

Research Article

Cite this article: Muhtarom, Nizaruddin, Nuvitalia, D., & Estrada, T. J. B. (2025). Rasch Calibration on the Belief about Mathematics Teaching and Learning Scale Development. *Educational Process: International Journal*, 16, e2025273.
<https://doi.org/10.22521/edupij.2025.16.273>

Received April 24, 2025

Accepted June 2, 2025

Published Online June 22, 2025

Keywords: Learning beliefs, mathematical beliefs, mathematics teaching, measurement tool development, Rasch model

Author for correspondence:

Muhtarom

✉ muhtarom@upgris.ac.id

✉ Universitas PGRI Semarang, Indonesia



OPEN ACCESS

© The Author(s), 2025. This is an Open Access article, distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction, provided the original article is properly cited.

Rasch Calibration on the Belief about Mathematics Teaching and Learning Scale Development

Muhtarom, Nizaruddin, Duwi Nuvitalia, Tjaart Jan B. Estrada

Abstract

Background. Evaluating beliefs is essential, as the concept has a significant influence on mathematics learning in the classroom. Therefore, this research aimed to provide an instrument for assessing beliefs regarding mathematics teaching and learning.

Methods. The survey consisted of 311 females and 96 males, totaling 407 respondents, and a validated questionnaire was used to collect the data. The reliability and validity of the instrument were evaluated using the Rasch Model to analyze Cronbach's Alpha, item reliability, and person reliability. Additionally, validity was assessed through item fit order, item unidimensionality, item measurement, partial-credit scale, and bias testing. This statistical process provided strong support for the reliability and validity of the instrument.

Results. The results showed the alignment between the research instrument items and responses. This is evident from the Cronbach's Alpha value, which has an excellent interpretation. The person's reliability value, with fairly good interpretation, indicates the consistency of their answers. Meanwhile, the item reliability value indicates the quality of the instrument, providing an excellent understanding.

Conclusion. A total of thirty statements were effective out of sixty-four developed to measure beliefs about mathematics teaching and learning. This research makes an original contribution to the field of mathematics education by developing a valid and reliable measurement instrument for mathematical beliefs. The implementation of the instrument evaluated students' and teachers' beliefs, enabling targeted professional development interventions.

1. Introduction

Teachers are essential in teaching, training, directing, evaluating, and assessing students in the capacity of professional educators and subject matter experts. In this context, an understanding of mathematical knowledge directly influences classroom instruction (Matitaputty et al., 2024). Although knowledge influences learning, teachers with similar expertise may approach instruction differently. According to Ernest, beliefs are important factors used to shape teaching practices (Hibi, 2017; Siswono, Kohar, Hartono et al., 2019; Yang et al., 2020). Varying beliefs lead to diverse methods of mathematics teaching practices in the classroom (Ferguson & Braten, 2018; Lui & Bonner, 2016; Muhtarom et al., 2019; Purnomo et al., 2020). These attitudes are composed of various cognitive structures and serve as the foundation for guiding behavior and shaping the understanding of individuals (Belbase, 2024; Buehl & Fives, 2009; Wang et al., 2022). Beliefs are located between the cognitive and affective domains, called the “twilight zone.” Beliefs have components in both domains. Every individual has perceptions, attitudes, beliefs, and values/ethics/morals. The distinction between these concepts can be characterized symbolically as “hot” perceptions, “cool” attitudes, and “cold” beliefs (Pehkonen & Pietilä, 2003).

Differences in teacher beliefs also impact students' beliefs in learning (Ariza Muñoz et al., 2023; Muhtarom et al., 2017). What students believe is largely based on their experiences while learning mathematics in school. Thus, teachers play a crucial role in shaping students' beliefs about mathematics. Furthermore, mathematics education students who will become successors to professional teachers must also have beliefs about mathematics teaching and learning. These beliefs must be instilled in children early on, as they can serve as the foundation for disposition, action, change, and learning (Muhtarom et al., 2018; Perera & John, 2020). This is reinforced by research results, which suggest that differences in teaching experience influence the differences in beliefs they hold (Arikan, 2016; Buehl & Fives, 2009; Muhtarom et al., 2020; Siswono, Kohar, & Hartono, 2019).

Measuring beliefs in learning is essential. Unfortunately, there is very little research related to mathematical beliefs. Several studies are primarily concentrated on the qualitative analysis of beliefs (Bal, 2015; Buehl & Fives, 2009; Muhtarom et al., 2019; Siswono, Kohar, & Hartono, 2019; Siswono, Kohar, Hartono, et al., 2019). Limited analyses focus on developing an instrument to measure teachers' beliefs about mathematics teaching and learning. However, several efforts have described the creation of measurement scales (Masal & Takunyaci, 2012) and a survey to measure the early childhood beliefs of teachers in mathematics learning (Platas, 2015). The data analysis used a classical test theory, where the validity and reliability of the instrument were examined utilizing Split-Half, item-total correlation, and Cronbach's Alpha. Another research was conducted by Purnomo discussing the scale of beliefs in learning (Purnomo, 2017). Data analysis using Exploratory Factor Analysis reported that the instrument had a sufficient fit index. The results were only conducted on teachers and required an improvement with more significant respondents from different locations (Muhtarom, 2024). This research lacks specific guidance on measuring teachers' prospective beliefs about mathematics learning. A broader range of respondents should be included, beyond prospective teachers, to enhance the person-separation value and more accurately capture diversity levels.

Considering previous results, further investigation on the development of measuring instruments in mathematics teaching and learning should be conducted with a larger sample size and demographics of respondents. Furthermore, item response theory, as the core component of the Rasch Model, is used to examine data to gain a better understanding of assessing views about mathematics instruction. Item response theory provides benefits over conventional test theory, where the chance of the subject responding to the item correctly relies on the skills of the subject and the qualities of the question (Chan et al., 2021; Rahim & Haryanto, 2021; Sumintono, 2018). Instrument validity can be assessed using a variety of criteria to improve the reliability of the final instrument (Andrich & Marais, 2019; Julianto et al., 2025; Medvedev & Krägeloh, 2022; Sumintono,

2018). Therefore, this research uses the Rasch model to investigate the reliability and validity of the belief measuring instrument in mathematics teaching and learning.

2. Literature Review

2.1. Beliefs

Beliefs are a person's natural trait (Yang et al., 2020). Mathematical beliefs relate to how a person behaves towards mathematical objects and believes in the existence of those objects (Belbase, 2019; Fives & Buehl, 2016; Muhtarom & Nizaruddin, 2024). Beliefs and perceptions are interchangeable in the context of mathematics. When asked what mathematics is, people answer by stating their views, which can also be called beliefs about mathematics. Belief is the basis for driving a person to behave and the understanding that an individual has of an event. A person uses beliefs to predict what will happen (Bal, 2015; Fives & Buehl, 2016).

2.2. Beliefs about Nature Mathematics

Mathematics can be viewed as a dynamic field of human creation, continuously evolving according to patterns of discovery and the results remain open to revision; or mathematics is static but integrated, consistent, connected, and objective; or mathematics consists of some tools made of a set of facts, rules that are not interrelated but useful (Ariza Muñoz et al., 2023; Belbase, 2019; Siswono, Kohar, Hartono, et al., 2019). Beliefs about mathematics are divided into three orientations, namely: 1) mathematics can be viewed as static knowledge involving a series of rules and procedures, 2) mathematics is a set of rules and does not change, is proven to be true and stands on proven theories, and 3) continuous research is carried out on problems in life based on experience, and observations of the regularity of social context phenomena so that mathematics is constantly changing, never static (Belbase, 2024).

2.3. Beliefs about Mathematics Teaching

Beliefs about mathematics impact how they play a role in mathematics teaching (Ariza Muñoz et al., 2023; Li et al., 2024; Lui & Bonner, 2016; Yang et al., 2020). Mathematics teaching beliefs are individual perspectives on engaging in mathematical tasks and pedagogical practices. Teachers' beliefs about how knowledge is developed relate to their beliefs about how to teach science and how students learn (Ferguson & Braten, 2018; Li et al., 2024). Teachers' beliefs reflect their philosophies about the nature of knowledge and how to acquire it, influencing decision-making in learning. There is a degree of consistency between beliefs and teaching practices implemented by teachers in the classroom (Belbase, 2024; Purnomo et al., 2020; Siswono, Kohar, & Hartono, 2019). Teachers with instrumentalist beliefs emphasize that mathematics teaching focuses on correct answers and views learning as passive knowledge acquisition. Teachers with Platonic beliefs emphasize students' understanding of mathematical material and the relationships between concepts. Constructivist teachers emphasize mathematics learning as an autonomous exploration, with the learner being the focus rather than the material's content (Siswono et al., 2016).

2.4. Beliefs about Learning Mathematics

Students' beliefs about mathematics help increase motivation and learning achievement. Beliefs motivate students to learn problem solving, giving them more effort in solving (Ozturk & Guven, 2016). Students' beliefs about mathematics significantly influence their learning achievement and increase motivation. As an illustration, students can learn to understand mathematical problems by solving several problems. Previous experiences determine what information students think is relevant and what concepts are appropriate to solve problems (Arikan, 2016; Belbase, 2024). This experience plays a role in shaping students' behavior in learning mathematics. What is believed is primarily based on the experiences gained during learning mathematics. This shows a relationship

between beliefs and learning, where learning experiences contribute to a person's beliefs about the meaning of learning mathematics. In turn, beliefs about mathematics tend to influence how they approach new knowledge about mathematics (Buehl & Fives, 2009; Muhtarom et al., 2017; Spruce & Bol, 2015).

3. Methodology

3.1. Type of Research

This development research aims to create a measurement instrument for assessing beliefs about mathematics teaching and learning. The analysis was conducted in three stages, namely introduction, prototype, as well as evaluation (Akker et al., 2006; Nieveen & Folmer, 2013). In the introduction stage, a review of previous research was conducted. The questionnaire framework and belief statements were designed for the mathematics learning instrument during the prototype stage. In the assessment stage, data were collected using a questionnaire distributed through Google Forms.

3.2. Research Respondents

The research respondents were teachers and students. The research was conducted using a purposive sampling technique, which involves selecting respondents who are deliberately chosen to provide research data. A total of 407 individuals were included, comprising 96 males and 311 females. Approximately 158 respondents were teachers, and 249 were students completing professional education. Table 1 presents the characteristics of the respondents in this research.

Table 1. Research Respondents

Attribute	Category	Code	Person (n)	Proportion
Gender	Male	L	96	23.59%
	Female	P	311	76.41%
Respondent Age	21 – 30 years	A	253	62.16%
	31 – 40 years	B	97	23.83%
	41 – 50 years	C	30	7.37%
	51 – 60 years	D	27	6.63%
Respondents' Occupation	Bachelor of Education Students	M	141	34.64%
	Pre-Service Teacher	P	108	26.54%
	Teacher	G	158	38.82%
Number of Students	Small schools (up to 300 students)	K	67	16.46%
	Medium schools (between 300 and 1000 students)	S	94	23.10%
	Large schools (more than 1000 students)	B	49	12.04%
	Irrelevant question (not taught in school)	T	197	48.40%
Respondents' Demographics	Rural	D	61	14.99%
	Small towns and semi-densely populated areas	K	94	23.10%
	Large cities	B	55	13.51%
	Irrelevant question (not taught in school)	T	197	48.40%

3.3. Research Data Collection

Data for the research were gathered through a questionnaire on beliefs regarding mathematics teaching and learning, which utilized a four-point Likert scale. The questionnaire consists of both positive and negative statements. Each item uses a 4-point Likert scale to assess beliefs. The assessment for positive statement items, the numerical scale, is presented as follows: 1 = strongly disagree, 2 = disagree, 3 = agree, and 4 = strongly agree. While for negative statement items, the numerical scale is as follows: 1 = strongly agree, 2 = agree, 3 = disagree, and 4 = strongly disagree. Positive statements are statement items that support the scale measurement attribute, and negative statements are statements with characteristics that do not support the scale measurement attribute. With the presence of negative statements, respondents are expected to read each statement item more carefully and not provide responses carelessly.

Table 2. Mathematics Beliefs Questionnaire

Research focus	Research sub-focus	Description	Positive items	Negative items
Beliefs on the nature of mathematics	Definition of mathematics	Subject views on mathematics	1, 8	5, 33
	Relationship between mathematics and life	Subject views concerning the advantages of mathematics in daily life	13, 20	55, 61
	Development of mathematical knowledge	Subject beliefs on mathematical knowledge development	9, 47	15, 54
Beliefs about mathematics learning	Designing learning activities	Respondents' beliefs regarding designing learning activities that are appropriate to students' needs	2, 49	14, 34
	Teaching approaches	Respondents' beliefs regarding appropriate approaches to mathematics teaching	26, 36, 52	27, 48, 56
	The role of teachers in teaching	Respondents' beliefs regarding the role of teachers in mathematics teaching	3, 21, 31, 39, 45, 53	6, 11, 19, 37, 51, 63
	Problem-solving	Respondents' beliefs regarding problem-solving activities in mathematics learning	7, 10, 50, 59	4, 23, 30, 58
	Assessing learning activities	Respondents' beliefs about assessments used in learning	22, 42, 62, 64	16, 25, 28, 41
Beliefs about mathematics learning	Student activities in mathematics learning	Respondents' beliefs about how students learn mathematics	18, 24, 32, 38	12, 46, 57, 60
		Respondents' beliefs about the basic things that need to be mastered when learning mathematics	29, 43, 44	17, 35, 40

The questionnaire instrument was developed based on previous preliminary research (Muhtarom et al., 2019). Before obtaining research data, the instrument was validated to ensure content validity. The validation results indicated that the questionnaire was suitable for measuring beliefs related to mathematics teaching and learning. Table 2 presents the instrument grid, which provides readers with a comprehensive overview of the problems addressed in this research.

3.4. Research Data Analysis

Testing was carried out using the Rasch Model to obtain a valid and reliable instrument (Adi et al., 2022; Andrich & Marais, 2019; Sumintono, 2018). The software used is Winsteps version 3.73. Instrument reliability testing uses output from summary statistics. Summary statistics provide overall information about the quality of respondents, the quality of the instruments used, and the interaction between persons and items (Muhtarom, 2024; Oktaviyanti et al., 2024). The separation value showed the grouping of items and persons. The greater the separation value, the better the quality in terms of total items and respondents. Meanwhile, unidimensionality, fit order, rating (partial-credit) scale, and item measures were employed to evaluate the instrument's validity (Andrich & Marais, 2019; Chan et al., 2021; Sumintono, 2018). Using the Rasch model for instrument validation yields more comprehensive information and better aligns with the definition of measurement: 1) Unidimensionality assesses the instrument's ability to accurately measure the intended variables (Andrich & Marais, 2019; Muhtarom, 2024). 2) The item fit order uses INFIT MNSQ to identify outliers or misfits by adding the mean and standard deviation. A misfit item is a logit value more significant than the sum of the mean and standard deviation. Another method is to apply the following criteria: 1) Accepted Outfit Mean Square (MNSQ) value: $0.5 < \text{MNSQ} < 1.5$; 2) Accepted Outfit Z-Standard (ZSTD) value: $-2.0 < \text{ZSTD} < 2.0$; and 3) Received Point Measure Correlation (Pt Mean Corr) value: $0.4 < \text{Pt Measure Corr} < 0.85$. Valid items satisfy at least one of the three requirements (Medvedev & Krägeloh, 2022; Sumintono, 2018). 3) Validity of the partial-credit scale is determined by all ratings reaching a peak point (Julianto et al., 2025; Muhtarom, 2024). 4) The use of item measures and maps identifies the most difficult and easiest variables in agreement with respondents (Chan et al., 2021; Sumintono, 2018), and 5) Differential Item Functioning (DIF) testing shows that the mathematics learning belief instrument is not biased (Adi et al., 2022; Rahim & Haryanto, 2021). This testing was conducted considering the respondents' gender, age, occupation, number of students, and demographics.

4. Results

4.1. Instrument Reliability

The overall reliability of the interaction between the person and the item is 0.81, which falls in the very good category. Furthermore, the person and item reliability values are 0.77 and 1.00, respectively. The quality of the belief measurement instrument is in a particular category. Meanwhile, the consistency of the respondents' answers is a fairly good category. Figure 1 shows that OUTFIT MNSQ and INFIT MNSQ values are obtained in the item and person measurements, providing good measurement results. OUTFIT MNSQ and INFIT MNSQ in the person measurement showed a mean value of 0.98 and 1.01, respectively. Similarly, INFIT MNSQ and OUTFIT MNSQ had a mean value of 0.98 for the item measurement.

Considering the results, the separation value shows the grouping of items and people. The greater the separation value, the higher the quality of the instrument in terms of all items and respondents. The person separation value is 1.83, but $H = ((4 \times 1.83) + 1)/3 = 2.77$, approximated to 3 since there are three groups of respondents. The item separation value is 19.02, but $H = ((4 \times 19.02) + 1)/3 = 25.69$, approximated to 26 since the instrument is made up of 26 groups of items.

SUMMARY OF 407 MEASURED Person								
	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	174.7	64.0	.54	.20	1.01	-.7	.98	-.7
S.D.	11.4	.0	.48	.01	.65	3.6	.60	3.4
MAX.	215.0	64.0	2.37	.24	3.63	9.1	3.37	9.2
MIN.	153.0	64.0	-.34	.20	.27	-6.3	.29	-6.1
REAL RMSE	.23	TRUE SD	.42	SEPARATION	1.83	Person	RELIABILITY	.77
MODEL RMSE	.20	TRUE SD	.43	SEPARATION	2.13	Person	RELIABILITY	.82
S.E. OF Person MEAN = .02								
Person RAW SCORE-TO-MEASURE CORRELATION = 1.00								
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .81								
SUMMARY OF 64 MEASURED Item								
	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	1110.9	407.0	.00	.08	.98	-.6	.98	-.6
S.D.	268.2	.0	1.61	.01	.22	3.4	.23	3.4
MAX.	1466.0	407.0	3.13	.10	1.56	6.7	1.58	7.0
MIN.	592.0	407.0	-2.37	.07	.61	-6.9	.61	-6.9
REAL RMSE	.08	TRUE SD	1.60	SEPARATION	19.02	Item	RELIABILITY	1.00
MODEL RMSE	.08	TRUE SD	1.60	SEPARATION	19.74	Item	RELIABILITY	1.00
S.E. OF Item MEAN = .20								
UMEAN=.0000 USCALE=1.0000								
Item RAW SCORE-TO-MEASURE CORRELATION = -1.00								
26048 DATA POINTS. LOG-LIKELIHOOD CHI-SQUARE: 47223.98 with 25576 d.f. p=.0000								
Global Root-Mean-Square Residual (excluding extreme scores): .6139								

Figure 1. Output Reliability of Respondents and Instruments

4.2. Instrument Validity

The item unidimensionality and fit order show that the belief measurement instrument items are valid. The total raw variance value in the observations is 55.3%, which is a good category. The unexplained variance for the mathematics learning belief instrument is 12.7%, with an eigenvalue of 18.2 (see Figure 2). Therefore, there is a problematic statement item since the eigenvalue exceeds 3.

Table of STANDARDIZED RESIDUAL variance (in Eigenvalue units)				
		-- Empirical --	Modeled	
Total raw variance in observations	=	143.0	100.0%	100.0%
Raw variance explained by measures	=	79.0	55.3%	55.1%
Raw variance explained by persons	=	5.6	3.9%	3.9%
Raw Variance explained by items	=	73.4	51.3%	51.2%
Raw unexplained variance (total)	=	64.0	44.7%	100.0%
Unexplned variance in 1st contrast	=	18.2	12.7%	28.4%
Unexplned variance in 2nd contrast	=	5.8	4.1%	9.1%
Unexplned variance in 3rd contrast	=	1.7	1.2%	2.6%
Unexplned variance in 4th contrast	=	1.6	1.1%	2.5%
Unexplned variance in 5th contrast	=	1.4	1.0%	2.2%

Figure 2. Unidimensionality Items

The subsequent step views the Misfit and Fit items using INFIT MNSQ value. The procedure includes the addition of standard deviation and mean, as well as comparing the result to the logit value. Figure 3 shows that MEAN + SD = 1.22, and this criterion obtains ten items with INFIT MNSQ values greater than 1.22.

Item STATISTICS: MISFIT ORDER													
ENTRY	TOTAL	TOTAL		MODEL	INFIT	OUTFIT	PT-MEASURE	EXACT	MATCH				
NUMBER	SCORE	COUNT	MEASURE	S.E.	MNSQ	ZSTD	MNSQ	ZSTD	CORR.	EXP.	OBS%	EXP%	Item
7	1161	407	-1.18	.08	1.56	6.7	1.58	7.0	A .03	.28	46.2	60.0	item7
5	953	407	1.00	.07	1.44	6.1	1.47	6.4	B .47	.30	43.5	51.9	item5
41	1160	407	-1.17	.08	1.46	5.6	1.45	5.5	C .61	.28	49.4	60.0	item41
45	1210	407	-.49	.08	1.35	4.4	1.37	4.7	D .00	.27	60.7	62.0	item45
14	1164	407	-.20	.08	1.35	4.4	1.34	4.3	E .62	.28	55.8	60.2	item14
24	1274	407	-.91	.08	1.33	4.3	1.35	4.6	F .12	.27	57.5	61.6	item24
34	1156	407	-.15	.08	1.34	4.3	1.33	4.3	G .61	.28	52.1	59.7	item34
6	1193	407	-.38	.08	1.29	3.7	1.27	3.5	H .54	.28	57.7	61.5	item6
55	1012	407	-.68	.07	1.25	3.6	1.27	3.8	I .51	.29	52.6	51.0	item55
30	1261	407	-.82	.08	1.24	3.1	1.21	2.9	J .55	.27	58.0	61.9	item30
25	1157	407	-.16	.08	1.22	2.9	1.22	2.9	K .57	.28	57.0	59.8	item25
3	1466	407	-2.37	.10	1.15	2.3	1.22	3.2	L .17	.21	62.7	61.4	item3
12	980	407	.86	.07	1.20	2.9	1.22	3.2	M .43	.30	44.2	51.3	item12
54	841	407	1.60	.07	1.21	3.3	1.22	3.3	N .41	.31	52.6	53.8	item54
8	1271	407	-.89	.08	1.15	2.0	1.20	2.7	O .04	.27	65.6	61.7	item8
60	934	407	1.11	.07	1.15	2.3	1.17	2.6	P .42	.30	50.9	52.4	item60
16	1137	407	-.03	.08	1.12	1.6	1.11	1.5	Q .60	.28	58.5	58.5	item16
18	1277	407	-.93	.08	1.05	.7	1.08	1.1	R .00	.27	67.8	61.5	item18
17	693	407	2.44	.08	1.07	1.1	1.04	.7	S .00	.31	53.1	49.8	item17
48	655	407	2.68	.08	1.06	1.0	1.03	.5	T .22	.30	42.3	51.4	item48
4	819	407	1.72	.07	1.06	.9	1.06	.9	U .25	.31	55.0	53.7	item4
28	961	407	.96	.07	1.04	.7	1.06	.9	V .54	.30	52.3	51.7	item28
58	844	407	1.59	.07	1.04	.7	1.05	.9	W .17	.31	58.0	53.8	item58
11	592	407	3.13	.09	1.03	.5	1.04	.6	X .25	.29	41.5	60.4	item11
46	873	407	1.43	.07	1.03	.5	1.04	.6	Y .54	.31	51.6	53.6	item46
57	916	407	1.20	.07	1.01	.1	1.02	.3	Z .54	.30	52.1	52.8	item57
BETTER FITTING OMITTED													
32	1331	407	-1.30	.08	.91	-1.4	.90	-1.5	z .29	.26	64.1	58.9	item32
37	741	407	2.15	.08	.90	-1.7	.88	-2.0	y .00	.31	65.4	51.3	item37
59	1304	407	-1.11	.08	.89	-1.7	.89	-1.6	x .18	.26	67.8	60.4	item59
27	1346	407	-1.41	.09	.84	-2.5	.88	-2.0	w .18	.25	64.9	58.1	item27
15	708	407	2.35	.08	.86	-2.4	.84	-2.7	v .04	.31	64.1	49.3	item15
35	931	407	1.12	.07	.84	-2.6	.86	-2.3	u .54	.30	58.2	52.5	item35
26	1438	407	-2.13	.09	.84	-2.7	.81	-3.2	t .43	.22	67.8	57.0	item26
36	1335	407	-1.33	.08	.82	-2.8	.84	-2.7	s .25	.26	70.5	58.7	item36
10	1464	407	-2.36	.10	.83	-2.9	.80	-3.3	r .36	.21	62.2	61.1	item10
9	1375	407	-1.63	.09	.82	-3.1	.83	-3.0	q .27	.25	62.7	56.6	item9
2	1391	407	-1.75	.09	.81	-3.3	.81	-3.4	p .33	.24	66.3	55.8	item2
40	704	407	2.37	.08	.81	-3.3	.79	-3.7	o .07	.31	67.8	49.7	item40
38	1309	407	-1.15	.08	.80	-3.1	.81	-3.1	n .24	.26	70.3	60.1	item38
42	1348	407	-1.43	.09	.77	-3.9	.79	-3.6	m .32	.25	69.0	58.0	item42
62	1321	407	-1.23	.08	.75	-4.0	.77	-3.8	l .20	.26	71.3	59.5	item62
31	1410	407	-1.90	.09	.75	-4.6	.74	-4.7	k .37	.24	68.6	55.5	item31
20	1350	407	-1.44	.09	.74	-4.4	.74	-4.5	j .31	.25	69.0	57.8	item20
52	1360	407	-1.51	.09	.71	-5.0	.71	-5.1	i .33	.25	68.3	57.3	item52
39	1349	407	-1.43	.09	.70	-5.1	.71	-5.0	h .25	.25	70.8	57.9	item39
13	1359	407	-1.51	.09	.69	-5.4	.70	-5.3	g .30	.25	69.8	57.4	item13
64	1362	407	-1.53	.09	.67	-5.8	.67	-5.9	f .36	.25	71.5	57.3	item64
22	1312	407	-1.17	.08	.67	-5.5	.67	-5.6	e .22	.26	74.7	60.0	item22
44	1364	407	-1.54	.09	.67	-6.0	.67	-6.0	d .39	.25	70.5	57.2	item44
53	1350	407	-1.44	.09	.67	-5.9	.66	-6.0	c .36	.25	71.5	57.8	item53
43	1366	407	-1.56	.09	.66	-6.1	.66	-6.2	b .36	.25	70.0	57.1	item43
50	1331	407	-1.30	.08	.61	-6.9	.61	-6.9	a .34	.26	73.5	58.9	item50
MEAN	1110.9	407.0	.00	.08	.98	-.6	.98	-.6			60.8	56.3	
S.D.	268.2	.0	1.61	.01	.22	3.4	.23	3.4			8.5	3.9	

Figure 3. Output item fit order

The following criteria to indicate Outliers or Misfits items are 1) The Outfit MNSQ value of $0.5 < \text{MNSQ} < 1.5$, 2) the Outfit ZSTD value of $-2.0 < \text{ZSTD} < +2.0$, and 3) the PT Measure Corr value of $0.4 < \text{PT Measure Corr} < 0.85$. The second row shows the output results for item number 5 on the belief instrument in mathematics teaching and learning, with scores of 1.47, 6.4, and 0.47, respectively. However, number 7 did not meet the three validity criteria, making the item invalid and unsuitable for use. The other items were Fit and could still be maintained for use since one of the three criteria was met. The complete results of the item fit order output are shown in Figure 3.

The value data, arranged from highest to lowest, shows that the 11th item has a 3.13 logit. This is the most challenging item to agree upon among respondents. Meanwhile, the 3rd item has a -2.37 logit, and the range of difficulty levels is fairly broad without correction due to a standard deviation value of 1.61. INFIT MNSQ value of each person is employed to determine the level of fit or misfit. The mean value and standard deviation are combined ($1.01 + 0.65 = 1.66$). In this context, a logit number greater than the value suggests a misfit person. The result is 65 persons with a more excellent

INFIT MNSQ value. The person logit value data, sorted from highest to lowest, shows that 084LAPTT is a respondent with the highest tendency to believe in mathematics learning compared to others. Furthermore, person 284LBPKD disagrees with the belief instrument item in mathematics learning and teaching.

Validity testing, which involves analyzing the rating results, reveals that each rating (1, 2, 3, 4) has a distinct peak (Figure 4). This report shows that respondents can rate the instrument of mathematics teaching and learning beliefs differently.

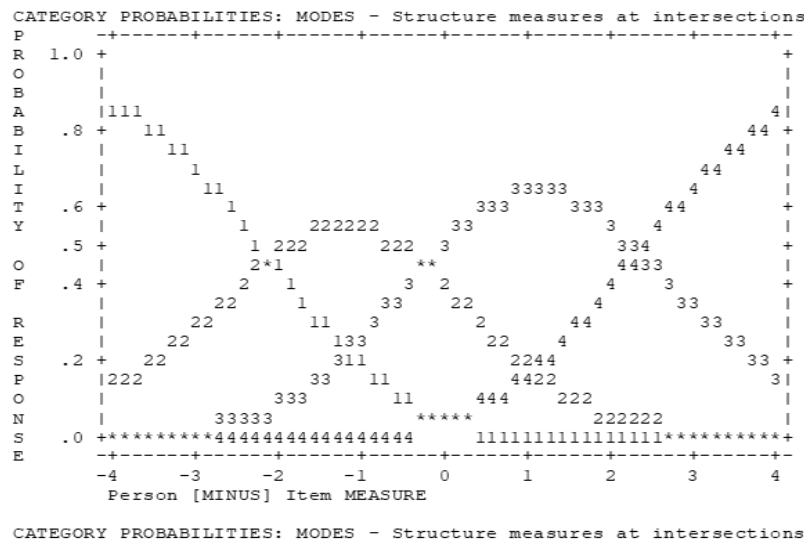


Figure 4. Partial-Credit Scale

Rating scale analysis is conducted to determine whether the rating of the answer choices is confusing to respondents. Figure 5 clearly shows that the mean observation starts at logit values of -1.33, -1.05, 1.26, and 2.04 for score options 1 (strongly disagree), 2 (disagree), 3 (agree), and 4 (strongly agree). Therefore, there is a rise in the logit value from options 1 to 4. Another recommended measure is the Andrich Threshold to test the correctness of the polytomy. The Andrich Threshold value, which progresses from 'NONE' to negative and then to positive, sequentially indicates that the provided options are valid for respondents. In addition, the threshold distance between each rating is also known. Category 1-2 ranges from -3.31 to -1.20, with a difference of 2.11; Category 2-3 starts from -1.20 to 1.09, with a distance of 2.29; and Category 3-4 starts from 1.09 to 3.50, with a distance of 2.41. It can be seen that the adjusted threshold distance meets the criteria, which fall within the interval of 1.4-5.0.

SUMMARY OF CATEGORY STRUCTURE. Model="R"									
CATEGORY	OBSERVED	OBSVD	SAMPLE	INFIT	OUTFIT	ANDRICH	CATEGORY		
LABEL	SCORE	COUNT	%	AVRGE	EXPECT	MNSQ	MNSQ	THRESHOLD	MEASURE
1	1	2926	11	-1.33	-1.72	1.51	1.52	NONE	(-3.31)
2	2	6671	26	-1.05	-1.67	.72	.73	-2.10	-1.20
3	3	10977	42	1.26	1.12	.74	.74	-1.25	1.09
4	4	5474	21	2.04	2.07	1.06	1.04	2.35	(3.50)

OBSERVED AVERAGE is mean of measures in category. It is not a parameter estimate.

Figure 5. Rating Scale Analysis

Differential Item Functioning (DIF) testing results were obtained with probability values on Winstep, where the difference in mathematics teaching and learning beliefs was declared significant when the p-value was less than 0.05. DIF is shown through a picture that shows the differences in grain characteristics between groups. If an item works the same for all groups, then the curves from each group will overlap or be very similar. This research has several attributes, including gender,

respondents' age, respondents' occupation, the number of students, and respondents' demographics. Therefore, the four items must be corrected because the values contain bias based on the respondents' gender.

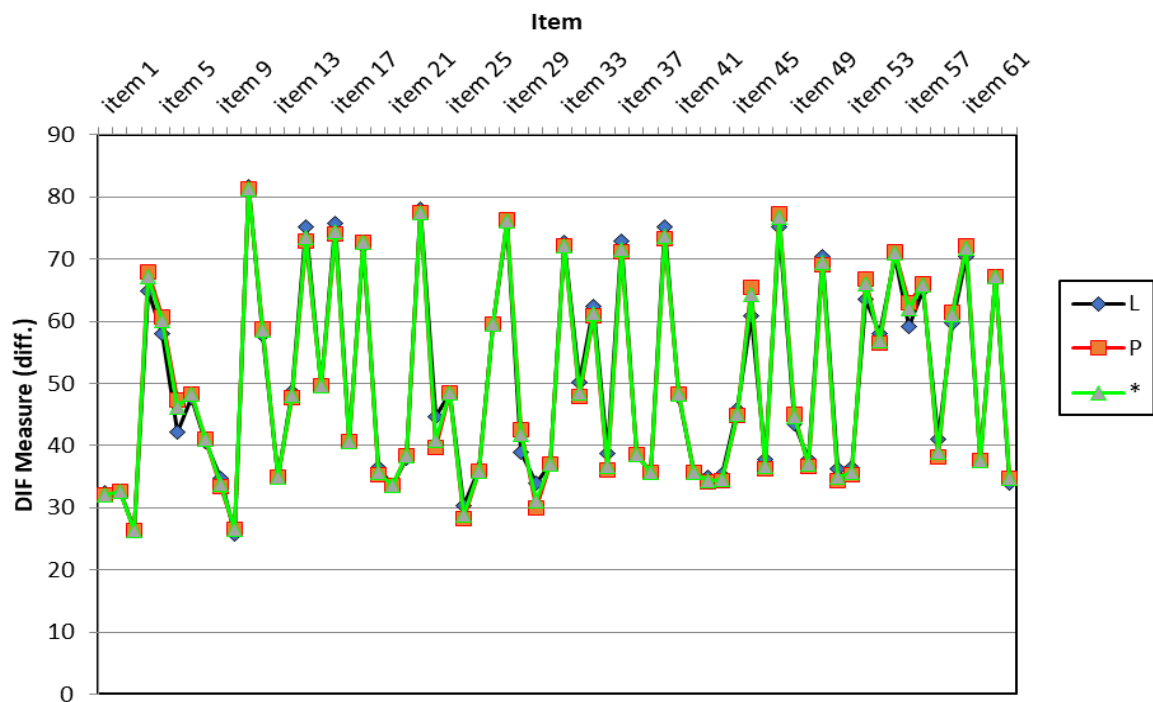


Figure 6. DIF by Respondent's Gender Category

Figure 6 shows that four items are biased toward the respondent's gender category, namely items 6, 24, 46, and 57. Subsequently, DIF testing was conducted, taking into account the age differences. The ages were divided into four groups: 21-30, 31-40, 41-50, and 51-60 years. The results showed that five items were biased towards the age category, namely items 6, 7, 10, 20, and 55. DIF testing was conducted considering the occupations of the respondents, who were divided into three groups: bachelor's degree students in education, pre-service teachers, and teachers. The results showed that 22 items were biased towards the job category, specifically items 3, 10, 11, 13, 15, 16, 17, 19, 21, 29, 32, 33, 34, 35, 37, 40, 41, 45, 48, 56, 59, and 62. The subsequent test was conducted in the schools, taking into account the number of students. The schools contained (1) 300 students, (2) 300–1,000 students, (3) more than 1,000 students, and (4) irrelevant responses since some students did not work. The results showed that four items were biased towards the category of the number of students, namely items 6, 13, 25, and 41. DIF testing was conducted considering the demographics of the areas, including rural, small towns, semi-densely populated areas, and large cities. The results showed that seven items were biased towards the demographics of the areas, namely items 13, 14, 24, 28, 33, 35, and 41. Items affected by DIF required correction due to measurement-biased factors (Appendix 1).

5. Discussion

The person measure value of 0.54 shows that respondents agree with statements in diverse belief instrument items (Rahim & Haryanto, 2021; Sumintono, 2018). Table 3 summarizes the data and shows the alignment between the research instrument items and responses. This is evident from the Cronbach's Alpha value, which has an excellent interpretation. The person's reliability value, with fairly good interpretation, indicates the consistency of their answers. Meanwhile, the item reliability value indicates the quality of the instrument, providing an excellent understanding. These results show that belief measurement instruments are reliable and deliver consistent results (Adi et al., 2022; Andrich & Marais, 2019; Medvedev & Krägeloh, 2022; Rahim & Haryanto, 2021; Sumintono, 2018).

Table 3. Summary of Instrument Reliability Test Analysis

No.	Reliability Test Analysis	Value	Interpretation	Conclusion
1	Cronbach Alpha	0.81	Excellent	Reliable
2	Item Reliability	1.00	Excellent	
3	Person Reliability	0.77	Fairly Good	

Instrument validity shows that a measuring instrument can assess mathematics teaching and learning beliefs (Perera et al., 2018; Rahim & Haryanto, 2021; Sumintono, 2018). Validity tests based on item unidimensionality indicate that the total variance in observations accounts for 55.3% of the raw variance. Interpretation of item unidimensionality based on the value of raw variance explained by measures is indicated by a score of > 20% being fulfilled, > 40% being good, and > 60% meeting special criteria. This shows that the unidimensionality value in instrument development can be met because the minimum unidimensionality requirement is 20% (Medvedev & Krägeloh, 2022). The eigenvalue must be less than 3 to indicate the absence of problematic instrument items, and the observed value must be less than 15% to indicate the presence of appropriate instrument items (Sumintono, 2018). Research data indicate a match between instrument items, as the variance that the mathematics teaching and learning belief instrument cannot explain accounts for 12.7%. The eigenvalue of 18.2 indicates problematic instrument items, so an item fit order analysis is carried out to determine which items should be discarded.

Item fit or suitability explains the functionality of the instrument. Several criteria are used to determine item fit, including Point Measure Correlation, INFIT MNSQ, Outfit MNSQ, and Outfit ZSTD. A rating (partial-credit) scale analysis was also carried out, while variable maps and item measures were used to define the instrument. DIF testing showed that the belief instrument was not biased towards the categories of gender, age, occupation, number of students, and demographics of respondents. Several items are problematic and must be discarded (Misfit) because they do not meet predetermined criteria. Table 4 summarises the analysis results of the Misfit items, and Table 5 shows a summary of the Fit items.

Table 4. Misfit Items

No	Misfit Item Analysis	Misfit Items	Misfit Item Number
1	INFIT MNSQ	10 Item	5, 6, 7, 14, 24, 30, 34, 41, 45, 55
2	OUTFIT MNSQ; OUTFIT ZSTD; Pt- Correlation	1 Item	7
3	DIF by gender	4 Item	6, 24, 46, 57
4	DIF by respondent age	5 Item	6, 7, 10, 20, 55
5	DIF by respondent occupation	22 Item	3, 10, 11, 13, 15, 16, 17, 19, 21, 29, 32, 33, 34, 35, 37, 40, 41, 45, 48, 56, 59, 62
6	DIF by the number of students	4 Item	6, 13, 25, 41
7	DIF by Respondent demographics	7 Item	13, 14, 24, 28, 33, 35, 41
Item Misfit		34 Item	3, 5, 6, 7, 10, 11, 13, 14, 15, 16, 17, 19, 20, 21, 24, 25, 28, 29, 30, 32, 33, 34, 35, 37, 40, 41, 45, 46, 48, 55, 56, 57, 59, 62

Table 5. Fit Items Mathematics Beliefs Questionnaire

Research focus	Research sub-focus	Description	Fit items
Beliefs on the nature of mathematics	Definition of mathematics	Subject views on mathematics	1, 8
	Relationship between mathematics and life	Subject views concerning the advantages of mathematics in daily life	61
	Development of mathematical knowledge	Subject beliefs on mathematical knowledge development	9, 47, 54
Beliefs about mathematics learning	Designing learning activities	Respondents' beliefs regarding designing learning activities that are appropriate to students' needs	2, 49
	Teaching approaches	Respondents' beliefs regarding appropriate approaches to mathematics teaching	26, 27, 36, 52
	The role of teachers in teaching	Respondents' beliefs regarding the role of teachers in mathematics teaching	31, 39, 51, 53, 63
	Problem-solving	Respondents' beliefs regarding problem-solving activities in mathematics learning	4, 23, 50, 58
	Assessing learning activities	Respondents' beliefs about assessments used in learning	22, 42, 64
Beliefs about mathematics learning	Student activities in mathematics learning	Respondents' beliefs about how students learn mathematics	12, 18, 38, 60
		Respondents' beliefs about the basic things that need to be mastered when learning mathematics	43, 44

A good instrument must meet the requirements of reliability and validity (Adi et al., 2022; Chan et al., 2021; Sumintono, 2018). Testing using the Rasch model gives better results and considers various criteria than the Pearson product-moment correlation (Medvedev & Krägeloh, 2022; Oktaviyanti et al., 2024). A total of 30 fit items may be used to assess mathematics learning beliefs. These items have met several established criteria with a good level of reliability. The results also show that each rating has a separate peak (Figure 4). Therefore, respondents can rate mathematics teaching and learning belief instruments differently (Sumintono, 2018). Furthermore, the research results also show that the rating of the answer choices given does not confuse respondents. This can be seen from the Andrich Threshold value, which moves sequentially from 'NONE' to negative and then to positive, indicating that the options provided are valid for the respondents. This means that respondents can be sure of each choice they make (Andrich & Marais, 2019; Chan et al., 2021). Thirty Fit items for measuring mathematics teaching and learning beliefs are feasible because the fewer instrument items, the better they are for carrying out measurements. Respondents may become fatigued with a long instrument, which can reduce reliability (Perera et al., 2018). The results are an

improvement from previous research in measuring beliefs to solve mathematical problems, because only a separation value was obtained (Muhtarom, 2024). In this research, the separation value $H = 2.77$ was rounded to 3. This aligns with the initial concept of developing an instrument to categorize respondents' beliefs into constructivist, Platonic, and instrumentalist categories.

These results indicate that the scale of beliefs in mathematics learning analyzed using the Rasch model has good psychometric characteristics. The developed instrument shows an adequate level of item fit and respondent fit, as well as an even distribution of item difficulty levels, so that it is able to measure the construct of beliefs in various categories of respondents. This finding provides an important contribution when compared to previous studies using the classical test theory approach (Masal & Takunyaci, 2012; Platas, 2015). The classical approach relies heavily on sample characteristics and is unable to separate item parameters from respondent characteristics. In contrast, the Rasch model allows researchers to identify misfit items, detect bias between groups, and produce linear measures of latent construct beliefs, which cannot be achieved through classical test theory (CTT).

The novelty of this research is demonstrated by developing an instrument using Rasch analysis, which can provide more objective and accurate measurement results than conventional approaches. With Rasch analysis, the instrument developed can ensure that each item has high validity and reliability and can precisely distinguish the level of teacher beliefs. This research is necessary because teachers' beliefs in mathematics teaching directly impact teaching practices and student learning outcomes.

6. Conclusion

The research results from Rasch modeling have developed a reliable and valid mathematics teaching and learning belief measurement instrument. Thirty instrument statement items meet the established criteria. However, some items still require improvement because they still contain bias and do not encompass the full range of variables studied. The availability of valid and reliable instruments is needed to support continuous evaluation and professional development of teachers. This research contributes to the development of measurement theory and practice in education, especially in understanding the psychological factors that influence the effectiveness of mathematics learning.

7. Suggestion

In the future, it is necessary to develop separate mathematics teaching and learning belief instruments for teachers and students, as the results of this study still contain many items biased towards the respondent's job category.

Declarations

Author Contributions. MM: Conceptualisation, Investigation, Formal Analysis, and Writing - Review & Editing; NN: Methodology, and Writing - Original Draft; DN: Validation, Supervision, and Project Administration; and TJBE: Writing - Review & Editing.

Conflicts of Interest. The authors state that there is no conflict of interest.

Funding. Directorate of Research, Technology, and Community Service, Indonesian Ministry of Research, Technology, and Higher Education, supported and sponsored this research.

Data Availability Statement. The datasets produced in this research can be obtained from the author upon reasonable request.

Acknowledgements. The authors are grateful to the Directorate of Research and Community Service, Kemdiktisaintek, for granting the Fundamental research grants under contract 108/E5/PG.02.00.PL/2024. Additionally, the authors express their gratitude to all teachers and students willing to become research respondents.

References

- Adi, N. R. M., Amaruddin, H., Adi, H. M. M., & A'yun, I. L. Q. (2022). Validity and Reliability Analysis Using the Rasch Model to Measure the Quality of Mathematics Test Items of Vocational High Schools. *Journal of Research and Educational Research Evaluation*, 11(2), 103–112.
- Akker, J. V., Gravemeijer, K., McKenney, S., & Nieveen, N. (2006). *Educational Design Research* (1st ed.). Taylor and Francis Group.
- Andrich, D., & Marais, I. (2019). *A Course in Rasch Measurement Theory: Measuring in the Educational, Social and Health Sciences*. Springer Nature Singapore. <https://doi.org/10.1007/978-981-13-7496-8>
- Ariza Muñoz, E. D. C., González-Calero, J. A., De Oro Aguado, C. M., & Cózar-Gutiérrez, R. (2023). Colombian future teachers' beliefs about mathematics and its learning. *International Journal of Mathematical Education in Science and Technology*, 54(8), 1474–1489. <https://doi.org/10.1080/0020739X.2023.2170834>
- Arıkan, E. E. (2016). Prospective teachers' beliefs about problem solving in multiple ways. *Universal Journal of Educational Research*. <https://openaccess.izu.edu.tr/xmlui/handle/20.500.12436/5297>
- Bal, A. P. (2015). Examination of the Mathematical Problem-Solving Beliefs and Success Levels of Primary School Teacher Candidates Through the Variables of Mathematical Success and Gender. *Educational Sciences: Theory and Practice*, 15(5), 1373–1390.
- Belbase, S. (2019). Meanings, dimensions, and categories of mathematics teacher beliefs: A navigation through the literature. *International Journal on Emerging Mathematics Education*, 3. <https://zuscholars.zu.ac.ae/works/2345/>
- Belbase, S. (2024). Teacher belief, knowledge, and practice: A trichotomy of mathematics teacher education. *Pragya Ratna प्रज्ञारत्न*, 6(2), 186–209.
- Buehl, M. M., & Fives, H. (2009). Exploring Teachers' Beliefs About Teaching Knowledge: Where Does It Come From? Does It Change? *The Journal of Experimental Education*, 77(4), 367–408. <https://doi.org/10.3200/JEXE.77.4.367-408>
- Chan, S.-W., Looi, C.-K., & Sumintono, B. (2021). Assessing computational thinking abilities among Singapore secondary students: A Rasch model measurement analysis. *Journal of Computers in Education*, 8, 213–236.
- Ferguson, L., & Braten, I. (2018). Student teachers' beliefs about learning, teaching, and teaching knowledge. *Teacher Education Practice*, 31(3), 348.
- Fives, H., & Buehl, M. M. (2016). Teachers' Beliefs, in the Context of Policy Reform. *Policy Insights from the Behavioral and Brain Sciences*, 3(1), 114–121. <https://doi.org/10.1177/2372732215623554>
- Hibi, W. (2017). Beliefs and attitudes about mathematics among students and their relationship to their achievement in mathematics. *Journal of Harbin Institute of Technology*, 49(11), 47–57.
- Julianto, V., Sumintono, B., Almakhi, N. P. Z., Avetazain, H., Wilhelmina, T. M., & Wati, D. A. (2025). Academic Motivation Scale's psychometric attribute: Analysis using Rasch measurement model. *Current Psychology*, 44(1), 114–124. <https://doi.org/10.1007/s12144-024-07142-7>
- Li, R., Cevikbas, M., & Kaiser, G. (2024). Mathematics teachers' beliefs about their roles in teaching mathematics: Orchestrating scaffolding in cooperative learning. *Educational Studies in Mathematics*, 117(3), 357–377. <https://doi.org/10.1007/s10649-024-10359-9>

- Lui, A. M., & Bonner, S. M. (2016). Preservice and inservice teachers' knowledge, beliefs, and instructional planning in primary school mathematics. *Teaching and Teacher Education*, 56, 1–13.
- Masal, E., & Takunyaci, M. (2012). The Turkish adaptation of mathematics belief scale: The validity and reliability study. *Procedia-Social and Behavioral Sciences*, 64, 123–132.
- Matitaputty, C., Nusantara, T., & Hidayanto, E. (2024). How Mathematics Teachers' Special Knowledge Changing: A Case Study in the Professional Teacher Education Program. *Journal on Mathematics Education*, 15(2), 545–574.
- Medvedev, O. N., & Krägeloh, C. U. (2022). Rasch Measurement Model. In O. N. Medvedev, C. U. Krägeloh, R. J. Siegert, & N. N. Singh (Eds.), *Handbook of Assessment in Mindfulness Research* (pp. 1–18). Springer International Publishing. https://doi.org/10.1007/978-3-030-77644-2_4-1
- Muhtarom, Juniati, D., & Siswono, T. Y. (2017). Consistency and inconsistency of prospective teachers' beliefs in mathematics, teaching, learning and problem solving. *AIP Conference Proceedings*, 1868, 050014.
- Muhtarom, M. (2024). Developing an instruments to measure prospective teacher beliefs about mathematical problem-solving using the Rasch model. *Jurnal Elemen*, 10(2), 274–288.
- Muhtarom, M., Juniati, D., & Siswono, T. Y. E. (2019). Examining prospective teacher beliefs and pedagogical content knowledge towards teaching practice in mathematics class: A case study. *Journal on Mathematics Education*, 10(2), 185–202.
- Muhtarom, M., Juniati, D., Siswono, T. Y. E., & Rahmatika, I. (2018). Teachers' and students' beliefs in mathematics at State Senior High School 5 Semarang. *Jurnal Riset Pendidikan Matematika*, 5(1), 64–72.
- Muhtarom, M., Siswono, T. Y. E., & Juniati, D. (2020). Climber prospective teacher: Relationship beliefs and mathematics teaching practice. *Al-Jabar: Jurnal Pendidikan Matematika*, 11(1), 83–92.
- Muhtarom, S., & Nizaruddin, Y. H. M. (2024). Research on mathematical beliefs: Systematic literature review. *Int J Eval & Res Educ*, 13(2), 693–704.
- Nieveen, N., & Folmer, E. (2013). Formative evaluation in educational design research. *Design Research*, 153(1), 152–169.
- Oktaviyanthi, R., Agus, R. N., Garcia, M. L. B., & Lertdechapat, K. (2024). *Cognitive load scale in learning formal definition of limit: A rasch model approach*. <https://archium.ateneo.edu/mathematics-faculty-pubs/265/>
- Ozturk, T., & Guven, B. (2016). Evaluating students' beliefs in problem solving process: A case study. *Eurasia Journal of Mathematics Science and Technology Education*, 12(3). <https://avesis.ktu.edu.tr/yayin/1cd209ab-9adc-4ca8-9ca0-97ee45572eea/evaluating-students-beliefs-in-problem-solving-process-a-case-study>
- Pehkonen, E., & Pietilä, A. (2003). On relationships between beliefs and knowledge in mathematics education. *Proceedings of the CERME-3 (Bellaria) Meeting*, 1–8. https://www.academia.edu/download/41357702/TG2_pehkonen_cerme3.pdf
- Perera, C. J., Sumintono, B., & Jiang, N. (2018). The psychometric validation of the principal practices questionnaire based on item response theory. *International Online Journal of Educational Leadership*, 2(1), 21–38.

- Perera, H. N., & John, J. E. (2020). Teachers' self-efficacy beliefs for teaching math: Relations with teacher and student outcomes. *Contemporary Educational Psychology*, 61, 101842.
- Platas, L. M. (2015). The Mathematical Development Beliefs Survey: Validity and reliability of a measure of preschool teachers' beliefs about the learning and teaching of early mathematics. *Journal of Early Childhood Research*, 13(3), 295–310. <https://doi.org/10.1177/1476718X14523746>
- Purnomo, Y. W. (2017). A Scale for Measuring Teachers' Mathematics-Related Beliefs: A Validity and Reliability Study. *International Journal of Instruction*, 10(2), 23–38.
- Purnomo, Y. W., Pramudiani, P., Aziz, T. A., Kaur, A., Ismail, S. N., & Nuriadin, I. (2020). Indonesian teachers beliefs on the gap between educational research and practice. *Australian Journal of Teacher Education (Online)*, 45(12), 24–42.
- Rahim, A., & Haryanto, H. (2021). Implementation of Item Response Theory (IRT) Rasch Model in Quality Analysis of Final Exam Tests in Mathematics. *Journal of Research and Educational Research Evaluation*, 10(2), 57–65.
- Siswono, T. Y. E., Kohar, A. W., & Hartono, S. (2019). Beliefs, knowledge, teaching practice: Three factors affecting the quality of teacher's mathematical problem-solving. *Journal of Physics: Conference Series*, 1317, 012127.
- Siswono, T. Y. E., Kohar, A. W., Hartono, S., Rosyidi, A. H., Kurniasari, I., & Karim, K. (2019). Examining teacher mathematics-related beliefs and problem-solving knowledge for teaching: Evidence from Indonesian primary and secondary teachers. *International Electronic Journal of Elementary Education*, 11(5), 493–506.
- Siswono, T. Y. E., Kohar, A. W., Kurniasari, I., & Astuti, Y. P. (2016). An investigation of secondary teachers' understanding and belief on mathematical problem solving. *Journal of Physics: Conference Series*, 693(1), 012015. <https://iopscience.iop.org/article/10.1088/1742-6596/693/1/012015/meta>
- Spruce, R., & Bol, L. (2015). Teacher beliefs, knowledge, and practice of self-regulated learning. *Metacognition and Learning*, 10, 245–277.
- Sumintono, B. (2018). Rasch model measurements as tools in assesment for learning. *1st International Conference on Education Innovation (ICEI 2017)*, 38–42.
- Wang, Y., Qin, K., Luo, C., Yang, T., & Xin, T. (2022). Profiles of Chinese mathematics teachers' teaching beliefs and their effects on students' achievement. *ZDM – Mathematics Education*, 54(3), 709–720. <https://doi.org/10.1007/s11858-022-01353-7>
- Yang, X., Kaiser, G., König, J., & Blömeke, S. (2020). Relationship between pre-service mathematics teachers' knowledge, beliefs and instructional practices in China. *Zdm*, 52, 281–294.

About the Contributor(s)

Muhtarom, is an Associate Professor and Teacher Educator at the Universitas PGRI Semarang. His research interests focus on issues in teacher and teacher education, mathematics education, higher education, 21st-century teaching and learning, school-based assessment, beliefs, problem-solving, and metacognition.

Email: muhtarom@upgris.ac.id

ORCID: <http://orcid.org/0000-0001-8824-9166>

Nizaruddin received a Ph.D. degree in education from Universitas Negeri Semarang. He has over 20 years of experience as an academician at Universitas PGRI Semarang. His current research interests

include students' learning and development at various levels and in various areas of education. His publication topics include educational courses, math translation, multi-representation mathematics concepts, and problem-solving.

Email: nizaruddin@upgris.ac.id

ORCID: <http://orcid.org/0000-0002-1166-6006>

Duwi Nuvitalia is a lecturer in the Elementary School Teacher Education Department, Universitas PGRI Semarang. Her main research interests are on learning, critical thinking skills and STEAM approaches.

Email: duwinuvitalia@upgris.ac.id

ORCID: <http://orcid.org/0000-0001-9447-6349>

Tjaart Jan B. Estrada is a lecturer in the Mathematics and Allied Disciplines Department, Don Mariano Marcos Memorial State University.

Email: tjestrada@dmmmsu.edu.ph

ORCID: <https://orcid.org/0000-0003-0255-8334>

Publisher's Note: The opinions, statements, and data presented in all publications are solely those of the individual author(s) and contributors and do not reflect the views of Universitepark, EDUPIJ, and/or the editor(s). Universitepark, the Journal, and/or the editor(s) accept no responsibility for any harm or damage to persons or property arising from the use of ideas, methods, instructions, or products mentioned in the content.

Appendix 1

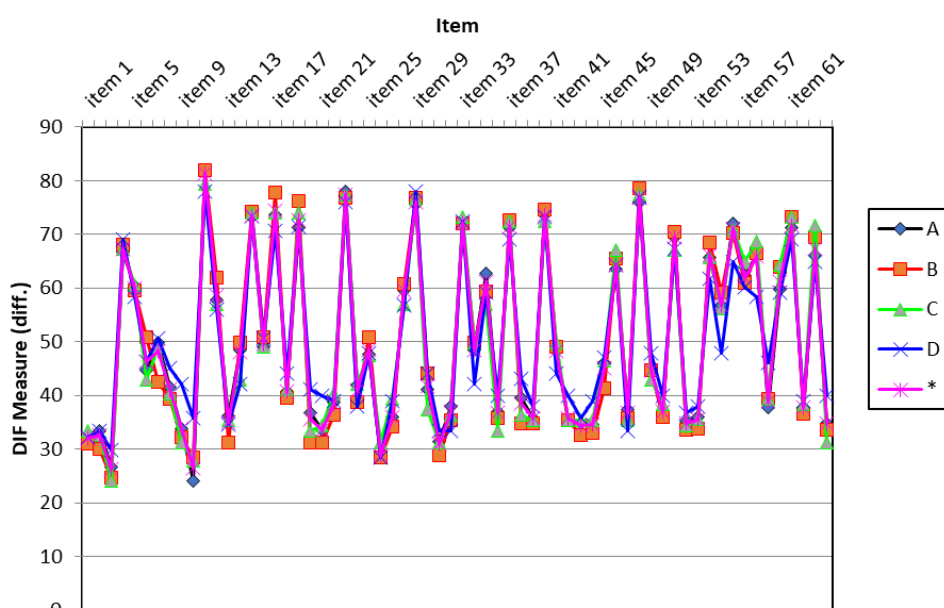


Figure 7. DIF by Respondents' Age Category

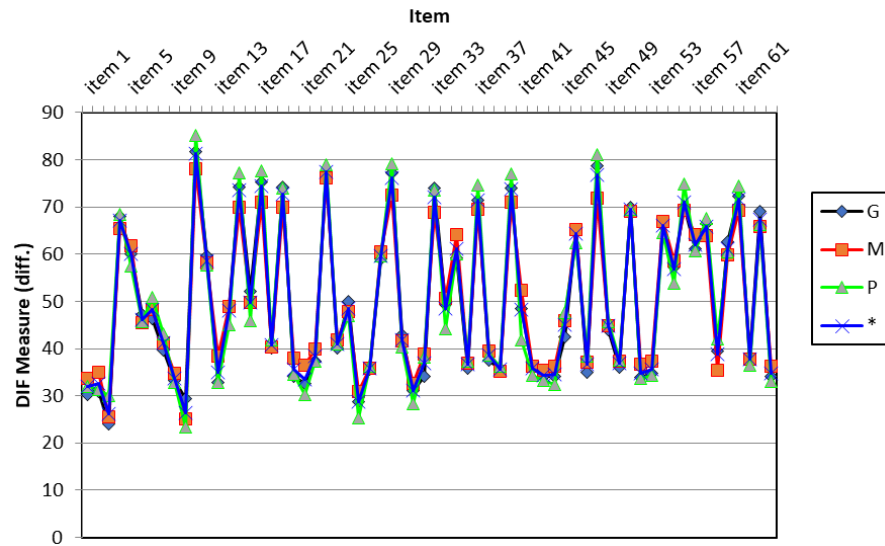


Figure 8. DIF by Respondents' Occupation Category

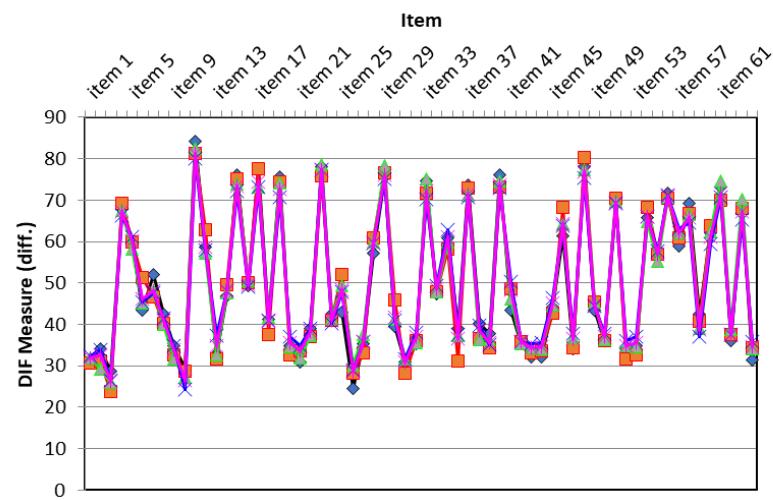


Figure 9. DIF by Number of Students Category

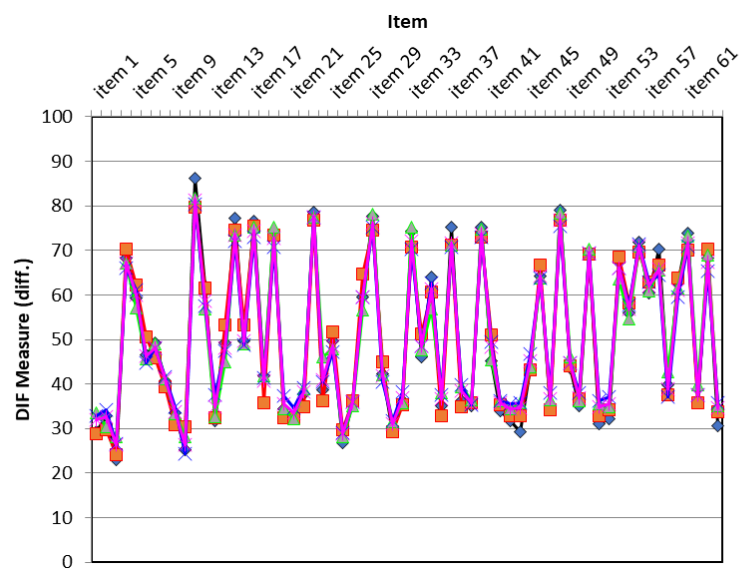


Figure 10. DIF by Respondents' Demographics Category