\text{I-CVI} = \frac{\text{Total of experts give score 3 and 4}}{\text{Total experts}} = \frac{4}{4} = 1.00$
22	$\text{I-CVI} = \frac{\text{Total of experts give score 3 and 4}}{\text{Total experts}} = \frac{4}{4} = 1.00$
23	$\text{I-CVI} = \frac{\text{Total of experts give score 3 and 4}}{\text{Total experts}} = \frac{4}{4} = 1.00$
24	$\text{I-CVI} = \frac{\text{Total of experts give score 3 and 4}}{\text{Total experts}} = \frac{3}{4} = 0.75$
25	$\text{I-CVI} = \frac{\text{Total of experts give score 3 and 4}}{\text{Total experts}} = \frac{4}{4} = 1.00$
TOTAL	24

The results showed that 90% of the experts agreed and strongly agreed for each item. In addition, all content validity (I-CVI, S-CVI and Kappa) were used to test the content validity of instruments. The I-CVI result referred to formats of the item including a 4-point Likert scale from strongly disagree, disagree, agree and strongly agree to test the content validity. The items generated were based on relevant literature and research on mathematical modelling. The results showed that I-CVI for each question was 1.00, but three items acquired a result of 0.75 I-CVI. Table 1 shows the I-CVI for each item.

The S-CVI/UA = 0.88 and the S-CVI/Ave = 0.93. The Universal Agreement was calculated by adding all I-CVI's equal to 1.00 (22 items) divided by 25, while the Average took the sum of all I-CVI (23.25) divided by 25. Overall, the Universal Agreement method and the Average approach showed high content validity of development of stem-based modelling instrument for pre-service teachers of mathematics education in Malaysia. Although CVI was extensively used to estimate content validity, according to chance agreement, this index did not consider the possibility of inflated values. Kappa provides the degree of agreement beyond chance, as calculated using the following formula: K = (I-CVI – Pc)/ (1- Pc), where Pc = [N! /A! (N-A)!]* 0.5^N. In this formula Pc = the probability of chance agreement; N = number of experts; and A = number of experts who agreed that the item was relevant. Kappa values above 0.74 were considered excellent, between 0.60 to 0.74 good and 0.40 to 0.59 fair (Landis & Kosh, 1977). The value of Kappa in this research was 0.89 which indicated excellent validity instruments.

4.2. Descriptive Analysis

The sub-dimensions for mean, standard deviation (SD), skewness, kurtosis, and inter-correlation were calculated (Table 2).

Table 2: Sub-Dimensions with Their Mean Values, SD, Skewness, and Kurtosis

Sub-dimensions	Mean	SD	Skew	Kurtosis	1	2	3	4	5
Simplifying	2.80	.83	.092	-.806	1	.207*	.210*	.211*	.349**
Mathematising	2.50	.92	.565	-.914		1	.361**	.268**	.214*
Computing	2.87	.66	-.387	.722			1	.351**	.193*
Interpreting	2.65	.85	.236	-1.115				1	.318**
Validating	2.76	.75	.123	-.512					1

Based on Table 2, all sub-dimensions and items had kurtosis and skewness values between 3 and +3. (Brown & Greene, 2006). The four sub-dimensions had modest to strong correlations with all scale items, and all scale items had substantial relationships (ranging from $r = .19$ to $r = .36$, p < =.05). Since none of the connections were greater than 90, multicollinearity was not present (Kline, 2005). The mean score differed in each sub-dimension, with M = 2.80 and SD = .83 for simplifying, M = 2.50 and SD = .92 for mathematising, M = 2.87 and SD = .66 for computing, M = 2.65 and SD = .85 for interpretating, and M = 2.76 and SD = .75 for validating. At the same time, given that the values for kurtosis and skewness ranged from -1.96 to +1.96, the data met the assumption of normality.

4.3. Exploratory Factor Analysis

EFA was utilized as the first phase of the empirical technique to examine the pattern and linkages among the components. The EFA's findings suggested that four factors accounted for 70.86% of the variance (Table 3). The Kaiser-Meyer-Olkin Measure of Sampling Adequacy was .81, showing that the sample was suitable, and Bartlett's Test of Sphericity provided a p-value of <0.001. These scores were not enough (Osbourne, 2015). Therefore, KMO, factor loading, Bartlett, scree plot, eigenvalue, and varimax rotation were used in the current work.

Table 3: The Results of the EFA

Sub-Dimensions	Items	Communalities	Eigenvalue	% of Variance	Loading Factor				
					1	2	3	4	5
Simplifying	A1	.80			.87				
	A2	.74			.84				
	A3	.70	7.13	28.52	.82				
	A4	.72			.76				
	A5	.74			.83				
Mathematising	B1	.72				.82			
	B2	.76				.85			
	B3	.68	3.37	13.47		.80			
	B4	.69				.80			
	B5	.76				.82			
Computing	C1	.53							.67
	C2	.69							.76
	C3	.61	1.88	7.54					.74
	C4	.66							.78
	C5	.60							.70
Interpreting	D1	.65						.78	
	D2	.69						.80	
	D3	.71	2.35	9.41				.80	
	D4	.65						.73	
	D5	.65						.77	
Validating	E1	.64					.76		
	E2	.76					.82		
	E3	.75	2.73	10.92			.83		
	E4	.73					.83		
	E5	.74					.82		

The communalities for these 25 questions ranged from .54 to .80 based on Table 3. The first factor, which accounted for 28.52% of the variance, was the simplifying factor. The second factor, which accounted for 13.47% of the variance, was the factor of mathematising. The third factor, which accounted for 10.92% of the variance, was the validating. The fourth factor, which accounted for 9.41% of the variance, was the factor of interpretation. The final factor, which accounted for 7.54% of the variance, was the computing factor. The item with the highest loading factor was A1 (.87), whereas the items with the lowest loading factors were C1 (.67). All of the items' factor loadings, however, exceeded .50. Cross-loadings were absent from the current work. The scree-plot test supported the decision to preserve four components, hence the research kept five elements overall (Figure 2).

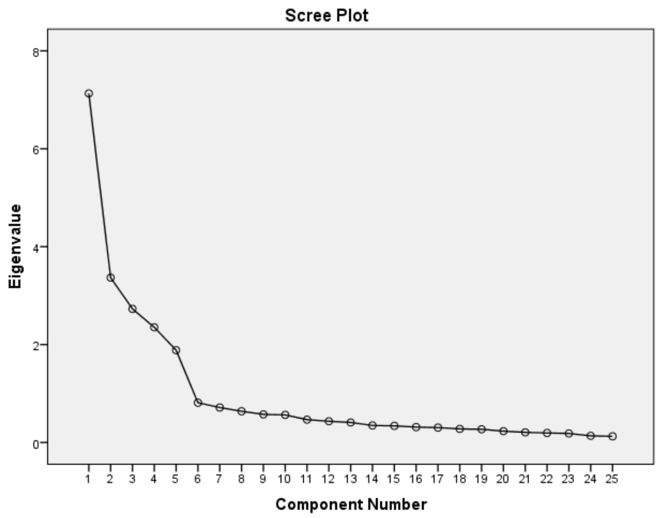

Figure 2: Scree plot of the 25 items for STEM-based Modelling Instrument

4.4. Reliability of Instrument

Reliability is defined as the consistency and stability of the results obtained (Creswell, 2012). When researchers administer the test numerous times during different eras, scores should be nearly comparable. We examined the reliability of the STEM-based modelling scale (simplifying, mathematising, computing, interpreting and validating) and overall STEM-based modelling items for the total respondents (N = 135) (Table 4). In the current work, internal consistency values were: a) simplifying: $\alpha = .92$, b) mathematising: $\alpha = .96$, c) computing: $\alpha = .91$, d) interpreting: $\alpha = .91$, and e) validating: $\alpha = .97$. The STEM-based modelling scale was a good Cronbach alpha coefficient (Hair et al., 2015). The AVE values varied from .67 to .79, all over 0.5, indicating that each dimension had good internal consistency (Hair et al., 2017) and supporting convergent validity (Fornell & Larcker, 1981). Furthermore, MSV and ASV scores were lower than AVE ratings, showing that the STEM-based modelling had strong discriminant validity. Composite reliability for the STEM-based modelling ranged from .91 to .97, indicating good internal consistency.

5. Discussion

This study developed and provided content validity of the STEM-based modelling instrument for pre-service teachers of mathematics education which consisted of five sub-competencies related to Science, Technology, Engineering and Mathematics (STEM). Nowadays, most of the pre-service teachers focus on non-subjects related to STEM because their mindset was STEM subjects and question would be much more difficult than subjects which were not related to STEM. Unfortunately, they did not realize the importance of STEM in this era. As such, this research developed an instrument which could be used to test pre-service teachers answering the items related to STEM consisting of five sub-competencies from easy to difficult. The five sub-competencies were simplifying, mathematising, computing, interpreting and validating. This strategy may also increase high order thinking skills for pre-service teachers when they answer the items. Before the real experiment was conducted, the instrument must go through validation with three steps (I-CVI, S-CVI and Kappa). Face validity was conducted to test to ascertain whether the items met specifications related to STEM. Most of the experts chose scale 3 (agree) and 4 (strongly agree) for overall items.

Calculating the item level CVI (I-CVI) is the most popular way for gauging content validity. Scale-level CVI (S-CVI), which can be calculated using S-CVI/UA or S-CVI/Ave and lead to different values, is an alternative, unrecognized approach to quantify content validity. The I-CVI assesses the content validity of individual items, while the S-CVI assesses the total scale's content validity. The I-CVI or the S-CVI are usually reported in most studies, but not both. As the S-CVI is an average score that might be influenced by outliers, this study used both the I-CVI and the S-CVI. An I-CVI of 1.00 is regarded as excellent. All items had I-CVIs ranging from 0.75 to 1.00, with only three having an I-CVI less than 0.78. This supports the assertion that each item was significant and pertinent when assessing each sub-competence. Any value between 0.80 and 0.90 was considered the minimum acceptable S-CVI. S-CVI/UA and S-CVI/Ave values were computed. The universal agreement method indicated moderate overall content validity (S-CVI/UA = 0.88), whereas the average method indicated high content validity (S-CVI/Ave = 0.93). While the average approach may be more thorough and only take into account items with an I-CVI of 1.00, the universal agreement method may undervalue the content validity of the entire questionnaire because it was less likely to get 100% agreement across the board as the number of experts increased.

The kappa statistic is widely used to assess interdependence. The importance of rater dependability is that it demonstrates how accurately the study's data reflect representations of the variables under investigation. Interrater reliability is defined as the extent to which data collectors (experts) award the same score to the same variable. Due to the possibility that different data collectors may experience and interpret the phenomena of interest, interrater reliability is an issue in the majority of large studies to some level. In this research, the value kappa was higher at 0.89 which indicated that consider that the validity of the instrument can be used. The closer the value to 1.00, the better the validity of the instruments.

EFA is a statistical technique for revealing the underlying organization of a sizable collection of variables. EFA is a method for factor analysis whose main objective is to discover the underlying connections between measured variables. This work computed a reliable and valid STEM-based modelling instrument for pre-service teachers of mathematics education. The result of EFA revealed that the teacher data involved a five-dimension structure which included simplifying, mathematising, computing, interpreting and validating. The result showed that sampling adequacy was higher enough but the p-value for Bartlett's Test of Sphericity was 0.001. Overall, the findings of this study showed that the STEM-based modelling instrument's components were generally regarded as useful and applicable for assessing modelling proficiency among Malaysian respondents who were pre-service mathematics teachers. The STEM-based modelling construct successfully and internally consistent caught the primary five categories of STEM-based modelling instrument. The study's findings were in line with earlier research by Maaß (2006), which proved that the MMAS were very good for the five sub-competencies.

The findings of this investigation are consistent with those of earlier studies, as documented by Haines and Crouch (2001), Izard et al. (2003), and Lingefjärd and Holmquist (2005). These results underscore the suitability of the mathematical modeling test for prospective teachers in the current study. Our conclusion is drawn from the observed parallels between our research and prior studies, particularly concerning the sub-constructs of STEM-based modelling. This congruence is attributed to the commonality in the educational backgrounds of the populations under scrutiny, characterized by the need for nuanced perspectives in dealing with complex mathematical concepts. The shared patterns across these studies, including ours, point to the significance of considering higher education contexts when assessing the proficiency of individuals in mathematical modeling. The complexity of opinions required within higher education settings seems to contribute to the consistency in findings across various research endeavors. In light of these consistent outcomes, we advocate for the inclusion of a mathematical modeling test in future research endeavors. This recommendation is based on the belief that such assessments have proven effective in gauging the competency levels of individuals, especially those within higher education contexts. By incorporating mathematical modeling tests, future research can continue to contribute to the evolving understanding of competency in this field, fostering advancements in educational strategies and practices.

One of the most significant and prevalent statistics in research concerning test development and application is Cronbach's alpha (Cortina, 1993). We found that all of the sub-dimensions showed satisfactory internal consistency. The reliability of the mathematical modeling test aligns with earlier research, as indicated by Lingefjärd and Holmquist (2005). The results furnish compelling support for the widely acknowledged STEM-based modelling, underscoring its robust global applicability.

Cronbach's alpha value of each domain or construct was all over .80. It is important to remember that Cronbach's alpha was rated similarly to scale reliability, with scores between .70 and.90 being considered as good. Then, the Cronbach's alpha values higher than .80, were considered acceptable. The scale reliability, also known as the construct of reliability, was assessed using the findings of EFA (Dillon et al., 1984; Joreskog, 1971). Our research contributed to the body of evidence by demonstrating the reproducibility of the STEM-based modelling tool for future pre-service mathematics teachers. While the current findings validate the Modeling-based STEM instrument's reliability and validity, it is crucial to acknowledge that the study sample exclusively comprised female students. This limitation highlights the imperative for future research to broaden participant representation, incorporating individuals from diverse demographic backgrounds. Encompassing a wider range of participants, including male students and those from varied cultural and educational contexts, would bolster the applicability and generalizability of study outcomes. Moreover, exploring the instrument's efficacy across diverse student populations holds promise for uncovering potential variations in STEM learning outcomes and instructional requirements. This, in turn, can inform the development of more inclusive and equitable educational strategies.

6. Limitations and Recommendations

The design of any preliminary questionnaire had several limitations. The following limitations apply to this study: probable lack of generalizability; online survey may lead respondents to not read properly; and questionnaire length. Although this instrument was intended for pre-service teachers, it may be useful in senior communities; nonetheless, its generalizability to other teacher demographics is uncertain and needs to be tested. There is a possibility of recollection bias or inflated responses in an online survey. It also takes roughly 20 to 30 minutes to finish the questionnaire. The time may lead respondents to not answer the questionnaire seriously. In terms of data analysis, the current research only calculated validity using I-CVI and EFA. Advanced analysis employing confirmatory factor analysis (CFA) and Rasch analysis should be conducted in different settings. The integration of EFA, CFA and Rasch analysis have been widely employed in mathematics education context (Qudratuddarsi et al., 2022) to validate instrument for diverse settings.

Regarding the composition of our sample, it is noteworthy to highlight that all participants in the current study were female. While this allowed for a focused exploration of certain aspects, it is imperative that future research endeavors prioritize achieving gender balance to ensure the robustness, validity, and reliability of the findings. By including participants of diverse genders, researchers can gain a more comprehensive understanding of the phenomenon under investigation and account for potential gender-specific differences in responses. Moreover, striving for gender balance aligns with principles of inclusivity and equity in research, fostering a more representative and nuanced interpretation of results. Additionally, addressing gender imbalance in research samples contributes to broader efforts aimed at promoting gender equality and diversity in academia and society at large. Therefore, future studies should actively consider and implement strategies to recruit and engage participants from a variety of gender identities to enrich the research landscape and enhance the overall quality of findings. Given that our research was conducted through an online survey, we encountered challenges related to time constraints. We were unable to control respondents' completion of the questionnaire within a specific timeframe. However, we made efforts to mitigate this by implementing follow-up procedures. Despite these challenges, the online survey methodology allowed for flexibility in data collection and facilitated participation from a wide geographic area. Moving forward, future studies could explore strategies to encourage timely completion of surveys, such as setting clear deadlines or providing incentives for participation. Additionally, considering alternative data collection methods may offer solutions to address time constraints while maintaining research integrity.

7. Implication

The development and validation of the STEM-Based Modelling instrument not only established substantial item-content validity but also affirmed the overall scale validity for assessing sub-competency in pre-service teachers. The exploratory factor analysis identified five distinct sub-components—simplifying, mathematizing, computing, interpreting, and validating—highlighting the multifaceted nature of STEM-based modeling skills. These results underscore the significance of developing and validating focused modeling tools tailored for pre-service teachers in the realm of mathematics education. The implications extend to the body of knowledge, providing a nuanced understanding of the specific competencies involved in STEM-based mathematical modeling. For pre-service teachers, the tool serves as a valuable assessment and development resource, offering insights into their strengths and areas for improvement. Curriculum developers in higher education can leverage this instrument to refine and enhance programs, ensuring that the essential skillset related to mathematical modeling is effectively integrated into STEM education, ultimately benefiting both educators and students.

In addition, exploring how the developed tool aligns with broader goals and challenges in higher education settings is highly relevant for enhancing educational practices and results. By ensuring that the tool meets the specific needs and objectives of higher education, educators can improve teaching effectiveness, encourage student engagement, and conduct more precise assessments of learning outcomes. Understanding how the tool corresponds with broader goals helps institutions better prepare students for academic and professional success in a rapidly changing global environment. Furthermore, aligning the tool with the challenges encountered in higher education, such as promoting inclusivity, adapting to technological advancements, and nurturing critical thinking skills, enables educators to customize their teaching approaches effectively. Ultimately, this examination not only elevates educational quality but also fosters overall progress and innovation within higher education institutions. The insights gleaned from the data generated by these instruments provide educators with valuable information regarding student performance and misconceptions, empowering them to customize instructional approaches and interventions to effectively target specific needs. Moreover, the utilization of STEM-based modeling tools plays a role in shaping evidence-driven educational policies, directing the development of curricula, refining assessment methods, and allocating resources efficiently to maximize learning achievements and equip students for success in a progressively STEM-oriented society.

8. Conclusion

Students use mathematical modelling as a technique to consider and make sense of a real-world problem that will be examined using mathematics in order to understand, clarify or forecast something. The STEM tool is used to assess the five sub-competencies of mathematical modelling competency. This questionnaire's design employed a one-method strategy to select items required to comprehend STEM-based modelling devices. The tool demonstrated strong content validity of individual items and overall questionnaire content validity. The value of Kappa, which was 0.89, was likewise outstanding. As a result, the instrument had a greater face validity, content validity, and Kappa value. EFA revealed that STEM-based modelling instrument had five sub-components—simplifying, mathematising, computing, interpreting and validating. This suggests that the instrument can be extensively used to assess mathematical modelling proficiency using a STEM setting. Moreover, an actual experiment will be carried out in the future.

References

Albarracín, L., & Gorgorió, N. (2020). Mathematical modeling projects oriented towards social impact as generators of learning opportunities: A case study. *Mathematics*, *8*(11), 2034. https://doi.org/10.3390/math8112034

Anhalt, C. O., Cortez, R., & Bennett, A. B. (2018). The emergence of mathematical modeling competencies: An investigation of prospective secondary mathematics teachers. *Mathematical Thinking and Learning, 20*(3), 202-221. https://doi.org/10.1080/10986065.2018.1474532

Anhalt, C., Staats, S., Cortez, R., & Civil, M. (2018). Mathematical modeling and culturally relevant pedagogy. In Y. J. Dori, Z. Mevarech & D. Baker (Eds.), *Cognition, metacognition, and culture in STEM education* (pp. 174–195). Springer.

Arikan, S., Erktin, E., & Pesen, M. (2020). Development and validation of a STEM competencies assessment framework. *International Journal of Science and Mathematics Education*, 1-24. https://doi.org/10.1007/s10763-020-10132-3

Blomhøj, M. (2009). Different perspectives on mathematical modelling in educational research - categorising the TSG21 papers. In M. Blomhøj & S. Carreira (Eds.), *Mathematical applications and modelling in the teaching and learning of mathematics/* Vol. 11. 1–13. Roskilde University.

Blum, W., & Leiss, D. (2007). How do students and teachers deal with modeling problems? In C. Haines, P. Galbraith, W. Blum & S. Khan (Eds.), *Mathematical modeling: Education, engineering and economics–ICTMA 12*.

Brand, S. (2014). Effects of a holistic versus an atomistic modelling approach on students' mathematical modelling competencies. In P. Liljedahl, C. Nicol, S. Oesterle & D. Allan (Eds.), *Proceedings of the Joint Meeting of PME 38 and PME-NA 36*, Vol. 2. 185–192. PME

Carreira, S., & Baioa, A. M. (2011). Students' modelling routes in the context of object manipulation and experimentation in mathematics. In G. Kaiser, W. Blum, R. B. Ferri, & G. Stillman (Eds.), *Trends in teaching and learning of mathematical modelling, international perspectives on the teaching and learning of mathematical modelling*. 211–220. New York: Springer.

Cevikbas, M., Kaiser, G., & Schukajlow, S. (2021). A systematic literature review of the current discussion on mathematical modeling competencies: State-of-the-art developments in conceptualizing, measuring, and fostering. *Educational Studies in Mathematics*. 1-32. https://doi.org/10.1007/s10649-021-10104-6

Chamberlin, S., Payne, A. M., & Kettler, T. (2020). Mathematical modeling: A positive learning approach to facilitate student sense making in mathematics. *International Journal of Mathematical Education in Science and Technology*, 1-14. https://doi.org/10.1080/0020739X.2020.1788185

Cooper, R., & Heaverlo, C. (2013). Problem solving and creativity and design: What influence do they have on girls' interest in STEM subject areas? *American Journal of Engineering Education, 4*(1), 27-38. https://doi.org/10.19030/ajee.v4i1.7856

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78*(1), 98–104. https://doi.org/10.1037/0021-9010.78.1.98

Creswell, J. W. (2012). *Educational research: planning, conducting, and evaluating quantitative and qualitative research. Educational Research* (Vol. 4). Pearson.

Dillon, W., & Goldstein, M. (1984). *Multivariate analysis: Method and applications*. John Wiley & Sons.

English, L. (2003). Mathematical modeling with young learners. In S. J. Lamon, W. A. Parker, & S. K. Houston (Eds.), *Mathematical modeling: A way of life*. Horwood Publishing.

English, L. D. (2017). Advancing elementary and middle school STEM education. *International Journal of Science and Mathematics Education, 15*(S1), 5–24. https://doi.org/10.1007/s10763-017-9802-x

Fan, S. -C., Yu, K. -C., & Lin, K. -Y. (2020). A framework for implementing an engineering-focused STEM curriculum. *International Journal of Science and Mathematics Education, 19*(8), 1523–1541. https://doi.org/10.1007/s10763-020-10129-y

France, B. (2018). Modeling in technology education: A route to technological literacy. In M. J. de Vries (Ed.), *Handbook of Technology Education*. Springer.

Gravemeijer, K. (2008). RME theory and mathematics. In D. Tirosh & T. Wood (Eds.), *Tools and processes in mathematics teacher education*. 283–302. Sense Publishers.

Haines, C., & Crouch, R. (2001). Recognizing constructs within mathematical modelling. *Teaching Mathematics and Its Applications, 20*(3), 129–138. https://doi.org/10.1093/teamat/20.3.129

Hallström, J., & Schönborn, K. J. (2019). Models and modelling for authentic STEM education: Reinforcing the argument. *International Journal of STEM Education, 6*(1), 1-10. https://doi.org/10.1186/s40594-019-0178-z

Hankeln, C., Adamek, C., & Greefrath, G. (2019). Assessing sub-competencies of mathematical modelling—Development of a new test instrument. In *Lines of inquiry in mathematical modelling research in education* (pp. 143-160). Springer, Cham. https://doi.org/10.1007/978-3-030-14931-4_8

Herrington, J., & Parker, J. (2013). Emerging technologies as cognitive tools for authentic learning. *British Journal of Educational Technology, 44*(4), 607–615. https://doi.org/10.1111/bjet.12048

Hidayat, R., Adnan, M., Abdullah, M. F. N. L., & Safrudiannur. (2022). A systematic literature review of measurement of mathematical modeling in mathematics education context. *Eurasia Journal of Mathematics, Science and Technology Education, 18*(5), em2108. https://doi.org/10.29333/ejmste/12007

Hidayat, R., & Wardat, Y. (2023). A systematic review of Augmented Reality in Science, Technology, Engineering and Mathematics education. *Education and Information Technologies*. https://doi.org/10.1007/s10639-023-12157-x

Izard, J., Haines, C., Crouch, R., Houston, K., & Neill, N. (2003). Assessing the impact of teachings mathematical modelling: Some implications. In S. J. Lamon, W. A. Parker, & S. K. Houston (Eds.), *Mathematical modelling: A way of life ICTMA 11*. 165–177. Horwood.

Kaiser, G., & Stender, P. (2013). Complex modelling problems in co-operative, self-directed learning environments. In *Teaching mathematical modelling: Connecting to research and practice*. 277-293. Springer.

Kaiser, G., Blum, W., Ferri, R. B., & Stillman, G. (2011). Trends in Teaching and Learning of Mathematical Modelling – Preface. *Trends in Teaching and Learning of Mathematical Modelling*, 1–5. https://doi.org/10.1007/978-94-007-0910-2_1

Kertil, M., & Gurel, C. (2016). Mathematical modeling: A bridge to STEM education. *International Journal of Education in Mathematics, Science and Technology*, *4*(1), 44-55. https://doi.org/10.18404/ijemst.95761

Kurniadi, E., Suganda, V. A., & Harini, B. (2022). The pedagogical content knowledge dimensions of mathematics teacher in mathematics modeling learning. *Prima: Jurnal Pendidikan Matematika,* *6*(2), 90. https://doi.org/10.31000/prima.v6i2.5381

Lingefjärd, T., & Holmquist, M. (2005). To assess students' attitudes, skills and competencies in mathematical modelling. *Teaching Mathematics and Its Applications*, *24*(2–3), 123–133. https://doi.org/10.1093/teamat/hri021

Maaβ, K. (2006). What are modeling competencies? *Zentralblatt FüR Didaktik Der Mathematik, 38*(2), 113–142. https://doi.org/10.1007/BF02655885

Minarni, A., & Napitupulu, E. E. (2020). The role of constructivism-based learning in improving mathematical high order thinking skills of Indonesian students. *Infinity Journal*, *9*(1), 111-132. https://doi.org/10.22460/infinity.v9i1.p111-132

Osbourne, J. W. (2015). What is rotating in exploratory factor analysis? *Practical Assessment, Research, and Evaluation*, *20*(2), 1-7.

Qudratuddarsi, H., Hidayat, R., Nasir, N., Imami, M. K. W., & Mat Nor, R. (2022). Rasch validation of instrument measuring Gen-Z Science, Technology, Engineering, and Mathematics (STEM) application in teaching during the pandemic. *International Journal of Learning, Teaching and Educational Research, 21*(6). https://doi.org/10.26803/ijlter.21.6.7

Rodrigues, I. B., Adachi, J. D., Beattie, K. A., & MacDermid, J. C. (2017). Development and validation of a new tool to measure the facilitators, barriers and preferences to exercise in people with osteoporosis. B*MC Musculoskeletal disorders, 18*(1), 1-9. https://doi.org/10.1186/s12891-017-1914-5

Salingkat, S., & Bilalu, A. (2021). The influence of treffinger learning models on mathematics communication skills of students. *Jurnal Pendidikan Matematika Dan IPA, 12*(2), 97. https://doi.org/10.26418/jpmipa.v12i2.40814

Sanders, M. (2009). STEM, STEM Education, STEM mania. The Technology Teacher, 20–26. Retrieve from https://www.teachmeteamwork.com/files/sanders.istem.ed.ttt.istem.ed.def.pdf

Shi. J., Mo, X., & Sun., Z. (2012). Content validity index in scale development. Journal of Central South University. *Medical Sciences, 37*(2):152-155. https://doi.org/10.3969/j.issn.1672-7347.2012.02.007

Stillman, G., Galbraith, P., Brown, J., & Edwards, I. (2007). A framework for success in implementing mathematical modelling in the secondary classroom. In J. Watson & K. Beswick (Eds.), *Proceedings of the 30th annual conference of the mathematics education research group of Australasia*. Vol. 2. 688–697. MERGA Inc.

Williams, J. (2017). The vocational goals of STEM education: Is that enough? Paper presented at PATT 2017. *Philadelphia: Technology & Engineering Education – Fostering the Creativity of Youth Around the Globe* 10-14 July 2017.

Xu, T., Wu, X., Sun, S., & Kong, Q. (2023). Cognitive diagnostic analysis of students' mathematical competency based on the dina model. *Psychology in the Schools, 60*(9), 3135-3150. https://doi.org/10.1002/pits.22916

Zammanzadeh, V., Ghahramanian, A., Rassouli, M., Abbaszadeh, A., Alavi-Majd, H., & Nikanfar, A. R. (2015). Design and implementation content validity study: development of an instrument for measuring patient-centered communication. *Journal of caring sciences, 4*(2), 165. https://doi.org/10.15171/jcs.2015.017

Zöttl, L., Ufer, S., & Reiss, K. (2011). Assessing modelling competencies using a multidimensional IRT approach. In G. Kaiser, W. Blum, R. B. Ferri, & G. Stillman (Eds.), *Trends in teaching and learning of mathematical modelling ICTMA 14*. Vol. 1. Dordrecht: Springer. https://doi.org/10.1007/978-94-007-0910-2