Journal of Mathematics Science and Computer Education
https://ppjp.ulm.ac.id/journals/index.php/jmsc-edu/index
e-ISSN: 2807-9167 p-ISSN: 2807-9329
Vol 6 No 1 2026
DOI: https://doi.org/10.20527/jmscedu.v6i1.18388

Rasch Validation of a Self-Regulation Instrument in Mathematics Learning for Junior High School Students

Aminah Ekawati[1*] and Isna Wardiah[2]
[1]Mathematics Education, Universitas PGRI Kalimantan, Banjarmasin, Indonesia
[2]Informatics Engineering, Politeknik Negeri Banjarmasin, Banjarmasin, Indonesia
* aminahekawati@upk.ac.id

Article Info

Article history:
Received: March 2, 2026
Accepted: April 20, 2026
Published: May 24, 2026

Keywords:
Instrument
Rasch model
Self-regulation
Validation

Abstract

Although self-regulation instruments have been widely developed, differences in respondent characteristics and the context of the research area require the development of instruments appropriate to the conditions. Nonetheless, few studies have particularly validated self-regulation tools for junior high school children in mathematics learning employing a Rasch measuring approach. This study focuses on developing and validating a self-regulation instrument for junior high school students in mathematics learning through the Rasch model, which involves analyzing reliability, validity, item difficulty levels, and the alignment between respondent abilities and item difficulty levels. The validity consists of item fit and unidimensionality. The instrument was developed through a synthesis of several theoretical frameworks and simplified into three main dimensions, namely cognitive and metacognitive regulation, motivational and emotional regulation, and behavioral and social regulation, which consisted of 12 four-choice Likert scale statements. The subjects of this research were 60 eighth-grade students at a public junior high school in Banjarmasin City. Data analysis was performed using Winsteps 4.7.0 and RStudio 4.5.2 software. The results of the analysis showed that the instrument had good reliability, namely Cronbach's Alpha = 0.869, WLE Reliability = 0.873, and EAP Reliability = 0.872, and satisfactory construct validity, supported by acceptable item fit and unidimensionality based on PCA of residuals. These findings support the unidimensionality of self-regulation within a Rasch framework in mathematics learning and present a context-specific validated instrument with three dimensions, operationalized into 12 theoretically grounded items for diagnosis.

To cite this article: Ekawati, A., & Wardiah, I. (2026). Rasch validation of a self-regulation instrument in mathematics learning for junior high school students. *Journal of Mathematics, Science, and Computer Education*, 6(1), 89-103.

INTRODUCTION

Mathematics encompasses procedural elements and higher-order thinking, necessitating that students possess logical, analytical, and systematic thinking skills as fundamental competencies for problem-solving. The results of international surveys like the Programme for International Student Assessment (PISA) show that Indonesian students' mathematical problem-solving abilities are still relatively low (Maharani et al., 2019; Mashuri & Jahring, 2024; Pramesti & Rosyidi, 2020; Putri & Zulkardi, 2020). These low scores are not solely caused by academic aspects but may also be due to psychological aspects, such as self-regulation, which shows a relationship or influence on mathematics learning outcomes (Arsyad et al., 2022; Dzahabiyyah et al., 2024; Saputri et al., 2022). Furthermore, self-regulation also affects students' independence in carrying out tasks, their ability to plan and manage the learning process, and their competence in managing time effectively (Qadaristin, 2021).

Self-regulation is defined as the process by which individuals proactively control their thoughts, emotions, behavior, and motivation to achieve specific learning goals (Zimmerman, 2000). Self-regulation is essential in mathematics education, as students must comprehend concepts and devise solution strategies, monitor and assess their solutions, and manage their emotions when confronted with difficult problems (Landa et al., 2024, 2025; Torregrosa & Albarracín, 2025). Students who are skilled in self-regulation can set learning goals, choose appropriate strategies, and maintain focus throughout the problem-solving process. Mathematical problem-solving abilities are highly dependent on the application of self-regulation strategies, which are theoretically viewed as essential characteristics and contribute to the effectiveness of the problem-solving process (Noor et al., 2025).

Numerous studies have underscored self-regulation as a multifaceted and contextual construct, incorporating cognitive, metacognitive, motivational, and behavioral dimensions of students' learning management (Ekawanti & Mulyana, 2016; Perdana & Wijaya, 2021). In the realm of mathematical problem-solving, self-regulation is intricately linked to the comprehension of the problem, the formulation of a solution strategy, the execution of the solution steps, and the assessment of the outcomes achieved, all underpinned by student motivation and accountability for their learning (Nurvicalesti & Ratnasari, 2023). In the context of online learning, self-regulation also includes the ability to set learning goals, manage the learning environment, manage time, choose task strategies, seek assistance, and conduct self-evaluations, as findings recorded during online learning studies and research on adaptations of self-regulation instruments show (Kholisna & Fatimah, 2022; Mutiara & Rifameutia, 2021; Siregar, 2021). These various views demonstrate that self-regulation is a complex construct and is strongly influenced by the learning context.

Previous studies have conceptualized self-regulation using different theoretical perspectives, resulting in variations in the dimensions proposed. For example, Landa et al. (2024) emphasize aspects such as self-perception of ability, problem-solving strategies, and social influences, while Panadero (2017) describes self-regulation in terms of cognitive, metacognitive, motivational, emotional, and behavioral processes. In addition, López-Angulo et al. (2024) view self-regulation as a cyclical process consisting of learning dispositions, performance, and self-evaluation, whereas Dai et al. (2023) highlight practical regulatory strategies such as goal setting and time management.

These differences indicate that self-regulation is a complex construct with no single standardized dimensional structure. As a result, determining a framework that is both comprehensive and practical for instrument development remains a challenge, particularly in mathematics learning at the junior high school level. Therefore, synthesizing these diverse perspectives into a more manageable and representative dimensional structure becomes an important step in developing context-appropriate measurement instruments. Research shows that in recent years, numerous self-regulation instruments have been developed and validated across various learning contexts and educational levels. These instruments were created utilizing different theoretical frameworks and tested with both CTT and current psychometric methods, such as the Rasch Model (Fitria et al., 2026; Marweli et al., 2025; Parantika et al., 2022; Siregar, 2021). Research results generally report adequate reliability and excellent item fit, thus deeming the instruments suitable for measuring student self-regulation. This report demonstrates the development of research on self-regulation measurement, both theoretically and methodologically. Nonetheless, this study illustrates the diversity in the evolution of self-regulation tools. The varying theoretical frameworks employed by each study have led to discrepancies in the number of dimensions and indicators assessed. Each instrument is developed with a specific context in mind and has its strengths. However, this variation indicates that there is currently no consensus on the most representative structure of a self-regulation instrument for mathematics learning at various levels of education, particularly junior high school.

Several studies emphasize that the development and use of self-regulation instruments must undergo an adequate validation process, taking into account the characteristics of the subjects and the context in which they are applied. Empirical data from testing an instrument on a sample accurately representing the target population determines its quality and validity (American Educational Research Association, 2014). Similarly, Fitria et al., (2026) also state that instrument quality is strongly influenced by the characteristics of the respondents and the educational context in which the instrument is used. These limitations indicate the importance of developing and validating a self-regulation instrument that is aligned with students’ characteristics and the local learning context.

Although various studies have developed self-regulation instruments, the diversity in theoretical frameworks, respondent characteristics, and research contexts indicates that existing instruments may not fully capture students’ self-regulation in specific learning settings. This limitation is particularly important in mathematics learning at the junior high school level, where self-regulation plays a critical role in problem-solving processes. Without context-appropriate and empirically validated instruments, it is difficult to accurately assess students’ self-regulation and to design effective learning interventions. Therefore, there is a need to develop and validate a self-regulation instrument that is aligned with students’ characteristics and the local learning context. Accordingly, this study aims to develop and validate a self-regulation instrument for junior high school students using the Rasch Model by analyzing reliability, item fit, unidimensionality, item difficulty, and the alignment between respondent ability and item difficulty.

METHODS

Research Design

The research design used was instrument development research aimed at developing and validating a self-regulation instrument for junior high school students, following established practices of instrument development and Rasch-based validation (Saputra et al., 2023; Wiyono et al., 2025). Instrument validation used the Rasch model assisted by Winsteps 4.7.0 and RStudio 4.5.2 software. The Rasch model was chosen because it adheres to the principle of specific objectivity, namely, its ability to produce independent estimates of respondent ability and item parameters without being influenced by sample or assessor characteristics (Kassiavera et al., 2024; Verdú-Soriano & González-de la Torre, 2024).

Subject of Research

The study was conducted in two eighth-grade classes at a public junior high school in Banjarmasin, with 60 students participating, as shown in Table 1. All participants participated voluntarily, and the data collection process was conducted in accordance with ethical research principles. The sample size was determined based on recommendations in Rasch measurement, which indicate that there is no fixed requirement for sample size; however, a commonly used guideline suggests 5–20 respondents per item (Verdú-Soriano & González-de la Torre, 2024). Given that this study employed 12 items, the sample size of 60 participants meets the minimum requirement for stable estimation.

Table 1 Number of research participants

Grade	Gender	
	Male	Female
VIII A	15	15
VIII B	15	15

Instruments

The instrument developed synthesizes various dimensions from several literature sources, such as Landa et al. (2024) who examined self-regulation through a multidimensional approach that includes self-perception of ability, attitudes toward problems, problem-solving methods, self-efficacy beliefs, and social environmental influences. (Panadero, 2017) understands self-regulated learning as a multidimensional construct involving cognitive, metacognitive, motivational, emotional, and behavioral processes. Meanwhile, López-Angulo et al. (2024) developed and validated a self-regulated learning instrument based on a self-regulation phase approach that reflects learning dispositions, learning performance, and self-evaluation. This approach shows that self-regulation is a dynamic and cyclical process, where cognitive, motivational, and reflective aspects interact during the learning process. In the context of online learning, Dai et al. (2023) identified six dimensions of self-regulation, namely goal setting, environment structuring, task strategies, time management, help-seeking, and self-evaluation.

Furthermore, based on the results of the synthesis of the four theoretical frameworks, the self-regulation instrument in this study was adapted and simplified into three main dimensions, namely, (1) cognitive and metacognitive regulation, (2) motivational and emotional regulation, and (3) behavioral and social regulation. Each dimension consists of four statements, resulting in a total of twelve. The statements were constructed using a four-point Likert scale: Strongly Disagree, Disagree, Agree, and Strongly Agree, with no neutral option included. All statements are positive, so the scores are 1, 2, 3, and 4, respectively. The development of this instrument was based on a well-established theoretical framework, which provided a strong conceptual foundation for content representation. Prior to empirical testing, the instrument was validated by a mathematics education expert experienced in self-regulation research. Validation was based on examining the alignment between the proposed dimensions and their corresponding items. The expert evaluation revealed that all items were deemed appropriate and consistent with the intended dimensions, making the instrument suitable for use. The self-regulation instrument developed based on the three dimensions is shown in Table 2.

Table 2 Self-regulation instrument developed

Dimensions	Item
Cognitive and metacognitive regulation	I read a math problem repeatedly until I fully understand it (Var 1). I plan the steps to solve it before starting the problem (Var 2). I am able to adjust my strategy when the first method doesn't work (Var 3). I always double-check my results or predictions to ensure they meet the problem's requirements (Var 4).
Motivation and Emotional Regulation	I am able to control my emotions when faced with difficult math problems (Var 5). I remain motivated to solve math problems even when they seem difficult or boring (Var 6). I remind myself that solving math problems is important for my academic success (Var 7). I prefer challenging math problems to ones that are too easy (Var 8).
Behavioral and Social Regulation	I divide my math study time each day so that it doesn't pile up on one day (Var 9). I choose a study place and time that minimizes distractions (Var 10).

Dimensions	Item
	If I can't solve a math problem, I'm willing to seek help from a friend or teacher (Var 11). I feel responsible for solving a math problem on my own before asking for help (Var 12).

Data Analysis Technique

Instrument validity in the Rash model was analyzed through item fit statistics and unidimensional testing. Item fit was evaluated using Mean_Infit and Mean_Outfit values. Mean_Infit and Mean_Outfit values ranging from 0.5 to 1.5 are considered adequate and indicate that the item fits the Rasch model (Davis & Boone, 2021; Sari et al., 2026).Furthermore, the Outfit Z-Standardized (ZTSD) and Point Measure Correlation (PTMEA Corr) indicators were used to strengthen item validity. A good ZSTD value is between -2.0 < ZSTD < +2.0, while an acceptable PTMEA Corr value is between 0.40 and 0.85 (Davis & Boone, 2021). The PTMEA Corr value measures the direction and strength of the relationship between an item and a latent construct.

Furthermore, in the Rasch model, construct validity is also assessed through unidimensional analysis using Principal Component Analysis (PCA) on normalized Rasch residuals. This analysis aims to ensure that the developed instrument measures only one primary latent dimension. Unidimensional is demonstrated if the variance explained by the primary dimension is sufficiently dominant and there are no significant secondary dimensions expressed in eigenvalues. If the eigenvalue is < 3, no items are difficult (Marweli et al., 2025). Furthermore, in this study, unidimensionality was analyzed through item fit using infit and outfit mean squares (Park, 2025).

Three reliability measures were used: Cronbach's alpha, WLE Reliability, and EAP reliability. Cronbach's alpha is a reliability test used in classical test theory (CTT), with the requirement that a Cronbach's alpha ≥0.7 be reliable (Lacave et al., 2018). A WLE reliability value ranging from 0.60 to 0.74 is considered moderate and acceptable for educational research, although an ideal reliability value is recommended at ≥ 0.80 (Davis & Boone, 2021; Rentiana et al., 2024). An EAP reliability value >0.7 indicates adequate measurement consistency (Bean & Bowen, 2021; Kılıç et al., 2023; Wess et al., 2021).

Item difficulty is determined based on the item parameter value in logit units. Higher logit values indicate relatively more difficult items, while lower logit values indicate easier items. Item difficulty is analyzed based on the position of the logit values relative to the average respondent's ability. An instrument is considered appropriate if item difficulty values are evenly distributed and do not show significant gaps on the Wright Map (Park, 2025; Sari et al., 2026).

Research Procedures

The research consisted of four stages. The first stage began with developing a self-regulation instrument based on a synthesis of the theoretical framework from previous research. The instrument statements were formulated in the form of positive statements tailored to the characteristics of junior high school students. Next, the researcher reviewed the conceptually developed instrument to ensure the representation of each established dimension of self-regulation. The second stage involved pilot testing the developed instrument on eighth-grade students at a public junior high school in Banjarmasin City. Data collection was conducted directly in class during mathematics lessons with the approval of the subject teacher and the school. Before students completed the instrument, they were given an explanation of the research objectives and how to complete the statements, and they were informed that student participation was voluntary and that the collected data would be kept confidential. The third stage involved students completing the self-regulation instrument, which consisted of 12 statements with a four-choice Likert scale, within a maximum of 30 minutes. After all instruments were collected, the data were checked to ensure completeness of responses and then coded for analysis. The fourth stage involved data analysis using the Rasch model with the assistance of Winsteps software version 4.7.0 and RStudio 4.5.2. The results of the analysis are used as a basis for assessing the feasibility of the self-regulation instrument developed.

RESULTS AND DISCUSSION

The reliability results obtained from Winsteps and RStudio are presented in Table 3. The Cronbach's Alpha value in Table 3 of 0.869 indicates that the items in the instrument correlate well with each other in measuring the self-regulation construct so that students' responses to these items are relatively stable and consistent. This means that based on the CTT, the instrument is categorized as reliable. In CTT, reliability refers to the level of consistency and stability of measurement results (Defrijon et al., 2025; Muharromah et al., 2025). A reliable instrument will provide consistent information when used under equivalent conditions (Defrijon et al., 2025; Zayrin et al., 2025).

Table 3 Instrument reliability results

No.	Type of Reliability	Value	Conclusion
1	Cronbach Alpha	0.869	Good
2	WLE Reliability	0.873	Acceptable
3	EAP Reliability	0.872	Acceptable

Furthermore, the reliability based on the Rasch model is shown by the WLE reliability and EAP reliability in Table 3. This further strengthens the findings regarding the instrument's reliability based on the CTT. Both reliability values reflect the accuracy of the respondent's ability estimates generated by the model, taking into account item characteristics and student response patterns. WLE and EAP values above 0.70 indicate that the instrument is able to accurately distinguish students' levels of self-regulation ability (Aini et al., 2020; Sagan, 2025). Therefore, it can be said that the instrument has good reliability not only in terms of score consistency (CTT) but also in terms of the accuracy of ability estimation (Rasch).

Furthermore, the results in Table 2 demonstrate not only technical consistency but also that each item makes a significant contribution to representing the construct of self-regulation. The relationships between items indicate the integration of cognitive-metacognitive, motivational-emotional, and socio-behavioral dimensions into a comprehensive self-regulation system, rather than simply a set of isolated indicators. Acceptable WLE and EAP values indicate model stability at the score level and precision in estimating individual abilities.

These results align with Chang et al. (2015), who stated that instrument evaluation through a combination of CTT and Rasch analysis provides a more comprehensive picture of reliability because it integrates internal consistency with the precision of respondent ability estimation. In the Rasch model, reliability is not only understood as score stability but also as measurement precision at every level of ability and the model's ability to detect response inconsistencies through integration with the model fit (Sagan, 2025). Therefore, the reliability of this instrument is due to the instrument design being based on theory and being integrative.

The results of the item fit analysis using the Rasch model assisted by Winsteps software are shown in Figure 1, where all self-regulation instrument items had Infit Mean Square (MNSQ) values ranging from 0.69 to 1.44, while Outfit MNSQ values ranged from 0.69 to 1.42. The highest Infit and Outfit MNSQ values were for statement Var 8 and the lowest for statement Var 2. Var 8 shows a more varied and less predictable response (underfit). This phenomenon occurs possibly because it is interpreted differently by respondents or requires more complex thinking, so that Var 8 is thought to measure a more contextual or higher-level aspect of self-regulation. Conversely, Var 2 shows a very consistent response (overfit), meaning that the item represents an aspect of self-regulation that is relatively straightforward for respondents to understand and recognize. Furthermore, all infit and outfit scores were within acceptable limits in the Rasch analysis, indicating that all items had a good fit with the model and were suitable for use in measuring the construct of self-regulation (Davis & Boone, 2021; Ding et al., 2023; Sari et al., 2026). Furthermore, these results not only indicate statistical fit but also indicate that each item plays a distinct role in representing the construct of self-regulation. These variations actually enrich the instrument in distinguishing levels of self-regulation ability more sensitively.

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTMEASUR-AL CORR.	PTMEASUR-AL EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	ITEM
8	148	60	1.02	.24	1.44	2.21	1.42	1.91	A .44	.63	53.3	65.0	VAR00008
12	178	60	-.75	.25	1.24	1.28	1.21	1.13	B .46	.65	70.0	66.3	VAR00012
11	164	60	.09	.24	1.19	1.06	1.16	.90	C .60	.64	61.7	64.2	VAR00011
7	176	60	-.63	.25	1.07	.42	1.04	.26	D .78	.65	60.0	65.9	VAR00007
9	167	60	-.09	.24	1.07	.43	1.07	.42	E .54	.65	53.3	64.5	VAR00009
6	163	60	.15	.24	.99	.00	.98	-.02	F .67	.64	58.3	64.2	VAR00006
10	154	60	.67	.24	.98	-.05	.98	-.05	f .77	.64	65.0	64.4	VAR00010
5	159	60	.38	.24	.92	-.42	.92	-.39	e .66	.64	66.7	63.9	VAR00005
4	165	60	.03	.24	.76	-1.41	.91	-.43	d .68	.64	71.7	64.2	VAR00004
1	178	60	-.75	.25	.89	-.58	.88	-.65	c .66	.65	66.7	66.3	VAR00001
3	166	60	-.03	.24	.72	-1.73	.70	-1.77	b .69	.64	65.0	64.4	VAR00003
2	167	60	-.09	.24	.69	-1.89	.69	-1.85	a .77	.65	76.7	64.5	VAR00002
MEAN	165.4	60.0	.00	.24	1.00	-.1	1.00	.0			64.0	64.8	
P.SD	8.7	.0	.52	.00	.21	1.2	.20	1.1			6.8	.8	

Figure 1 Item fit based on the rasch model

Figure 1 shows a ZSTD value greater than 2 for Var 8. However, this value is still within the recommended tolerance limits for Rasch-based measurement. This outcome is because the infit and outfit mean square values for this item are within the acceptable range recommended in the Rasch literature (Ding et al., 2023; Rentiana et al., 2024). These results indicate that each item adequately measures the construct of self-regulation according to the respondent's ability level. Based on the item fit analysis, all items demonstrated infit and outfit mean square values within the acceptable range. Although one item had a ZSTD value slightly exceeding the ±2 limit, its mean square value still indicates a good fit with the model and does not indicate a serious misfit. These results indicate that the instrument meets the unidimensional assumption, where all items function consistently in measuring the same construct, namely self-regulation, as recommended in measurements using the Rasch model (Ding et al., 2023; Pelikan et al., 2022; Rentiana et al., 2024; Zenger & Bitzenbauer, 2022).

Figure 1 also shows the Point-Measure Correlation (PTMEA Corr.); the PRMEA Corr. value ranges from 0.44 to 0.78, indicating a positive correlation with the latent ability measures across all items, meaning the direction of measurement for each item is in line with the construct being measured (Davis & Boone, 2021). Items with an MNSQ value close to 1.0 reflect optimal fit to the model, while items with an MNSQ value slightly above 1.0 indicate acceptable response variation that does not interfere with the validity of the measurement. Thus, the item fit results provide empirical evidence that the self-regulation instrument developed has adequate internal validity and is suitable for measuring the self-regulation abilities of junior high school students.

In addition, Figure 2 shows the results of a Principal Component Analysis (PCA) of the Rasch model residuals as a test of unidimensionality. The raw variance explained by measures was 44.3%, meeting the minimum criteria for Rasch measurements to support the assumption of unidimensionality. This variance consisted of respondents' ability (29.6%) and item difficulty (14.7%). Meanwhile, the eigenvalue for the first contrast, 2.65, was still below the critical threshold of 3.0 Marweli et al. (2025); thus, not indicating a strong second dimension. These results indicate that the structure of the instrument is dominated by one main construct, namely self-regulation.

```
     Table of STANDARDIZED RESIDUAL variance in Eigenvalue units = ITEM information units
                                           Eigenvalue   Observed   Expected
Total raw variance in observations     =      21.5493 100.0%         100.0%
  Raw variance explained by measures   =       9.5493  44.3%          44.1%
    Raw variance explained by persons  =       6.3852  29.6%          29.5%
    Raw Variance explained by items    =       3.1641  14.7%          14.6%
  Raw unexplained variance (total)     =      12.0000  55.7% 100.0%   55.9%
    Unexplned variance in 1st contrast =       2.6453  12.3%  22.0%
    Unexplned variance in 2nd contrast =       1.8555   8.6%  15.5%
    Unexplned variance in 3rd contrast =       1.5462   7.2%  12.9%
    Unexplned variance in 4th contrast =       1.1742   5.4%   9.8%
    Unexplned variance in 5th contrast =       1.0783   5.0%   9.0%
```

Figure 2 PCA of standardized residuals for assessing unidimensionality

Furthermore, Table 4 displays the item difficulty for the developed instrument. Item difficulty is expressed in logit units, where higher logit values indicate more difficult items, while lower logit values indicate easier items. Based on Table 4, logit values range from -0.75 to 1.02, resulting in three easy items, two difficult items, and the remainder moderate. These results indicate that the instrument's items have adequate difficulty variation and are distributed proportionally. Furthermore, the distribution of item difficulty levels also indicates that the instrument is capable of measuring students' self-regulation at various ability levels, ranging from low to high. No items were found with extreme difficulty levels; thus, all items were deemed to function optimally in differentiating students' self-regulation abilities.

Table 4 Item difficulty levels

Item	**Measure (Logit)**	**Category**
Var 1	−0.75	Easy
Var 2	−0.09	Moderate
Var 3	−0.03	Moderate
Var 4	0.03	Moderate
Var 5	0.38	Moderate
Var 6	0.15	Moderate
Var 7	−0.63	Easy
Var 8	1.02	Hard
Var 9	−0.09	Moderate
Var 10	0.67	Hard
Var 11	0.09	Moderate
Var 12	−0.75	Easy

Furthermore, when viewed from a dimensional structure, the variation in difficulty levels is also consistent with the multidimensional character of self-regulated learning, as emphasized in the literature (Panadero, 2017). In this study, items in the cognitive and metacognitive regulation dimensions tended to be in the easy to moderate category, which can be interpreted as procedural learning strategies and basic monitoring that are relatively more operational in the context of student learning. Meanwhile, the items with the highest level of difficulty were in the motivation and emotion regulation dimension, which in the self-regulated learning literature involves affective control, resilience to difficulties, and more complex goal orientations. In the behavioral and social regulation dimensions, the varying distribution of difficulty indicates that indicators of learning environment regulation and academic responsibility were proportionally represented across the respondents' ability range. This pattern strengthens the fit between the empirical structure of the Rasch analysis results and the underlying theoretical framework of the self-regulation construct. Although this instrument is structured around three dimensions, the distribution of item difficulty levels indicates that these three dimensions operate in an integrated manner within a single latent construct. This reinforces the notion that self-regulation is an integrated system, in which cognitive, motivational, and behavioral aspects interact dynamically, rather than existing in isolation.

For a more comprehensive analysis, the Wright Map shown in Figure 3 was used. The Wright Map represents the distribution of respondent ability and item difficulty on the same logit scale (Akour, 2022; Aliffiansyah et al., 2025; Planinic et al., 2019). The left side of Figure 3 shows the distribution of respondent ability, while the right side shows the difficulty levels of the self-regulation instrument items. The distribution of respondent ability appears quite wide, ranging from approximately −6 logit to +4 logit, indicating significant variation in self-regulation ability among students. The majority of respondents fell within the −1 to +2 logit range, indicating that most students were at low to moderate ability levels. There were also some respondents with very low ability (below −4 logit) and high ability (above +3 logit), although their numbers were relatively limited. On the item side, item difficulty levels were spread over a narrower range, approximately −1 to +1 logit. This data is consistent with Table 2, which shows that most items are at a moderate level of difficulty, with only a few items classified as hard or easy. Most items are at a moderate level of difficulty and appropriate for the abilities of the majority of students. This allows the instrument to differentiate self-regulation skills at the most common level in the classroom, which is essential for tailoring instruction to meet the needs of the majority of students. However, the consistent variation in difficulty levels, ranging from low to high, indicates that the instrument is still capable of differentiating students' overall abilities.

The overall fit between the distribution of respondents' abilities and the item difficulty levels can be said to be excellent, especially in the medium ability range, where most items fall. However, at high ability levels (above +2 logit), the number of available items is relatively limited, so the instrument's ability to differentiate respondents with very high levels of self-regulation is still less than optimal. Similarly, at very low ability levels, only a few items specifically target respondents at that level. Overall, this Wright Map indicates that the self-regulation instrument adequately measures respondents' abilities at low-to moderate-levels, but there is still room for further development by adding more difficult items to optimize measurement coverage at high-ability levels. However, previous research tends to be limited to describing logit ranges and identifying the easiest and most difficult items, without categorizing difficulty levels or deeply linking them to the distribution of respondents' abilities (Sujatmika et al., 2025). Therefore, this study complements these studies by classifying item difficulty levels and interpreting them in the context of students' self-regulation abilities.

The results of the Wright map distribution (People-item map) are listed in Figure 3.

```
MEASURE PERSON - MAP - ITEM
        <more>|<rare>
    4       X  +
             T|
           XX  |
               |
            X  |
               |
    3      XX  +
               |
            X  |
               |
             S|
    2   XXXXX  +
               |
          XXX  |
               |
         XXXX  |
        XXXXX  |
    1          +
        XXXXX  |T VAR00008
             M|
           XX  |  VAR00010
       XXXXXX  |S
               |  VAR00005
    0   XXXXX  +  VAR00006
         XXXX  |M VAR00003  VAR00004  VAR00011
               |  VAR00002  VAR00009
         XXXX  |
         XXXX  |S
   -1          +  VAR00001  VAR00007  VAR00012
           XX S|
               |T
               |
           XX  |
               |
   -2          +
               |
            X  |
               |
             T|
   -3          +
               |
               |
               |
               |
               |
   -4          +
               |
               |
               |
               |
               |
   -5          +
               |
               |
               |
               |
               |
   -6          +
            X  |
   -6          +
        <less>|<freq>
```

Figure 3 Wright map distribution (Person–item map)

CONCLUSION

Based on the results regarding reliability, item fit, unidimensionality, item difficulty level, and the alignment between respondent ability and item difficulty level as shown in the Wright Map, it was concluded that the developed self-regulation instrument meets the psychometric feasibility criteria. The instrument showed good internal consistency, adequate item fit to the Rasch Model, and good ability to measure student self-regulation at various ability levels. Beyond these findings, this study contributes theoretically by providing empirical evidence that self-regulation can be consistently measured as a unidimensional construct within the Rasch framework. Methodologically, this study highlights the importance of evaluating instruments using multiple indicators rather than relying on a single statistic, thereby strengthening instrument validation practices. This study has limitations, particularly regarding the research subjects, which are currently restricted to junior high school students in the context of mathematics learning. Additionally, the number of statement items in this instrument can be further developed to better represent the construct and enhance measurement precision. Therefore, further research is recommended to test this instrument on a more diverse sample, covering different educational levels and learning contexts.

Practically, the validated self-regulation instrument is not only relevant for instrument development research but also has the potential to be used in future research examining the role of self-regulation in mathematics learning. It can be used as a reliable tool to assess students' self-regulation levels and support the design of learning strategies and interventions that promote independent learning. Furthermore, this study opens opportunities for future research exploring the role of self-regulation in improving learning outcomes, particularly in mathematics education.

CONFLICTS OF INTEREST

The authors declare that there is no conflict of interest.

AUTHOR CONTRIBUTIONS STATEMENT

AE: Conceptualization, Methodology, Formal Analysis, and Writing–Original Draft. IW: Methodology, Investigation, and Writing-Review & Editing. All authors have read and approved the final version of the manuscript.

AI DISCLOSURE STATEMENT

The authors used WinSteps 4.70 and RStudio 4.5.2 during the preparation of this work to analyze the data obtained in this study. After using the tool/service, the author thoroughly reviewed and edited the content as needed and takes full responsibility for the content of the publication. Additionally, the authors used QuillBot for grammar checking and language refinement.

ACKNOWLEDGMENTS

The authors would like to express sincere gratitude to Universitas PGRI Kalimantan and Politeknik Negeri Banjarmasin for their support and assistance in facilitating this research.

REFERENCES

Aini, R. Q., Rachmatullah, A., Roshayanti, F., Shin, S., Lee, J. K., & Ha, M. S. (2020). Testing validity inferences of science motivation questionnaire (SMQ-II) instrument: Rasch-based analysis with Indonesian secondary students. *Journal of Physics: Conference Series*, *1521*(4), 042101. https://doi.org/10.1088/1742-6596/1521/4/042101

Akour, M. M. (2022). Rasch rating scale analysis of the survey of attitudes toward statistics. *Eurasia Journal of Mathematics, Science and Technology Education*, *18*(12), em210. https://doi.org/10.29333/ejmste/12646

Aliffiansyah, A. N., Ekarini, F., Djuniadi, D., Artikel, I., & Artikel, A. H. (2025). Implementasi model Rasch pada analisis kualitas butir soal jaringan dasar pada jenjang pendidikan vokasi. *Jurnal Metaedukasi*, *7*(2), 144–154. https://doi.org/10.37058/metaedukasi

American Educational Research Association, the A. P. A. and the N. C. on M. in E. (2014). *Standards for Educational and Psychological Testing*. American Educational Research Association.

Arsyad, R. N., Pomalato, S. W. Dj., Abbas, N., & Achmad, N. (2022). Hubungan antara self regulated learning dengan hasil belajar matematika pada materi trigonometri. *Jambura Journal of Mathematics Education*, *3*(1), 48–56. https://doi.org/10.34312/jmathedu.v3i1.12423

Bean, G. J., & Bowen, N. K. (2021). Item response theory and confirmatory factor analysis: Complementary approaches for scale development. *Journal of Evidence-Based Social Work (United States)*, *18*(6), 597–618. https://doi.org/10.1080/26408066.2021.1906813

Chang, C. C., Su, J. A., Tsai, C. S., Yen, C. F., Liu, J. H., & Lin, C. Y. (2015). Rasch analysis suggested three unidimensional domains for affiliate stigma scale: Additional psychometric evaluation. *Journal of Clinical Epidemiology*, *68*(6), 674–683. https://doi.org/10.1016/j.jclinepi.2015.01.018

Dai, Y., Luan, L., & Lin, X. (2023). The effects of online learning readiness on self-regulated learning for the first-time online learning students. *Asian Journal of Distance Education*, *18*(2), 42–62. https://doi.org/https://doi.org/10.5281/zenodo.8194076

Davis, D. R., & Boone, W. (2021). Using Rasch analysis to evaluate the psychometric functioning of the other-directed, lighthearted, intellectual, and whimsical (OLIW) adult playfulness scale. *International Journal of Educational Research Open*, *2*, 100054. https://doi.org/10.1016/j.ijedro.2021.100054

Defrijon, Matra, E., Lazuardi, M. R., Kurniadi, Y., Yardi, L., & Julhadi. (2025). Reliabilitas alat ukur, jenis-jenis dan faktor-faktor yang mempengaruhinya. *Jurnal Studi Multidisipliner*, *9*(1), 77–86. https://oaj.jurnalhst.com/index.php/jsm/article/view/8747

Ding, C., Lammert, C., Fulmer, G. W., Hand, B., & Suh, J. K. (2023). Refinement of an instrument measuring science teachers' knowledge of language through mixed method. *Disciplinary and Interdisciplinary Science Education Research*, *5*(1), 1–16. https://doi.org/10.1186/s43031-023-00080-7

Dzahabiyyah, M., Hikmah, N., & Ahludin, D. (2024). Pengaruh self regulation learning terhadap kemampuan pemahaman konsep matematika kelas VII. *Prosiding DIskusi Panel Nasional Pendidikan Matematika*, 583–590.

Ekawanti, S., & Mulyana, O. P. (2016). Regulasi diri dengan Burnout pada guru. *Jurnal Psikologi Teori Dan Terapan*, *6*(2), 113–118. https://doi.org/10.26740/jptt.v6n2.p113-118

Fitria, M., Fauziah, N., Hidayati, N., & Biologi, P. (2026). Validitas dan reliabilitas instrumen regulasi diri untuk siswa kelas XII SMA. *IDENTIK: Jurnal Ilmu Ekonomi, Pendidikan Dan Teknik*, *3*(1), 219–224. https://doi.org/https://doi.org/10.70134/identik.v3i1.1256

Kassiavera, S., Suparmi, A., Cari, C., & Sukarmin, S. (2024). Application of Rasch model in two-tier test for assessing critical thinking in physics education. *Journal of Baltic Science Education*, *23*(6), 1227–1242. https://doi.org/10.33225/jbse/24.23.1227

Kholisna, T., & Fatimah, S. (2022). Regulasi diri belajar online dimasa pandemi Covid-19. *Nusantara of Research : Jurnal Hasil-Hasil Penelitian Universitas Nusantara PGRI Kediri*, *9*(2), 108–115. https://doi.org/10.29407/nor.v9i2.17690

Kılıç, A. F., Koyuncu, İ., & Uysal, İ. (2023). Scale development based on item response theory: A systematic review. *International Journal of Psychology and Educational Studies*, *10*(1), 209–223. https://doi.org/10.52380/ijpes.2023.10.1.982

Lacave, C., Molina, A. I., & Redondo, M. A. (2018). A preliminary instrument for measuring students' subjective perceptions of difficulties in learning recursion. *IEEE Transactions on Education*, *61*(2), 119–126. https://doi.org/10.1109/TE.2017.2758346

Landa, J., Berciano, A., & Marbán, J. M. (2024). Self-regulation profiles of pre-service mathematics teachers for primary education in mathematical problem-solving contexts. *Education Sciences*, *14*(9), 1018. https://doi.org/10.3390/educsci14091018

Landa, J., Berciano, A., & Marbán, J. M. (2025). Moderating effect of perceived self-efficacy on university students'self-regulation in mathematics problem solving. *International Journal of Science and Mathematics Education*, *23*(8), 3815–3839. https://doi.org/10.1007/s10763-025-10597-0

López-Angulo, Y., Sáez-Delgado, F., Gaeta, M. L., Mella-Norambuena, J., González-Robaina, Y., & Muñoz-Inostroza, K. (2024). Validation of the self-regulation of learning instrument for undergraduates. *Frontiers in Education*, *9*. https://doi.org/10.3389/feduc.2024.1464424

Maharani, L., Putri, R. I. I., & Hartono, Y. (2019). Aquatic in Asian Games: Context of Pisa-like Mathematics Problem. *Journal on Mathematics Education*, *10*(3), 459–470. https://doi.org/10.22342/jme.10.3.5252.459-470

Marweli, Rahayu, W., & Aurelia Hidajat, F. (2025). Validitas dan reliabilitas angket self-regulated learning matematis siswa SMA. *SUPERMAT Jurnal Pendidikan Matematika*, *9*(1), 29–42. https://doi.org/https://doi.org/10.33627/sm.v9i1.2941

Mashuri, S., & Jahring. (2024). Test instrument development PISA-Based mathematical problem solving ability. *Scientific Research Journal of Education and Literature*, *4*, 1–6. https://doi.org/DOI:10.47310/srjehl.2024.v04.i02.003

Muharromah, F., Al Fani, F., & Ridho, A. R. (2025). Reliabilitas alat ukur: Jenis-jenis, cara pengukuran, dan faktor-faktor yang mempengaruhinya dalam pembelajaran. *JIPDAS: Jurnal Ilmiah Pendidikan Dasar*, (1), 63–73. https://doi.org/https://doi.org/10.65946/232k8198

Mutiara, T., & Rifameutia, T. (2021). Adaptasi alat ukur regulasi diri dalam belajar daring. *Edcomtech: Jurnal Kajian Teknologi Pendidikan*, *6*(2), 301–309. https://doi.org/10.17977/um039v6i12021p301

Noor, N. L., Waluya, S. B., & Widodo, S. A. (2025). Self-regulated learning for solving mathematical problems: A systematic literature review. *Jurnal Riset Pendidikan Matematika*, *12*(2), 235–246. https://doi.org/10.21831/jrpm.v12i2.89026

Nurvicalesti, N., & Ratnasari, R. (2023). Self-regulated learning dalam pembelajaran matematika pada peserta didik dalam menyelesaikan soal literasi numerasi. *Journal of Mathematics Education and Science*, *6*(2), 159–165. https://doi.org/10.32665/james.v6i2.1460

Panadero, E. (2017). A Review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*, *8*, 422. https://doi.org/10.3389/fpsyg.2017.00422

Parantika, I. W. A., Sariyasa, & Astawan, I. G. (2022). Pengembangan instrumen self regulated learning dan kecerdasan emosional pada pembelajaran matematika kelas V Sekolah Dasar. *PENDASI: Jurnal Pendidikan Dasar Indonesia*, *6*(2), 133–140. https://doi.org/10.23887/jurnal_pendas.v6i2.1357

Park, E. Y. (2025). Rasch analysis of general self-efficacy among individuals with intellectual disabilities. *Behavioral Sciences*, *15*(12), 1639. https://doi.org/10.3390/bs15121639

Pelikan, J. M., Link, T., Straßmayr, C., Waldherr, K., Alfers, T., Bøggild, H., Griebler, R., Lopatina, M., Mikšová, D., Nielsen, M. G., Peer, S., & Vrdelja, M. (2022). Measuring comprehensive, general health literacy in the general adult population: The development and validation of the HLS19-Q12 instrument in seventeen countries. *International Journal of Environmental Research and Public Health*, *19*(21), 14129. https://doi.org/10.3390/ijerph192114129

Perdana, K. I., & Wijaya, H. E. (2021). Regulasi diri dalam belajar sebagai prediktor resiliensi akademik mahasiswa yang sedang menyusun skripsi. *Psycho Idea*, *19*(2), 186–198. https://doi.org/https://doi.org/10.30595/psychoidea.v19i2.10711

Planinic, M., Boone, W. J., Susac, A., & Ivanjek, L. (2019). Rasch analysis in physics education research: Why measurement matters. *Physical Review Physics Education Research*, *15*(2), 020111. https://doi.org/10.1103/PhysRevPhysEducRes.15.020111

Pramesti, P., & Rosyidi, A. H. (2020). Profil argumentasi siswa dalam memecahkan masalah PISA-like berdasarkan model Toulmin. *Jurnal Riset Pendidikan Dan Inovasi Pembelajaran Matematika (JRPIPM)*, *3*(2), 92–101. https://doi.org/10.26740/jrpipm.v3n2.p92-101

Putri, R. I. I., & Zulkardi. (2020). Designing PISA-like mathematics task using Asian games context. *Journal on Mathematics Education*, *11*(1), 135–144. https://doi.org/10.22342/jme.11.1.9786.135-144

Qadaristin, D. F. L. (2021). Assessing self-regulated learning in primary school: A systematic literature review. *OPTIMA: Journal of Guidance and Counseling*, *1*(1), 55–72.

Rentiana, L. H., Dhini, U. R., Yuliawati, L., & Wulandari, E. T. (2024). Item fit analysis for evaluating academic writing performance with Rasch measurement. *Attractive : Innovative Education Journal*, *6*(1). https://doi.org/https://doi.org/10.51278/aj.v6i1.1144

Sagan, A. (2025). IRT reliability of coffee ritual behavior scale - a new metrics. *Coffee & Tea Marketing Journal*, *2*(2), 9–28. https://doi.org/10.69102/CTMJ.2025.020201

Saputra, W. N. E., Hidayah, N., Ramli, M., & Atmoko, A. (2023). Development and validation of Indonesian peace of mind scale: The Rasch analysis. *International Journal of Evaluation and Research in Education*, *12*(2), 877–885. https://doi.org/10.11591/ijere.v12i2.24847

Saputri, V., Juandi, D., Herlina, S., & Anwar, V. N. (2022). Self-regulated learning dan motivasi belajar dalam pembelajaran matematika secara online: Systematic literature review. *ANARGYA: Jurnal Ilmiah Pendidikan Matematika*, *5*(1), 53–60. https://doi.org/10.24176/anargya.v5i1.7208

Sari, S., Sari Dewi, R., Pribadi, P., & Sulistyaningrum, I. H. (2026). Psychometric validation of a culturally adapted medication literacy questionnaire in Indonesian communities: evidence from EFA and Rasch modeling. *Pharmacia*, *73*(73), 1–10. https://doi.org/10.3897/pharmacia.73.e176368

Siregar, H. M. (2021). Pengembangan angket self-regulation mahasiswa pendidikan matematika di masa pandemi Covid-19. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, *10*(3), 1685–1695. https://doi.org/10.24127/ajpm.v10i3.3870

Sujatmika, S., Masykuri, M., & Prayitno, B. A. (2025). Applying the Rasch model to measure students' critical thinking skills on the science topic of the human circulatory system. *Eurasia Journal of Mathematics, Science and Technology Education*, *21*(4), em2622.

Torregrosa, A., & Albarracín, L. (2025). Characterizing metacognitive processes and mathematical skills during problem solving by using a non-linear orientation base as a self-regulation tool. *Educational Studies in Mathematics*, 1–32. https://doi.org/10.1007/s10649-025-10451-8

Verdú-Soriano, J., & González-de la Torre, H. (2024). Rasch analysis implementation in nursing research: A methodological approach. *Enfermeria Clinica*, *34*(6), 493–506. https://doi.org/10.1016/j.enfcli.2024.10.005

Wess, R., Klock, H., Siller, H.-S., & Greefrath, G. (2021). *Measuring Professional Competence for the Teaching of Mathematical Modelling: A Test Instrument*. Springer. http://www.springer.com/series/10093

Wiyono, B. D., Hidayah, N., Ramli, M., Atmoko, A., & Shafie, A. A. H. (2025). Development and instrument validation of Indonesian achievement motivation scale using the Rasch model. *International Journal of Evaluation and Research in Education*, *14*(2), 1427–1435. https://doi.org/10.11591/ijere.v14i2.29374

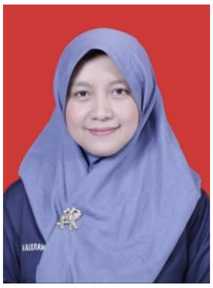
Zayrin, A. A., Nupus, H., Maizia, K. K., Marsela, S., Hidayatullah, R., & Harmonedi. (2025). Analisis instrumen penelitian pendidikan (uji Validitas dan relibilitas instrumen penelitian). *Jurnal QOSIM Jurnal Pendidikan Sosial & Humaniora*, *3*(2), 780–789. https://doi.org/10.61104/jq.v3i2.1070

Zenger, T., & Bitzenbauer, P. (2022). Exploring German secondary school students' conceptual knowledge of density. *Science Education International*, *33*(1), 86–92. https://doi.org/10.33828/sei.v33.i1.9

Zimmerman, B. J. (2000). Attaining Self-Regulation. In *Handbook of Self-Regulation* (pp. 13–39). Elsevier. https://doi.org/10.1016/B978-012109890-2/50031-7

AUTHOR PROFILES

Aminah Ekawati, aminahekawati@upk.ac.id, Universitas PGRI Kalimantan. Google Scholar: mteCG4QAAAAJ, Sinta: 6136742, Orcid: 0009-0008-5596-9579, Scopus: 5954820000



Isna Wardiah, isnawardiah@poliban.ac.id, Politeknik Negeri Banjarmasin. Google Scholar: o0QBZGkAAAAJ , Sinta: 6011162, Orcid: 0009-0008-6666-2273, Scopus: 58080438800