Volume 6 Number 1 (2026)
January-June 2026
Page: 380-401

E-ISSN: 2797-0965
P-ISSN: 2809-9710
DOI: 10.37680/basica.v6i1.9751

BASICA Journal of Primary Education

The Rasch Approach in the Evaluation of Reasoning Instruments for Elementary School Students: A Literature Review

Sri Wulan Dewi[1], Purwo Susongko[2], Nuni Widiarti[3], Arif Widiyatmoko[4]
[1234] Universitas Negeri Semarang; Indonesia
Correspondence Email; sriwulandewi91@students.unnes.ac.id

Submitted: 03/04/2025 Revised: 20/04/2026 Accepted: 24/05/2026 Published: 21/06/2026

Abstract

This study aims to analyze the application of the Rasch Model in evaluating reasoning assessment instruments for elementary school students through a systematic literature review. Reasoning ability is a fundamental component of Higher Order Thinking Skills (HOTS) and plays an important role in developing critical thinking and problem-solving skills. This study employed a Systematic Literature Review (SLR) approach guided by the PRISMA 2020 framework. Literature was collected from Google Scholar, Scopus, ERIC, and Garuda databases. The review included peer-reviewed journal articles published between 2020 and 2025 that focused on the application of the Rasch Model in educational assessment, reasoning assessment, HOTS, and instrument validation. Following the identification, screening, eligibility, and inclusion stages, 15 articles met the inclusion criteria and were analyzed using thematic content analysis. The findings indicate that the Rasch Model provides more objective, valid, reliable, and stable measurements than Classical Test Theory (CTT). The model effectively evaluates item validity, reliability, item difficulty levels, person-item maps, and item fit statistics, thereby improving the accuracy of reasoning assessment. The study concludes that the Rasch Model is highly relevant for developing HOTS-based reasoning instruments in elementary education and contributes significantly to improving the quality of educational assessment and supporting 21st-century learning evaluation.

Keywords

Rasch Model, Reasoning Assessment, HOTS Evaluation.

INTRODUCTION

The quality of education in Indonesia remains a significant concern, particularly in relation to students' higher-order thinking and reasoning abilities. International large-scale assessments indicate that Indonesian students continue to face challenges in applying knowledge, reasoning logically, and solving complex problems. The Programme for International Student Assessment (PISA) 2022 reported that Indonesian students achieved an average mathematics score of 366, substantially below the OECD average of 472. In reading and science, Indonesian students scored 359 and 383 points respectively, compared to OECD averages of 476 and 485. Furthermore, only 18% of Indonesian students reached at least Level 2 proficiency in mathematics, while the OECD average was 69%, indicating that many students still experience difficulties in applying mathematical reasoning to real-world situations (PISA, 2022).

The findings of international assessments are consistent with previous evidence from Trends in International Mathematics and Science Study (TIMSS), which has repeatedly shown that Indonesian students perform below the international average in mathematics and science achievement. These results suggest that students' reasoning, analytical thinking, and problem-solving abilities require further improvement. The ability to reason effectively is essential because it enables students to connect concepts, interpret information, evaluate evidence, and construct logical conclusions in various learning contexts (Mullis et al., 2020).

At the national level, the implementation of the National Assessment has shifted the focus of educational evaluation from content mastery toward literacy and numeracy competencies. This policy emphasizes students' ability to reason, analyze information, and solve contextual problems rather than merely recalling factual knowledge. The increasing emphasis on literacy and numeracy highlights the importance of developing assessment instruments capable of accurately measuring reasoning abilities and higher-order thinking skills in elementary education.

Several studies have also reported that elementary school students still demonstrate relatively low levels of reasoning and critical thinking skills. Students frequently encounter difficulties in interpreting information, constructing logical arguments, identifying relationships among concepts, and solving non-routine problems. Research on HOTS-based learning in elementary education indicates that many students remain dependent on procedural approaches and have limited opportunities to engage in analytical and reflective thinking processes. These findings demonstrate the need for valid and reliable assessment instruments that can accurately

measure reasoning abilities and provide meaningful information regarding students' cognitive development.

Elementary education plays a strategic role in building the quality of human resources as it serves as the foundational stage for developing students' intellectual, social, emotional, and character skills. The elementary school stage is a critical period in children's cognitive development, as it is during this phase that students begin to develop the ability to understand concepts, solve problems, and build logical thinking patterns. 21st-century education prioritizes critical, creative, collaborative, and communicative thinking as core competencies that students must possess starting from the elementary level. This shift means that the learning process is no longer solely focused on mastering subject matter but also on developing higher-order thinking skills (HOTS). These skills are closely related to the reasoning process, which forms the foundation for understanding relationships between concepts and making logical decisions. Studies on elementary education indicate that strengthening critical thinking and reasoning skills from an early age influences students' readiness to face learning challenges at subsequent educational levels (Syadiah & Hamdu, 2020).

Reasoning skills are a crucial component of elementary students' intellectual development. Reasoning can be defined as the cognitive process used to draw conclusions based on facts, data, or specific logical relationships. In elementary school students, reasoning skills develop through activities such as observing, comparing, classifying, analyzing, and systematically drawing conclusions from information. These skills form the foundation for learning mathematics, science, literacy, and contextual problem-solving. The development of reasoning in elementary school students must be carried out systematically because these skills are closely linked to critical thinking skills and the ability to solve problems in daily life. Research on critical thinking among elementary school students indicates that students' reasoning abilities still require reinforcement through appropriate instruction and assessment so that students can understand concepts deeply rather than merely memorizing information (Purnamasari et al., 2026).

The development of modern curricula encourages the implementation of Higher-Order Thinking Skills (HOTS)-based learning assessments as an effort to improve educational quality. HOTS-based assessments emphasize the ability to analyze, evaluate, and create solutions based on conceptual understanding. Assessment tools used in 21st-century learning must be able to measure higher-order thinking skills, including students' reasoning abilities. Good instruments do not

merely assess memorization but also identify students' abilities to connect concepts, construct arguments, and solve problems logically. Developing HOTS-based instruments presents a unique challenge for teachers and researchers as it requires high levels of validity, reliability, and objectivity. Research on critical thinking and HOTS instruments indicates that the quality of the instruments significantly determines the accuracy of measurements of students' abilities (Ocy et al., 2023).

Learning assessment instruments play a crucial role in determining the quality of students' learning processes and outcomes. A good instrument must provide accurate information regarding students' ability levels so that evaluation results can serve as a basis for educational decision-making. Measuring reasoning abilities requires instruments that are valid, reliable, objective, and capable of accurately distinguishing students' abilities. Errors in instrument development can lead to measurement bias and result in inaccurate interpretations of student abilities. Therefore, instrument quality analysis is a crucial step in educational evaluation development. Studies on educational instrument development indicate that the quality of test items must be comprehensively analyzed to ensure the instrument truly aligns with the construct of the ability being measured (Nuryanti et al., 2018).

The evaluation approach widely used in education to date is Classical Test Theory (CTT). Although this theory is relatively simple and easy to apply, this approach has several limitations in producing objective measurements. Classical Test Theory is highly dependent on the characteristics of the respondent group, so the reliability and validity of the instrument may vary when used with different groups. Furthermore, Classical Test Theory has not been able to distinguish between individual ability and item characteristics independently. Item analysis in conventional approaches is also limited because it only produces total scores without providing detailed information regarding item difficulty levels and student response patterns. Research on educational instrument analysis indicates that Classical Test Theory has limitations in producing stable and in-depth measurements, necessitating a more modern evaluation approach (Purba, 2018).

The Rasch model emerges as one of the modern approaches in the analysis of educational assessment instruments capable of addressing various weaknesses of classical test theory. The Rasch model is part of Item Response Theory (IRT), which focuses on the relationship between individual ability and item difficulty levels within a single measurement scale. This approach allows for more objective analysis because item parameters and respondent ability can be separated. The Rasch model provides detailed information regarding item validity, instrument reliability, item difficulty

levels, student response patterns, and the detection of biased items. The use of the Rasch model in educational evaluation is growing because it is considered more modern, efficient, and offers higher measurement accuracy compared to conventional approaches. Various studies indicate that the Rasch model is effective for analyzing Higher-Order Thinking Skills (HOTS), critical thinking, and students' reasoning abilities (Prayoga et al., 2024).

The implementation of the Rasch model in elementary education research has yielded positive results regarding the quality of assessment instruments. Research on the analysis of critical thinking questions for elementary school students indicates that the Rasch model can identify item-student fit and student ability in greater detail, enabling teachers to identify weaknesses in the instruments as well as students' abilities. Other studies also indicate that the Rasch approach is effective in developing HOTS and literacy instruments as it yields higher validity and reliability. Furthermore, the use of person-item maps in Rasch analysis helps researchers and teachers understand the relationship between students' ability levels and item difficulty levels on a single measurement scale. These factors make the Rasch model relevant for evaluating the reasoning abilities of elementary school students, a task that requires high precision and objectivity (Wandyka & Aini, 2025).

Although various studies have discussed the use of the Rasch model in educational assessment, research specifically addressing the Rasch approach in evaluating reasoning instruments for elementary school students remains relatively limited. Most previous studies have focused more on the development of specific instruments without conducting a comprehensive synthesis of the Rasch model's effectiveness in measuring the reasoning abilities of elementary school students. Other studies tend to discuss the use of the Rasch model at the secondary and higher education levels, so the context of elementary education has not been studied in depth. This situation indicates a research gap that needs to be addressed through a literature review to gain a comprehensive understanding of the implementation of the Rasch model in the evaluation of elementary school students' reasoning instruments. Based on these conditions, this study aims to analyze the Rasch approach in the evaluation of elementary school students' reasoning instruments through a literature review by examining the concept of reasoning, evaluation instruments, the basic principles of the Rasch model, and its implementation in the development of HOTS-based learning assessments.

METHOD

This study employed a Systematic Literature Review (SLR) approach guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) framework (Page et al., 2021). The PRISMA framework was adopted to ensure a transparent, systematic, and reproducible process in identifying, screening, selecting, and synthesizing relevant studies concerning the application of the Rasch Model in evaluating reasoning assessment instruments for elementary school students.

Data Collection

The literature search was conducted from January to February 2025 using four scientific databases, namely Google Scholar, Scopus, ERIC, and Garuda. The search process employed combinations of the keywords "Rasch Model", "reasoning ability", "reasoning assessment", "Higher Order Thinking Skills (HOTS)", "assessment instrument", and "elementary school students". The review focused on studies published between 2020 and 2025 to ensure that the selected literature reflected recent developments in educational assessment and Rasch measurement applications.

Table 1. Results of Literature Search Across Databases

Database	Records Identified
Google Scholar	28
Scopus	12
ERIC	10
Garuda	12
Total	62

Article Selection

The selection of articles followed predefined inclusion and exclusion criteria. Studies were included if they: (1) were published between 2020 and 2025; (2) were peer-reviewed journal articles; (3) discussed the application of the Rasch Model in educational assessment, instrument validation, reasoning assessment, critical thinking, HOTS, literacy, or learning evaluation; (4) were available in full-text form; and (5) contained empirical findings relevant to the research objectives. Studies were excluded if they: (1) were conference abstracts, editorials, proceedings papers, book reviews, or opinion articles; (2) did not apply the Rasch Model; (3) were unrelated to educational assessment; (4) contained incomplete methodological information; or (5) were duplicate publications retrieved from multiple databases.

Table 2. Inclusion and Exclusion Criteria

Criteria	Inclusion Criteria	Exclusion Criteria
Publication Year	2020–2025	Before 2020
Publication Type	Peer-reviewed journal articles	Editorials, proceedings, book reviews
Research Topic	Rasch Model, reasoning, HOTS, critical thinking, educational assessment	Unrelated topics
Accessibility	Full-text available	Full-text unavailable
Methodology	Empirical studies and instrument development studies	Conceptual papers without empirical findings
Context	Educational assessment and learning evaluation	Non-educational contexts

The identification stage yielded 62 records consisting of 28 articles from Google Scholar, 12 articles from Scopus, 10 articles from ERIC, and 12 articles from Garuda. After removing 8 duplicate records, 54 articles remained for title and abstract screening. During the screening stage, 24 articles were excluded because they did not directly address the Rasch Model, reasoning assessment, educational evaluation, or HOTS-related constructs. The remaining 30 full-text articles were assessed for eligibility. Subsequently, 15 articles were excluded because they lacked sufficient empirical evidence, did not meet the inclusion criteria, or provided inadequate information regarding Rasch-based instrument evaluation. Finally, 15 articles fulfilled all eligibility requirements and were included in the qualitative synthesis. The article selection process is presented in Figure 1 following the PRISMA 2020 framework.

.

.

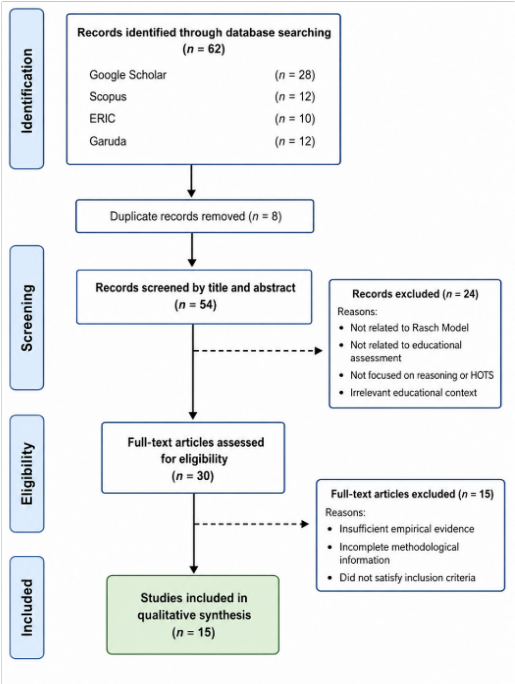

Figure 1. PRISMA Flow Diagram

Data Analysis

Data extraction was conducted by documenting the author(s), publication year, research objectives, research design, instrument characteristics, Rasch analysis indicators, and key findings from each selected study. The extracted data were analyzed using thematic content analysis to identify recurring themes, patterns, similarities, differences, and research trends concerning the effectiveness of the Rasch Model in evaluating reasoning assessment instruments. The synthesis focused on four major themes: (1) instrument validity, (2) instrument reliability, (3) item difficulty and person-item mapping, and (4) the contribution of the Rasch Model to Higher Order Thinking Skills (HOTS) and reasoning assessment in elementary education. This thematic synthesis enabled a comprehensive understanding of how Rasch analysis contributes to improving the quality, objectivity, and accuracy of educational assessment instruments.

FINDINGS AND DISCUSSION

Findings

1. The Concept of Reasoning in Elementary School Students

Reasoning ability is a crucial aspect of elementary school students' cognitive development. Reasoning can be understood as the logical thinking process individuals use to draw conclusions based on facts, concepts, experiences, and specific relationships. In the context of elementary

education, reasoning ability is closely linked to students' capacity to understand cause-and-effect relationships, connect information, analyze situations, and systematically solve problems. These abilities form the primary foundation for the development of higher-order thinking skills (HOTS). Research on elementary education indicates that students with strong reasoning skills tend to grasp abstract concepts more easily and are able to apply their knowledge in daily life. Developing reasoning skills starting in elementary school also influences critical thinking and problem-solving abilities in subsequent educational levels (Saraswati & Agustika, 2020).

Elementary school students' reasoning develops in accordance with the stages of children's cognitive development. Based on the developmental characteristics of elementary school-aged children, students tend to be in the concrete operational stage, so logical thinking skills begin to develop through real-world experiences and contextual activities. At this stage, students begin to be able to classify, group, sequence, and draw simple conclusions based on direct observation. The development of reasoning is influenced by the learning environment, learning strategies, and students' experiences in solving problems. Exploration- and problem-solving-based learning activities have been proven to enhance students' reasoning abilities because they provide opportunities for students to actively develop their thinking processes. Research on elementary school students' critical thinking skills indicates that Higher-Order Thinking Skills (HOTS)-based learning activities can help students develop logical and argumentative reasoning abilities (Maylia et al., 2024).

The types of reasoning that develop in elementary school students include inductive and deductive reasoning. Inductive reasoning is the ability to draw general conclusions based on specific facts, while deductive reasoning involves applying general concepts to specific situations. Both types of reasoning are crucial in mathematics, science, and literacy education because they help students understand patterns, relationships, and concepts more deeply. Inductive reasoning typically develops earlier through observation and pattern recognition, while deductive reasoning develops through practice in analyzing and solving more complex problems. Research on the development of reasoning skills in elementary school students indicates that the use of contextual and Higher-Order Thinking Skills (HOTS) -based problems can enhance students' ability to draw logical conclusions and construct arguments systematically (Anggraeni et al., 2022).

The urgency of developing reasoning skills in elementary school students is increasing in line with the demands of 21st-century education. Reasoning skills form the foundation for the

development of critical thinking, creativity, communication, and collaboration. Students with strong reasoning skills tend to be more adaptable in the face of change and capable of solving problems effectively (Anggraini et al., 2020). The development of reasoning skills also supports improvements in numeracy literacy and rational decision-making abilities. These factors necessitate that schools develop learning and assessment strategies capable of measuring and enhancing students' reasoning skills optimally.

2. Reasoning Assessment Instruments

An evaluation instrument is a tool used to gather information about students' abilities in specific areas. In educational evaluation, instruments play a crucial role as they form the basis for determining the level of students' learning outcomes. Reasoning evaluation instruments are designed to measure students' ability to think logically, analyze information, draw conclusions, and solve problems. The development of reasoning assessment instruments must consider the cognitive developmental characteristics of elementary school students so that the measurement results truly reflect students' abilities objectively. Research on learning assessment indicates that a good instrument must be able to measure students' thinking abilities in depth and not merely focus on rote memorization (Natoil, 2021).

Characteristics of a good assessment instrument include validity, reliability, objectivity, practicality, and discriminative power. Validity indicates the instrument's accuracy in measuring the intended ability, while reliability relates to the consistency of measurement results. Objectivity indicates that evaluation results are not influenced by the evaluator's subjectivity. Reasoning assessment instruments must also have a difficulty level appropriate to students' abilities so they can accurately distinguish between high- and low-ability students. Research on the development of HOTS instruments shows that the quality of the instrument significantly determines the accuracy of interpreting students' learning outcomes. An invalid instrument can produce measurement bias, leading to errors in educational decision-making (Antari et al., 2021).

Forms of reasoning assessment instruments can include essay questions, HOTS-based multiple-choice questions, diagnostic tests, or problem-based assessments. HOTS-based instruments are designed to encourage students to analyze, evaluate, and synthesize information, thereby allowing reasoning abilities to be measured more comprehensively. Reasoning questions are typically framed within real-life contexts so that students can connect learning concepts to concrete situations. The use of contextual instruments has been proven to help students develop

critical thinking and logical reasoning skills. Research on HOTS-based assessment shows that contextual questions are more effective in measuring higher-order thinking skills than conventional questions that merely assess concept recall (Sabir et al., 2021).

HOTS-based evaluation is a key approach in modern education because it encourages students to think critically and creatively. Implementing HOTS evaluation in elementary education requires high-quality instruments to ensure accurate measurement results that can serve as a basis for improving learning. Teachers must understand techniques for developing reasoning instruments that align with the characteristics of elementary school students to ensure evaluations are neither too difficult nor too easy. Additionally, the development of reasoning instruments must be supported by systematic item quality analysis to ensure the validity and reliability of the instruments. These factors make modern analytical approaches, such as the Rasch model, essential in the development of HOTS-based assessment instruments.

3. Basic Concepts of the Rasch Model

The Rasch model is one of the modern measurement approaches used in the analysis of educational instruments. This model was developed by Georg Rasch and falls under the category of Item Response Theory (IRT). The Rasch model is used to analyze the relationship between individual ability and item difficulty on a single measurement scale called the logit. This approach allows for more objective measurement because student ability and item characteristics are analyzed separately. The use of the Rasch model in educational evaluation is growing because it produces more accurate and stable measurement data compared to classical test theory approaches (Purba, 2018).

The basic principle of the Rasch model states that the probability of a student answering an item correctly is determined by the difference between individual ability and item difficulty. The higher the student's ability relative to the item's difficulty, the greater the likelihood of the student answering correctly. This approach allows researchers to obtain detailed information about student ability and item quality simultaneously. The Rasch model also incorporates the principle of unidimensionality, which emphasizes that an instrument must measure a single specific construct for measurement results to be valid. In the context of evaluating elementary school students' reasoning skills, the Rasch model is used to ensure that the instrument truly measures reasoning ability in accordance with learning objectives (Ramadhan & Handayani, 2025).

Analysis within the Rasch model includes several key indicators such as item fit, person and

item reliability, the separation index, item difficulty levels, and the person-item map. Item fit values are used to assess how well items align with the measurement model, while reliability indicates the instrument's consistency in measuring student ability. The separation index is used to determine the instrument's ability to distinguish between students' ability levels. Item difficulty level analysis helps researchers understand the distribution of questions ranging from easy to difficult categories. Meanwhile, the person-item map is used to illustrate the relationship between student ability and item difficulty levels on a single measurement scale. Research on educational instrument analysis indicates that these indicators are highly beneficial for developing more objective and high-quality instruments (Prayoga et al., 2024).

The advantage of the Rasch model over classical test theory lies in its ability to produce measurements that are independent of the sample and the items. In classical test theory, measurement results are heavily influenced by the characteristics of the respondent group, so the reliability of the instrument can vary when used with different groups. Conversely, the Rasch model allows for more stable measurements because item parameters and student ability are analyzed separately. The Rasch model is also capable of detecting biased items, analyzing patterns of unusual responses, and producing more precise interval data compared to the ordinal scores in classical test theory. These factors have led to the Rasch model being increasingly used in the development of evaluation instruments based on Higher-Order Thinking Skills (HOTS) and the reasoning abilities of elementary school students.

Table 3. Comparison of Classical Test Theory and the Rasch Model

Analysis Aspects	Classical Test Theory	Rasch Model
Measurement Basis	Total score	Response Probability
Sample Dependency	High	Low
Item Analysis	General	Details
Difficulty Level	Relative to the group	Objective
Reliability	Depends on the respondents	More stable
Item bias detection	Limited	More accurate
Data Type	Ordinal	Interval
Analysis of Individual Abilities	Limited	In-Depth

The table shows that the Rasch model has significant advantages over classical test theory in producing objective and in-depth measurements. Rasch analysis allows researchers to obtain detailed information regarding item quality and student ability, thereby making evaluation instruments more valid and reliable.

4. Implementation of the Rasch Approach in Reasoning Assessment

The implementation of the Rasch model in elementary education evaluation is growing, particularly in the development of instruments for Higher-Order Thinking Skills (HOTS), critical thinking, and mathematical reasoning among elementary school students. The use of the Rasch model helps researchers objectively evaluate instrument quality through item validity and reliability analysis. Instruments analyzed using the Rasch model generally have better measurement quality because each item is evaluated based on its fit to the measurement model. Research on HOTS instrument development indicates that the Rasch model is effective for improving item quality and enhancing the accuracy of student ability measurements.

Validity analysis in the Rasch model is conducted using the outfit mean square and infit mean square values. Items with values outside the ideal range are considered inconsistent with the model and thus need to be revised or removed. This approach helps researchers ensure that each item truly measures reasoning ability in accordance with the established construct. Reliability analysis is conducted through person reliability and item reliability to determine the instrument's consistency and the stability of measurement results (Prayoga et al., 2024). Research on Rasch-based educational instrument analysis indicates that instruments with high reliability tend to produce more consistent and accurate measurements.

Item difficulty levels in the Rasch model are analyzed using logit values. This analysis helps researchers understand the distribution of items ranging from easy to difficult categories, allowing the instrument to be structured proportionally according to the abilities of elementary school students. Additionally, the Rasch model provides a person-item map analysis that illustrates the relationship between student ability and item difficulty on a single measurement scale. This map helps teachers understand whether the instrument aligns with students' abilities or requires revision. The use of person-item maps in evaluating elementary school students' reasoning is considered effective for identifying gaps between student ability and instrument characteristics (Purnamasari et al., 2026).

The development of reasoning instruments based on the Rasch model makes a significant contribution to improving the quality of elementary education assessment. Instruments developed using the Rasch approach tend to have better construct validity, high reliability, and a more accurate level of measurement objectivity. The use of the Rasch model also helps teachers and researchers systematically revise items based on empirical analysis results. These conditions indicate that the

Rasch model is highly relevant for use in developing HOTS-based reasoning assessment instruments for elementary school students and 21st-century learning.

Table 4. Summary of Previous Research on the Implementation of the Rasch Model

No	Researcher	Research Focus	Key Findings
1.	Uli Sihombing et al., (2020)	Analysis of science literacy using the Rasch model	Item reliability is excellent and the instrument fits the model
2.	Oktaviani et al., (2021)	Validity and reliability of the coping strategy instrument	The Rasch model is effective for analyzing the validity and reliability of the instrument
3.	Afandi et al., (2025)	Development of a digital literacy instrument	Person reliability and item reliability are classified as good
4.	Ahmad et al., (2021)	Development of an elementary school mathematics learning outcome instrument	Item difficulty levels can be objectively measured using Rasch
5.	Supriyadi et al., (2026)	Validation of the critical consciousness scale	The instrument's reliability is very high based on Rasch analysis
6.	Piliyang et al., (2025)	Implementation of the Rasch model for modern test evaluation	Rasch produces a more objective and systematic test evaluation
7.	Qadriyyah et al., (2025)	Analysis of the validity and reliability of psychological instruments	Rasch is capable of identifying item fit and instrument reliability
8.	Chan et al., (2024)	Analysis of statistical reasoning ability	Rasch effectively maps students' reasoning abilities
9.	Pratama & Husnayaini, (2020)	Measuring students' reading ability using the Rasch model	The Rasch model effectively maps students' abilities and the quality of reading test instruments
10.	Ramdani et al., (2020)	Validation of the academic resilience scale	The instrument has good construct validity
11.	Sadhu & Laksono, (2021)	Development of an assessment of critical thinking and chemistry literacy	The instrument is valid and reliable based on Rasch analysis
12.	Aminudin et al., (2021)	Development of a Rasch-based multitiered instrument	Rasch is effective in detecting student misconceptions
13.	Difa, (2026)	Development of a self-efficacy scale	Rasch improves the accuracy of scale development
14.	Bond et al., (2020)	Fundamental measurement in the human sciences	Rasch produces more objective measurements
15.	Wandyka & Aini, (2025)	Linking Rasch-scaled tests	Rasch provides high measurement stability

Based on the results of the aforementioned review of previous research, the implementation of the Rasch model demonstrates consistent findings that this approach is capable of improving the quality of educational assessment instruments through the objective analysis of validity, reliability, item difficulty levels, and the mapping of students' abilities. Most studies affirm that the Rasch model is superior to classical test theory approaches because it is capable of producing stable measurements that are independent of sample characteristics. The use of the Rasch model in evaluating reasoning and HOTS in elementary education is becoming increasingly relevant because it supports the development of 21st-century learning assessments that demand critical thinking, logical reasoning, and comprehensive problem-solving skills.

Discussion

1. The Effectiveness of the Rasch Model in Instrument Evaluation

The Rasch model is highly effective for evaluating educational instruments because it yields measurements that are more objective, accurate, and stable than classical test theory approaches. Rasch analysis allows researchers to separate student ability from item characteristics, ensuring that measurement results are not overly influenced by respondent group characteristics. These conditions make the Rasch model more relevant for evaluating the reasoning abilities of elementary school students, which requires high measurement precision. Various studies indicate that the Rasch model can enhance instrument quality through analyses of validity, reliability, item difficulty levels, and the identification of items inconsistent with the measurement construct. Research by Syadiah & Hamdu, (2020) shows that Rasch analysis of STEM-based critical thinking questions can produce more valid and reliable instruments for elementary school students. Additionally, the Rasch model helps researchers understand students' response patterns in greater detail, allowing instruments to be developed systematically and empirically.

The effectiveness of the Rasch model is also evident in its ability to generate interval data through logit scores. Interval data allows for a more precise interpretation of student ability compared to the ordinal scores used in classical test theory. The use of the Rasch model in the development of educational instruments helps teachers and researchers objectively understand the distribution of student ability and empirically determine item difficulty levels. Purba, (2018) research on Rasch model analysis in achievement test instruments demonstrates that the Rasch approach yields more consistent and in-depth measurement results compared to conventional analysis. These findings indicate that the Rasch model functions not only as a statistical analysis tool

but also as a modern evaluation approach capable of supporting the comprehensive development of HOTS-based learning assessments.

A synthesis of various studies indicates that the Rasch model is effective for developing reasoning instruments because it can identify items that are too easy, too difficult, or do not fit the model. This information is crucial for designing evaluation instruments for elementary school students, as students' abilities at this level develop gradually and require a proportional distribution of item difficulty levels. Research by Prayoga et al., (2024) explains that the use of the Rasch model helps researchers understand the instrument's construct more deeply, thereby improving the instrument's quality through item revisions based on empirical analysis results.

2. The Relevance of the Rasch Model to Measuring Elementary School Students' Reasoning

Elementary school students' reasoning ability is a complex skill that cannot be optimally measured through conventional, memorization-based assessments. Reasoning requires logical, analytical, and reflective thinking processes; therefore, assessment instruments must be able to objectively capture these abilities. The Rasch model is relevant for measuring elementary school students' reasoning because it maps the relationship between student ability and item difficulty on a single measurement scale. Through the person-item map, teachers can determine whether the difficulty level of the questions aligns with students' abilities or requires adjustment. This insight helps teachers develop more adaptive assessment instruments that align with the cognitive development of elementary school students. Research by Purnamasari et al., (2026) demonstrates that the validation of a Rasch-based critical thinking instrument provides a more detailed understanding of students' abilities in solving Higher-Order Thinking Skills (HOTS) questions.

The relevance of the Rasch model to measuring elementary students' reasoning is also evident in its ability to support 21st-century learning assessment. Modern education emphasizes the development of HOTS, which includes critical thinking, creativity, and problem-solving skills. Measuring these abilities requires instruments that can truly identify students' thinking skills in depth. The Rasch model supports this process through item fit and reliability analysis, ensuring the quality of the instrument before it is used in learning evaluation. Research by Ocy et al., (2023) on the development of a HOTS-based mathematical abstraction ability instrument demonstrates that the Rasch model is effective in ensuring instrument quality and enhancing the accuracy of measuring students' thinking abilities.

The characteristics of elementary school students, who are still in the concrete operational stage, necessitate that reasoning assessments be designed contextually and aligned with children's cognitive development. The Rasch model helps ensure that the items used are truly appropriate for students' ability levels. Analyses of item difficulty levels and person reliability allow researchers to determine whether the instrument is too difficult or too easy for students. Research by Ahmad et al., (2021) on the development of elementary school mathematics achievement instruments shows that the Rasch model can help researchers produce instruments that are more proportionate and appropriate for the characteristics of elementary school students.

3. Advantages of the Rasch Model Over Conventional Approaches

The classical test theory approach has several limitations in educational evaluation, particularly in measuring higher-order thinking skills such as reasoning. Evaluation scores in classical test theory are heavily influenced by respondent group characteristics, so the reliability and validity of the instrument may vary when used with different groups. The Rasch model serves as a modern approach capable of addressing these weaknesses through more objective and sample-independent analysis. The primary advantage of the Rasch model lies in its ability to separate individual ability from item characteristics, resulting in more stable measurement outcomes. Research by Uli Sihombing et al., (2020) demonstrates that Rasch analysis of science literacy tests yields instruments that better fit the model and exhibit higher reliability compared to conventional approaches.

The Rasch model also excels at detecting biased items and unusual response patterns. Conventional approaches typically only produce total scores without providing detailed information about item quality. In contrast, the Rasch model can identify items that do not align with the measurement construct, allowing for more systematic instrument revision. Item fit analysis in the Rasch model helps researchers determine which items should be retained, revised, or removed. Research by Nuryanti et al., (2018) on the development of critical thinking instruments demonstrated that the Rasch model significantly improves item quality compared to standard item analysis.

Another advantage of the Rasch model lies in its ability to produce interval-scaled measurements using logit units. The use of interval data provides more accurate interpretations compared to ordinal scores in classical test theory. The Rasch model also supports the development of adaptive assessment and competency-based measurement, which are demands of modern

education. Research by Ramdani et al., (2020) on the validation of the academic resilience scale shows that the Rasch model yields better construct validity and instrument reliability, allowing the instrument to be used more widely and consistently.

4. Challenges in Implementing the Rasch Model

Although the Rasch model has various advantages, its implementation in elementary education still faces a number of challenges. One of the main obstacles is the limited understanding of teachers and researchers regarding Rasch analysis. The use of the Rasch model requires statistical skills and proficiency in specialized software such as Winsteps, Ministep, or R- and Python-based applications. This situation causes some teachers to still prefer using conventional approaches, which are considered simpler. Research by Piliyang et al., (2025) indicates that using Python-based Rasch models requires specific technical skills to ensure optimal analysis.

Another challenge relates to the development of HOTS-based instruments suitable for the characteristics of elementary school students. The development of reasoning questions requires teachers to have the ability to design contextual items capable of measuring students' logical thinking skills. In practice, some teachers still face difficulties in developing HOTS-based questions because they are accustomed to using memorization-based questions. Research by Sabir et al., (2021) indicates that the implementation of HOTS questions still faces challenges regarding the formulation of indicators and the alignment of questions with students' abilities.

The availability of a sufficient number of respondents also poses a challenge in implementing the Rasch model. Rasch analysis requires a certain amount of data to ensure stable and valid measurement results. In some elementary schools with limited student populations, the analysis process becomes less than optimal. Additionally, technological facilities and access to analysis software remain obstacles in some regions. These conditions highlight the need for training and mentoring for teachers, as well as support from educational institutions, to develop Rasch-based assessment more broadly.

5. Implications for Elementary Education

The implementation of the Rasch model in elementary education assessment has significant theoretical and practical implications. Theoretically, the Rasch model reinforces the modern assessment paradigm that emphasizes the objectivity, validity, and reliability of measurement. This approach supports the development of Higher-Order Thinking Skills (HOTS)-based instruments capable of measuring students' critical thinking and reasoning abilities more deeply. The use of the

Rasch model also demonstrates that educational assessment serves not only to assign scores but also as a diagnostic tool to comprehensively understand students' abilities. Research by Syadiah & Hamdu, (2020) indicates that Rasch analysis helps teachers identify the strengths and weaknesses of assessment instruments, thereby enabling systematic improvements to the learning process.

Practically, the Rasch model helps teachers and researchers develop higher-quality assessment instruments that align with the characteristics of elementary school students. Analysis of item difficulty levels and person-item maps allows teachers to design more proportionate questions, making evaluations fairer and more objective. Additionally, the use of the Rasch model supports the implementation of diagnostic assessments that can be used to identify students' reasoning abilities on an individual basis. Research by Wandyka & Aini, (2025) demonstrates that the Rasch model is effective in the development of science process skills instruments because it produces valid and reliable instruments.

Another implication is evident in the enhancement of 21st-century learning quality in elementary schools. Rasch-based assessment enables the development of instruments that are better aligned with the needs of higher-order thinking skills (HOTS) and numeracy literacy. These conditions support improvements in the quality of learning oriented toward critical thinking, problem-solving, and logical decision-making. Therefore, the implementation of the Rasch model needs to be continuously developed through teacher training, the development of modern assessment tools, and further research on the assessment of elementary school students' reasoning.

CONCLUSION

The Rasch approach in evaluating elementary school students' reasoning instruments demonstrates a significant contribution to improving the quality of learning assessment. The Rasch model is capable of producing evaluation instruments that are more objective, valid, reliable, and stable compared to the classical test theory approach because it can separate student ability and item characteristics independently. Rasch analysis also provides deeper insights into item difficulty levels, instrument reliability, item fit to the model, and the relationship between student ability and instrument characteristics through the person-item map. Findings from various studies indicate that the implementation of the Rasch model is effective in developing instruments based on Higher-Order Thinking Skills (HOTS), critical thinking, literacy, and reasoning skills for elementary school students, thereby supporting a more comprehensive and accurate 21st-century learning assessment.

The implementation of the Rasch model in elementary education also has significant implications for the development of more modern and adaptive learning assessments. The use of the Rasch model helps teachers and researchers design assessment instruments aligned with the cognitive developmental characteristics of elementary school students and supports the enhancement of critical thinking and problem-solving skills. However, the implementation of the Rasch model still faces challenges, including limited teacher understanding of Rasch analysis, the need for proficiency in statistical software, and technological infrastructure limitations in some educational institutions. Therefore, training, mentoring, and the development of teachers' competencies in modern assessment are necessary so that the Rasch model can be more widely applied in elementary education assessment. Future research is expected to develop Rasch-based reasoning instruments across various subjects and learning contexts so that the quality of elementary education assessment continues to improve.

REFERENCES

Afandi, A., Sri Wahyuni, E., & Rosyadi, A. R. (2025). Developing Instrument of Digital Age Literacy Inventory: Validity and Reliability Analysis Using Rasch Model. *JTP - Jurnal Teknologi Pendidikan*, *27*(3), 784–798. https://doi.org/10.21009/jtp.v27i3.58103

Ahmad, Izmaimusah, D., & Idris, M. (2021). Instrument Development Of Mathematics Learning Outcomes By The Rasch Model In Elementary School To Support The Implementation Of The 2013 Curriculum. *JISAE: Journal of Indonesian Student Assessment and Evaluation*, *7*(2 SE-Articles), 88–102. https://doi.org/10.21009/jisae.v7i2.21830

Aminudin, A. H., Adimayuda, R., Kaniawati, I., Suhendi, E., Samsudin, A., & Coştu, B. (2021). Rasch analysis of Multitier Open-ended Light-Wave Instrument (MOLWI): Developing and assessing second-years sundanese-scholars alternative conceptions. *Journal for the Education of Gifted Young Scientists*, *7*(3), 557–579. https://doi.org/10.17478/jegys.574524

Anggraeni, N., Rustini, T., & Wahyuningsih, Y. (2022). Keterampilan Berpikir Kritis Siswa Sekolah Dasar Pada Mata Pelajaran Ips Di Kelas Tinggi. *Jurnal Review Pendidikan Dasar : Jurnal Kajian Pendidikan Dan Hasil Penelitian*, *8*(1), 84–90. https://doi.org/10.26740/jrpd.v8n1.p84-90

Anggraini, V. S., Fanani, A., & Juniarso, T. (2020). Media Math Scrabble Untuk Kemampuan Berpikir Kritis Siswa Sekolah Dasar: Studi Kuasi Eksperimental. *Journal of the Indonesian Mathematics Education Society*, *1*(2), 38–42. https://doi.org/10.26740/jimes.v1n2.p38-42

Antari, C. J., Triyogo, A., Egok, & Sukenda, A. (2021). Penerapan Model Outdoor Learning pada Pembelajaran Tematik Siswa di Sekolah Dasar. *Jurnal Basicedu*, *5*(5), 3(2), 524–532.

Bond, T., Yan, Z., & Heene, M. (2020). *Applying the Rasch Model:Fundamental Measurement in the Human Sciences*. Routledge. https://doi.org/10.4324/9780429030499

Chan, S. W., Ismail, Z., & Sumintono, B. (2024). A Rasch Model Analysis on Secondary Students' Statistical Reasoning Ability in Descriptive Statistics. *Procedia - Social and Behavioral Sciences*, *129*, 133–139. https://doi.org/https://doi.org/10.1016/j.sbspro.2014.03.658

Difa, A. (2026). Aplikasi Model Rasch pada Pengembangan Skala Efikasi Diri dalam Pengambilan Keputusan Karir Siswa [Application of the Rasch Model on the Development of Self-Efficiency

Scale in Student Career Decision Making]. *Jurnal Psikologi*, *43*(3), 248.

Maylia, E. C., Amelia, A. P., Suwarna, D. M., Muyassaroh, I., & Jenuri, J. (2024). Strategi Pembelajaran Inkuiri Terhadap Kemampuan Berpikir Kritis Siswa SD. *Jurnal Review Pendidikan Dasar : Jurnal Kajian Pendidikan Dan Hasil Penelitian*, *10*(1), 32–41. https://doi.org/10.26740/jrpd.v10n1.p32-41

Mullis, I. V. S., Martin, M. O., Foy, P., Kelly, D. L., & Fishbein, B. (2020). *TIMSS 2019 International Results in Mathematics and Science*. TIMSS & PIRLS International Study Center.

Natoil, N. (2021). Penggunaan Media Video dalam Pembelajaran Managemen Pemasaran untuk Meningkatkan Pemahaman Mahasiswa di Masa Pandemi Covid-19. *Jurnal Educatio FKIP UNMA*, *7*(2), 432–438. https://doi.org/10.31949/educatio.v7i2.1073

Nuryanti, S., Masykuri, M., & Susilowati, E. (2018). Analisis Iteman dan model Rasch pada pengembangan instrumen kemampuan berpikir kritis peserta didik sekolah menengah kejuruan. *Jurnal Inovasi Pendidikan IPA*, *4*(2 SE-Articles), 224–233. https://doi.org/10.21831/jipi.v4i2.21442

Ocy, D. R., Rahayu, W., & Makmuri. (2023). Rasch Model Analysis: Development Of Hots-Based Mathematical Abstraction Ability Instrument According To Riau Islands Culture. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, *12*(4), 3542–3560.

Oktaviani, M., Hasanah, U., Faesal, M., Elmanora, & Zulfa, V. (2021). the Use of the Rasch Model in Analysing the Validity and Reliability of Students' Coping Strategy Scale. *JISAE: Journal of Indonesian Student Assessment and Evaluation*, *7*(2), 110–119. https://doi.org/10.21009/jisae.v7i2.22009

Piliyang, Y., Arafiyah, R., & Rahayu, W. (2025). Implementing Rasch Model for Modern Test Evaluation Using Python Prototype. *J-KOMA : Jurnal Ilmu Komputer Dan Aplikasi*, *8*(1), 27–32. https://doi.org/10.21009/j-koma.v8i1.04

PISA. (2022). PISA Result. *Journal Pendidikan*, 10. https://www.oecd.org/publication/pisa-2022-results/country-notes/malaysia-1dbe2061/

Pratama, D., & Husnayaini, I. (2020). Applying Rasch Model To Measure Students` Reading Comprehension. *JISAE: Journal of Indonesian Student Assessment and Evaluation*, *6*(2), 203–209. https://doi.org/10.21009/jisae.v6i2.14920

Prayoga, K. P., Suryana, D., Supriatna, M., & Budiman, N. (2024). Penggunaan Rasch Model Untuk Menganalisis Konstruk Instrumen Kontrol Diri Pada Siswa Sekolah Menengah. *G-Couns: Jurnal Bimbingan Dan Konseling*, *9*(1), 367–381. https://doi.org/10.31316/gcouns.v9i1.4459

Purba, S. E. D. (2018). Analisis model Rasch instrumen tes prestasi pada mata pelajaran dasar dan pengukuran listrik. *Wiyata Dharma: Jurnal Penelitian Dan Evaluasi Pendidikan*, *6*(2), 142–147. https://doi.org/10.30738/wd.v6i2.3393

Purnamasari, R., Maftuh, B., Hidayat, M., Warlim, W., Dahliyana, A., Ruyadi, Y., Syarief Nurdin, E., & Fauziah, N. (2026). Rasch Model Validation of Critical Thinking Assessment in Elementary Education: Validasi Instrumen Penilaian Berpikir Kritis Siswa Sekolah Dasar Berbasis Rasch. *Indonesian Journal of Innovation Studies*, *27*. https://doi.org/10.21070/ijins.v27i1.1842

Qadriyyah, Q., Suprananto, S., & Hayat, B. (2025). The Analysis of Validity and Reliability of the Indonesian's SCARED Instrument Using the Rasch Model. *JISAE: Journal of Indonesian Student Assessment and Evaluation*, *11*(1), 1–12. https://doi.org/10.21009/jisae.v11i1.54614

Ramadhan, A. S., & Handayani, I. (2025). Analisis Kualitas Instrumen Minat Belajar Dalam Pembelajaran Matematika: Pendekatan Model Rasch. *SIGMA: JURNAL PENDIDIKAN MATEMATIKA*, *17*, 353–367.

Ramdani, R., Hanurawan, F., Ramli, M., Lasan, B. B., & Afdal, A. (2020). Development and Validation of Indonesian Academic Resilience Scale Using Rasch Models. *International Journal*

of Instruction, *14*(1), 105–120. https://doi.org/10.29333/IJI.2021.1417A

Sabir, A., Mayong, M., & Usman, U. (2021). Analisis Soal Higher Order Thinking Skills (Hots) Berdasarkan Dimensi Kognitif. *INDONESIA: Jurnal Pembelajaran Bahasa Dan Sastra Indonesia*, *2*, 117. https://doi.org/10.26858/indonesia.v2i3.23971

Sadhu, S., & Laksono, E. W. (2021). Development and validation of an integrated assessment for measuring critical thinking and chemical literacy in chemical equilibrium. *International Journal of Instruction*, *11*(3), 557–572. https://doi.org/10.12973/iji.2018.11338a

Saraswati, P. M. S., & Agustika, G. N. S. (2020). Kemampuan Berpikir Tingkat Tinggi dalam Menyelesaikan Soal HOTS Mata Pelajaran Matematika. *Jurnal Ilmiah Sekolah Dasar*, *4*(2 SE-Articles), 257–269. https://doi.org/10.23887/jisd.v4i2.25336

Supriyadi, Haenilah, E. Y., & Maulina, D. (2026). Development and validation of the critical consciousness scale of environmental issues (CCS-EI): A RASCH model analysis. *Biosfer: Jurnal Pendidikan Biologi*, *19*(1), 382–395.

Syadiah, A. N., & Hamdu, G. (2020). Analisis rasch untuk soal tes berpikir kritis pada pembelajaran STEM di sekolah dasar. *Premiere Educandum : Jurnal Pendidikan Dasar Dan Pembelajaran*, *10*(2), 138–148. https://doi.org/10.25273/pe.v10i2.6524

Uli Sihombing, R., Naga, D. S., & Rahayu, W. (2020). A Rasch Model Measurement Analysis On Indonesian Science Literacy Test: Smart Way To Improve The Learning Assessment. *IJER - Indonesian Journal of Educational Review*, *6*(2 SE-Articles).

Wandyka, A., & Aini, F. Q. (2025). Penggunaan Model Rasch untuk Mengembangkan Instrumen Penilaian Keterampilan Proses Sains pada Materi Gaya Antarmolekul. *JagoMIPA: Jurnal Pendidikan Matematika Dan IPA*, *5*(1), 12–24. https://doi.org/10.53299/jagomipa.v5i1.1041