

Developing a Science Process Skills Assessment on Cell and Microscope: A Rasch Analysis

Della Pusvita Anggraeni, Risa Dwita Hardianti*

Department of Science Education, Universitas Negeri Semarang, Semarang, Indonesia

*e-mail: risadwita@mail.unnes.ac.id

Received: May 20, 2026. Accepted: June 18, 2026. Published: August 30, 2026

Abstract: Science process skills (SPS) are essential competencies in science education; however, various studies indicate that students' SPS achievement remains relatively low. Furthermore, existing assessment instruments do not fully align with the learning outcome indicators. This study aims to develop a valid and reliable SPS test instrument for cell and microscope topics through Rasch model analysis. The research method employs a Research and Development (R&D) approach using the ADDIE model, which includes the Analyze, Design, Develop, Implement, and Evaluate stages. The developed instrument consists of 25 multiple-choice items covering 6 SPS indicators: observing, questioning and predicting, planning and conducting investigations, processing and analyzing data, evaluating and reflecting, and communicating results. Content validation was conducted by 5 validators, yielding an average Aiken's V of 1.00, indicating excellent alignment between the items and the measured indicators. Rasch analysis using the Winsteps applications showed that all items met the fit criteria based on outfit MNSQ values, with difficulty levels ranging from very easy to very difficult. The instrument demonstrated high reliability ($\alpha = 0.97$) and adequate item separation, indicating strong consistency and discriminative power. Implementation involving 32 seventh-grade students revealed that most students were categorized as having moderate SPS level, suggesting that their SPS had developed but was not yet optimal across all indicators. These results confirm the instrument's validity and reliability in assessing junior high school students' SPS. This study contributes a Rasch-based assessment instrument that can objectively map students' SPS abilities and support the evaluation of science learning.

Keywords: Rasch Analysis; Science Process Skills; Test Instrument; Winstep.

Introduction

Students need to possess strong skills to overcome future challenges. One such skill is the ability to support the scientific learning process [1]. Scientific learning is inseparable from scientific and systematic inquiry. The set of skills used for conducting scientific investigations is known as the 'Science Process Skills' (SPS). [2]. SPS enables students to learn the concepts that are the focus of science education while also developing basic science skills, a scientific mindset, and critical thinking skills [3]. SPS are the skills one must possess to conduct scientific investigations in a more systematic manner [4]. Decision Head of the Agency for Standards, Curriculum, and Educational Assessment, Ministry of Basic and Secondary Education No. 046/H/KR/2025, states that the Merdeka Curriculum, particularly in Phase D of the Natural Science learning, emphasizes conceptual understanding, process skills, and learning outcomes that integrate attitudes, knowledge, and skills [5].

This study focuses on the process skills element, which in the Merdeka Curriculum includes activities such as observing, asking questions and making predictions, planning and conducting investigations, processing, analyzing data and information, evaluating and reflecting, and communicating results. SPS are competencies that need to be developed in education and are crucial for every student as a foundation for acquiring new knowledge and expanding existing knowledge [6], [7]. Learning activities that

emphasize SPS also promote innovation and creative thinking, as they involve students' skills in carrying out scientific projects [8]. Therefore, SPS is a key indicator of success in science education.

Recent research indicates that junior high school students continue to face various challenges related to SPS. A literature review reveals that the level of SPS remains relatively low [9], [10]. This is because students are not sufficiently engaged directly in the investigative process; rather, they tend to be passive when performing laboratory tasks [11], [12]. International and national studies consistently show that students have difficulty with complex aspects of SPS, such as experimental design, data interpretation, problem formulation, and communicating results. The SPS indicator regarding hypothesis formulation is often overlooked [13]. Consequently, the assessment process requires instruments capable of comprehensively measuring SPS, including its psychomotor aspects. However, many teachers have not yet developed SPS-oriented assessments, and those that do have outcomes. Therefore, an appropriate instrument is needed to accurately measure students' SPS.

Interview results with science teachers indicate that the available assessment instruments are not fully aligned with learning or meet quality assessment criteria. This issue highlights a gap between the assessment instruments and the SPS achievement indicators that should be measured [14]. Research on the development of SPS tests in science education has been extensively conducted. It demonstrates

How to Cite:

D. P. Anggraeni and R. D. Hardianti, "Developing a Science Process Skills Assessment on Cell and Microscope: A Rasch Analysis", *J. Pijar.MIPA*, vol. 21, no. 4, pp. 788–795, Aug. 2026. <https://doi.org/10.29303/jpm.v21i4.12166>

that SPS tests can be used as both pre-tests and post-tests to measure student proficiency [15], [16]. The development of the SPS test has also been applied to various topics, such as vibrations, waves, and light; elements, compounds, and mixtures; chemical equilibrium; and ecosystems. The results of these studies indicate that the developed tests effectively measure students' SPS [17], [18], [19], [20]. However, no researchers have yet developed an SPS test on cells and microscopes.

Developing an assessment instrument for SPS on cells and microscopes is a crucial step toward improving the quality of learning assessment. A good instrument must be valid and reliable, and must measure what it is intended to measure [21]. The Rasch model is an effective analytical tool for evaluating the suitability of a measurement instrument. The Rasch model can measure reliability indices, conduct item-level analysis, evaluate respondent reliability and instrument dimensionality, and identify potential item bias [22]. Additionally, the Rasch model can determine student ability, such that each time a student answers an item correctly, their ability level increases continuously in line with their existing proficiency [23], [24]. Therefore, the Rasch model is considered appropriate for objectively and comprehensively evaluating the quality of SPS assessment instruments.

This research focuses on developing an SPS test for the cell and the microscope, which will be analyzed using the Rasch model. This research aims to provide an overview of the validation results and Rasch model analysis, including item validity, reliability, fit, and difficulty levels. Additionally, the study aims to determine students' SPS levels on cell and microscope topics and serve as a valid and reliable measurement tool for assessing junior high school students' SPS.

Research Methods

The study employs a Research and Development (R&D) approach using the ADDIE model, which consists of five stages: Analyze, Design, Develop, Implement, and Evaluate. The ADDIE model was chosen because it provides a systematic process for developing and testing the feasibility of an educational product, including valid and reliable testing instruments [25]. It also facilitates continuous evaluation and revision at each stage to improve the quality of the resulting instrument. The research was conducted at SMP Negeri 2 Semarang during the second semester of the 2025/2026 academic year. The population consisted of all seventh-grade students, totalling 288, who had received instruction on cells and microscopes. The subjects of this study were seventh-grade students who had already received instruction on cells and microscopes. Data were collected using non-test techniques through interviews and test techniques through multiple-choice questions. Figure 1 describes the development process and activities based on the ADDIE model.

In the analyzed stage, observations and interviews were conducted with science teachers to identify issues related to the SPS assessment instrument, particularly regarding the cell and the microscope. The design stage involved developing an instrument framework based on SPS indicators, which included skills in observing, questioning and predicting; planning and conducting investigations;

processing, analyzing data and information; evaluating and reflecting; as well as communicating results [26]. The development stage involved transforming the SPS instrument framework into 25 multiple-choice items. The specifications of the SPS instrument are presented in Table 1. The instrument was then validated by 5 validators using Aiken's V index and revised based on their suggestions. Next, a small-scale trial test and an item characteristic analysis were conducted using the Rasch model. This was based on a table compiled by Krijcie and Morgan (1970) and involved a sample size of 184 students from the entire population.

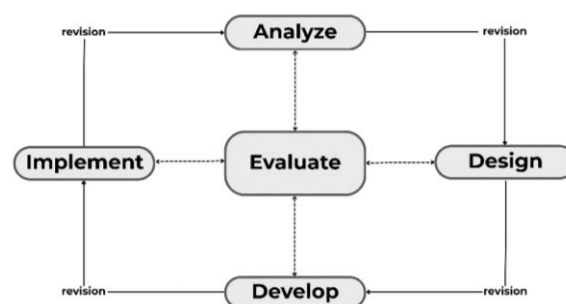


Figure 1. The ADDIE Model by Branch (2009)

Table 1. Specifications of The SPS Instrument for Cell and Microscope Material

Science Process Skills Indicators	Item Number	Psychomotor Level
Observing	1, 2, 3, 4, 5, 6	P1
Questioning and predicting	7, 8, 9, 10, 11	P2, P3
Planning and conducting investigations	12, 13, 14, 15	P2, P3, P4
Processing, analyzing data, and information	16, 17, 18, 19, 20	P4, P5
Evaluating and reflecting	21, 22	P5
Communicating results	23, 24, 25	P4, P5

The implementation stage involved analyzing the instrument using the Rasch model via the Winstep application. The analysis included reliability, item fit, and item measure, Wright Map, and person measure. A reliability analysis was conducted to assess the instrument's consistency. The item fit analysis examined how well each item conformed to the Rasch model, while item measure analysis identified the difficulty level of each on the logit scale. The Wright Map was analyzed to describe the distribution of item difficulty levels on the logit score.

Following the instrument's validation and reliability assessment through trial testing with 184 students, the students' SPS was measured in a class of 32 students, selected using cluster random sampling. This technique was chosen because the population had naturally formed into groups. Selecting one class at random gave each class an equal chance of being selected for the research sample, thereby reducing sampling bias and producing a sample representative of the study population. The data obtained were analyzed using the Rasch model, particularly through person measure analysis, to determine the profile of students' levels of SPS. Subsequently, the evaluation stage was carried

out comprehensively and continuously at every stage of development, from analysis, design, and development through implementation, to improve the quality of the research and ensure that each stage met the criteria before proceeding to the next.

Results and Discussion

Content Validity of the Instrument

The instrument that had been developed was validated by 5 validators, comprising 3 science teachers and 2 expert lecturers. Validation was conducted in two stages using Aiken's V index to assess the alignment between test items and SPS indicators. In the first validation stage, most test items aligned with the indicator outcomes; however, several items required revisions based on the validators' suggestions. After the instrument was revised, a second validation was conducted, yielding an average Aiken's V score of 1.00. All assessment indicators fall into the high category ($0.8 < V \leq 1.0$) [27]. This value indicates that each test item has a very strong relevance to the SPS indicators being measured. The closer the Aiken index value is to 1.00, the better the level of item fit with the established indicators [28]. Thus, the developed instrument has met content validity and is suitable for use in the empirical trial stage.

Item Fit Analysis of the Instrument

Student response scores from the trial test were analyzed using the Rasch model with the support of the Winstep application to evaluate the instrument's empirical quality. In Rasch analysis, item validity is examined through the Output Table: Item Fit Order menu. Item fit describes the extent to which a sample's response pattern to a particular item aligns with the responses of other individuals to different items. In the Rasch model, item fit is evaluated using several statistics, including infit and outfit values, Mean Squared Error (MNSQ), Z-standardized (ZSTD), and Point Measure Correlation (PTMEA Corr). These indices provide complementary information about how well observed responses conform to the Rasch model's expectations [29]. In this study, item fit was evaluated primarily using Outfit MNSQ values. This approach is considered acceptable because the MNSQ statistic is regarded as the most informative and stable for assessing the magnitude of the misfit in a practical sense [30]. Unlike ZSTD, which is sensitive to sample size and can indicate significant misfit when the magnitude is trivial, MNSQ reflects the degree to which item responses deviate from model expectations [31], [32], [33].

According to Linarche's (2006) criteria, an item is considered to fit if its MNSQ outfit value falls within the range $0.5 < \text{MNSQ outfit} < 1.5$. The results show that the MNSQ outfit values for all test items range from 0.76 to 1.29, indicating that all items fall into the fit category. A summary of the item fit analysis results is presented in Table 2. As no items showed misfit, all test items were able to consistently represent the SPS construct [34], [35]. Additionally, the relatively stable distribution of MNSQ outfit values suggests that the items in the instrument functioned consistently in measuring students' SPS abilities. These findings demonstrate that students' response patterns aligned well

with the Rasch measurement model's expectations. In line with previous research, well-designed science assessment instruments with strong content validity tend to produce response patterns consistent with Rasch model expectations, indicating good construct representation [36], [37].

Table 2. Summary of Item Characteristics

Item	Outfit MNSQ	Description
1	0.90	Fit
2	0.92	Fit
3	1.16	Fit
4	0.98	Fit
5	1.05	Fit
6	1.14	Fit
7	1.34	Fit
8	0.85	Fit
9	1.06	Fit
10	1.06	Fit
11	0.95	Fit
12	1.15	Fit
13	0.90	Fit
14	0.99	Fit
15	1.10	Fit
16	0.95	Fit
17	1.13	Fit
18	0.88	Fit
19	0.89	Fit
20	0.75	Fit
21	0.99	Fit
22	0.88	Fit
23	1.16	Fit
24	0.93	Fit
25	0.99	Fit

Instrument Reliability

Reliability in the Rasch model was analyzed using the Output Summary Statistic menu. The reliability of a test indicates that a test instrument can provide the same or consistent information when administered repeatedly [38]. The Cronbach's Alpha coefficient ranges from 0.00 to 1.00. If $r \geq 0.70$, the instrument is considered highly reliable (=reliable), while if $r \leq 0.70$, it is not yet highly reliable (=unreliable). A reliability value of 0.97 is shown in Figure 3, as indicated by the analysis results. The value indicates that the instrument has high consistency and is capable of producing a stable level of difficulty [39], [40].

SUMMARY OF 25 MEASURED Item									
	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFINIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD	
MEAN	107.0	184.0	.00	.18	.99	-.1	1.00	.1	
S.D.	36.1	.0	1.06	.03	.09	1.5	.13	1.3	
MAX.	164.0	184.0	1.61	.24	1.29	4.8	1.34	3.7	
MIN.	50.0	184.0	-1.90	.16	.83	-3.2	.76	-2.8	

REAL RMSE	.18	TRUE SD	1.05	SEPARATION	5.72	Item	RELIABILITY	.97	
MODEL RMSE	.18	TRUE SD	1.05	SEPARATION	5.80	Item	RELIABILITY	.97	
S.E. OF Item MEAN = .22									

Figure 3. Reliability Item

The quality of the instrument was also assessed using item separation values, which indicate the discriminative power of test items. In the Rasch model, item discrimination refers to an item's ability to distinguish between individuals

with high and low abilities [29]. Higher item separation values indicate that the instrument can more accurately and systematically classify item difficulty levels [29]. In the Rasch model, high reliability and separation value confirm that the measurement results are relatively free from the influence of a specific student sample and are not due to chance. This indicates that the instrument developed not only has content validity but has also been empirically proven to be reliable. Therefore, the instrument can be appropriately utilized as an objective tool measuring SPS.

Item Measure Analysis of the Instrument

The instrument's quality was further examined by difficulty level using the Output Table: Item Measure menu. The measure value represents a logit score that reflects the level of item difficulty, arranged from most to least difficult. In the Rasch model, the analysis of item difficulty using item measure values is carried out by referring to the mean and standard deviation of the item measures [29]. The theoretical mean item value in Rasch modeling is 0.00 with a standard deviation of 1.06. Using these criteria, items are classified into five categories, as shown in Table 3.

Table 3. Item Difficulty Criteria

Criteria	Description
$b > 1.06$ logits	Very difficult
$0.00 < b < 1.06$ logits	Difficult
$b = 0.00$ logits	Moderate
$-1.06 < b < 0.00$ logits	Easy
$b < -1.06$ logits	Very easy

The results indicate that the instrument is of good quality, with a balanced distribution of difficulty levels for measuring SPS, as shown in Table 4.

Table 4. Summary of Item Difficulty Levels

Item	Item Measure (b)	Description
1	-0.24	Easy
2	-0.80	Easy
3	1.58	Very difficult
4	-0.41	Easy
5	0.64	Difficult
6	0.48	Difficult
7	0.10	Difficult
8	-1.00	Very easy
9	-1.90	Very easy
10	-0.49	Easy
11	-1.31	Very easy
12	0.92	Difficult
13	0.02	Moderate
14	-1.19	Very easy
15	1.61	Very difficult
16	1.37	Very difficult
17	0.87	Difficult
18	0.59	Difficult
19	-1.84	Very easy
20	-0.05	Easy
21	0.36	Difficult
22	1.22	Very difficult
23	-1.78	Very easy
24	0.00	Moderate
25	1.25	Very difficult

These findings are reinforced by the Wright Map: Person Item in Figure 2, which illustrates the distribution of item difficulty levels arranged sequentially from highest to lowest (top to bottom), indicating a relatively balanced distribution of items. The Wright Map also serves to compare theoretical predictions with empirical difficulty, thereby providing evidence of construct validity and ensuring the instrument measures in accordance with theory [41]. A balanced distribution of item difficulty levels is essential because it allows the instrument to accurately assess students' abilities at various proficiency levels [42]. Variation in item difficulty levels is also essential for accurately assessing students' abilities through Rasch scaling [37]. Overall, the characteristics of the item difficulty levels meet the criteria for an effective instrument [35].

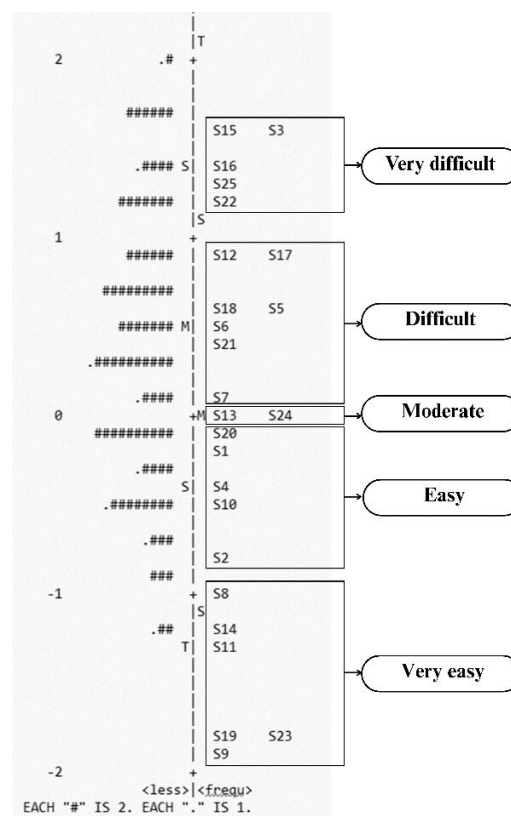


Figure 2. Wright Map Item Measure

To provide a clearer overview of the distribution of item difficulty, Table 5 presents the classification of instrument items by difficulty category based on the Rasch logit measure.

Table 5. Classification of Sample Items by Difficulty Level

No.	Question	Category
9	During a plant cell observation lab, a student wanted to examine the structure of the stomata on a <i>Rhoeo discolor</i> leaf. The student applied the same technique he used when making a thin section of a red onion—that is, removing the outermost layer of skin or membrane. What type of incision is most appropriate to use, and why?	Very difficult

- A transverse section, so that all tissues from the upper to the lower epidermis are visible vertically
- A longitudinal (paradermal) section, so that the shape of the stomata can be clearly seen
- An oblique section, to reveal more chloroplasts in neighbouring cells
- A perpendicular section, so that the leaf vein structure and transport tissues (xylem/phloem) are clearly visible

12 Follow these steps for using an electric microscope: Difficult

- Use the micrometer (fine adjustment knob) to obtain a sharp and clear image.
- Place the microscope on the table by holding the microscope arm.
- Clean the microscope after use.
- Turn the macrometer (coarse adjustment knob) to bring the objective lens closer to the specimen.
- Attach the eyepiece to the correct tube of the microscope.
- Place the specimen on the stage and secure it with the specimen clamp.
- Turn on the light and adjust the illumination and diaphragm to allow light to enter.
- Rotate the revolver to select the objective lens with the lowest magnification.

The correct sequence for using a microscope is ...

- 2-5-8-7-6-4-1-3
- 2-7-5-8-6-4-1-3
- 5-2-7-8-4-6-1-3
- 2-5-7-8-6-4-1-3

24 Examine the data from the microscopic observation of cells in the following table. Based on this data, the most accurate conclusion is

Number	Observed item	Cell A	Cell B
1.	Cell wall	Present	Absent
2.	Chloroplast	Present	Absent
3.	Cell shape	Fixed	Not fixed
4.	Centriole	Absent	Present

- Cells A and B are plant cells because they have different shapes.
- Cell A is a plant cell and Cell B is an animal cell.
- Cell A is an animal cell and Cell B is a plant cell.

- Cell A and Cell B are plant cells because they have irregular shapes.

2 Look at the microscope images below! What is the difference between the two types of microscope shown in images (A) and (B)? Easy

Light Microscope



- Microscope (A) has two eyepieces, while microscope (B) has only one.
- Microscope (A) has a higher magnification than microscope (B).
- Microscope (A) has one eyepiece, while microscope (B) has two.
- Microscope (A) has two observation tubes, while microscope (B) has one.

9 Two students conducted a laboratory experiment observing onion cells as follows: Very easy

- Student A peeled a very thin layer of onion.
- Student B peeled a fairly thick layer of the onion.

Based on the students' observations, the most accurate prediction regarding the results of their experiment is ...

- Student B's observations are clearer because the layer is thicker.
- Student A obtained clearer cell boundaries because the peeling was done correctly.
- Both students observed the same object because the object being observed was the same.
- Student B observed a greater number of cells, resulting in better observations.

Implementation of the Instrument and Categorization of SPS

The SPS instrument, having met the criteria for content validity and empirical quality, was subsequently implemented to describe students' SPS. As this study focused on instrument development, the primary objective was to produce an SPS instrument that is valid, reliable and possesses good item characteristics. Therefore, a larger sample involving five classes was utilized during the

instrument trial stage. After the instrument satisfied the established quality criteria, its implementation was conducted in a single measurement class consisting of 32 students to obtain an overview of students' SPS. The selected class was determined through cluster random sampling, in which each class in the population had an equal chance of being selected. Consequently, the selected class was considered representative of the population within the scope of the study.

The implementation stage was intended to demonstrate the application of the developed instrument rather than to generalise findings broadly to the entire population. The measurement results represent students' SPS as analyzed through person measure scores, which serve as indicators of individual ability. Based on the Rasch analysis results, students' SPS can be categorized using a logit measure score with a reference mean of 0.442 logits and a standard deviation of 1.04 logits [29].

Table 6. Categorization of Science Process Skills

Criteria	Logit Measure Value (<i>b</i>)	Description
$b < (b_{\text{mean}} - b_{\text{stddev}})$	$b < -0.62$	Low
$(b_{\text{mean}} - b_{\text{stddev}}) < b < (b_{\text{mean}} + b_{\text{stddev}})$	$-0.62 < b < 1.46$	Moderate
$b > (b_{\text{mean}} + b_{\text{stddev}})$	$b > 1.6$	High

A Rasch analysis of students' SPS indicates that students' abilities can be classified into three categories based on their measure scores: low, moderate, and high. Table 6 outlines the categorization, with the category boundaries determined based on the mean and standard deviation of logit scores. According to the analysis results, four students had measured scores below -0.62 logits, placing them in the low category. This group demonstrated limited SPS and struggled to answer questions at an appropriate level of difficulty. Twenty-four students had scores between -0.62 and 1.46 logits, placing them in the moderate category. This group showed developing SPS, though not yet optimal across all indicators. Four students had scores above 1.46 logits, placing them in higher-difficulty items and indicating stronger mastery of SPS indicators. This group demonstrated consistent ability and outperformed the other groups.

As Figure 4 shows, the majority of students fall into the moderate category. This suggests that students' SPS have generally developed well, though not yet optimally, across all SPS indicators. Furthermore, categorizing students' abilities based on logit scores is supported by the instrument's high quality. This finding is consistent with previous research emphasizing that high reliability and item separation support the instrument's ability to detect differences in student ability. Instruments that can produce categorization based on the mean and standard deviation, assessing science process skills [39], [40], [43]. Figure 4 shows the person measure, which is based on the results of the Rasch analysis.

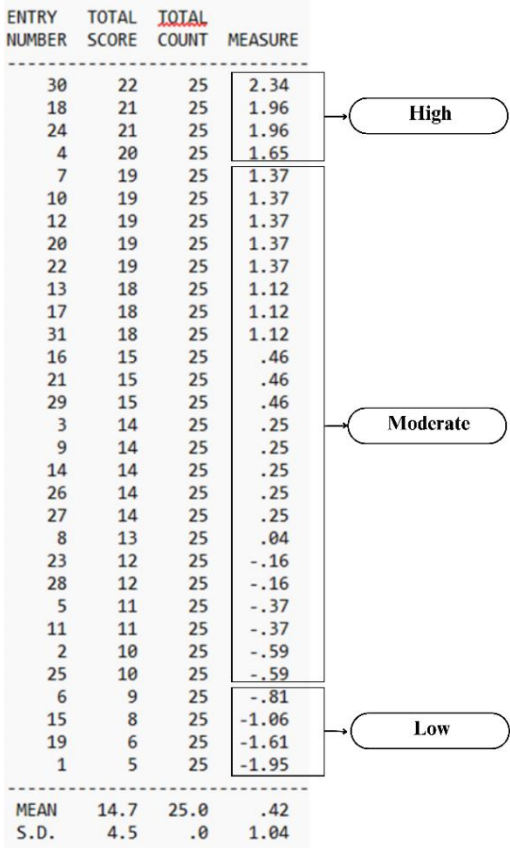


Figure 4. Person Measure SPS Test

Conclusion

This study successfully developed a science process skills (SPS) test instrument on cell and microscope topics using the ADDIE development model and Rasch analysis. The instrument consists of 25 multiple-choice items covering 6 indicators. Content validation by 5 validators yielded an average Aiken's V value of 1.00, indicating a high level of content validity. Rasch analysis showed that all items met the fit criteria, with outfit MNSQ values ranging from 0.75 to 1.34. This indicates that all items consistently measured the intended construct. The instrument demonstrates excellent reliability ($\alpha = 0.97$) and adequate item discrimination, confirming high consistency and strong discriminative power. Furthermore, item-item analysis revealed a balanced distribution of difficulty levels, ranging from very easy to very difficult, indicating that this instrument can measure students' SPS across a wide range of ability levels. Wright's map analysis further supports the instrument's construct validity by demonstrating alignment between theoretical expectations and empirical data. Applying the instrument to a group of 32 seventh-grade students revealed that most students were categorized at the moderate SPS level. This finding suggests that students' SPS have developed, but not yet optimally across all indicators. This study's main contribution is the development of a Rasch-based SPS assessment instrument that objectively measures psychomotor-oriented SPS and accurately maps variations in student ability. While the instrument demonstrated strong validity, reliability, and item characteristics, it was only implemented in one class with a small sample size. Therefore, further research involving larger, more diverse

samples is recommended to strengthen the generalizability of the findings and support the development of broader science assessment instruments integrated with project-based learning and guided inquiry approaches.

Author's Contribution

D. P. Anggraeni: contributed to the research, product development, data collection, data analysis, and manuscript writing; R.D. Hardianti: contributed to supervising, guiding, and revising the research.

Acknowledgements

Sincere gratitude is expressed to the Science Education Study Program, Faculty of Mathematics and Natural Science, Semarang State University. Appreciation is also given to the supervisor for the valuable guidance and support throughout the research process. Special thanks are dedicated to family members for their encouragement and motivation, as well as to friends and all related parties who contributed to the completion of this article. Hopefully, this article will be beneficial to readers and broaden their horizons.

References

- [1] F. Widiati, I. Ibrohim, and A. Tenzer, "Improvement of science process skills and students cognitive learning outcomes through implementation of guided inquiry learning model on biology subject," *AIP Conf. Proc.*, vol. 2330, 2021, doi: 10.1063/5.0043410.
- [2] I. W. Suja, *Keterampilan Proses Sains dan Instrumen*. Depok: PT RajaGrafindo Persada, 2023. doi: <https://journal.global.ac.id/index.php/sisfotek/article/view/16008/0>.
- [3] I. A. Wulandari, M. B. Mu'min, and M. G. Firdaus, "Peningkatan Keterampilan Berpikir Kritis (KBKr) Melalui Pembelajaran Biologi Berbasis Keterampilan Proses Sains," *J. BIOEDUIN Progr. Stud. Pendidik. Biol.*, vol. 11, no. 1, pp. 63–69, 2021, doi: 10.15575/bioeduin.v11i1.12081.
- [4] G. D. Masruhah, R. Rusdianto, and S. Wahyuni, "Pengembangan E-LKPD Berbasis Inkuiri Terbimbing untuk Meningkatkan Keterampilan Proses Sains Siswa SMP," *SAP (Susunan Artik. Pendidikan)*, vol. 7, no. 1, 2022, doi: 10.30998/sap.v7i1.12935.
- [5] M. A. Mahdiannur, "Analisis Keterampilan Praktik Ilmiah Siswa dalam Pembelajaran IPA SMP berciri Inkuiri Induktif," *J. Pendidik. MIPA*, 2022, [Online]. Available: <https://www.ejournal.tsb.ac.id/index.php/jpm/article/view/705>
- [6] Maison *et al.*, "Science process skill in science program higher education," *Univers. J. Educ. Res.*, vol. 8, no. 2, pp. 652–661, 2020, doi: 10.13189/ujer.2020.080238.
- [7] M. Ekici and M. Erdem, "Developing Science Process Skills through Mobile Scientific Inquiry," *Think. Ski. Creat.*, vol. 36, no. December 2019, p. 100658, 2020, doi: 10.1016/j.tsc.2020.100658.
- [8] M. S. Fredagsvik, "Student Approaches to Creative Processes When Participating in an Open-Ended Project in Science," *Int. J. Sci. Educ.*, vol. 44, no. 10, p. 18, 2022, [Online]. Available: <https://eric.ed.gov/?q=Student+Approaches+to+Creati>
- ve+Processes+When+Participating+in+an+Open-Ended+Project+in+Science&pr=on&id=EJ1367208
- [9] I. A. Libata, M. N. Ali, and H. N. Ismail, "Fostering science process skills through constructivist-based module among form two students of different cognitive levels," *Contemp. Math. Sci. Educ.*, vol. 4, no. 1, p. ep23005, 2022, doi: 10.30935/conmaths/12747.
- [10] O. B. Mirian, "an Assessment of Science Process Skills Among Science Students in River State, Nigeria," *J. Adv. Res. Appl. Sci. (ISSN 2208-2352)*, vol. 9, no. 3, pp. 1–5, 2023, doi: 10.53555/nnas.v9i3.1591.
- [11] T. D. Jaya, M. B. Tukan, and F. Komisia, "Penerapan Pendekatan Inkuiri Terbimbing Untuk Melatih Keterampilan Proses Sains Siswa Materi Larutan Penyangga," *Educ. J. Pendidik.*, vol. 1, no. 2, pp. 359–366, 2022, doi: 10.56248/educativo.v1i2.44.
- [12] N. Atikah and Toni Haryanto, "Analisis Keterampilan Proses Sains melalui Penerapan Metode Praktikum di SMAN 6 Kerinci," *Biosf. J. Biol. dan Pendidik. Biol.*, vol. 10, no. 1, pp. 1–6, 2025, doi: 10.23969/biosfer.v10i1.26360.
- [13] T. Hunegnaw and S. Melesse, "An evaluative study of the experimental tasks of the Ethiopian grade 12 chemistry textbook considering developing 'science process skills,'" *Cogent Educ.*, vol. 10, no. 1, 2023, doi: 10.1080/2331186X.2023.2208944.
- [14] S. N. Habibah and S. Fuadiyah, "Validitas Instrumen Tes Keterampilan Proses Sains Pada Pembelajaran," *J. Jendela Pendidik.*, vol. 5, no. 02, pp. 353–360, 2025.
- [15] B. K. Temiz, M. F. Taşar, and M. Tan, "Volume 7 Number 7 December 2006," *Int. Educ. J.*, vol. 7, no. 7, pp. 1007–1027, 2006.
- [16] A. Koomson, C. Y. Kwaah, and C. Adu-Yeboah, "Effect of science process skills and entry grades on academic scores of student teachers," *J. Turkish Sci. Educ.*, vol. 21, no. 1, pp. 118–133, 2024, doi: 10.36681/tused.2024.007.
- [17] H. Hasanah and F. Q. Aini, "Pengembangan Instrumen Tes Keterampilan Proses Sains pada Topik Kesetimbangan Kimia: Analisis dengan Model Rasch," *J. Pendidik. Mat. dan IPA*, 2025, [Online]. Available: <https://bimaberilmu.com/jurnal/index.php/jagomipa/article/view/1043>
- [18] N. M. F. Goga, E. Salma, T. W. Nurhuda, R. Zahra, and F. A. Sari, "Analisis Kelayakan Instrumen Keterampilan Proses Sains pada Uji Coba Soal Materi Ekosistem Pendahuluan Pendidikan memiliki peran penting dalam mengembangkan potensi peserta didik," vol. 1, no. 2, pp. 39–50, 2025.
- [19] P. Anggelina, H. Akhsan, and K. A. Winarni, "pengembangan instrumen tes kps stem inklusif," vol. 9, no. 2, pp. 445–459, 2025.
- [20] I. Hasmiah, W. P. Hadi, M. C. Sutarja, and A. Rakhmawan, "Analisis Kps Siswa Menggunakan Soal Kps Berbasis Hots Pada Materi Getaran, Gelombang Dan Cahaya," *J. Nat. Sci. Educ. Res.*, vol. 8, no. 2, pp. 130–139, 2025.
- [21] A. Riyadi, P. Susongko, and M. Munadi, "Model Asesmen Literasi Sains pada peserta Didik Sekolah Dasar dengan Aplikasi Model Rasch," *J. Educ. Res.*, vol. 5, no. 3, pp. 3044–3054, 2024, doi: 10.37985/jer.v5i3.1411.

- [22] J. M. Linarche, *A user's guide to Winsteps: Rasch-model computer programs*. 2006.
- [23] A. Y. M. A. Islam, X. Gu, C. Crook, and J. M. Spector, "Assessment of ICT in Tertiary Education Applying Structural Equation Modeling and Rasch Model," *SAGE Open*, vol. 10, no. 4, 2020, doi: 10.1177/2158244020975409.
- [24] Y. Xiao, J. Han, K. Koenig, J. Xiong, and L. Bao, "Multilevel Rasch modeling of two-Tier multiple choice test: A case study using Lawson's classroom test of scientific reasoning," *Phys. Rev. Phys. Educ. Res.*, vol. 14, no. 2, p. 20104, 2018, doi: 10.1103/PhysRevPhysEducRes.14.020104.
- [25] R. M. Branch, *Instructional design: The ADDIE approach*. 2009. doi: 10.1007/978-0-387-09506-6.
- [26] S. Apeadido, W. School, and G. Kumasi, "Assessment Rubric for Evaluating Students' Science Process Skills," *African J. Educ. Stud. Math. Sci.*, pp. 134–136, 2024.
- [27] L. R. Aiken, "Three Coefficients for Analyzing the Reliability and Validity of Ratings," *Educ. Psychol.*, vol. 45, pp. 131–142, 1985.
- [28] H. Retnawati, "Proving Content Validity of Self-Regulated Learning Scale (The Comparison of Aiken Index and Expanded Gregory Index)," *Res. Eval. Educ.*, vol. 2, no. 2, pp. 155–164, 2016.
- [29] B. Sumintono and W. Widhiarso, *Aplikasi Pemodelan Rasch pada Assessment Pendidikan*. Cimahi: Trim Komunikata, 2015.
- [30] M. A. Shahat, W. J. Boone, K. A. Al-Alawi, and S. M. Al-Balushi, "Rasch analysis in developing a questionnaire to measure self-efficacy beliefs of Omani preservice science teachers for teaching through engineering design," *Humanit. Soc. Sci. Commun.*, vol. 11, no. 1, pp. 1–17, 2024, doi: 10.1057/s41599-024-04087-x.
- [31] M. Suseno, B. Hayat, M. D. K. Putra, J. K. Bien, R. Rachmawati, and H. Hartanto, "A differential item functioning (DIF) analysis of the mobile phone problem use scale in Indonesian schools with and without smartphone banned policy," *Cogent Psychol.*, vol. 9, no. 1, 2022, doi: 10.1080/23311908.2022.2137306.
- [32] H. Abdellatif, "Test results with and without blueprinting: Psychometric analysis using the Rasch model," *Educ. Medica*, vol. 24, no. 3, p. 100802, 2023, doi: 10.1016/j.edumed.2023.100802.
- [33] L. Tesio, A. Caronni, A. Simone, D. Kumbhare, and S. Scarano, "Interpreting results from Rasch analysis 2. Advanced model applications and the data-model fit assessment," *Disabil. Rehabil.*, vol. 46, no. 3, pp. 604–617, 2024, doi: 10.1080/09638288.2023.2169772.
- [34] A. Darmana, A. N. Pulungan, H. Annazili Nasution, M. Agustina Nst, and P. Faradilla, "Test Instrument Validation with Rasch Model," 2024, doi: 10.4108/eai.24-10-2023.2342115.
- [35] D. R. Ocy, W. Rahayu, and Makmuri, "Rasch Model Analysis: Development of HOTS-Based Mathematical Abstraction Ability Instrument According to Riau Islands Culture," *J. Progr. Stud. Pendidik. Mat.*, vol. 12, no. 4, pp. 3542–3560, 2023.
- [36] M. Wahyu Setiyadi, N. Ketut Rapi, P. Budi Adnyana, and S. Syamratulangi, "Development of a Science Process Skills Instrument for Ecology: A Validity and Reliability Study," vol. 10, no. 1, pp. 98–114, 2025, [Online]. Available: <http://ejournal.uki.ac.id/index.php/edumatsains98ThisisanopenaccessarticleundertheHYPERLINK%22https://creativecommons.org/licenses/by-sa/4.0/%22CCBY-SALicense.https://doi.org/10.33541/edumatsains.v10i1.6642>
- [37] A. Ilhami, T. Hidayat, R. Riandi, and S. Sriyati, "Psychometric Validation of A Research Skills Assessment For Pre-Service Science Teachers Using The Rasch Model," *Thabiea J. Nat. Sci. Teach.*, vol. 8, no. 2, p. 149, 2025, doi: 10.21043/thabiea.v8i2.34359.
- [38] S. Arikunto, *Prosedur Penelitian Suatu Pendekatan Praktik*. Jakarta: Rineka Cipta, 2013.
- [39] S. L. Dewi and T. Lestari, "analysis of junior high school students level on science process skills," 2024.
- [40] M. Z. Rozali, S. Puteh, F. A. N. Yunus, N. H. Hamdan, and H. F. M. Latif, "Reliability and Validity of Instrument on Academic Enhancement Support for Student-Athlete Using Rasch Measurement Model," *Asian J. Univ. Educ.*, vol. 18, no. 1, pp. 290–299, 2022, doi: 10.24191/ajue.v18i1.17199.
- [41] W. J. Boone, "Rasch analysis for instrument development: Why,when,and how?," *CBE Life Sci. Educ.*, vol. 15, no. 4, 2016, doi: 10.1187/cbe.16-04-0148.
- [42] M. F. Resma Wahyuni, Firdaus L.N, Riki Apriyandi Putra, Mariani Natalina Lingasari, Putri Adita Wulandara, "Design of Science Process Skills Instruments Reaction Rate Material: Rasch Model Approach," *J. Educ. Sci.*, vol. 9, no. 2, pp. 876–885, 2025, [Online]. Available: <https://doi.org/10.31258/jes.6.3.p.444-458>
- [43] S. N. T. M. Yasin, M. F. M. Yunus, and I. Ismail, "The use of rasch measurement model for the validity and reliability," *J. Couns. Educ. Technol.*, vol. 1, no. 2, p. 22, 2018, doi: 10.32698/0111.