



Integrating *Batik Ghentongan* Tanjung Bumi Madura into Science Education: Development of Student Response Instrument

Wiwin Puspita Hadi* and Aida Fikriyah

Department of Science Education, Universitas Trunodjoyo Madura, Bangkalan, Indonesia

* Email: wiwin.puspitahadi@trunojoyo.ac.id

Abstract

Batik Ghentongan Tanjung Bumi Madura, as a representation of local cultural wisdom, offers meaningful potential for contextual integration into science education. The aim of this research was to analyze the validity and reliability of an instrument designed to assess student responses to Batik Ghentongan Tanjung Bumi Madura. The study was conducted in Tanjungbumi Subdistrict, Bangkalan Regency, East Java, Indonesia, involving 248 junior high school pupils, including of 119 males and 129 females, selected through simple random sampling. A quantitative survey design was employed using 8 items on a 4 point Likert scale. Data were analyzed using common factor analysis and the Rasch model with Winsteps software version 4.7.0.0. The results show that 6 items met the validity criteria, while 2 items require revision; the raw variance explained by measures reached 24 percent, indicating acceptable unidimensionality, and the Cronbach Alpha value of 0.95 demonstrates excellent internal consistency reliability. In conclusion, the instrument is adequate valid and reliable for identification students' responses to Batik Ghentongan Tanjung Bumi Madura, although further refinement of several items is recommended to strengthen construct validity and improve measurement precision. This study contributes to Sustainable Development Goal 4, Quality Education, and Sustainable Development Goal 8, Decent Work and Economic Growth, by promoting culturally responsive science learning and strengthening awareness of local cultural industries as a component of sustainable economic development.

Keywords: *Batik Ghentongan*, Local Wisdom, Science Education, Student Responses

SDGs: Goal 4 (Quality Education), Goal 8 (Decent Work and Economic Growth)

Manuscript History

Received: October 28, 2025

Revised: February 19, 2026

Accepted: February 22, 2026

How to cite:

Hadi, W.P. & Fikriyah, A. (2026). Integrating *Batik Ghentongan* Tanjung Bumi Madura into Science Education: Development of Student Response Instrument. *International Journal of Research and Community Empowerment*, 4(1), 59-71. DOI: <https://doi.org/10.58706/ijorce.v4n1.p59-71>.

INTRODUCTION

Context based learning has been applied in science education. Context based learning frames science concepts within meaningful real-world situations (local environment) and teaches content on a “need-to-know” basis so concepts are learned to make sense of that context (Giamellaro et al., 2025; Podschuweit & Bernholt, 2018; Zuleni & Marsidin, 2025). Context based education is posited to enhance students’ motivation and interest by situating new ideas in familiar, real-world situations, thereby rendering learning more meaningful (Picardal & Sanchez, 2022; Puspitasari et al., 2024). Context based education connects with constructivist and sociocultural theories that promote the development of knowledge through experience, social interaction, and cultural (Mishra, 2023; Rai, 2025; Sulindra et al., 2024). The implementation of a context-rich instructional method has the potential to overcome these challenges by clearly tying concepts to practical applications,

linking factual content with conceptual understanding, and providing chances for knowledge transfer across tasks.

This issue leads to science learning often being perceived as abstract and disconnected from the context of students' lives, resulting in low engagement and negative attitudes toward science (Coffie et al., 2025; Muñoz et al., 2025). To address these issues, context is needed that can connect scientific concepts to students' daily lives so that they can learn science more easily. One potential context is local wisdom or culture in society, which is expected to make students more interested in learning science (Usman et al., 2024). Integrating local wisdom into learning is seen as a strategy to bridge this gap, enhancing the significance, relevance, and contextuality of science for pupils. The linking of local wisdom and cultural aspects into science learning is expected to make science more significance, relevance, and contextuality of science for pupils. This context can be helpful and practical for students in improving their ability to think concretely to an abstract level and transform science from the context of local wisdom to the context of science or vice versa (Erman et al., 2022).

Cultural product *batik Ghentongan* Tanjung Bumi Madura is one of the unique types of *batik* because it is dyed using natural dyes, namely *Indigofera tinctoria* and *Garcinia dulcis*, and soaked in barrels (*Gentong* : Indonesian) for several months. The batik process is carried out on both sides, which makes it very expensive (Bawono et al., 2023; Sari, 2022; Sari & Miftah, 2020). *Batik Ghentongan* Tanjung Bumi, Bangkalan Madura, represents a cultural heritage rich in techniques and empirical knowledge that is expected to connect with scientific concepts. The integration of *batik Ghentongan* into science concepts and science classes has the potential to strengthen culture, entrepreneurship, and environmental awareness (Kusumastuti et al., 2023; Purwastuti, 2022). The research activities conducted on *batik Ghentongan* included research on the analysis of requirements for creating learning media with the context of *batik Ghentongan* (Afrida et al., 2025; Martareza et al., 2025) and the exploration of scientific concepts in *batik Ghentongan* (Hadi & Fikriyah, 2024). However, the sustainability of *batik Ghentongan* Tanjung Bumi faces challenges amid social, economic, and technological changes that require a comprehensive study within the framework of sustainable development goals in particular. SDGs 4 (quality education) dan SDGs 8 (decent work and economic growth) (Prasetiyo et al., 2022). The process of passing down batik-making is still done informally based on family traditions, so integrating it with education can ensure the sustainability of *batik Ghentongan*. Additionally, this batik is related to SDG 8 because it serves as a source of income for the community, thus actions are needed, one of which is understanding students' responses to *batik Ghentongan* so that it can serve as a reference for the preservation of *batik Ghentongan* Tanjung Bumi.

Nevertheless, *batik Ghentongan* Tanjung Bumi is famous, integrated batik into science lessons is still limited. Student responses are still limited to the need for learning media that uses *batik Ghentongan* Tanjung Bumi, and the number of respondents involved is also very limited. Based on these conditions, there is a gap in previous research, namely the lack of data on students' knowledge of *batik Ghentongan*, their perception of the relevance of *batik Ghentongan* to scientific concepts, and their attitudes towards the existence of *batik Ghentongan* Tanjung Bumi Madura (Hadi et al., 2025). Based on this urgency, this study focuses on describing student responses to the implementation of *batik Ghentongan* Tanjung Bumi in science education.

METHOD

Research Design

The research method used was quantitative descriptive to provide an overview of information about student responses to *batik Ghentongan* Tanjung Bumi obtained during data collection.

Research Procedure

Figure 1 illustrates the research procedures. Based on Figure 1, the research was carried out systematically through five continuous stages. The first stage involved questionnaire development, in which the items were constructed based on the conceptual framework of indigenous knowledge integration in science education. The second stage was questionnaire distribution to the selected respondents. The third stage consisted of data collection from 248 junior high school students who completed the instrument. Subsequently, the data collected were processed and assessed utilizing common factor analysis. and the Rasch model to examine validity and reliability. Finally, the analysis results were used to assess the instrument's quality and to identify issues for modification. This structured procedure ensured that each stage was carried out rigorously to produce a valid and reliable measurement tool.

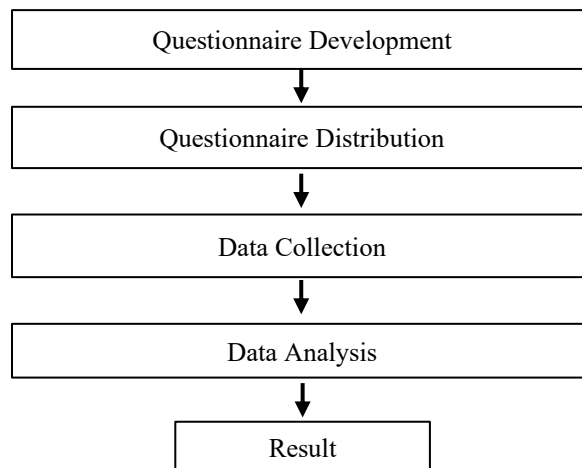


Figure 1. The Research Procedure

Data Sources or Participants

The study was conducted in Tanjungbumi Subdistrict, Bangkalan Regency, East Java, Indonesia. Population all 8th-grade students in junior high schools in Tanjungbumi with purposive sampling technique, included 248 students from two state junior high schools, aged 12-14 years. The Table 1 is the data of the respondents.

Table 1. Respondents

	Male	Female	Total
Junior high school A	76	87	163
Junior high school B	43	42	85
Total	119	129	248

Data Collection Techniques

The data collection instrument was a questionnaire, and the primary data were obtained directly from students’ responses. The questionnaire was developed using a Likert scale with four response options: strongly agree (SA), agree (A), disagree (DA), and strongly disagree (SDA). The questionnaire given to students included statements about: (1) students' knowledge of the process of making *batik Ghentongan* Tanjung Bumi, Madura; (2) *batik Ghentongan* differs from other hand-drawn batik primarily in its dyeing techniques; (3) The application of natural colors in *Batik Ghentongan* Tanjung Bumi; (4) the integration of *batik Ghentongan* in books, worksheets, and science videos; (5) Science topics that can be linked to *batik Ghentongan* Tanjung Bumi , such as natural and synthetic dyes; (6) Science topics that can be linked to *batik Ghentongan* Tanjung Bumi, such as temperature and heat; (7) *batik Ghentongan* Tanjung Bumi, Madura, can also be taught in subjects other than science; (8) *batik Ghentongan* Tanjung Bumi is increasingly rare and thus requires preservation.

Data Analysis Techniques.

The analytical technique was executed utilizing the Rasch Model with the assistance of Winstep 4.7.0.0 software. The Rasch model is a mathematical analysis model capable of describing the relationship between individuals and how they respond to items in the instrument (Muthia et al, 2021). Validity data derived from the Rasch model, utilizing item unidimensionality, can assess if the instrument effectively measures its intended construct. An instrument can be considered unidimensional when two requirements are met, namely that the raw variance explained by the measure is not less than 20%. Table 2 shows the criteria in detail.

Table 2. Raw Variance Explained Category

Raw Variance Explained	Criteria
< 20%	Low
20% – 40%,	Moderate
40% – 60%	Good
> 60%	Very Good

The second requirement is that the unexplained variance in the first construct is no more than 15%, If the unexplained variance in the first contrast exceeds 15%, the instrument may be considered multidimensionality, indicating that it measures more than one dimension, including potential dimensions not captured by this analysis (Sumintono & Widhiarso, 2014).

Table 3. Item Reliability Alpha Cronbach

Reliability	Criteria
< 0.5	Very poor
0.5 – 0.6	Poor
0.6 – 0.7	Quite
0.7 – 0.8	Good
> 0.8	Very Good

Table 4. Person Reliability Criteria

Person Reliability	Criteria
< 0.67	Weak
0.67 – 0.80	Enough
0.81 – 0.90	Good
0.91 – 0.94	Very Good
> 0.94	Excellent

The validity of the student response instrument was determined based on item fit and person fit analyses. The reliability of the instrument (item reliability) was analysed based on the Cronbach's Alpha value criteria (measuring reliability, i.e. the connection between the person and the item as a whole) with criteria as listed in in Table 3, meanwhile the person reliability criteria (measuring the consistency of student responses and items in measurement or analysis) as listed in Table 4. The testing was conducted with the help of Winsteps software and through Cronbach's Alpha (α) value analysis to assess the internal consistency of the instrument.

RESULTS AND DISCUSSION

Based on the results of processing student response data, the following results were obtained.

The Validity of Instrument

The term used in instrument validity testing using the Rasch model is instrument unidimensionality item. Unidimensionality items can be used to test and evaluate whether the instruments used are capable of measuring what they are supposed to measure and accurately represent the data from the variables so that the items can be considered valid (Nurhudaya et al., 2019). The results of the unidimensionality test can be seen in Figure 2.

Table of STANDARDIZED RESIDUAL variance in Eigenvalue units = ITEM information units

	Eigenvalue	Observed	Expected
Total raw variance in observations =	10.5293	100.0%	100.0%
Raw variance explained by measures =	2.5293	24.0%	24.5%
Raw variance explained by persons =	.7526	7.1%	7.3%
Raw Variance explained by items =	1.7768	16.9%	17.2%
Raw unexplained variance (total) =	8.0000	76.0%	100.0% 75.5%
Unexplnd variance in 1st contrast =	1.5898	15.1%	19.9%
Unexplnd variance in 2nd contrast =	1.2937	12.3%	16.2%
Unexplnd variance in 3rd contrast =	1.2533	11.9%	15.7%
Unexplnd variance in 4th contrast =	1.1679	11.1%	14.6%
Unexplnd variance in 5th contrast =	1.0328	9.8%	12.9%

Figure 2. Unidimensionality Test

From the unidimensionality test table 2, the following data was obtained raw variance explained measure 24.0% which means that the minimum requirement of 20% is met and is classified as sufficiently valid, meaning that this instrument is capable of capturing some of the variation in students' understanding or response to batik Ghentongan. Unexplnd variance in 1st to 5th construct, 15.1% unexplnd variance in 1st construct; 12.3% unexplnd variance in 2nd construct; 11.9% unexplnd variance in 3rd construct; 11.1% unexplnd variance in 4th construct; dan 9.8% unexplnd variance in 5th construct. unexplained variance value

in 1st (15.1%) or more than 15% indicates that the items may measure more than one construct. This can be summarized in Table 5.

Table 5. Unidimensionality Test

	Score	Criteria
Raw variance explained by measures	24%	Medium
Unexplned variance in 1 st contrast	15.1%	Unidimensionality

Unexplained variance in 1st contrast (15.1%) the items still measure one construct although the variance explained by the measures was relatively low (15.1%), the instrument still satisfied the minimum criteria for unidimensionality. This means that the completion of the student questionnaire was relevant to the topic of batik Ghentongan Tanjung Bumi in science learning. The questionnaire was designed to determine students' responses to the integration of batik Ghentongan Tanjung Bumi in science learning and science concepts. However, if there were questions related to motivation to learn about batik and its connection to the preservation of batik culture, students who liked batik tended to give higher responses to those questions, which could lead to errors in interpretation and reduce the validity of the instrument. Efforts that can be made to ensure that the instrument used is valid include focusing on one relevant dimension, such as ensuring that the instrument only measures one relevant dimension, such as knowledge of batik Ghentongan Tanjung Bumi, without mixing in other factors such as interest in art or general skills in art that are not directly related. The next step is to remove irrelevant items if the instrument contains items that measure other things that are not related to batik Ghentongan Tanjung Bumi (such as general interest in art or attitudes towards art). These items should be removed or modified to maintain focus on one dimension. Retest the instrument for unidimensionality: After the instrument has been modified, retest it to ensure that it measures only one dimension (unidimensionality) and that the students' responses better reflect their knowledge and understanding of batik Ghentongan Tanjung Bumi. Meanwhile, the unexplained variance in the first contrast (15.1%) indicates the presence of minor secondary factors influencing students' responses to Batik Ghentongan; however, the instrument remains predominantly unidimensional and suitable for measuring students' overall responses.

Item Level of Difficulty

The level of difficulty, items can be classified into four groups according to number of SD values with the average in the measurement column, very difficult (more than +1 SD), difficult (0.00 logit + 1SD), easy (0.0 logit-1SD), and very easy (less than -1SD) (Sumintono & Widhiarso, 2014). The analysis of item difficulty levels was derived from the item measure values calculated using the Rasch model, expressed in logit units. The item logit describes the difficulty or ease of a test item for participants according to the Rasch model. A higher item logit value implies that the thing is more difficult, while a lower item logit value means that the item is easier. This means that if the item logit value is high, then the item is the most difficult for respondents to agree with. The results of the item logit analysis are shown in Figure 3.

ITEM STATISTICS: MEASURE ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S. E.		INFIT		OUTFIT		PTMEASUR-AL		EXACT OBS%	MATCH EXP%	ITEM
				MNSQ	ZSTD	MNSQ	ZSTD	CORR.	EXP.					
1	651	248	.47	.07	1.24	3.02	1.25	2.98	.38	.46	43.5	40.7	1	
6	670	248	.36	.07	.89	-1.44	.91	-1.18	.51	.45	35.0	41.1	6	
2	692	248	.24	.08	.77	-3.17	.77	-3.04	.41	.44	50.4	41.4	2	
5	701	248	.19	.08	.90	-1.33	.90	-1.30	.46	.44	44.3	41.4	5	
7	743	248	-.06	.08	.99	-.05	.98	-.24	.39	.42	52.0	43.3	7	
4	750	248	-.11	.08	.96	-.52	.94	-.71	.44	.42	51.2	43.4	4	
3	779	248	-.29	.08	1.05	.64	1.05	.62	.41	.40	37.4	45.3	3	
8	845	248	-.79	.09	1.36	3.34	1.32	2.79	.38	.35	47.2	52.1	8	
MEAN	728.9	248.0	.00	.08	1.02	.1	1.01	.0			45.1	43.6		
P. SD	59.5	.0	.38	.01	.18	2.1	.17	1.9			6.0	3.5		

Figure 3. Item Difficulty Test

Based on Figure 3, it can be seen that the SD value is 0.38, It can be argued that the most difficult category has a score more than 0.38, the difficult category has a score range of 0.00 to 0.38, the easy category has a score range of -0.38 to less than 0.00, and the very easy category has a score of less than -0.38. In addition, Figure 2 also shows the level of difficulty of the items in sequence, from the item with the highest level of difficulty, item no. 1, with a logit value of 0.47, to the easiest item to answer, item no. 8, with a logit value of -0.79. Item number 1 is the most difficult item for respondents to agree with, while item no. 8 is the easiest item for respondents to agree with. Students know the procedure of creating *batik Ghentongan* Tanjung Bumi, indicating that most students lack knowledge regarding this process, furthermore item number 8 is that *batik Ghentongan* Tanjung Bumi is becoming increasingly scarce and hence requires preservation.

Question number 1 regarding the process of making *batik Ghentongan*, many respondents answered that they did not know, indicating that students' interest in regional culture is still low. Students are less interested due to several factors, including the fact that students tend to be more interested in developing trends that originate from other countries' cultures rather than local cultures (Aisara et al., 2020). The low interest of the younger generation in *batik* is also influenced by factors such as uncertain income, lack of artistic spirit, low income, and lack of business capital (Suliyanto et al., 2016) Therefore, various approaches must be taken towards primary school students to enhance their engagement with the culture of the Indonesian nation, especially *batik Ghentongan* Tanjung Bumi (Wulandari et al., 2023). Number 8 is easy to agree with because it relates to the preservation of *batik Ghentongan* Tanjung Bumi. *batik Ghentongan* Tanjung Bumi batik is becoming increasingly rare due to several factors, including the complicated and lengthy production process, the lack of regeneration of *batik Ghentongan* makers, competition with stamped or printed batik, and its high price compared to other types of *batik tulis*

Item Conformity Level

An evaluation of the appropriateness of each item was performed to confirm that the items operated well within the framework of the measurement objectives and did not induce confusion or misinterpretation among respondents (Ilfiandra et al., 2022). To view the level of suitability of items, you can use the menu *output tables* 10. ITEM (column): *fit order* with a focus on the value OUTFIT MNSQ, OUTFIT ZSTD, and PT-MEASURE CORR. An item can be said to be suitable when one of the three criteria for checking the level of item suitability (fit) and item unsuitability (outlier or misfit) is fulfilled. Each of the three criteria used can be seen in Table 6.

Table 6. Item Fit Criteria

Criteria	Fit item criteria
Outfit MNSQ	0.5 < MNSQ < 1.5
Outfit ZSTD	-2.0 < ZSTD < +2.0;
PT-Measure Corr	0.4 < PT-MEASURE CORR < 0.85

Based on the criteria in Table 6, the analysis findings are presented in Table 7.

Table 7. Item Fit Analysis

Statement	Outfit MNSQ	Outfit ZSTD	PT. Measure Corr	Criteria
P1	1.25	2.98	0.38	Misfit
P2	0.77	-3.04	0.41	Fit
P3	1.05	0.62	0.41	Fit
P4	0.94	-0.71	0.44	Fit
P5	0.90	-1.30	0.46	Fit
P6	0.91	-1.18	0.51	Fit
P7	0.98	-0.24	0.39	Fit
P8	1.32	2.79	0.38	Misfit

Based on Table 7, there are three misfit statements 1 and 8. Items with misfit can be deleted or corrected to obtain appropriate results. Misfit means that students gave different answers, indicating variations in interpretation of *batik Ghentongan* Tanjung Bumi. For the fit/misfit item analysis, the three criteria mentioned earlier were still used. However, The criteria for item deletion are predicated on analytical data that unequivocally demonstrate the item's inconsistency, namely two of the three criteria above, one of which is a negative point measure correlation value (Wibisono, 2016).

Wright Map Analysis (Person-Item Map)

The map variable will show the distribution of student abilities and the level of difficulty of the items on the same logit scale. Student abilities are listed on the left side of the map, while the level of difficulty of the items is on the right side of the map. A higher logit represents students with higher abilities (left side) and more difficult items (right side), so the purpose of this variable map is to identify whether the items are appropriate for the students' abilities. Logit 0 is the average value of the test items). Figure 4 shows that the person (left) is the distribution of participants' abilities and the item (right) is the distribution of the difficulty levels of the items addressed by the respondents. Based on the overall variable map test results, there is good alignment between respondent ability and item difficulty, especially in the range of -1 to +1 logit. The statement items already represent each category (low, medium, and high), and the students are diverse.

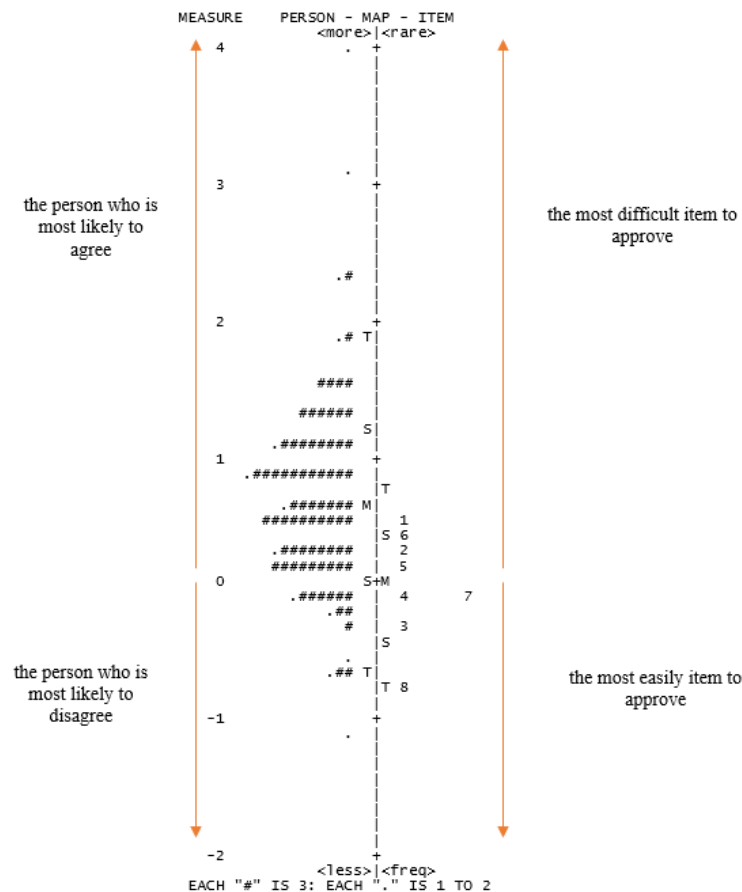


Figure 4. Variable Map Test

Rating Scale Diagnostic

The level of a measuring scale or alternative response possibilities for an instrument can be evaluated using a rating scale diagnostic (RSD) in Rasch modelling, with the aim of determining whether the measurement scale is appropriate (can distinguish between respondents with different abilities), reliable (consistently distinguishes between respondents with different abilities), and well-defined (respondents understand the differences between answer options). Scale validity analysis is a test conducted to verify whether the ratings used confuse respondents or not. Student responses were measured using a four-point Likert scale, with categories from strongly disagree (1) to strongly agree (4). Respondents might be said to comprehend an alternative answer when the observed average and Andrich threshold values increase as the level of the alternative response increases (Sumintono & Widhiarso, 2014). To see the RSD, you can use the output tables menu 3.2 Rating (partial credit) scale can be shown Figure 5.

SUMMARY OF CATEGORY STRUCTURE. Model="R"

CATEGORY LABEL	OBSERVED SCORE	OBSVD COUNT	SAMPLE %	INFINIT AVRGE	OUTFIT EXPECT	MNSQ MNSQ	ANDRICH THRESHOLD	CATEGORY MEASURE
1	1	165	8	.10	-.06	1.15 1.21	NONE	(-2.23)
2	2	437	22	.20	.26	.94 .95	-.88	-.65
3	3	736	37	.58	.61	.93 .87	-.09	.62
4	4	646	33	1.08	1.04	.95 .96	.97	(2.28)

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

Figure 5. Rating Scale Diagnostic

Figure 5 is described in accordance with Table 8.

Table 8. Rating Scale Diagnostic Analysis

Category Label	Description	Observed Average	Andrich Threshold
1	Strongly disagree	0.10	None
2	Disagree	0.20	-0.88
3	Agree	0.58	-0.09
4	Strongly agree	1.08	0.97

Table 8 shows that the average observation starts from a logit of +0.10 and rises sequentially to +1.80, indicating an increase that shows that the rating scale does not confuse respondents. Respondents can clearly distinguish between the ratings of strongly disagree, disagree, agree, and strongly agree. However, in addition to the direction of the logit from negative to positive, the consistency of the Andrich threshold value range also needs to be considered, where the standard logit value range for the Andrich threshold is greater than 1.4 and less than 5.0. The Andrich threshold value moves from none towards negative and continues to move towards positive sequentially, signifying that the options presented are legitimate for replies. Figure 6 presents a category probability graph based on Andrich thresholds. This graph is crucial for evaluating whether the scale categories (ratings) are well understood by respondents.

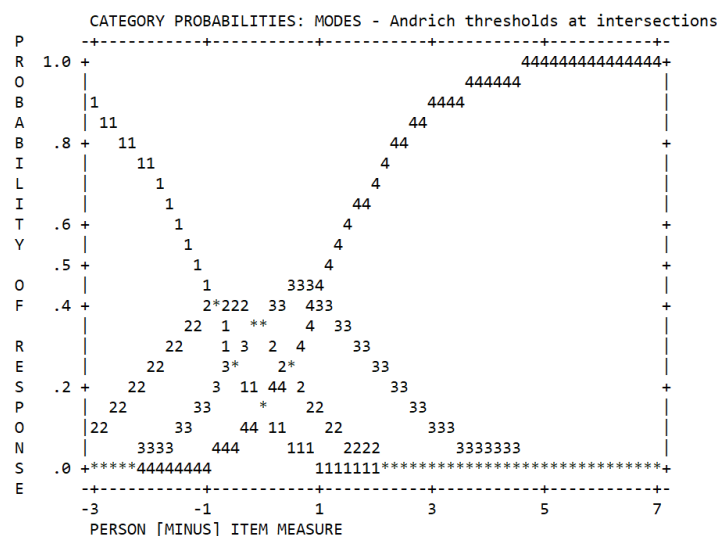


Figure 6. Category Probability Curves for Student Responds Items

Figure 6 shows the category probability curve from the Rasch model analysis with a rating scale of 1 to 4. This graph is used to see whether each category (e.g., disagree to strongly agree) is actually used and well understood by respondents. The results show that the four categories each have their own peaks, meaning that there are certain parts of the logit scale where that category is most frequently selected by respondents. This indicates that each category can be clearly distinguished and used consistently.

Differential Item Function (DIF) Analysis

Differential Item Function (DIF) is an important technique in conducting survey and test data analysis, and it is also necessary when conducting data analysis using the Rasch model (Boone et al., 2014). In addition, DIF can also be used to measure whether an item is biased or more favourable to a particular characteristic (Sumintono & Widhiarso, 2014). If the probability value of a statement item is less than 0.05, it is considered to have bias (Ilfiandra et al., 2022). Bias is only apparent from the perspective of gender in this study. Based on the data analysis results, the following table shows the results Figure 7 indicates that items 1 and 5 show evidence of item bias.

DIF class/group specification is: DIF=\$S5W1

PERSON CLASSES	SUMMARY DIF			BETWEEN-CLASS/GROUP ITEM		
	CHI-SQUARED	D.F.	PROB.	UNWTD MNSQ	ZSTD	Number Name
2	13.5070	1	.0002	14.0391	3.47	1 1
2	2.9500	1	.0859	2.9904	1.41	2 2
2	1.6277	1	.2020	1.6315	.85	3 3
2	.0388	1	.8438	.0695	-.78	4 4
2	9.7813	1	.0018	10.0707	2.93	5 5
2	.7918	1	.3736	.7889	.31	6 6
2	.1859	1	.6664	.1850	-.44	7 7
2	.3209	1	.5711	.3240	-.19	8 8

Figure 7. DIF Class Specification

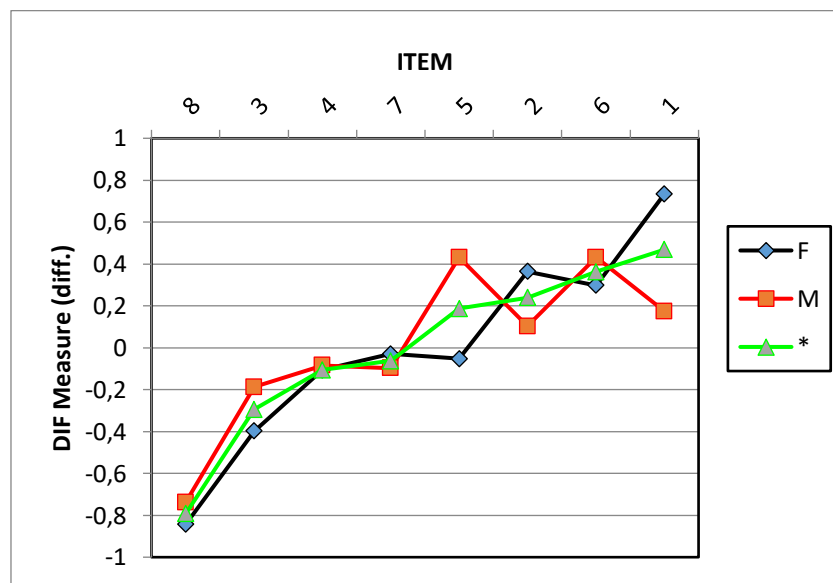


Figure 8. Graphic of Differential Item Function (DIF) based on Gender

The results show in Figure 8 that item 1 has a DIF Contrast value = +0.56 (p = 0.0002), which means there is a significant difference between the responses of male (M) and female (F). A positive value indicates that the item is relatively more difficult for male, so it can be said to be easier for female. Conversely, item 8 shows a DIF Contrast value of -0.48 with p = 0.0018, which means that the item is more difficult for female, so it is easier for male. Meanwhile, the other six items (items 2, 3, 4, 6, 7, and 8) did not show significant differences (p > 0.05), so it can be concluded that these items function fairly for both gender groups. The differences in the two items (items 1 and 5) indicate gender bias that needs to be considered in the instrument revision process. Item 1, which was easier for women, may be due to women's tendency to have perceptions or experiences that are more relevant to the context of the item related to the process of making batik Ghentongan. Conversely, item 5, which was easier for men, may be related to a style of thinking or empirical experience that is more in line with the content of the question, namely the concept of science in batik Ghentongan from Tanjung Bumi. Therefore, researchers are advised to review the wording, context, or stimulus of items that show significant differences so as not to cause gender bias.

Thus, in general, the instrument can be said to function well and fairly, as 75% of the items (6 out of 8) did not show significant DIF. Only two items require further evaluation to ensure equal functioning between respondent groups. DIF analysis based on gender shows that most items (items 2, 3, 4, 6, 7, and 8) do not show significant differences between male and female respondents, so the instrument can be categorized as invariant to gender. However, two items (items 1 and 5) need to be reviewed because they show significant differences in function. Revisions to these two items are expected to improve the overall validity and fairness of the measurement instrument.

Reliability Instrument

Reliability testing is a way to determine the level of consistency and stability of a questionnaire as a tool for measuring variables. A questionnaire is considered reliable if respondents' answers to questions remain stable and consistent over a long period of measurement (Azwar, 2016). The benchmark for reliability testing in using RASCH modelling is the Cronbach's alpha value (Sumintono & Widhiarso, 2014). The data obtained from summary statistics is information that shows the reliability measurement values, both from the respondent side (person reliability), item review (item reliability), and the interaction between participants and the statement (Muntazhimah et al., 2020).

Table 9. Person Reliability Instrument

	Reliability	Separation	Cronbach's Alpha
Person	0.34	0.72	0.41
Item	0.95	4.50	

The Table 9 shows two types of reliability: person reliability and item reliability, as well as separation values and Cronbach's Alpha. Person reliability value: 0.34 (low category), indicating that the instrument is not yet able to distinguish respondents' abilities well. Possible causes include students answering randomly, not fully understanding the questionnaire statements, or tending to give uniform answers (for example, choosing the middle option or agreeing with almost all statements). Person separation value: 0.72 (below the ideal number (2.0), meaning that the instrument is not yet able to group respondents into two or more ability groups). Cronbach's Alpha value: 0.41 (indicates low internal consistency, meaning that the data from respondents is unreliable). Item reliability value: 0.95 (very high category), this indicates that the items in the instrument are very consistent in measuring the intended construct. The results of the Rasch analysis reveal a discrepancy between item and person measurement in the student response questionnaire on *batik Ghentongan* Tanjung Bumi. The high item reliability indicates that the questionnaire items were well constructed and consistently assessed a single construct related to students' responses toward *batik Ghentongan* Tanjung Bumi.

This finding is reinforced by the high item separation index, which demonstrates that the instrument is capable of distinguishing statements across multiple levels of difficulty or agreement. However, the low person reliability and person separation values suggest that students' responses tended to be relatively homogeneous and inconsistent, limiting the instrument's ability to differentiate students based on their levels of response toward *batik Ghentongan* Tanjung Bumi. Therefore, the main limitation of the measurement lies not in the quality of the questionnaire items, but in the response patterns of the students, which may be influenced by factors such as limited familiarity with *batik Ghentongan* Tanjung Bumi, insufficient comprehension of the questionnaire statements, or a tendency to provide socially desirable responses. Item separation value: 4.50 (indicates that the instrument is able to distinguish items in more than four levels of difficulty, which means that the quality of the items is very good). The questionnaire is able to distinguish between the level of difficulty or intensity of response statements, ranging from statements that are easy to agree with to those that demand a stronger response. This confirms that the structure of the student response questionnaire is psychometrically adequate.

This study has limitations in that the respondents were limited to students in the Tanjung Bumi sub-district, so it is not yet possible to determine the response of students who are far from the Tanjung Bumi batik-making process. The study of *batik Ghentongan* Tanjung Bumi in learning activities has not been analysed. For future development, it is advisable that subsequent studies encompass not only the Tanjung Bumi Sub-district but also include pupils from across Bangkalan Regency. In addition, future research should examine the effectiveness of integrating *batik Ghentongan* into science learning through experimental or quasi-experimental designs to evaluate its impact on students' scientific literacy and cultural awareness. A more

diverse sample and mixed-method approaches would also provide deeper insights into students' perceptions and learning outcomes.

The implications of this research finding for the Sustainable Development Goals (SDGs) are highly relevant, particularly in supporting Goal 4 and 8. SDG 4 (Quality Education) is intricately connected to students' reactions to *batik Ghentongan* Tanjung Bumi via the use of local wisdom as a relevant and significant educational resource. *Batik Ghentongan* Tanjung Bumi is used in learning, students not only gain cognitive knowledge about culture and science (e.g. the dyeing process, motifs, and philosophical values), but also develop an appreciative attitude towards local culture. The positive response of students shows that science learning integrated with *batik Ghentongan* is able to increase learning engagement, conceptual understanding, and character building, which is in line with the objectives of SDG 4 in ensuring inclusive, quality education that is relevant to the context of students' lives. SDG 8 (Decent Work and Economic Growth) relates to students' response to *batik Ghentongan* Tanjung Bumi through strengthening awareness of the potential of creative economy based on local culture. Students' positive response reflects growing interest, appreciation, and concern for the batik-making profession as part of decent and sustainable work. Through early understanding of *batik Ghentongan*, students are encouraged to appreciate local products, preserve cultural heritage, and see creative economic opportunities that can support the sustainable economic growth of the Tanjung Bumi community, in line with the objectives of SDG 8.

CONCLUSION

The questionnaire revealed several items that needed improvement so that the students' responses would focus on measuring their responses to the integration of *batik Ghentongan* Tanjung Bumi in science learning. The findings of this study will hopefully act as a benchmark for assessing students' understanding of *batik Ghentongan*, enabling educators to implement integrated batik instruction more effectively. Based on the results of the validity and reliability analyses using the Rasch model, several questionnaire items require refinement to strengthen the measurement of students' responses to the integration of *batik Ghentongan* Tanjung Bumi in science learning. Although the high item reliability indicates that the items were generally well constructed and capable of measuring the intended construct, the low person reliability and person separation values suggest that the instrument has not yet optimally captured variability in students' responses. Therefore, item revisions are necessary to improve clarity, contextual relevance, and alignment with students' levels of understanding, thereby strengthening the construct validity and reliability of the instrument of the questionnaire. Further research that can be conducted is to create science learning tools using the context of *batik Ghentongan* Tanjung Bumi.

ACKNOWLEDGEMENT

We would like to acknowledge the Research and Community Service Institution Universitas Trunodjoyo Madura for the funding of the Research Grant, contract number 372/UN46.4.1/PT.01.03/RISMAN/2024.

AUTHOR CONTRIBUTIONS

Wiwini Puspita Hadi: Conceptualization, Visualization, Formal Analysis, Methodology, Writing – Original Draft; **Aida Fikriyah:** Validation, Resources, Writing – Review & Editing. All authors have read and approved the final version of this manuscript.

DECLARATION OF COMPETING INTEREST

The authors declare no known financial conflicts of interest or personal relationships that could have influenced the work reported in this manuscript.

DECLARATION OF ETHICS

The authors declare that the research and writing of this manuscript adhere to ethical standards of research and publication, in accordance with scientific principles, and are free from plagiarism.

DECLARATION OF ASSISTIVE TECHNOLOGIES IN THE WRITING PROCESS

The authors declare that generative artificial intelligence (Gen AI) and other AI-assisted tools were used prudently, not excessively, during the research and preparation of this manuscript. Specifically, ChatGPT was used for brainstorming ideas and DeepL was used for grammar and style correction. All AI-generated material

was reviewed and edited for accuracy, completeness, and compliance with ethical and scholarly standards. The authors accept full responsibility for the final content of the manuscript.

REFERENCES

- Afrida, W., Hadi, W.P., & Fikriyah, A. (2025). Pengembangan majalah IPA tema batik *gentongan* tanjung bumi madura terintegrasi literasi sains. *LENSA (Lentera Sains): Jurnal Pendidikan IPA*, **15**(1), 37–44. DOI: <https://doi.org/10.24929/lensa.v15i1.576>.
- Aisara, F., Nursaptini, N., & Widodo, A. (2020). Melestarikan kembali budaya lokal melalui kegiatan ekstrakurikuler untuk anak usia sekolah dasar. *Cakrawala Jurnal Penelitian Sosial*, **9**(2), 149–166. Retrieved from: <https://ejournal.uksw.edu/cakrawala/article/view/4411>.
- Azwar, S. (2016). *Reliabilitas dan validitas*. Yogyakarta: Pustaka Pelajar.
- Bawono, Y., Edi, F.R.S., & Salam, M.I.A. (2023). Di balik keindahan batik *gentongan* Madura. *Dinamika Kerajinan Dan Batik: Majalah Ilmiah*, **40**(2), 203–212. DOI: <https://doi.org/10.22322/dkb.v40i2.7086.g6092>.
- Coffie, I.S., Frempong, B. B., Taylor, I., & Appiah, E. (2025). What matters to students in the teaching and learning of science: Perspectives of junior high school students from Ghana. *Research in Science Education*, inpress. DOI: <https://doi.org/10.1007/s11165-025-10290-w>.
- Erman, E., Pare, B., Susiyawati, E., Martini, M., & Subekti, H. (2022). Using scaffolding set to help student addressing socio-scientific issues in biochemistry classes. *International Journal of Instruction*, **15**(4), 871–888. DOI: <https://doi.org/10.29333/iji.2022.15447a>.
- Giamellaro, M., Buxton, C., Beaudry, M.C., Taylor, J., Philippe, J., Beudet, A., & Heures, K.L. (2025). The landscape of research on contextualized science learning: A bibliometric network review. *Science Education*, **109**(3), 851–875. DOI: <https://doi.org/10.1002/sci.21937>.
- Hadi, W.P. & Fikriyah, A. (2024). The ethnoscience study of “batik *gentongan* tanjung bumi madura” and its integration in science learning process. *The 12th International Conference on Educational Research and Innovation*, 921–931. Retrieved from: https://iceri.uny.ac.id/sites/iceri.uny.ac.id/files/download-file/Prosiding%20ICERI%202024_0.pdf.
- Hadi, W.P., Fikriyah, A., & Mu'aziyah, S.E.S. (2025). Analisis respons siswa smp terhadap batik *ghentongan* tanjung bumi madura dalam pembelajaran IPA. *Prosiding Semnasfip 2025*, **2**(2), 4720–4729. Retrieved from: <https://jurnal.umj.ac.id/index.php/SEMNASFIP/article/view/30027>.
- Ilfiandra, Nadhirah, N.A., Suryana, D., & binti Ahmad, A. (2022). Development and validation peaceful classroom scale: Rasch model analysis. *International Journal of Instruction*, **15**(4), 497–514. DOI: <https://doi.org/10.29333/iji.2022.15427a>.
- Kusumastuti, A., Atika, Achmadi, T.A., Phusavat, K., & Hidayanto, A.N. (2023). Assessment of producer's perspective on the production of environmentally friendly fashion products: A case study in Indonesian natural dyes batik craftsmen. *Environmental Science and Pollution Research*, **30**(60), 124767–124779. DOI: <https://doi.org/10.1007/s11356-022-23330-z>.
- Martareza, A.W., Hadi, W.P., & Fikriyah, A. (2025). Analisis kebutuhan pengembangan media ajar etno-vlog batik *gentongan* tanjung bumi madura pada siswa SMP. *Konstruktivisme*, **17**(1), 134–142. DOI: <https://doi.org/10.35457/konstruk.v17i1.3972>.
- Mishra, N.R. (2023). Constructivist approach to learning: An analysis of pedagogical models of social constructivist learning theory. *Journal of Research and Development*, **6**(1), 22-29. DOI: <https://doi.org/10.3126/jrdn.v6i01.55227>.
- Muñoz, A., Isaac, L., & Cuello, C. (2025). Impact of interactive science workshops participation on primary school children's emotions and attitudes towards science. *International Journal of Science and Mathematics Education*, **23**(7), 2689–2706. DOI: <https://doi.org/10.1007/s10763-024-10539-2>.
- Muntazhimah, M., Putri, S., & Khusna, H. (2020). Rasch model untuk memvalidasi instrumen resiliensi matematis mahasiswa calon guru matematika. *JKPM (Jurnal Kajian Pendidikan Matematika)*, **6**(1), 65-74. DOI: <https://doi.org/10.30998/jkpm.v6i1.8144>.
- Muthia, A.E., Nugraha, M.G., & Chandra, A.F. (2021). Pengembangan instrumen sustainability awareness dalam materi getaran harmonik sederhana untuk siswa SMA. *Wahana Pendidikan Fisika*, **6**(2), 204–210. DOI: <https://doi.org/10.17509/wapfi.v6i2.31538>.
- Nurhudaya, Taufik, A., Yudha, E.S., & Suryana, D. (2019). The Raven's advanced progressive matrices in education assessment with a Rasch analysis. *Universal Journal of Educational Research*, **7**(9), 1996–2002. DOI: <https://doi.org/10.13189/ujer.2019.070921>.

- Picardal, M.T. & Sanchez, J.M.P. (2022). Effectiveness of contextualization in science instruction to enhance science literacy in the Philippines: A meta-analysis. *International Journal of Learning, Teaching, and Educational Research*, **21**(1), 140–156. DOI: <https://doi.org/10.26803/ijlter.21.1.9>.
- Podschuweit, S. & Bernholt, S. (2018). Composition-effects of context-based learning opportunities on students' understanding of energy. *Research in Science Education*, **48**, 717–752. DOI: <https://doi.org/10.1007/s11165-016-9585-z>.
- Prasetyo, P., Sulaiman, A.I., & Prastyanti, S. (2022). Educational communication in learning Batik as preservation of local wisdom products for the young generation. *Technium Education and Humanities*, **3**(1), 1–15. DOI: <https://doi.org/https://doi.org/10.47577/teh.v3i1.7615>.
- Purwastuti, L.A. (2022). Batik as a craft teaching-learning medium to preserve values of local wisdom in elementary schools in Bantul, Indonesia. *Humanika*, **22**(2), 91–98. DOI: <https://doi.org/10.21831/hum.v22i2.53888>.
- Puspitasari, E., Krismanto, W., Tuken, R., & Pasinggi, Y.S. (2024). Using video animation in contextual teaching and learning to enhance student engagement in science learning. *Quantum: Jurnal Inovasi Pendidikan Sains*, **15**(1), 13–27. DOI: <https://dx.doi.org/10.20527/quantum.v15i1.17134>.
- Rai, S. (2025). Reimagining learning: The relevance and application of constructivist theory in 21st century classrooms. *International Journal for Multidisciplinary Research*, **7**(4), IJFMR250450004. DOI: <https://doi.org/10.36948/ijfmr.2025.v07i04.50004>.
- Sari, I.P. (2022). Batik tulis tanjung bumi: Ciri khas dan potensi. *Dinamika Kerajinan dan Batik: Majalah Ilmiah*, **39**(2), 159-168. DOI: <https://doi.org/10.22322/dkb.v39i2.7469>.
- Sari, I.P. & Miftah, Z. (2020). Exploratory research on the myth of batik gentongan in Tanjung Bumi. *Proceedings of the 1st International Conference on Folklore, Language, Education and Exhibition*, 36–39. DOI: <https://doi.org/10.2991/assehr.k.201230.007>.
- Sulindra, E., Cendra, A.N., & Hartani, T. (2024). Utilization of instructional technology in english language teaching (ELT) based on constructivism: A literature review. *English Education and Literature Journal*, **4**(02), 141–153. DOI: <https://doi.org/10.53863/ejou.v4i02.1150>.
- Suliyanto, S., Novandari, W., & Setyawati, S.M. (2016). Persepsi generasi muda terhadap profesi pengrajin batik tulis di Purbalingga. *Jurnal Ekonomi dan Bisnis*, **18**(1), 135-144. DOI: <https://doi.org/10.24914/jeb.v18i1.275>.
- Sumintono, B. & Widhiarso, W. (2014). *Aplikasi Model Rasch untuk penelitian ilmu-ilmu sosial*. Bandung: Penerbit Trim Komunikata.
- Usman, M., Hermansyah, S., Yusrianti, Said, S., Febrianti, D., & Rizal, A. (2024). holistic learning models in remote areas: Enhancing student motivation through local wisdom, parental collaboration, teacher capacity, and government support. *Journal of Ecohumanism*, **3**(7), 4548–4561. DOI: <https://doi.org/10.62754/joe.v3i7.4565>.
- Wibisono, S. (2016). Aplikasi Model Rasch untuk validasi instrumen pengukuran fundamentalisme agama bagi responden muslim. *Jurnal Pengukuran Psikologi dan Pendidikan Indonesia*, **5**(1), 9239. DOI: <https://doi.org/10.15408/jp3i.v5i1.9239>.
- Wulandari, D., Yuniarti, V.D., & Wahyuningsih, Y. (2023). Analisis ketertarikan siswa sekolah dasar terhadap kebudayaan Indonesia. *Journal on Education*, **5**(2), 2376–2382. DOI: <https://doi.org/10.31004/joe.v5i2.894>.
- Zuleni, E. & Marsidin, S. (2025). Decoding the landscape of contextual science education: A bibliometric approach to research trends and future directions. *International Journal of Innovative Research and Scientific Studies*, **8**(4), 1383–1397. DOI: <https://doi.org/10.53894/ijirss.v8i4.8081>.