An Analysis of Students’ Mathematical Self-Efficacy Instruments Using Rasch Model

Ahmad Mukhibin\textsuperscript{1a)}, Husnul Khatimah Rusyid\textsuperscript{1b)}, Ahmad Lutfi\textsuperscript{1c)}, Tatang Herman\textsuperscript{1d)}, Dwi Arianti Suryo Utomo\textsuperscript{2e)}

\textsuperscript{1}Universitas Pendidikan Indonesia, Jl. Dr. Setiabudi No. 229, Bandung, Indonesia
\textsuperscript{2}SMP Bumi Cendekia, Sleman, Yogyakarta, Indonesia

e-mail: \textsuperscript{a)}a.mukhibin@upi.edu, \textsuperscript{b)}husnulkhatimah.r@upi.edu, \textsuperscript{c)}ahmadlutfi@upi.edu, \textsuperscript{d)}tatangherman@upi.edu, \textsuperscript{e)}dwibumicendekia@gmail.com

Abstract

The purpose of this study was to determine the quality of mathematics self-efficacy instruments in terms of validity, reliability, undimensionality, ability distribution, and suitability of student responses. Data were obtained by distributing mathematics self-efficacy instruments consisting of 20 statements to 45 seventh-grade students of a private junior high school in Yogyakarta, Indonesia, at the 2022/2023 school year selected through purposive random sampling technique. The data were analyzed using Rasch model with the help of the Ministep application. The results of the analysis showed that the mathematics self-efficacy instrument used was declared valid, with an outfit Z standardized (outfit ZSTD) value of -0.16 for respondents and 0.20 for items, and could be used to measure students’ mathematics self-efficacy level with a reliability level reaching a score of 0.54 or in the medium category. The results of the distribution of item abilities show that most students admit to having the fighting power to solve problems that occur, especially in learning mathematics. From the analysis of the suitability of student responses, 18 statements show the suitability between the level of self-efficacy and the responses given by students. While two other statements did not show suitability. Therefore, teachers or other academics can use the self-efficacy questionnaire that has been prepared by researchers because it has met the four criteria for testing instruments.

Keywords: Mathematical self-efficacy, Ministep, Rasch model, reliability, validity

INTRODUCTION

Facing the rapid development of the digital era, students are required to have various abilities, starting from cognitive, affective, to psychomotor abilities (Agy & Kartono, 2021). These three abilities must go hand in hand and be formed starting from the most basic level to support student success in learning and in the future. Student success in learning is influenced by external factors, internal factors, and learning approach factors (Arianti, 2018). External factors are factors that come from outside themselves or environmental conditions around students (Nabillah & Abadi, 2019), and internal factors are factors that come from within students (Asriyanti & Purwati, 2020). Meanwhile, the learning approach factor is a learning strategy or method used by students to study subject matter (Nurfadilah & Hakim, 2019). One of the internal factors (factors from within students) that can affect learning outcomes is self-belief (Hamdi & Abadi, 2014; Martyanti, 2013). Self-belief in their abilities and potential is referred to as self-efficacy (Sutanto, 2018).

Bandura (1997) defines self-efficacy as an individual's belief in his or her ability to control and perform actions needed to deal with future situations. Meanwhile, Woolfolk (2009) explains that self-efficacy is a person's feeling that he can perform certain tasks effectively. This belief affects a person's thinking, feelings, motivation, and behaviour. Bartimote-Aufflick et al. (2016) revealed that self-efficacy is also related to other things, such as grades, metacognition, and the use of learning strategies. The way individuals view self-
efficacy affects how much effort will be given and how long individuals will persist when facing obstacles or unfavourable experiences (Ferdyansyah et al., 2020; Sunaryo, 2017). Adni et al. (2018) also said that self-efficacy will motivate students to assess their ability to solve problems. Students with a high level of self-efficacy will be better prepared to face various problems encountered (Rajagukguk & Hazrati, 2021). Meanwhile, students with a low level of self-efficacy will tend to give up easily when facing difficult problems (Fitriani, 2017; Mukhibin & Himmah, 2020).

In the learning domain, mathematics is considered one of the difficult subjects (Gazali, 2016) and can have negative effects on students' psychology such as anxiety and worry due to uncertainty about their ability to complete tasks (Sunaryo, 2017). Therefore, Oktariani (2018) argues that students need to have supportive self-efficacy so that students can utilize their potential and achieve optimal mathematics learning outcomes. Self-efficacy in learning mathematics is called mathematical self-efficacy. Jumroh et al. (2018) said that mathematical self-efficacy refers to students' ability to present and solve problems in the field of mathematics, how to master and understand mathematical concepts, and skills to interact effectively with friends and educators.

Student self-efficacy can be distinguished and measured through three dimensions, namely: magnitude, strength, and generality (Lianto, 2019). The magnitude dimension refers to the level of difficulty that students believe they can overcome. The strength dimension indicates the level of student confidence in the level of task difficulty that can be overcome. Meanwhile, the generality dimension describes whether beliefs will apply within the confines of a specific domain or have relevance in a variety of activities and situations (Moma, 2014).

Given the importance of self-efficacy in supporting optimal math learning outcomes (Indirwan et al., 2021) teachers need to help improve students' self-efficacy. To be able to measure student self-efficacy objectively, a self-efficacy instrument that is feasible to use is needed. Research on self-efficacy scale testing using the Rasch model has been conducted by Sudihartinih & Wahyudin (2019) with the results showing that the Rasch model proved effective for investigating the self-efficacy of prospective teacher students seen from gender and geometry thinking levels. The analysis of this study focused on the validity, reliability, and distribution of item and respondent reliability. The test results show that the items used are reliable and can be used to measure the self-efficacy of prospective teacher students. In addition, research by Setiadi (2021) also shows that the results of testing the self-efficacy scale using the V Aiken coefficient and LISREL show a match between the statement items and the mathematical self-efficacy indicators. However, the comparison of the number of items by the indicators of mathematical self-efficacy is still disproportionate so the researcher suggests that future researchers use the Rasch model to test the validity of the instrument.

Based on the description above, The researchers want to examine further the measurement of the self-efficacy scale carried out on junior high school students. The novelty of this research lies in the analysis used which includes validity, reliability, undimensionality, ability distribution, and student response suitability using the Rasch model. Rasch model is a modern test analysis technique that can overcome various limitations possessed by classical test theory (Widhiarso, 2016). Rasch developed an analytical model of item response theory formulated into a mathematical model that connects students and items alternately through the same interval scale (Sumintono, 2018). The Rasch model is capable of structurally ranking questions from the hardest to the easiest and respondents from high to low ability so that Rasch can detect inconsistencies in respondents' answers (misfits) or unusual patterns (outliers). Rasch model can determine the reliability and validity of research instruments (Sumintono & Widhiarso, 2014, 2015).
METHOD

This research is a descriptive quantitative study using a cross-sectional survey to collect data one by one at a time. Survey research is a procedure in quantitative research, in which researchers administer a survey of a sample to describe attitudes, opinions, behaviours, or characteristics of the existing population (Creswell, 2010). The population of this study were all students of a private junior high school in Yogyakarta, Indonesia, at the 2022/2023 academic year. By using a purposive sampling technique, the researchers chose class VII A consisting of 25 students and VII B consisting of 20 students so the sample in this study amounted to 45 students. The same teacher taught both classes. The total group of participants consisted of 28 males and 17 females.

The instrument used in this study is a mathematical self-efficacy questionnaire adapted from Mukhibin (2019) using Sutanto’s (2018) theory regarding the 3 dimensions of academic self-efficacy. This questionnaire consists of 20 statements, with 10 positive statements and 10 negative statements. The questionnaire assessment uses a Likert scale consisting of 4 answer options, namely: strongly agree, agree, disagree, and strongly disagree. The self-efficacy questionnaire lattice is presented in Table 1.

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Indicator</th>
<th>Item Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Magnitude</td>
<td>Interested in difficult tasks</td>
<td>1,2</td>
</tr>
<tr>
<td></td>
<td>View tasks as a challenge</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Ability to complete all assigned tasks</td>
<td>4,5</td>
</tr>
<tr>
<td></td>
<td>Committed to completing tasks</td>
<td>6</td>
</tr>
<tr>
<td>Strength</td>
<td>Persist in solving problems under any circumstances</td>
<td>7,8</td>
</tr>
<tr>
<td></td>
<td>Have a fighting spirit in the face of obstacles</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Perseverance on task</td>
<td>10,11</td>
</tr>
<tr>
<td></td>
<td>Overcoming learning difficulties</td>
<td>12,13</td>
</tr>
<tr>
<td></td>
<td>Confidence in one’s abilities</td>
<td>14,15</td>
</tr>
<tr>
<td></td>
<td>Learning from experiences</td>
<td>16</td>
</tr>
<tr>
<td>Generality</td>
<td>Planning for task completion</td>
<td>17,18</td>
</tr>
<tr>
<td></td>
<td>Knowledge of various materials</td>
<td>19,20</td>
</tr>
</tbody>
</table>

The research data obtained were then analyzed using the Rasch model to determine the results of the validity, reliability, dimensionality, ability distribution, and student response suitability test. Figure 1 shows the research process.

RESULT AND DISCUSSION

Instrument Validity and Reliability Test Results

The validity test is a test that aims to measure the extent of the accuracy of an instrument in measuring what should be measured. Meanwhile, the reliability test is used to show how consistent the instrument is to measure students' self-efficacy (Ghozali, 2006). Using the Ministep application, the results of the instrument validity and reliability tests can be seen in Table 2.

<table>
<thead>
<tr>
<th>Validity</th>
<th>Outfit ZSTD</th>
<th>Infit MNSQ</th>
<th>Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Person</td>
<td>-0.16</td>
<td>1.00</td>
<td>0.60</td>
</tr>
<tr>
<td>Item</td>
<td>0.20</td>
<td>1.01</td>
<td>0.97</td>
</tr>
<tr>
<td>Cronbach</td>
<td>-</td>
<td>-</td>
<td>0.54</td>
</tr>
<tr>
<td>Alpha</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 2. Validity and Reliability Test Results
Based on Table 2, the validity value of the instrument is shown in the Outfit Z Standardized (Outfit ZSTD) value of -0.16 for the person and 0.20 for item. This value is between the range $-2.0 < \text{ZSTD} < 2.0$ which means that all items have a rational value possibility. It means that overall the statement items or items are in accordance with the Rasch model and can be used as instruments to measure student self-efficacy. Meanwhile, the reliability test results show that the Cronbach alpha (KR-20) value is 0.54, this means that the overall instrument reliability value is in the medium category. Meanwhile, the respondent reliability value of the RMSE model is 0.60 in the weak category and item reliability is 0.97 in the very high category.

**Undimensionality**

The capacity of an instrument to estimate what the researcher wants to explore is measured by its undimensionality. This study aims to explore students' self-efficacy. The minimum raw variance explained is greater than 24% (Purnami et al., 2021). The Rasch model showed undimensionality through Principal Component Analysis (PCA) and local independence analysis. However, the study only reported PCA. The explained variance of the self-efficacy instrument exceeds the minimum score limit of 40%, meaning that the instrument is valid in measuring self-efficacy. The results of the undimensionality test are presented in Table 3.

<table>
<thead>
<tr>
<th>Explained Variance</th>
<th>Item</th>
<th>Person</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed</td>
<td>45.6%</td>
<td>7.2%</td>
<td>47.2%</td>
</tr>
<tr>
<td>Eigenvalue</td>
<td>43.45</td>
<td>6.82</td>
<td>45.00</td>
</tr>
</tbody>
</table>

The eigenvalue in the table above of 6.82 which is more than 3 indicates that there are problematic statement items so that further analysis can be carried out with item fit order analysis to determine whether an item can be retained or must be replaced (Muntazhimah et al., 2020).

**Results of Respondents’ Abilities Distribution**

The results of the distribution of students' self-efficacy levels can be seen from the distribution of respondents' abilities presented in the Ministep application. The results of the self-efficacy level distribution test are presented in Figure 2.

![Figure 2. Respondent Wright Map](image-url)
In Figure 2, the distribution of students' self-efficacy levels can be observed from the answers to each statement item. The distribution of self-efficacy levels is analyzed based on the logit measure value. The logit mean value is set with 0.0 as the standard of student ability. Students 21L and 33L are students who have the highest level of self-efficacy with a value of -0.1 logit among other students. This means that both students have a moderate level of self-efficacy. While students 37L and 44L have a very low level of self-efficacy with a value of -3.0 logit.

**Item Abilities Scatter Results**

Figure 3 shows the distribution of the items as indicated by the logit value. Items P2, P3, and P13 have a logit value of +2.00 which means that items 2, 3, and 13 are the most difficult items for students to agree with. Items 2 and 3 are items on the magnitude dimension where students find it difficult to agree with statements on indicators of being interested in difficult tasks and viewing tasks as a challenge. This is caused by external and internal factors from students which include interest in learning, motivation to learn, and environmental conditions that do not support student learning activities (Yanti, 2017). Meanwhile, item P9 has a logit value of -3.00 which means that item 9 is the easiest item for students to agree with. Item 9 is in the strength dimension with an indicator of having a fighting spirit in the face of obstacles. This shows that most students admit to having the fighting spirit to solve problems that occur, especially in learning.

The self-efficacy that a person has influences the individual in determining the actions that will be taken to achieve goals. Thus, students will continue to work on assignments and will not give up easily if they encounter difficulties in learning (Afriani et al., 2022).
Analysis of Student Statement Conformity

The suitability of student answers can be seen in the Expected Score ICC graph on the Ministep Application. If the ICC expected score value cuts the gray standard line then the item is outside the confidence space boundary, otherwise if the ICC expected score value does not cut the gray standard line then the item is still within the confidence limit (Risdianto et al., 2021). The following presents the results of the suitability of student answers.

In the Figure 4, it can be seen that in item 1 or with code P1, students' answers are as expected, this is indicated by the blue line of the item that does not intersect with the standard line. There are 18 items that are similar, namely items 1-7, items 9-13, and items 15-20, which are spread across three dimensions, namely magnitude, strength, and generality. It means that these items show the suitability between the level of self-efficacy and the answers given by students. So that the level of student self-efficacy can be described through the instrument developed by the researcher.

Figure 4. Corresponding Student Responses

Figure 5 shows item 8 with the indicator persist in solving problems under any circumstances that do not match the expected student answers, this can be seen from the blue item line that intersects with the standard line.

Figure 5. Discrepant Student Responses

In addition, there is also item 14 with the confidence in one's abilities indicator that depicts the same graph, this means that the two items do not match the description between the level of self-efficacy and student answers. Thus, item 8 and item 14 cannot be used as a basis for determining the high-low self-efficacy of students and requires revision in the instrument.

CONCLUSIONS

Based on the Rasch Model analysis of students’ mathematics self-efficacy questionnaire using the Ministep application, it can be concluded that the mathematics self-efficacy questionnaire is valid and can be used to measure students’ self-efficacy levels. While the overall reliability of the questionnaire is in the medium category. The results of the distribution of item abilities show that most students admit to having the fighting power to solve problems that occur, especially in learning mathematics. From the analysis of the suitability of student responses, 18 statements show the suitability between the level of self-efficacy and the answers given by students. While the other two statements did not show conformity. Therefore, teachers or other academics can use the self-efficacy questionnaire that has been prepared by researchers because it has met the four criteria
for testing instruments. In addition, teachers can also develop similar instruments by taking into account the recommendations of this study.

REFERENCES


https://journal.unsika.ac.id/index.php/sesiomadika/article/view/2990

Indonesian Journal of Mathematics Education, Vol. 6, No. 2, October 2023


