ISSN 2502-5872 (Print)
ISSN 2622-3627 (Elektronik)

Volume 10 Number 1, February 2025, 55-69

DEVELOPMENT AND VALIDATION OF A SELF-CONFIDENCE MEASUREMENT INSTRUMENT IN MATHEMATICS LEARNING USING THE RASCH MODEL

Sintha Sih Dewanti[1*], Adib Ashabulabib[2], Indriana Puspita Dewi[3], Aris Mislikhatun Islamiyah[4]
[1, 2, 3, 4]Departement of Mathematics Education, UIN Sunan Kalijaga, Daerah Istimewa Yogyakarta, Indonesia
*Correspondence: sintha.dewanti@uin-suka.ac.id

ABSTRACT

Self-confidence is a critical factor in mathematics learning, influencing students' engagement and performance. However, studies show that Indonesian students often lack confidence in mathematics subject, hindering their academic potential. Accurate instruments are needed to measure and address this issue effectively. This tudy aims to develop a self-confidence measurement instrument for mathematics learning among secondary school students. The instrument comprises 25 statements categorized into four main aspects. The research employs a descriptive quantitative method with the Rasch Model approach and utilizes exploratory and confirmatory factor analysis. The sample consists of 122 randomly selected tenth-grade students in Yogyakarta. The results indicate that the developed instrument has high validity, with an Aiken's V coefficient of 0.99, and very high reliability, evidenced by a Cronbach's Alpha coefficient of 0.845 with a total of 18 items used. Factor analysis shows that the items in the self-confidence questionnaire meet the unidimensionality requirement, although some items need correction to enhance measurement accuracy for one self-confidence indicator. Overall, this instrument provides a valid and reliable tool for measuring students' self-confidence in mathematics learning, potentially contributing to the enhancement of students' confidence in this subject. So that, 18 statements are more efficient to use, without reducing the scope of the self-confident aspect being measured. The results allow the use of instruments for various levels of education and support the implementation of more personalized learning strategies.
Keywords: Item Response Theory, Mathematics Learning, Measurement Instrument, Rasch Model, Self-Confidence.

How to Cite: Dewanti, S. S., Ashabulabib, A., Dewi, I. P., & Islamiyah, A. M. (2025). Development and Validation of A Self-Confidence Measurement Instrument In Mathematics Learning Using The Rasch Model. *Mathline: Jurnal Matematika dan Pendidikan Matematika*, *10*(1), 55-69. https://doi.org/10.31943/mathline.v10i1.697

PRELIMINARY

Education is one way to improve human ethics and knowledge. Education is defined as a process that allows learners to change their behaviour to become more mature, independent, and able to integrate in society, especially in developing their personality (Yusniawan et al., 2019). Education is also considered a process of instilling knowledge, moral values, culture, and thoughts that guide human life (Asmara, 2016).With so many positive impacts, education has become essential in life, including in the field of mathematics.

The word "mathematics" originally comes from the Greek word "mathematike" which means "studying" (Gusteti & Neviyarni, 2022). Mathematics is a science that is obtained through a thinking process (Rahmah, 2018). Mathematical concepts are generated from the thinking process, so logic is the basis for the formation of mathematics. Mathematics is a subject that trains critical, logical, and systematic thinking skills (Rachmantika & Wardono, 2019). The importance of mathematics in life encourages humans to learn it, one way is through learning mathematics.

Mathematics learning is a teaching and learning process that aims to develop students' creativity and critical thinking skills, and help them build new knowledge to improve their understanding of mathematics material (Anisa et al., 2020). The existence of mathematics learning in schools aims to make students understand mathematical concepts, be able to explain the relationship between concepts, and be able to use concepts and algorithms flexibly, accurately, efficiently, and precisely in solving problems (Pujiadi, 2016). According to Kholil & Zulfiani (2020), learning mathematics at school aims for students to have the ability to understand concepts, use reasoning, solve problems, appreciate the benefits of mathematics in life, and communicate ideas through symbols and tables. However, obstacles such as students' lack of confidence can cause failure in learning mathematics.

Self-confidence comes from the word ‘self’ which means self and ‘confidence’ which means trust, so self-confidence can be interpreted as confidence in one's own abilities. Self-confidence is a belief in one's own ability to solve a problem (Adhimah & Simatupang, 2014). According to Rohana et al. (2020), self-confidence is a belief in one's own ability to complete tasks, which includes self-efficacy, optimism, objectivity, responsibility, and realistic rationality. Individuals with high self-confidence can complete tasks or work effectively according to their stage of development.

Self-confidence is an essential requirement for individuals to develop activities and creativity in an effort to achieve optimal learning achievement and learning outcomes (Andayani & Amir, 2019). According to Amri (2018), self-confidence is the most valuable attribute of a person in social life, because with self-confidence, a person is able to actualise all the potential that exists within him. Self-confidence has a crucial role in helping learners become more active and confident in exploring their abilities during the learning process. Therefore, self-confidence needs to be considered in learning activities so that each individual can develop their potential optimally.

According to TIMSS research (Herdiana et al., 2019), Indonesian students' confidence in mathematics skills is still low, which is below 30% compared to other countries. Research conducted by Fardani et al., (2021). shows that out of 30 students, there are 6 students who have high category self-confidence, 20 students who have a medium category, and 4 students in the low category. Learners who lack confidence tend to be passive, undeveloped, and lack enthusiasm in learning mathematics, and are shy to ask questions, which ultimately leads to a lack of understanding of the material. Therefore, an accurate instrument is needed to measure students' confidence levels to identify areas that need to be improved.

Instruments are tools for collecting and measuring quantitative information about the variables to be studied (Nasution, 2016). Instruments need to be designed with precision to ensure ease of use and accurate results (Dewanti et al., 2020). Assessment instruments can measure cognitive, affective, and psychomotor abilities. The development of self-confidence instruments is important to determine the level of student self-confidence and its effect on learning. Research conducted by Qurni (2021) found that the self-confidence instrument developed has a high reliability value so it is very feasible to develop and use. Further research revealed that the self-confidence instrument developed was very reliable with an alpha Cronbach value of 0.956 (Nursanti et al., 2020). The results of some of these studies can be concluded that the self-confidence assessment instrument is very feasible to develop because it has high reliability.

Most previous research on measuring students' self-confidence in mathematics learning tends to use classical approaches, such as simple descriptive analysis or conventional validation methods, which are often unable to capture the complexity of the data or ensure optimal measurement accuracy. This study uses the application of the Rasch model and exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) which are known to be able to provide more accurate, valid and reliable analysis results, especially in the context of measuring psychological aspects such as self-confidence, which are still rarely the main focus in similar studies. This creates an alignment, where there is a need to develop measurement instruments that not only meet the requirements of high validity and reliability, but also use more sophisticated approaches such as Rasch models to improve the quality of measurement results.

Based on the above description, self-confidence is an important factor in learning, especially in learning mathematics, because it directly affects students' activity, creativity and learning achievement. The low level of student confidence in Indonesia, as shown in the TIMSS research and various other studies, highlights the need for serious attention to this aspect. An accurate and reliable instrument is needed to measure students' confidence level, so that it can help teachers and researchers in designing more effective learning strategies. Therefore, this study aims to develop a confidence measurement instrument for learners in the hope that it can make a significant contribution in improving the quality of mathematics learning.

METHODS

This study is a quantitative descriptive study focusing on the analysis of non-test mathematics self-confidence instruments based on IRT with the Rasch model approach. This instrument was developed based on a literature review of self-confidence, interviews with mathematics teachers, and input from expert lecturers from mathematics learning experts. The instrument was designed to reflect aspects of self-confidence which are divided into 4, namely believing in one's own abilities, acting independently in making decisions, having a positive sense of oneself and daring to express opinions. The sampling technique used was random sampling to select research subjects, where the total student population was 288 and respondents were randomly selected from each class to ensure proportional representation. The research subjects were 122 students of grade X of Senior High Schools (SMA) in Yogyakarta. The sample size of 122 students was chosen based on the results of a literature study on Rasch analysis, which suggests a minimum of 100-150 respondents to ensure the stability of item parameter estimates. In addition, it also considers the level of population heterogeneity, including variations in the level of self-confidence. There are 25 statements about mathematics self-confidence which are closed questionnaires. The statements include aspects of self-confidence in one's abilities, making decisions independently, having a positive self-image, and daring to express opinions. Respondents can choose 4 answers, namely very inappropriate, inappropriate, appropriate, and very appropriate, each of which is given a score of 1 - 4 for positive statements and vice versa for negative statements. Therefore, the data obtained are polytomous.

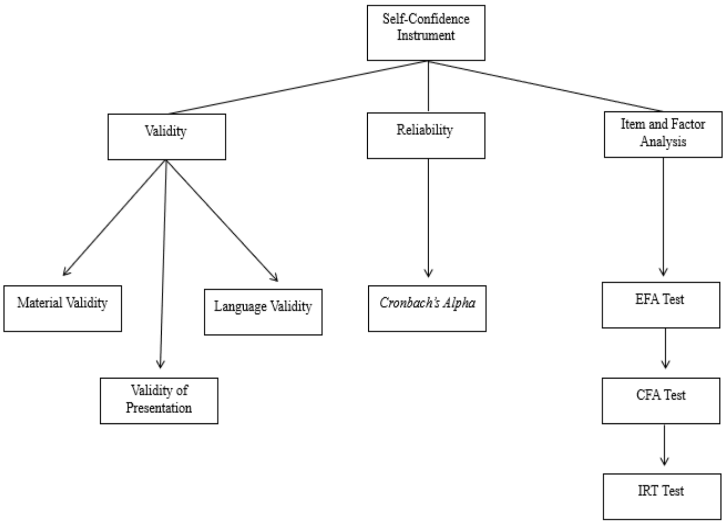

Figure 1. Flow of Self Confidence Measurement Instrument Development

The test results were analyzed using the Rasch Item Response Theory (IRT) model. There are several polynomial models in IRT including Rating Scale Model (RSM), Graded Response Model (GRM), Partial Credit Model (PCM), and Generalized Partial Credit Model (GPCM) (Retnawati, 2014). Related to discriminant parameters, the Rasch model (PCM) assumption that items have the same discrimination value in empirical data is usually violated (Diputera, 2018). The Rasch model is used because it has a simple but effective approach in measuring self-confidence. The Generalized Partial Credit Model (GPCM) is one of the IRT models designed for polytomous data, namely data with more than two answer categories, such as the Likert scale. GPCM is also able to capture transitions between answer categories accurately, thus providing more detailed information about self-confidence.

In this study, the item response model that can be used is GPCM, with data analysis techniques carried out using the IRT model with the GPCM approach and computing using R software. Quantitative data analysis techniques are employed to evaluate the validity of content and self-confidence instruments, which are validated by six validators, including mathematics education lecturers and experts. The validation technique used is the Delphi technique, which is an assessment conducted by a group of experts continuously through a questionnaire to facilitate the formation of a group decision, without the need to be entirely face-to-face (Lukman & Mulyanti, 2017). Validators chose answers by paying attention to the suitability of aspects, indicators, and statements.

RESULT AND DISCUSSION

This research is a descriptive quantitative research of self confidence non-test instruments. The non-test instrument developed is a questionnaire with a closed statement to measure the level of self-confidence of students in mathematics learning. The results of the study based on the stages carried out the first is testing validation and reliability. Instrument validation was carried out by 6 validators consisting of mathematics education lecturers and experts. Validators assess answer choices based on the appropriateness of aspects, indicators, and statements. All items in the questionnaire are considered valid due to an Aiken's V coefficient of 0.99, determined with input from 6 validators and using answer choices ranging from 1 to 5, indicating high validity of the instrument.

The reliability of the instruments, reflecting measurement error, is demonstrated by a Cronbach's Alpha coefficient of 0.845, indicating very high reliability. According to Ghozali in (Verdian, 2019), that a construct or variable is said to be reliable if it has a Cronbach's Alpha value > 0.6. The results of the corrected total correlation coefficient for each item are greater than 0.6, so it can be concluded that each item in the statement is reliable. Therefore, this self-confident instrument is suitable for use.

In analyzing data using item response theory, the first thing to do is to test the dimensions of empirical data. The testing process is carried out by Exploratory Factor Analysis (EFA) with the principal component method. The assumptions of the Kaiser-Meyer-Olkin (KMO) and Bartlett Tests that must be met in the EFA are shown in Figure 2.

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		,772
Bartlett's Test of Sphericity	Approx. Chi-Square	943,053
	df	276
	Sig.	,000

Figure 2. KMO and Bartlett's Test Results

The KMO test results are KMO > 0.50, according to Subhash Sharma in (Nafisah & Respatiwulan, 2019), if the KMO value is less than 0.50 then factor analysis is not suitable for these variables. So that from the test results above, it shows that the sample size of 122 already has sufficient data. According to Verdian (2019), that if the Bartlett's Test produces a significance value < 0.05, it is concluded that the indicators used are correlated and suitable for factor analysis. The above data forms a correlation matrix with a close relationship between variables, as evidenced by the significance of Bartlett's test (sig. < 0.05) which states that HO (correlation matrix is an identity matrix) is rejected. Based on the analysis results, the KMO value is 0.772 and the Bartlett's test significance value for sphericity is 0.000. These results indicate that the data used are suitable for factor analysis KMO. Therefore, factor analysis can be continued. According to Rahayu & Haryanto (2017), that if KMO-MSA > 0.5 then factor analysis can be continued and if furthermore, based on anti-image correlation, Measures of Adequate Sampling (MSA) > 0.5, so 24 statements are suitable for factor analysis and 1 statement, namely in one aspect of having a positive sense of self. The anti-image matrix output is presented in Figure 3.

Anti-image Matrices

		A1	A2	A3	A4	A5	A6	B1	B2	B3	B4	B5	B6	B7	C2	C3	C4	C5	C6	D1	D2	D3	D4	D5	D6
Anti-image Correlation	A1	,740[a]	-0,058	-0,083	-0,049	0,165	-0,295	-0,048	-0,037	-0,144	-0,046	-0,082	0,278	-0,102	-0,077	-0,263	-0,006	0,066	-0,036	-0,169	-0,042	0,012	-0,008	0,091	-0,048
	A2	-0,058	,752[a]	0,016	-0,301	-0,034	-0,003	-0,002	-0,242	-0,070	-0,025	-0,051	0,031	-0,149	-0,215	0,034	0,048	0,102	-0,124	-0,023	0,121	0,054	0,054	0,098	-0,032
	A3	-0,083	0,016	,702[a]	-0,037	0,173	-0,016	-0,020	0,026	0,116	0,036	0,077	0,067	0,019	-0,073	0,058	-0,335	-0,033	-0,023	-0,060	-0,371	-0,069	0,021	0,014	-0,038
	A4	-0,049	-0,301	-0,037	,751[a]	-0,273	0,042	0,135	0,037	-0,052	-0,176	-0,077	-0,075	0,200	-0,095	0,154	0,072	-0,233	0,103	-0,026	-0,013	-0,031	0,022	-0,104	-0,136
	A5	0,165	-0,034	0,173	-0,273	,751[a]	-0,094	-0,140	-0,118	0,009	0,077	0,009	0,182	-0,198	0,058	-0,187	0,022	-0,123	0,033	0,079	-0,060	0,020	-0,125	-0,085	0,131
	A6	-0,295	-0,003	-0,016	0,042	-0,094	,720[a]	0,055	-0,047	0,194	0,024	0,035	-0,331	0,063	0,000	0,170	-0,193	-0,402	0,055	0,011	0,212	-0,024	0,016	-0,114	-0,024
	B1	-0,048	-0,002	-0,020	0,135	-0,140	0,055	,848[a]	-0,138	-0,054	-0,075	0,066	0,012	0,109	-0,090	-0,060	-0,015	-0,154	0,028	-0,044	-0,014	0,080	-0,251	0,028	-0,226
	B2	-0,037	-0,242	0,026	0,037	-0,118	-0,047	-0,138	,833[a]	-0,205	0,052	-0,008	-0,237	-0,116	0,195	-0,076	-0,096	-0,088	-0,015	0,159	-0,118	0,017	-0,008	0,055	0,050
	B3	-0,144	-0,070	0,116	-0,052	0,009	0,194	-0,054	-0,205	,739[a]	-0,048	-0,101	0,009	-0,068	0,258	0,056	-0,223	0,072	-0,176	-0,096	0,000	0,144	0,049	-0,223	-0,143
	B4	-0,046	-0,025	0,036	-0,176	0,077	0,024	-0,075	0,052	-0,048	,841[a]	0,131	-0,231	-0,040	0,134	-0,160	-0,067	-0,067	-0,298	0,023	-0,021	0,006	-0,160	0,093	0,070
	B5	-0,082	-0,051	0,077	-0,077	0,009	0,035	0,066	-0,008	-0,101	0,131	,736[a]	-0,310	0,196	0,063	-0,060	-0,054	-0,125	-0,190	0,012	-0,107	-0,172	-0,037	0,182	0,030
	B6	0,278	0,031	0,067	-0,075	0,182	-0,331	0,012	-0,237	0,009	-0,231	-0,310	,710[a]	-0,258	-0,032	-0,243	0,130	0,109	0,125	0,008	0,018	-0,125	-0,053	-0,035	-0,108
	B7	-0,102	-0,149	0,019	0,200	-0,198	0,063	0,109	-0,116	-0,068	-0,040	0,196	-0,258	,773[a]	-0,068	0,103	-0,200	-0,014	0,005	-0,275	-0,132	-0,124	-0,024	-0,030	-0,030
	C2	-0,077	-0,215	-0,073	-0,095	0,058	0,000	-0,090	0,195	0,258	0,134	0,063	-0,032	-0,068	,666[a]	-0,023	-0,320	-0,148	-0,007	0,083	-0,132	0,025	-0,007	-0,115	0,069
	C3	-0,263	0,034	0,058	0,154	-0,187	0,170	-0,060	-0,076	0,056	-0,160	-0,060	-0,243	0,103	-0,023	,757[a]	0,066	-0,095	0,052	-0,262	0,010	-0,053	0,047	-0,142	0,139
	C4	-0,006	0,048	-0,335	0,072	0,022	-0,193	-0,015	-0,096	-0,223	-0,067	-0,054	0,130	-0,200	-0,320	0,066	,689[a]	-0,054	0,083	0,051	-0,009	-0,031	0,115	-0,001	0,079
	C5	0,066	0,102	-0,033	-0,233	-0,123	-0,402	-0,154	-0,088	0,072	-0,067	-0,125	0,109	-0,014	-0,148	-0,095	-0,054	,819[a]	-0,375	-0,205	0,057	-0,112	-0,051	0,038	0,071
	C6	-0,036	-0,124	-0,023	0,103	0,033	0,055	0,028	-0,015	-0,176	-0,298	-0,190	0,125	0,005	-0,007	0,052	0,083	-0,375	,801[a]	0,137	-0,072	0,016	-0,046	-0,095	-0,022
	D1	-0,169	-0,023	-0,060	-0,026	0,079	0,011	-0,044	0,159	-0,096	0,023	0,012	0,008	-0,275	0,083	-0,262	0,051	-0,205	0,137	,785[a]	0,145	-0,010	-0,031	-0,059	-0,194
	D2	-0,042	0,121	-0,371	-0,013	-0,060	0,212	-0,014	-0,118	0,000	-0,021	-0,107	0,018	-0,132	-0,132	0,010	-0,009	0,057	-0,072	0,145	,684[a]	-0,012	-0,208	0,012	-0,200
	D3	0,012	0,054	-0,069	-0,031	0,020	-0,024	0,080	0,017	0,144	0,006	-0,172	-0,125	-0,124	0,025	-0,053	-0,031	-0,112	0,016	-0,010	-0,012	,856[a]	-0,020	0,004	-0,115
	D4	-0,008	0,054	0,021	0,022	-0,125	0,016	-0,251	-0,008	0,049	-0,160	-0,037	-0,053	-0,024	-0,007	0,047	0,115	-0,051	-0,046	-0,031	-0,208	-0,020	,858[a]	-0,311	0,017
	D5	0,091	0,098	0,014	-0,104	-0,085	-0,114	0,028	0,055	-0,223	0,093	0,182	-0,035	-0,030	-0,115	-0,142	-0,001	0,038	-0,095	-0,059	0,012	0,004	-0,311	,799[a]	-0,234
	D6	-0,048	-0,032	-0,038	-0,136	0,131	-0,024	-0,226	0,050	-0,143	0,070	0,030	-0,108	-0,030	0,069	0,139	0,079	0,071	-0,022	-0,194	-0,200	-0,115	0,017	-0,234	,778[a]

a. Measures of Sampling Adequacy(MSA)

Figure 3. **Anti-image Matrix**

Maximum Likelihood (ML) can be used to reduce the amount of observed variability so that only a few principal components are formed from most of the observed variables. This can be achieved by selecting factors with initial eigenvalues greater than 1. The ML initial eigenvalues are shown in Figure 4.

Total Variance Explained

Factor	Initial Eigenvalues			Loadings			Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	5,743	23,930	23,930	2,583	10,764	10,764	2,613	10,887	10,887
2	2,369	9,872	33,803	3,563	14,844	25,608	2,525	10,522	21,410
3	1,822	7,591	41,393	1,558	6,494	32,102	2,000	8,335	29,745
4	1,435	5,981	47,374	1,423	5,927	38,029	1,988	8,284	38,029
5	1,326	5,527	52,901						
6	1,270	5,292	58,193						
7	1,100	4,584	62,777						
8	1,048	4,367	67,144						
9	0,839	3,496	70,640						
10	0,775	3,230	73,871						
11	0,736	3,067	76,938						
12	0,664	2,769	79,706						
13	0,623	2,598	82,304						
14	0,586	2,443	84,746						
15]	0,556	2,316	87,062						
16	0,504	2,098	89,160						
17	0,474	1,975	91,135						
18	0,396	1,650	92,785						
19	0,348	1,449	94,234						
20	0,326	1,357	95,591						
21	0,308	1,284	96,875						
22	0,291	1,211	98,086						
23	0,255	1,063	99,149						
24	0,204	0,851	100,000						

Extraction Method: Maximum Likelihood.

Figure 4. Results of Initial Eigenvalues of Maximum Likelihood

The results of factor analysis show that 8 factors have eigenvalues greater than 1. The percentage of construct variance that can be explained by the factors formed is the total variance explained. The cumulative sub column in the initial eigenvalue column shows that reducing 24 items to one factor accounts for 23.930% of the variance, two factors account for 33.803%, three factors explain 41.393%, and four factors account for 47.374% of the variance.

The decision on how many factors to take can be based on the eigenvalue. Generally, factors with eigenvalues above 1 are considered significant, so from the data it can be concluded that from 24 items, it can be reduced to 4 factors. The initial eigenvalue of the first factor is almost the same as the eigenvalue of the second factor, then the eigenvalue of the second factor, and so on are almost the same. Therefore, self-confidence meets the requirements of unidimensionality through non-mathematics test instruments (Friyatmi, 2018). Unidimensionality is needed to ensure that the self-confidence instrument truly measures the self-confidence construct in mathematics learning according to the aspects of self-confidence in this study, without being influenced by other factors that are not relevant to the aspects that have been determined. This is supported by the results of the exploratory factor analysis (EFA) showing that the KMO is 0.772, with a Bartlett's test significance of 0.000, indicating that the data is suitable for factor analysis. The first factor explained 23.93% of the total variance, while the eigenvalue of the second factor and so on showed a significant reduction in the contribution of variance, supporting the assumption of unidimensionality. Based on the Scree Plot, the inflection point occurred after the 8th factor, indicating that most of the variation was explained by one main dimension. In addition, the CFA results showed that all paths were significant after modification, with a p-value of 0.0047, thus supporting a data structure that was consistent with the construct of students' self-confidence in learning mathematics. These findings indicate that the instrument meets the requirements of unidimensionality and is in accordance with the Rasch Model. The unidimensionality property can be described more clearly by looking at the elbows (bends) in Figure 5.

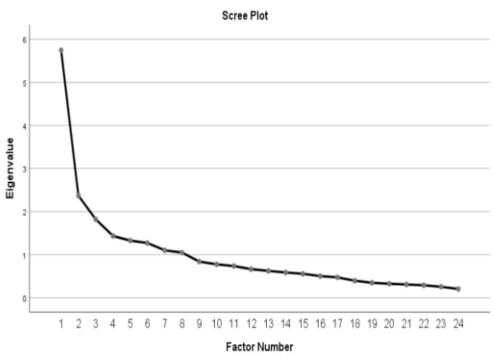

Figure 5. Scree Plot of Maximum Likelihood

Thus, based on empirical data, the mathematics self-confidence statement items can be analyzed to determine how many factors exist in the pattern on the scree plot. The point at which there is a change in the slope of the line indicates the limit to the number of factors to be taken, known as the inflection point. From the figure above, it can be seen that after reaching the 8th point, the line starts to show a change in slope and the variation explained decreases. Therefore, based on this analysis, it can be reduced that of the 24 items, the model can be reduced to 8 factors. The next step involves validating the theoretical construct of the non-test instrument for mathematics self-confidence through confirmatory factor analysis. The t-values from this analysis are illustrated in Figure 6.

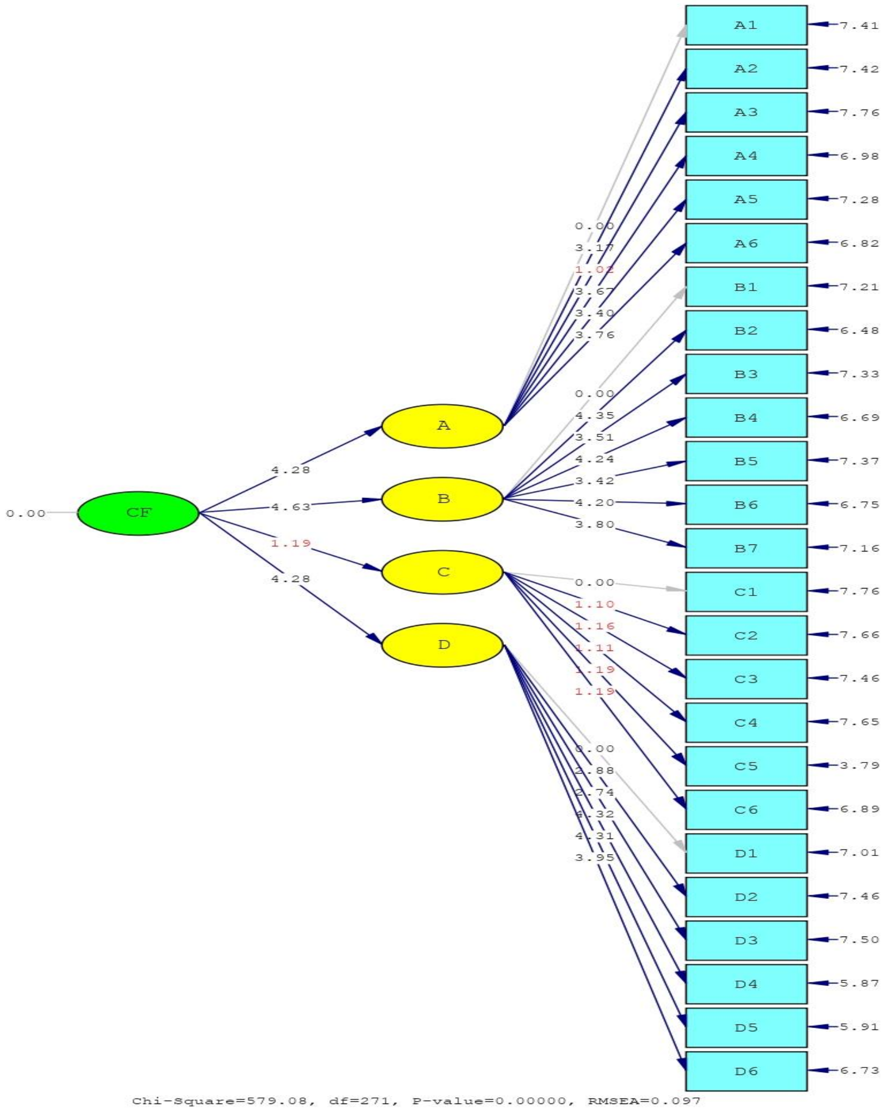

Figure 6. T-value of the Self Confidence Construct Theory

Figure 6 shows that the paths of C and A3 are not significant at p-value = 0.000. Therefore, researchers should focus on the modification index to identify suggestions for enhancing the model. Following adjustments, the finalized solution is depicted in Figure 7.

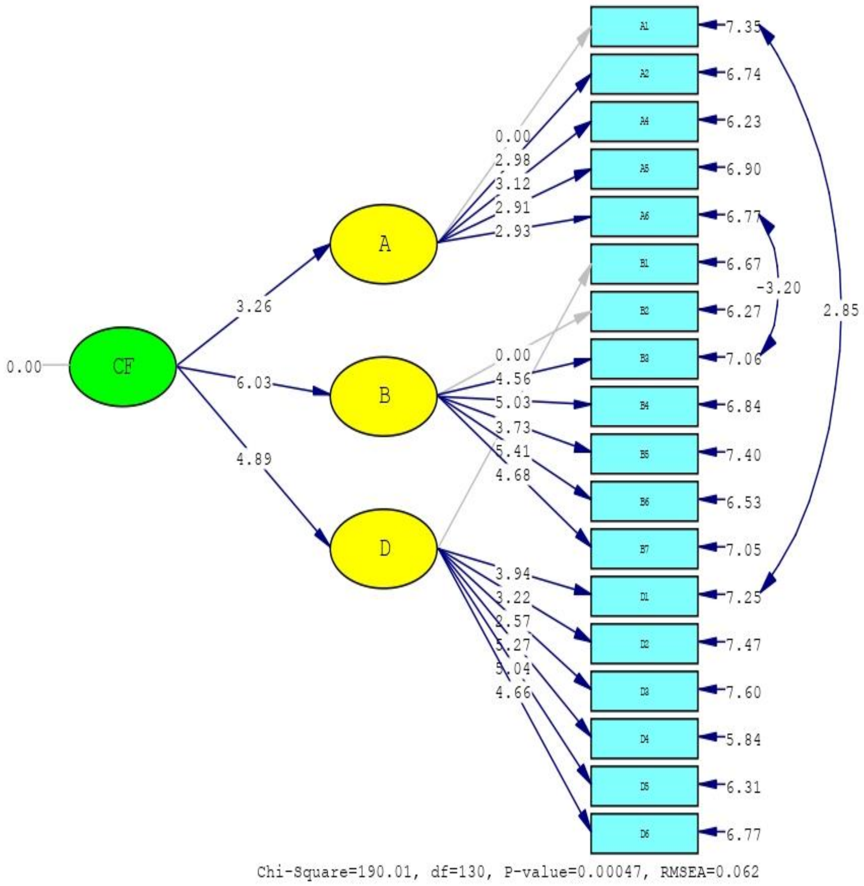

Figure 7. **Standardized solution of Self Confidence Construct Theory**

Figure 6 shows the results of the Self Confidence construct theory after being modified. The result is that all paths are significant with a p-value of 0.0047. Based on the results of the CFA analysis output, the data is collected in table 1 below.

Table 1. CFA Analysis Output Data

No	Goodness of Fit Model	Standart	Initial Result	Final Result
1	Goodness Fit Index (GFI)	≥ 0,90	0,72	0,85
2	Root Mean Square Error Of Approximation (RMSEA)	≤ 0,05	0,97	0,062
3	Comparative Fit Index (CFI)	≥ 0,90	0,81	0,92
4	Adjusted Goodness of Fit (AGFI)	≥ 0,90	0,67	0,80
5	Normed Fit Index (NFI)	≥ 0,95	0,69	0,80

From table 1, GFI (goodness fit index) is one of the indices often used to assess model fit. GFI measures how well the model explains the existing data. GFI can be used as one of the determinations of a fit model, namely the GFI value should be ≥ 0.90, with a range between 0.00 (no fit/poor fit) to 1.00 (good fit or perfect fit). The initial result of the

GFI of this model is 0.72. The final GFI result of this model is 0.85. From this GFI result, it increased by about 0.13. Although the GFI results have increased, the GFI value is still ≥0.90 so this model is not suitable enough as a reference for assessment.

Root Mean Square Error of Approximation (RMSEA) describes the model remaining error. The expected RMSEA value ≤ 0.05 indicates good fit, while the value of 0.05 between 0.08 the model can be said to be still accepted as a good fit model (Waluyo, 2016). The initial RMSEA result in this initial model is 0.97. The final result of RMSEA on this model is 0.062. This RMSEA result decreased by about 0.908. Even though the RMSEA results have decreased, they are still more than 0.05 and are in the range between 0.05 and 0.08, which means that this model can be accepted as a model that is quite suitable or a fit model.

Comparative Fit Index (CFI) is the value of comparing the model compiled with the ideal model. A CFI value above 0.90 is considered good (Waluyo, 2016). The initial CFI result in this initial model is 0.81. The final CFI result on this model is 0.92. This CFI result increased by about 0.11. The final CFI result of this model is ≥ 0.90, indicating that the revised model has been compiled with an ideal model.

Adjusted Goodness of Fit (AGFI) is one of the GFI developments which is the recommended fit index criteria for a suitable model which is ≥ 0.90 (Waluyo, 2016). The initial AGFI result in this initial model is 0.67. The final AGFI result on this model is 0.80. This AGFI result increased by about 0.13. The final AGFI result is ≤0.90, indicating that the revised model lacks fit.

NFI (Normed Fit Index) is a comparison between the proposed model and the null model, and the expected value is ≥ 0.95 (Waluyo, 2016). The initial NFI result in this initial model is 0.69. The final NFI result on this model is 0.80. This NFI result increased by about 0.11. The final NFI result is ≤0.90, indicating that the revised model lacks fit.

The threshold represents the point where each category intersects, indicating the minimum ability required to reach a specific level. Given the 4 answer choices on the attitude scale, the thresholds for each item are detailed in the following table.

Table 2. IRT Results R Program

Item	Location	b1	b2	b3
A1	-0.551	-2.875	1.773	NA
A2	0.413	-2.147	0.842	2.546
A4	0.971	-2.501	0.956	4.458
A5	0.271	-2.455	0.270	2.999
A6	0.253	-2.283	0.548	2.495
B1	-0.126	-3.491	-0.326	3.438

Item	Location	b1	b2	b3
B2	0.521	-3.060	0.824	3.797
B3	-0.813	-3.547	-1.119	2.228
B4	-0.220	-2.859	-0.487	2.685
B5	-0.964	-3.302	-1.531	1.942
B6	-0.661	-2.848	1.525	NA
B7	-1.229	-3.061	-1.840	1.214
D1	-1.674	-3.122	1.525	0.807
D2	-0.604	-3.267	-0.812	2.269
D3	-1.178	-4.051	-1.503	2.021
D4	-0.439	-2.749	-0.705	2.138
D5	-0.787	-2.018	-1.898	1.556
D6	-1.056	-3.537	-1.075	1.445

Out of the 18 items, 16 items feature 3 thresholds each, while 2 items have 2 thresholds. Typically, the sequence follows that threshold 1 < threshold 2 < threshold 3, reflecting a higher likelihood of agreement with "strongly agree" compared to "agree". Seventeen items exhibit thresholds that increase progressively from one category to the next. Item characteristic curves depicting this threshold pattern are shown in the following figure.

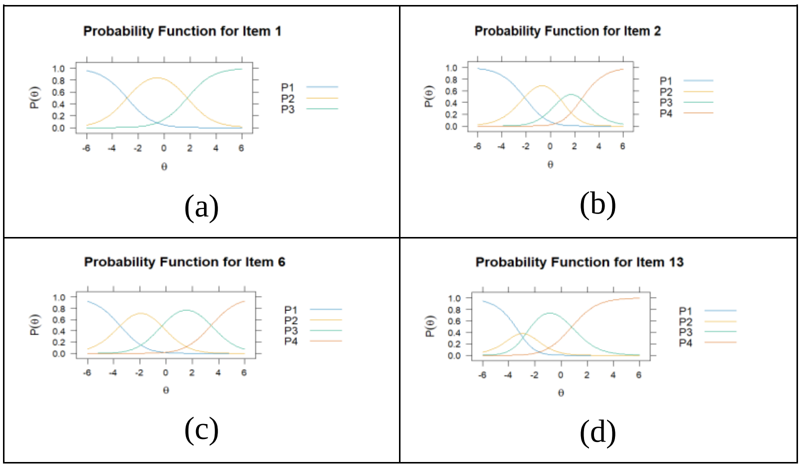

Figure 8. Function Response Category Graph

Based on the function graph above, three thresholds are obtained, the first threshold is the intersection between P1 and P2, the second threshold is the intersection between P2 and P3, and the third threshold is the intersection between P3 and P4. In figure (a) it can be seen that the threshold only reaches threshold 2 with threshold 1 having a negative value and threshold 2 having a positive value. The threshold pattern for Item 1 shows a regular probability curve for categories P1, P2, and P3. Each category has a clear probability peak at a certain ability value (θ), indicating that the category is functioning well.

In figure (b) it can be seen that the threshold reaches threshold 3 with threshold 1 having a negative value, thresholds 2 and 3 having positive values. The threshold pattern for Item 2 is also regular, with four categories (P1, P2, P3, and P4) each of which has a dominant probability area. This shows that the categories for Item 2 distinguish respondents' abilities well. In figure (c) it can be seen that the threshold reaches threshold 3 with the first threshold being negative, the second threshold being positive, and the third threshold being positive. The probability curve shows a pattern that is close to regular. However, there is little overlap between categories P2 and P3 in the range of θ between -2 and 2. This indicates that some respondents may have difficulty distinguishing between these two categories at that level of ability. In figure (d) it can be seen that the threshold reaches threshold 3 with threshold 1 being negative, threshold 2 being positive, and threshold 3 being positive, but from threshold 2 to threshold 3 there is a decrease. In Item 13, the probability curve shows that categories P3 and P4 have significant overlap in the range of θ around 2 to 4. This indicates a potential problem in discrimination between these categories, which may affect the function of the item. Some items deviate due to less specific statement formulation, unbalanced answer categories, respondents' perceptions of similar categories, and uneven distribution of respondents in the ability range (θ)

Based on the results of the description above, the quality of the statement items in the Self Confidence questionnaire is appropriate to be used to measure students' Self-Confidence. However, there are some statement items that are not appropriate to measure one of the indicators of Self Confidence. In a study conducted by Putri et al. (2021), the results were obtained that confirmed the validity and reliability of the self-confidence instrument in mathematics learning using the Rasch Model, which is different from the classical approach used in previous studies. In comparison, the study by (Putri et al., 2021) showed lower reliability ($\alpha = 0.51$) and 44 invalid subjects in their analysis. In addition, this study makes a unique contribution by showing that the instrument can meet the assumption of unidimensionality empirically, as supported by exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). Meanwhile, based on the findings of Khoiriyah et al., (2022), this study developed an instrument for assessing students' self-confidence attitudes in elementary schools, which utilizes a different approach from traditional instruments such as the Likert scale. In comparison, previous studies showed that the instruments used had lower reliability ($\alpha = 0.65$) and lower validity (Aiken's V < 0.85). This study makes an important contribution by showing that the developed instrument has higher validity and stronger reliability, and meets the criteria of unidimensionality based on factor analysis.

CONCLUSION

Based on the results of the research that has been conducted, this study produces an instrument to measure high school students' self-confidence in learning mathematics. The instrument that has been prepared is divided into four main aspects, namely having confidence in one's abilities, making decisions independently, maintaining a positive self-image, and daring to express opinions. However, after going through the analysis process, the aspects that can measure self-confidence are the first, second, and third aspects, while the third aspect does not meet the requirements for measuring self-confidence. Several analyzes used to obtain aspects that can measure self-confidence are: (a) the first validity and reliability test; (b) the second EFA test using SPSS 25; (c) the third analysis, namely the CFA test with the LISREL application and the results obtained meet the unidimensional requirements, meaning that the questions measure the same construct; (d) analysis of item response theory using the R program and producing varied questions, namely questions that have two thresholds and three thresholds. The instrument that has been analyzed can be applied in the classroom to assess students' self-confidence, helping teachers in providing appropriate interventions to increase their self-confidence. In developing educational policies, this instrument can be used to design programs that support students' psychological development. However, this instrument may be limited to certain contexts and may need adjustments to be applied in other areas or age groups. Further research can test this instrument in different areas or age groups to measure its reliability and validity more broadly.

SUGGESTIONS AND RECOMENDATION

This instrument can be improved by refining the statement items that overlap in response categories to ensure clearer and more accurate measurement of self-confidence. Training for teachers is also important to ensure they can effectively use and interpret the results of the instrument, enabling appropriate interventions to be implemented. Furthermore, further testing across various age groups and geographical regions is necessary to ensure the validity and reliability of the instrument in different contexts. Additional research is also recommended to explore the relationship between self-confidence and other variables, such as academic achievement and problem-solving abilities, in order to gain deeper insights. This instrument could also be utilized in educational policies to design programs that support the psychological and academic development of students in a personalized and inclusive manner.

REFERENCES

Adhimah, F. N., & Simatupang, N. D. (2014). Meningkatkan rasa percaya diri anak melalui cerita bergambar pada kelompok a di tk muslimat sidoarjo. *Jurnal PAUD Teratai*, *3*(3), 1–6. https://ejournal.unesa.ac.id/index.php/paud-teratai/article/view/8389

Amri, S. (2018). Pengaruh kepercayaan diri (self confidence) berbasis ekstrakurikuler pramuka terhadap prestasi belajar matematika siswa sma negeri 6 kota bengkulu. *Jurnal Pendidikan Matematika Raflesia*, *3*(2), 156–168. https://doi.org/10.31186/JPMR.V3I2.6309

Andayani, M., & Amir, Z. (2019). Membangun self-confidence siswa melalui pembelajaran matematika. *Desimal: Jurnal Matematika*, *2*(2), 147–153. https://doi.org/10.24042/djm.v2i2.4279

Anisa, R. N., Ambarwati, L., & Deasyanti. (2020). Pembelajaran matematika siswa sekolah dasar melalui kegiatan bermain. *JPD: Jurnal Pendidikan Dasar*, *1*(2), 1–11. https://doi.org/10.21009/10.21009/JPD.081

Asmara, A. P. (2016). Kajian integrasi nilai-nilai karakter islami dengan kimia dalam materi kimia karbon. *Jurnal Pendidikan Sains (Jps)*, *4*(2), 1–11. https://doi.org/10.26714/jps.4.2.2016.1-11

Dewanti, S. S., Ayriza, Y., & Setiawati, F. A. (2020). The application of item response theory for development of a students' attitude scale toward mathematics. *New Educational Review*, *60*, 108–123. https://doi.org/10.15804/tner.2020.60.2.09

Diputera, A. M. (2018). *Analisis irt menggunakan wingen 3 : teori respon & aplikasi*. Uwais Inspirasi Indonesia.

Fardani, Z., Surya, E., & Mulyono. (2021). Analisis kepercayaan diri (self confidence) siswa dalam pembelajaran matematika melalui model problem based learning. *Jurnal Pendidikan Matematika*, *14*(1), 39–51. https://doi.org/10.24114/paradikma.v14i1.24809

Friyatmi. (2018). Estimasi parameter tes dengan penskoran politomus menggunakan graded response model pada sampel kecil. *JIPE : Jurnal Inovasi Pendidikan Ekonomi*, *8*(1), 22–31. https://doi.org/doi.org/10.24036/01104490

Gusteti, M. U., & Neviyarni. (2022). Pembelajaran berdeferensiasi pada pembelajaran matematika di kurikulum merdeka. *Lebesgue: Jurnal Ilmiah Pendidikan Matematika, Matematika Dan Statistika*, *3*(3), 636–646. https://doi.org/10.46306/lb.v3i3.180

Herdiana, Y., Marwan, M., & Morina Zubainur, C. (2019). Kemampuan representasi matematis dan self confidence siswa smp melalui penerapan model problem based learning (pbl). *Al-Qalasadi : Jurnal Ilmiah Pendidikan Matematika*, *3*(2), 23–35. https://doi.org/10.32505/v3i2.1368

Khoiriyah, I., Haryanto, S., & Retnaningsih, R. (2022). Pengembangan instrumen penilaian sikap percaya diri siswa di sekolah dasar. *Tuladha : Jurnal Pendidikan Dasar*, *1*(2), 71–94. https://doi.org/10.30738/tuladha.v1i2.13498

Kholil, M., & Zulfiani, S. (2020). Faktor-faktor kesulitan belajar matematika siswa madrasah ibtidaiyah da'watul falah kecamatan tegaldlimo kabupaten banyuwangi. *EDUCARE: Journal of Primary Education*, *1*(2), 151–168. https://doi.org/10.35719/educare.v1i2.14

Lukman, H. S., & Mulyanti, Y. (2017). Analisis validitas bahan ajar struktur aljabar berbasis strategi abduktif-deduktif. *M A T H L I N E : Jurnal Matematika Dan Pendidikan Matematika*, *2*(2), 159–166. https://doi.org/10.31943/mathline.v2i2.43

Nafisah, N., & Respatiwulan, R. (2019). Analisis Faktor Indeks Harga Konsumen Kota Semarang. *Indonesian Journal of Applied Statistics*, *2*(2), 113. https://doi.org/10.13057/ijas.v2i2.34903

Nasution, H. F. (2016). Instrumen penelitian dan urgensinya dalam penelitian kuantitatif. *Analytical Biochemistry*, *11*(1), 59–75. https://doi.org/doi.org/10.24952/masharif.v4i1.721

Nursanti, G., Prihaswati, M., & Suprayitno, I. J. (2020). Pengembangan aplikasi self confidence test berbasis web sebagai alat ukur kepercayaan diri pada pembelajaran matematika. *Prosiding Seminar Nasional Unimus*, *3*, 133–139. https://prosiding.unimus.ac.id/index.php/semnas/article/view/655

Pujiadi. (2016). *Guru pembelajar modul matematika sma*. Direktorat Jenderal Guru dan Tenaga Kependidikan.

Putri, F. N., Haniyah, D., Rahmawati, W., & Faradillah, A. (2021). Validitas angket self-confidence peserta didik menengah pada pembelajaran matematika. *Prosiding Seminar Nasional Pendidikan Matematika Universitas Patimura*, *2*(3), 170–183. https://doi.org/doi.org/10.30598/snpmunpatti.2021.pp170-183

Qurni, K. Al. (2021). Pengembangan instrumen kepercayaan diri matematika pada materi program linear. *Prosiding Seminar Nasional Matematika Dan Pendidikan Matematika*, *6*, 175–186. https://conference.upgris.ac.id/index.php/senatik/article/view/1879

Rachmantika, A. R., & Wardono. (2019). Peran kemampuan berpikir kritis siswa pada pembelajaran matematika dengan pemecahan masalah. *Prosiding Seminar Nasional Matematika*, *2*(1), 439–443. https://journal.unnes.ac.id/sju/prisma/article/view/29029

Rahayu, A., & Haryanto, S. (2017). Pengembangan instrumen pengukuran nilai kedisiplinan siswa di sma negeri 2 bantul. *Jurnal Penelitian Dan Evaluasi Pendidikan*, *5*(2), 12–18. https://doi.org/doi.org/10.30738/wd.v5i2.3373

Rahmah, N. (2018). Hakikat pendidikan matematika. *Al-Khwarizmi: Jurnal Pendidikan Matematika Dan Ilmu Pengetahuan Alam*, *1*(2), 1–10. https://doi.org/10.24256/jpmipa.v1i2.88

Retnawati, H. (2014). *Teori respons butir dan penerapannya : untuk peneliti, praktisi pengukuran dan pengujian mahasiswa dan pascasarjana*. Nuha Medika.

Rohana, R., Wikanengsih, W., & Irmayanti, R. (2020). Profil kepercayaan diri peserta didik (studi deskriptif terhadap peserta didik kelas ix smp negeri 1 selawi tahun pelajaran 2019/2020). *FOKUS (Kajian Bimbingan & Konseling Dalam Pendidikan)*, *3*(6), 208. https://doi.org/10.22460/fokus.v3i6.6067

Verdian, E. (2019). Analisis faktor yang merupakan intensi. *Agora*, *7*(1).

Waluyo, M. (2016). Mudah cepat tepat penggunaan tools amos dalam aplikasi (SEM). *UPN Jatim Repository*.

Yusniawan, R., Isnaini, M., & Jayanti, E. (2019). Pengembangan lembar kerja peserta didik (lkpd) kimia terintegrasi nilai-nilai islam pada materi hidrokarbon. *Jurnal Intelektualita: Keislaman, Sosial Dan Sains*, *8*(2), 181–190. https://doi.org/10.19109/intelektualita.v8i2.4667