Jurnal Penelitian dan Evaluasi Pendidikan
Volume 29, No. 1, June 2025 (43-57)
Online: http://journal.uny.ac.id/index.php/jpep

# Literary comprehension competence among senior high school students in Purwokerto

**Memet Sudaryanto*; Aldi Nur Fadilah**
Universitas Jenderal Soedirman, Indonesia
*Corresponding Author. E-mail: memet.sudaryanto@unsoed.ac.id

**ARTICLE INFO**

**Article History**
**Submitted:**
13 February 2025
**Revised:**
21 April 2025
**Accepted:**
24 April 2025

**Keywords**
rasch model; item analysis; quest application

**Scan Me:**

**ABSTRACT**

Literature learning should be designed according to students' characteristics. However, there is still no available competency map that specifically measures students' understanding of literary competencies in schools. This study aims to measure and describe students' competencies in prose, poetry, and drama in Purwokerto. Quantitatively, this study employs a test method based on theoretical construction by developing 30 multiple-choice items that assess students' ability to understand intrinsic and extrinsic elements in prose, poetry, and drama. The study population comprised 5.187 10th-grade students in Purwokerto. A sample of 1.034 students was randomly selected from three public schools in Purwokerto using the PickerWheel application. Instrument validation was conducted using the Aiken Index, yielding a result above 0.75 with five experts. Ekspert merupakan guru dan dosen Bahasa Indonesia yang memenuhi kualifikasi dan kompetensi sebagai ahli di bidang sastra. Reliability was assessed using Cronbach's Alpha and the Test Information Function (TIF) to evaluate the items' ability to explain students' competencies. Through the Rasch Model, this study identified a range of item difficulty levels on a logit scale from -2.55 to 3.17, indicating that the difficulty level of the items falls within a normal range. Based on the analysis of student literature competencies at senior high schools in Purwokerto, the results indicate that the highest student ability score is 100, the lowest score is 0, with a median of 51.11, a standard deviation of 13.89, and the 75th percentile at 59.87.



**To cite this article (in APA style):**
Sudaryanto, M., & Fadilah, A. N. (2025). Literary comprehension competence among senior high school students in Purwokerto. *Jurnal Penelitian dan Evaluasi Pendidikan, 29(1)*, 43-57 doi: https://doi.org/10.21831/pep.v29i1.83343

## INTRODUCTION

The implementation of literary education in schools has been largely overlooked by key stakeholders, encompassing both school administrations and governmental bodies (Mirnawati, 2015). While assessments often focus on linguistic aspects, they neglect literature as a theoretical subject that students need to master. Literature strategically shapes students' personalities, knowledge, and language skills, which should be evaluated comprehensively (Aji, 2016; Ansari, 2020; Mansyur, 2016).

In 2024, Indonesia's Ministry of Education, Culture, Research, and Technology (Kemendikbudristek) launched the "Literature in Curriculum Program" as part of the Merdeka Curriculum implementation. This program aims to enhance students' reading interests and literacy skills by integrating literary works into school learning activities. Beginning in the 2024/2025 academic year, teachers should teach literature through a co-curricular approach across all levels of education, from elementary to senior high school.

As an initial step, Kemendikbudristek has prepared guidelines containing 177 recommended literary titles for use in teaching. The list includes 43 titles for elementary schools, 29 for middle schools, and 105 for high schools. These books have undergone a one-year


ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v29i1.83343

curation process and are expected to serve as relevant and engaging learning resources for students (Windiatmoko, 2024).

This context underscores the critical importance of evaluation as a foundation for policy-making before implementing literary learning materials in schools (Mustadi et al., 2021; Sunaryo et al., 2017). Currently, literary materials in schools are not systematically tailored to align with students' characteristics. Supported by the Merdeka Curriculum, which incorporates the concept of deep learning, literature holds a significant position to be developed specifically by teachers, enabling the optimal internalization of students' values and attitudes (Mellyzar et al., 2025). In order to meet the fundamental needs for literary materials development analysis, a competency map is required to measure and describe the literary profiles of high school students.

In high school literary education, students are expected to master three main aspects: literary appreciation, text analysis, and creative writing skills. Literary appreciation involves understanding various literary genres, including poetry, short stories, novels, and drama, as well as interpreting the meanings embedded in literary works. Text analysis focuses on understanding intrinsic elements (such as theme, plot, characters, setting, point of view, and language style) and extrinsic elements (including cultural, social, historical, and philosophical values) within a literary work (Gasong, 2019; Rachmah & Huda, 2021). Additionally, creative writing skills aim to train students to produce their own literary works by adhering to linguistic conventions and literary aesthetics. Mastery of these aspects equips students with the ability to understand, appreciate, and critically evaluate literary works more comprehensively.

The comprehension and analysis of literature in the high school curriculum can be reinforced by integrating relevant literary theories. For instance, Tzvetan Todorov's structuralism theory emphasizes text analysis based on its intrinsic elements, while Wolfgang Iser's reader-response theory highlights the reader's role in interpreting and assigning meaning to literary works. Additionally, M.H. Abrams' expressive theory asserts that literary works are expressions of the author's experiences, thoughts, and emotions, enabling students to grasp the context behind the creation of such works. Systematic evaluation is essential to ensure students' mastery of these literary aspects. Assessments can take various forms, including comprehension tests, literary text analyses, critical essays, and creative writing projects. These methods enable teachers to evaluate the extent to which students understand, analyze, and apply literary theories in their learning (Mardapi, 2012). Ultimately, such evaluations will enhance students' literacy skills and foster a deeper appreciation for literature in a comprehensive and meaningful way.

The theoretical framework for teaching prose, poetry, and drama in schools emphasizes the development of student's abilities to appreciate and analyze literary works with materials specifically designed by teachers. Based on preliminary studies, most analyses of the objectives of literature education include deepening students' appreciation of literary works, instilling moral values, and promoting local culture (Hambleton & Patsula, 1998; Kidwell, 2017). An effective teaching model, such as prose appreciation learning based on the inspiring experiences of literary figures, is used by teachers to help students understand and internalize the values embedded in these works. In practice, the teaching of prose, poetry, and drama can be carried out using various methods, including classroom discussions, text analysis, and creative writing assignments (Anderson & Krathwohl, 2001; Griffin, 1991; Nitko & Brookhart, 2011). The assessment in this study focuses on literary competence and is aligned with the material taught in class based on the lesson plans, learning resources, and core competencies outlined in the curriculum. The assessment of Grade X students in Purwokerto aims to evaluate their understanding of literature, including their ability to identify intrinsic and extrinsic elements of prose texts, as well as their capacity to interpret and appreciate literary works.

Research on the literary comprehension profiles of high school students in Indonesia indicates that the literacy levels in the Indonesian language remain low. For instance, a study conducted at SMA Muhammadiyah Maumere revealed that the average literacy proficiency of


ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v29i1.83343

students reached only 39%, which is classified as "Poor." In the context of literature assessment, evaluative approaches that focus on the aesthetic and qualitative aspects of literary works have become a primary focus (Fatmawati et al., 2024). This aligns with the implementation of the Kurikulum Merdeka (Independent Curriculum), which emphasizes literary appreciation through listening activities, whereby students are expected to evaluate and create insights from various literary works.

Other studies highlight the importance of literature education in shaping students' character, emphasizing the implementation of values derived from literary works through interactive teaching models (Oakleaf, 2009; Stiggins & Chappuis, 2012). Based on these findings, an ideal approach to literary assessment is one that not only measures students' understanding of the intrinsic and extrinsic elements of literary works but also evaluates their ability to appreciate, evaluate, and internalize the values contained within them. This comprehensive assessment approach is expected to enhance students' literary literacy skills while fostering better character development. This study aims to measure and describe students' competencies in prose, poetry, and drama in Purwokerto. To assess students' abilities in studying literature, this research constructed test indicators that were developed into a set of items. The test was administered to 10th-grade students to evaluate their competencies in prose, poetry, and drama. The measurement was conducted using the Rasch model by estimating the difficulty level of each item.

The Rasch model is one of the approaches within Item Response Theory (IRT) used to measure students' abilities more objectively and accurately. Developed by Georg Rasch, this model focuses on the relationship between a student's ability level and the difficulty level of test items. The Rasch model is based on the principle of one-parameter logistic probability (1PL), where the probability of a student answering an item correctly depends solely on the difference between the student's ability and the difficulty of the item (Widhiarso, 2013). In this model, both the student's ability parameter and the item's difficulty parameter are mapped onto the same scale (logit), enabling fairer and more meaningful comparisons.

The measurement process using the Rasch model involves analyzing students' response patterns to test items (Fayers, 2004; Woods & Baker, 1985). Each test item is analyzed for its difficulty level, and each student is assigned an estimated ability level based on their correct or incorrect answers. The data collected is analyzed using the R Program, which produces parameters for item difficulty levels and estimates of students' abilities. One of the main advantages of this model is that the measurement results are invariant, meaning that the difficulty level of the items does not depend on the group of students, and the student's ability levels do not depend on the type of items used.

Compared to Classical Test Theory (CTT), the Rasch model offers several significant advantages. In CTT, students' scores heavily depend on the number of items answered correctly, which can be influenced by test characteristics such as item difficulty and test length. In contrast, the Rasch model enables fairer measurement as each item is weighted according to its difficulty level, resulting in more accurate comparisons of students' abilities. Additionally, the Rasch model provides deeper analysis, such as detecting invalid items (misfits) and identifying biases in the measurement process (Hambleton et al., 1991; MW et al., 1997; Reise & Waller, 2009). For this reason, the model is widely used in educational research and the development of more adaptive and personalized computer-based assessment instruments.

## RESEARCH METHOD

The theoretical framework of this study guided the formulation of the following research questions: (1) What is the level of literary comprehension—covering prose, poetry, and drama—among Grade X high school students in Purwokerto? and (2) What is the profile of students'


ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v29i1.83343

literary comprehension in terms of test item characteristics? These questions were investigated using a quantitative approach, a survey method, and measurement techniques. Data were collected through a digitally administered dichotomous (true–false) test. The instrument was developed based on theoretically constructed indicators that were validated using Aiken's V. The research involved a randomly selected sample of 10th-grade students from three state high schools in Purwokerto, selected using the Picker Wheel randomization tool. Following the validation process, reliability estimation was carried out using Cronbach's alpha.

The data were analyzed to assess students' competencies in prose, poetry, and drama using a dichotomous scale. The analysis was conducted with the R program to examine sample adequacy, unidimensionality, item characteristics, and the test information function. Sample adequacy was determined using Slovin's formula, resulting in a minimum required sample size of 390 students. This study ultimately involved 1,020 high school students. A unidimensionality test was conducted to ensure that the literary competence test instrument measured a single construct, verified through the examination of scatterplots generated from Exploratory Factor Analysis (EFA). The study employed the Rasch measurement model, which was analyzed using the R programming language to describe item characteristics and evaluate the test information function.

The use of Aiken's Index in content validation offers advantages as it provides a quantitative measure of expert agreement, making it more objective compared to purely qualitative analysis. In this study, the calculation of Aiken's V will be carried out using formulas 1 and Formula 2, which take into account the number of raters and the rating scale employed. The results are then compared to critical values from Aiken's table to determine whether an item is valid or needs revision.

$$V = \sum S \; / \; [n(C-1)] \tag{1}$$

$$S = R - Lo \tag{2}$$

In the formulas, V represents Aiken's Index, S is the score obtained by subtracting the lowest possible rating score ($L_o$) from the score given by the rater (R). The variable $L_o$ denotes the lowest possible rating score (e.g., 1), while C is the highest possible rating score (e.g., 5), and n refers to the number of validators or raters involved. This method is widely applied in the development of research instruments such as questionnaires, academic assessments, and psychological scales, ensuring that each item within the instrument has strong relevance to the concept being measured. This method is widely applied in the development of research instruments such as questionnaires, academic tests, and psychological scales, ensuring that each item within the instrument has strong relevance to the measured concept**.**

## FINDINGS AND DISCUSSION

This study constructs a literary theory framework that integrates knowledge, skills, attitudes, and individual characteristics—dimensions which collectively enable individuals to optimize their potential in the domain of literature. Current theories on competence emphasize a holistic approach, which evaluates not only cognitive aspects but also affective and psychomotor domains. Competence measurement can be conducted through various methods, including performance-based tests, competency-based interviews, and observations in real or simulated environments. Standards for measuring competence must consider validity, reliability, and fairness to ensure that the results accurately reflect an individual's abilities in real-life situations (Retnawati, 2016, 2017).


ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v29i1.83343

Based on the findings from a literature review of instructional modules, it was identified that teachers deliver literary content encompassing three primary genres: prose, poetry, and drama. Prose includes works such as short stories and novels, focusing on narrative elements like plot, characterization, and setting. Poetry emphasizes the expression of emotions and ideas through condensed and rhythmic language, with attention to elements such as rhyme, rhythm, and figurative meaning. Drama, as a form of literature written for performance, highlights dialogue and action to bring stories to life on stage. Teaching these three genres aims to develop students' appreciation of literary works as well as their analytical and creative writing abilities.

Teachers' understanding of prose, poetry, and drama varies widely. Research indicates that some teachers face challenges in effectively teaching literary materials. For instance, a study identified issues in teaching literature in higher-grade classrooms, including a lack of in-depth understanding of the subject matter and limited teaching method variations. Additionally, a mapping of literature mastery among Indonesian language teachers in junior high schools in Sidoarjo Regency revealed that teachers' overall mastery of literary materials needs improvement.

Students' ability to comprehend and produce literary works also varies. A study conducted at MTs Nurul Khairiyah Sei Tuan found that the ability of eighth-grade students to write drama scripts was categorized as "sufficient," with an average score of 67.6. Another study at SMP Negeri 13 Bandar Lampung demonstrated improved skills in writing drama scripts after using short stories as a teaching aid (Priyandani et al., 2015; Sari & Lubis, 2017). However, overall, there is a pressing need to enhance students' abilities to appreciate and create literary works through more innovative and participatory teaching methods.

## Research Instrument Construction

The study of literature is identified through the materials taught in the curriculum and the teachers' development of those materials in the classroom. The materials are constructed based on literary research theories. This research draws on and constructs theories and findings from several sources: Siswanto (2008) on the introduction to literary theory; Purba (2010) on the introduction to literature; Lianawati (2019) on literature in Indonesia; Hermawan (2019) on the utilization of novel Seruni analysis; and Wuryani (2013) on literary works in the teaching of Indonesian language and culture (Hermawan, 2019; Lianawati, 2019; Purba, 2010; Siswanto, 2008; Wuryani, 2013). The materials are constructed based on literary research theories as outlined in Table 1.

The theoretical framework was developed based on a theoretical construction that was then extracted into additional indicators for confirmatory purposes. Based on the results of Focus Group Discussion (FGD), the construction was refined into: (1) intrinsic and extrinsic indicators for prose; (2) intrinsic and extrinsic indicators for poetry; and (3) intrinsic and extrinsic indicators for drama. These aspects were translated into measurable indicators and further developed into test items designed to assess these variables. Each indicator consists of 10 test items aimed at measuring the sub-indicators formulated during the expert validation process.

## Content Validation Using Aiken's V Index

Content validation is a process to assess the extent to which a measurement instrument reflects the construct or concept it aims to measure. One commonly used method in content validation is Aiken's Index. Aiken's Index is employed to measure expert agreement (expert judgment) on the relevance of items within an instrument based on an ordinal scale. The calculation of this index is based on experts' ratings of each item using a specified scale (e.g., 1–4 or 1–5), which are then analyzed using the Aiken's V formula to determine the level of


ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v29i1.83343

agreement or consistency among experts (Aiken, 1985; Retnawati, 2016). The closer the Aiken's V value is to 1, the more valid the item is in measuring the intended construct.

Table 1. Indicators for Assessing the Literary Competence of Senior High School Students

| | |
|---|---|
| Identifying intrinsic elements such as the setting in a novel. | Presented with a novel excerpt, students can identify the intrinsic element of the setting in the given novel. |
| Summarizing the theme of a poem. | Presented with a poem, students can infer the implied theme of the poem. |
| Identifying elements of drama performance. | Presented with characteristics of drama performance elements, students can identify the performance elements described. |
| Identifying the character traits of protagonists in a novel. | Presented with a novel excerpt, students can identify the character traits of the characters in the given novel. |
| Interpreting the meaning of song lyrics. | Presented with a lyric excerpt, students can interpret the meaning of the lyrics. |
| Identifying the moral message in a short story. | Presented with a short story, students can identify the moral message in the story. |
| Analyzing the structure of a short story. | Presented with a short story, students can analyze the structure of the story. |
| Identifying literary devices (figures of speech). | Presented with the definition of personification, students can identify answers containing this literary device. |
| Comparing rhyme schemes in poetry. | Presented with two poems, students can compare the rhyme schemes used. |
| Analyzing the structure of a short story. | Presented with a short story, students can analyze its structure. |
| Identifying the theme of a poem. | Presented with a poem, students can identify the theme of the given poem. |
| Determining extrinsic elements in a short story, such as social values. | Presented with a short story, students can determine the social value it contains. |
| Identifying types of drama. | Presented with a drama text, students can identify the type of drama. |
| Determining the narrative point of view. | Presented with a short story, students can determine the intrinsic element of point of view. |
| Comparing intrinsic elements of a short story and a novel. | Presented with characteristics of intrinsic elements in short stories and novels, students can compare the intrinsic elements of both. |
| Analyzing the values contained in a novel. | Presented with a novel excerpt, students can analyze the values contained within the novel. |
| Examining the physical elements of a poem. | Presented with a poem, students can examine the implicit elements, such as concrete words. |
| Summarizing elements of drama performance. | Students can distinguish the definitions of various drama performance elements. |
| Identifying methods of characterization in a novel excerpt. | Presented with a novel excerpt, students can identify methods of characterization based on the text's narration. |
| Determining types of imagery in poetry. | Presented with a poem, students can determine the type of imagery used in the poem. |
| Identifying character traits in a drama. | Presented with a drama text, students can identify character traits in the drama. |
| Predicting missing parts of a poem. | Presented with an incomplete poem, students can predict the missing parts. |
| Determining stage directions in a drama text. | Presented with a drama text, students can determine stage directions within the text. |
| Identifying gestures, facial expressions, and intonation in a drama text. | Presented with a drama text, students can identify the function of characters in the drama. |
| Analyzing intrinsic elements of a poem, such as theme and mood. | Presented with a poem, students can analyze the theme and mood of the poem. |
| Summarizing the moral message of a drama. | Presented with a drama text, students can infer the implied moral message of the drama. |
| Summarizing the content of a poem. | Presented with a poem, students can summarize its content. |
| Summarizing conflicts. | Presented with a drama excerpt, students can infer the conflict within the drama. |
| Identifying conflicts in a drama text. | Presented with a drama text excerpt, students can identify the conflict in the excerpt. |
| Comparing the theme, mood, and setting of a drama. | Presented with two drama texts, students can compare the themes, moods, and settings of the dramas. |


ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v29i1.83343

After entering the assessment scores provided by the experts, the difference between the highest score and the lowest score is calculated based on Aiken's formula. Table 2 presents the summary of Aiken's calculation for items 1 to 30.

Table 2. Distribution and Measurement of the Aiken Index

| | R1 | R2 | R3 | R4 | R5 | s1 | s2 | s3 | s4 | s5 | Σs | V |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| B1 | 5 | 5 | 5 | 5 | 5 | 4 | 4 | 4 | 4 | 4 | 20 | 1 |
| B2 | 5 | 5 | 5 | 5 | 5 | 4 | 4 | 4 | 4 | 4 | 20 | 1 |
| B3 | 5 | 2 | 3 | 5 | 5 | 4 | 1 | 2 | 4 | 4 | 15 | 0.75 |
| B4 | 5 | 5 | 5 | 3 | 5 | 4 | 4 | 4 | 2 | 4 | 18 | 0.9 |
| B5 | 4 | 5 | 5 | 4 | 5 | 3 | 4 | 4 | 3 | 4 | 18 | 0.9 |
| B6 | 5 | 4 | 5 | 4 | 5 | 4 | 3 | 4 | 3 | 4 | 18 | 0.9 |
| B7 | 5 | 4 | 4 | 5 | 5 | 4 | 3 | 3 | 4 | 4 | 18 | 0.9 |
| B8 | 5 | 5 | 4 | 5 | 5 | 4 | 4 | 3 | 4 | 4 | 19 | 0.95 |
| B9 | 5 | 5 | 5 | 3 | 5 | 4 | 4 | 4 | 2 | 4 | 18 | 0.9 |
| B10 | 5 | 4 | 5 | 5 | 5 | 4 | 3 | 4 | 4 | 4 | 19 | 0.95 |
| B11 | 5 | 5 | 5 | 3 | 5 | 4 | 4 | 4 | 2 | 4 | 18 | 0.9 |
| B12 | 5 | 2 | 5 | 3 | 5 | 4 | 1 | 4 | 2 | 4 | 15 | 0.75 |
| B13 | 5 | 5 | 5 | 5 | 5 | 4 | 4 | 4 | 4 | 4 | 20 | 1 |
| B14 | 5 | 4 | 5 | 5 | 5 | 4 | 3 | 4 | 4 | 4 | 19 | 0.95 |
| B15 | 5 | 4 | 5 | 4 | 4 | 4 | 3 | 4 | 3 | 3 | 17 | 0.85 |
| B16 | 5 | 4 | 5 | 5 | 4 | 4 | 3 | 4 | 4 | 3 | 18 | 0.9 |
| B17 | 5 | 4 | 5 | 5 | 5 | 4 | 3 | 4 | 4 | 4 | 19 | 0.95 |
| B18 | 5 | 5 | 4 | 3 | 4 | 4 | 4 | 3 | 2 | 3 | 16 | 0.8 |
| B19 | 5 | 4 | 5 | 3 | 5 | 4 | 3 | 4 | 2 | 4 | 17 | 0.85 |
| B20 | 5 | 5 | 5 | 5 | 5 | 4 | 4 | 4 | 4 | 4 | 20 | 1 |
| B21 | 5 | 5 | 5 | 5 | 5 | 4 | 4 | 4 | 4 | 4 | 20 | 1 |
| B22 | 5 | 4 | 5 | 5 | 5 | 4 | 3 | 4 | 4 | 4 | 19 | 0.95 |
| B23 | 5 | 5 | 5 | 5 | 5 | 4 | 4 | 4 | 4 | 4 | 20 | 1 |
| B24 | 5 | 5 | 5 | 5 | 5 | 4 | 4 | 4 | 4 | 4 | 20 | 1 |
| B25 | 5 | 3 | 5 | 3 | 5 | 4 | 2 | 4 | 2 | 4 | 16 | 0.8 |
| B26 | 5 | 1 | 4 | 5 | 5 | 4 | 0 | 3 | 4 | 4 | 15 | 0.75 |
| B27 | 5 | 4 | 5 | 5 | 4 | 4 | 3 | 4 | 4 | 3 | 18 | 0.9 |
| B28 | 5 | 4 | 5 | 3 | 4 | 4 | 3 | 4 | 2 | 3 | 16 | 0.8 |
| B29 | 5 | 4 | 5 | 5 | 5 | 4 | 3 | 4 | 4 | 4 | 19 | 0.95 |
| B30 | 5 | 1 | 5 | 5 | 5 | 4 | 0 | 4 | 4 | 4 | 16 | 0.8 |

In the calculation of Aiken's Index, B refers to the information identifying the item number, which helps distinguish each test item being evaluated. R represents the rater's assessment or score for the test item, reflecting their judgment of its relevance. S is the result of subtracting the lowest possible score in the rating scale from the score assigned by the rater, which standardizes the rating. Lastly, V denotes the final outcome of the Aiken's Index calculation, indicating the degree of agreement among raters regarding the validity of the item.

Based on the data distribution in Table 2, it can be concluded that all instrument items used to measure literary comprehension skills among tenth-grade high school students in Purwokerto are valid and suitable for use. This conclusion is drawn from both logical reasoning and rational analysis. The content validity test was conducted by examining whether the test items align with the specified test blueprint or table of specifications.

**Analysis of Students' Literary Comprehension Skills Using the Rasch Model**

Table 3 displays the estimation results for item information: items A1 through A10 measure the dimension of prose comprehension ability, items B1 through B10 measure the dimension of poetry comprehension ability, and items C1 through C10 measure the dimension of drama comprehension ability.


ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v29i1.83343

Table 3. Distribution of Difficulty Level Measurement

| Item | Difficulty Level |
|---|---|
| A1 | -0.02128550 |
| A2 | 0.91345089 |
| A3 | -1.11117678 |
| A4 | 2.02322319 |
| A5 | -0.85694866 |
| A6 | -0.43326388 |
| A7 | 2.82428330 |
| A8 | -1.47343699 |
| A9 | 1.42017786 |
| A10 | 0.66634352 |
| B1 | -2.55922374 |
| B2 | 0.68069625 |
| B3 | 0.03432282 |
| B4 | -0.36989525 |
| B5 | -0.96796204 |
| B6 | 3.13320968 |
| B7 | 2.71710823 |
| B8 | 2.32505955 |
| B9 | 0.55228157 |
| B10 | 3.17456691 |
| C1 | -2.15531220 |
| C2 | 1.86819671 |
| C3 | 1.01008241 |
| AdC4 | -0.99779575 |
| C5 | -1.17246393 |
| C6 | 1.14618848 |
| C7 | 1.14618848 |
| C8 | -1.59762484 |
| C9 | 0.31981620 |
| C10 | -0.06301518 |

In the context of test theory, the discrimination index for the literary competence test administered to Grade X high school students in Purwokerto indicates how effectively an item differentiates between students of lower and higher ability. In the Rasch model, the discrimination index is assumed to be equal for all items. This is because the Rasch model posits that all test items function uniformly in distinguishing students' abilities. In other words, the Rasch model emphasizes that the probability of correctly answering an item is solely determined by the gap between the difficulty level of the item and the student's ability.

The assumption of equal discrimination within the Rasch model has significant implications for educational measurement. It ensures that every test item contributes equally to distinguishing between varying levels of student ability, making the measurement process more


ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v29i1.83343

objective and independent of the tested population. This leads to fairer comparisons of students' abilities. However, if an item exhibits a significantly different discrimination index in practice (e.g., items that are either too easy or too difficult to effectively distinguish abilities), the Rasch model identifies such items as misfits. Misfit data can then be further analyzed to revise and improve the problematic items.

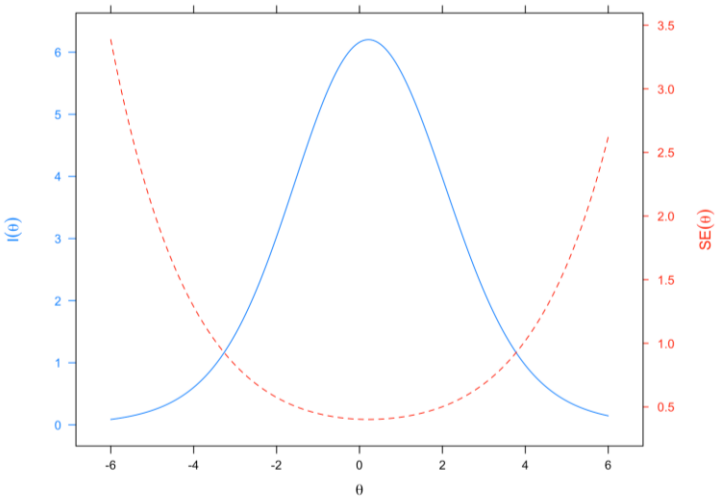


Figure 1. Test Information and Standard Errors

The Figure 1 illustrates Test Information and Standard Error. The blue line represents the information function for the abilities of grade 10 high school students, while the red line indicates the standard error. The relationship between item information and standard error is negative, meaning that as the information function increases, the standard error decreases.

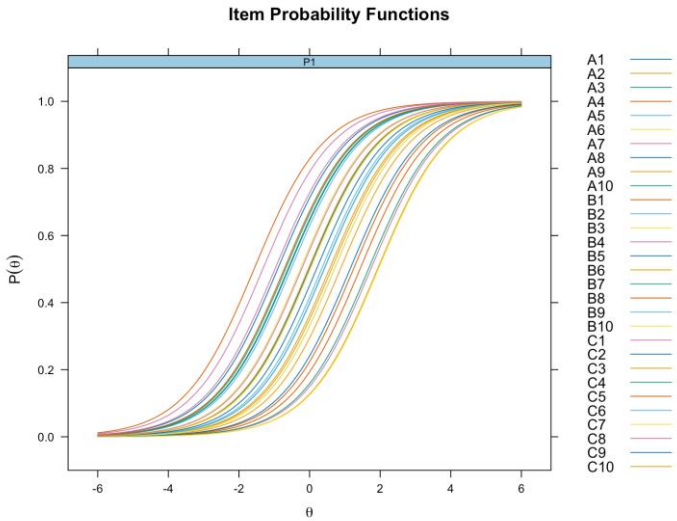


Figure 2. Item Probability Functions

Figure 2 shows the Item Probability Function (Item Characteristic Curve) in the Rasch Model. In the Rasch Model, the Item Characteristic Curve (ICC) is a graphical representation that illustrates the probability of a student with a specific ability level correctly answering an item. The curve depicts the relationship between the student's ability (θ) and the probability of a correct response (P(θ)) for an item with a certain difficulty level (b). The ICC in the Rasch Model takes the form of a one-parameter logistic curve (1PL) governed by the following formula


ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v29i1.83343

3. Formula 3 describes the probability that a student with a certain ability level (θ) will answer an item correctly. In this formula, P(X = 1 | θ) represents that probability, while e is the exponential constant used in the logistic function. The variable θ denotes the student's ability, expressed in logits, and b refers to the item's difficulty level, also expressed in logits.

$$P(X = 1|\theta) = \frac{e^{(\theta - b)}}{1 + e^{(\theta - b)}} \quad (3)$$

The ICC has a single inflection point at b=θ, where the probability of answering correctly is 0.5 (50%). In other words, when a student's ability θ matches the item's difficulty bbb, they have a 50% chance of answering correctly. All ICCs in the Rasch Model share the same shape, differing only in their horizontal position, which shifts based on the item's difficulty level. This is because, in the Rasch Model, the discrimination parameter is assumed to be constant across all items. Another critical curve in the Rasch Model is the Item Information Curve (IIC). This curve shows the amount of information provided by an item at various levels of ability (θ). The IIC is essential for evaluating how effectively an item measures individuals at specific ability levels. Items provide the most information when their difficulty (b) is close to the ability level (θ) of the students being measured. Figure 3 illustrates an example of an Item Information Curve in the Rasch Model.

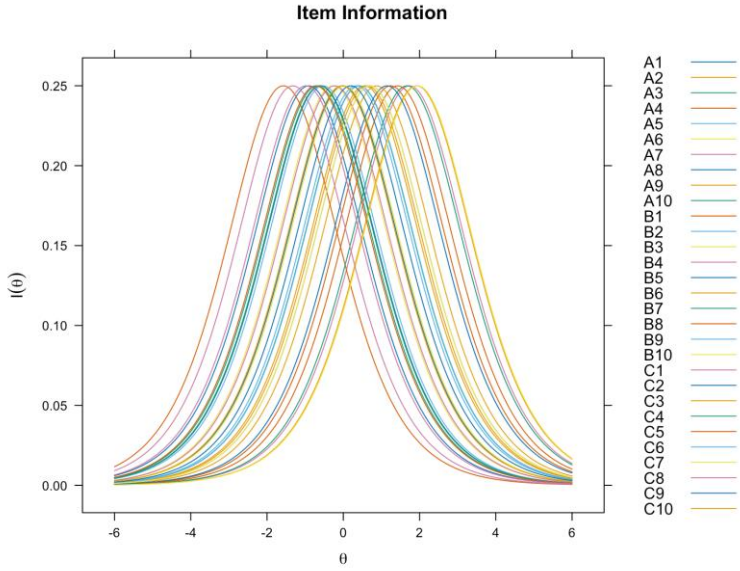


Figure 3. Item Information

The higher the peak of the Item Information Curve (IIC) at a specific ability level, the better the item is at providing information for individuals at that ability level. Below is the IIC for the literature competency test administered to 10th-grade high school students in Purwokerto. Based on the curve, each item reaches the same peak height, indicating that the items provide excellent information across the measured abilities.

**Interpretation of data analysis on literary comprehension ability of 10th-grade high school students**

Based on the analysis of literary comprehension abilities of 10th-grade high school students in Purwokerto using the Rasch model, it was found that the test-takers' abilities were highly diverse. From the conducted analysis and examination of the resulting curves, the


ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v29i1.83343

findings can be summarized in the Table 4. The data in the Table 4 illustrates that the highest score achieved by the test participants, specifically students of Grade X at a state high school in Purwokerto, was 100, while the lowest score was 0. The median, or the middle value after the data is arranged in order, was 51.1, which was also the same as the mode (the value that appears most frequently). The standard deviation, a statistical measure indicating how far the data in a set deviates from the mean, was calculated to be 13.89.

From the test results, the score of 42 was identified as the 25th percentile, meaning that 25% of the answers were below this score, while 75% were above it. On the other hand, the score of 59 was the 75th percentile, indicating that 75% of all measurements were below this score, and 25% were above it.

Table 4. Descriptive Statistics of Participants' Theta Scores

| Statistic | Value |
|---|---|
| Highest Score | 100 |
| Lowest Score | 0 |
| Median | 51, 11 |
| Mode | 51, 11 |
| Standard Deviation | 13, 89 |
| 25th Percentile | 42, 18 |
| 50th Percentile (Median) | 51, 11 |
| 75th Percentile | 59, 87 |

Some aspects of prose questions were relatively easy for students. Based on the level of difficulty, identifying characterization methods in novel excerpts was the easiest aspect. This is because characterization is often explicitly explained in the text's narrative. Additionally, characters in novels generally have more detailed backgrounds, aiding students in understanding and analyzing these characters. Other types of questions—such as identifying extrinsic elements of short stories (e.g., social values), determining their moral messages, and recognizing intrinsic elements like the setting in novels—were also found to be relatively easy. This ease stems from the clear presentation patterns in the stories and the frequent teaching of such elements in literature lessons, making students more familiar with identifying them. As noted by Hakim & Utami (2024), students who are accustomed to terms and theories taught in class find it easier to identify settings and other elements in novels or short stories, as they have clear analytical tools to apply (Hakim & Utami, 2024).

However, some prose questions were more challenging, such as analyzing the structure of short stories. This difficulty arises because the structure of short stories is not always presented explicitly, requiring deeper understanding. Barthes & Duisit (1975) explain that narratives operate on several levels or strata, and to comprehend a narrative, readers must navigate not only the plot but also various levels of understanding (Barthes & Duisit, 1975). For instance, elements such as character motivation, symbolism, or overarching themes may be embedded in the narrative and can only be uncovered through in-depth reading and structural analysis.

Comparing the intrinsic elements of short stories and novels was another challenging task because it required students to understand the distinct characteristics of these two forms of prose. Moussa and Amer (2024) assert that short stories and novels differ in the complexity of their intrinsic elements, such as narrative structure, character development, and themes (Moussa & Amer, 2024). Novels, with their broader scope, allow for deeper and more complex character development than short stories. Students need to grasp these differences to make effective


ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v29i1.83343

comparisons. Additionally, identifying characters' traits in novels and determining narrative perspectives were also quite challenging, as these often required an implicit understanding of the text and its narrative. Understanding elements such as setting, plot, characterization, and narrative perspective in novels demands high-level analytical skills. Students must be able to identify how these elements function in the context of each prose form, which poses a unique challenge. This process requires attention to detail and deep reflection on the literary work being analyzed.

Poetry questions are considered easier for students, based on their level of difficulty, and include concluding the theme of a poem and identifying its theme. The ease of concluding and identifying a poem's theme arises because the theme is often explicitly conveyed within the overall meaning of the poem. Blohm et al. (2022) explain that students use different strategies to comprehend poetry, often focusing on sound patterns and rhythms that help them recognize themes explicitly. Furthermore, poetry themes are easier to identify because familiar genre categories allow readers to form initial expectations about the poem's content, thereby speeding up the comprehension process (Blohm et al., 2022). Similarly, comparing rhyme patterns in poetry is relatively simple, as rhyme schemes can be recognized through the recurring sound patterns at the end of poetic lines. These types of questions are easier because the elements being tested are concrete and frequently taught in literature lessons at school.

Some poetry questions, however, present a higher level of difficulty, such as analyzing the physical elements of poetry that involve implicit concrete words and summarizing the poem's content. These challenges arise because concrete words in poetry are often symbolic, requiring a deeper understanding of implied meaningsIn line with Chaudhuri et al. (2024), understanding poetry relies not only on its literal meaning but also on interpreting nuances and deeper connotations (Chaudhuri et al., 2024). Additionally, the variability inherent in poetry adds complexity to the process of drawing conclusions, as readers must identify and interpret unexpected elements. Determining types of imagery in poetry and predicting missing sections of poems are also challenging tasks, as they require a deep understanding of figurative language and poetic structure, which do not always follow clear patterns.

Drama-related questions that are considered easier for students, based on difficulty levels, include identifying elements of drama staging, summarizing the moral messages of dramas, identifying character traits in dramas, summarizing drama staging elements, and identifying conflicts within drama texts. The ease of identifying staging elements stems from the fact that components such as set design and props are often explicitly described in the text. As proposed by Ifianti & Fitriani (2022), moral messages, character traits, and conflicts are conveyed through character interactions and storylines containing moral lessons, making it easier for students to grasp the characters and dynamics involved in dramas (Ifianti & Fitriani, 2022). Furthermore, summarizing the moral messages of dramas is relatively straightforward because such messages are usually tied to the moral lessons that can be derived from the storyline. Identifying character traits in dramas is also easier because dialogues and character interactions often clearly indicate their personalities.

Certain drama-related questions, however, are more difficult, such as comparing themes, atmospheres, and settings between dramas, identifying types of dramas, determining stage directions within drama texts, and identifying gestures, facial expressions, and intonations of characters in dramas. The difficulty in comparing themes, atmospheres, and settings arises because students must analyze two texts simultaneously and identify their differences. Identifying types of dramas is also challenging, as it requires an understanding of various drama genres and their characteristics. According to Marantika (2014), these elements are difficult for students because they demand in-depth analysis and an understanding of various drama genres and characteristics (Marantika, 2014). Moreover, the focus of current teaching methods remains theoretical, with limited role-playing experiences, which exacerbates the challenges. Meanwhile,


ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v29i1.83343

determining stage directions, gestures, facial expressions, and intonations requires a deeper understanding of script interpretation and non-verbal elements in drama performances, which are not always explicitly stated in the text.

## CONCLUSION

Using the Rasch model, this study produced varying difficulty levels of items, ranging from a logit scale of -2.55 to 3.17, indicating that the item difficulty levels fall within the normal range. Based on the analysis of student competencies in high schools in Purwokerto, the highest student score was 100; the lowest score was 0, with a median of 51.11, a standard deviation of 13.89, and the 75th percentile at 59.87. Comparing intrinsic elements of short stories and novels was one of the most challenging tasks, as it required students to understand the unique characteristics of both prose forms presented in the questions. Some poetry questions were also more difficult, such as analyzing physical elements involving implicit concrete words and summarizing the content of the poem. Similarly, certain drama questions were highly challenging, such as comparing themes, atmospheres, and settings, identifying types of dramas, determining stage directions, and analyzing gestures, facial expressions, and intonations of characters. This study serves as a reference for Indonesian language teachers in Purwokerto to gauge students' abilities—particularly their literary competence—enabling teachers to develop instructional materials and assessments aligned with each student's proficiency level.

### Conflict of interests

There are no known conflicts of interest associated with this publication.

## REFERENCES

Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. Educational and Psychological Measurement, 45(1), 131–142. https://doi.org/10.1177/0013164485451012

Aji, W. N. (2016). Model pembelajaran Dick and Carrey dalam pembelajaran bahasa dan sastra Indonesia. Kajian Linguistik Dan Sastra, 1(2), 119–126.

Anderson, L. W., & Krathwohl, D. R. (2001). A Taxonomy for Learning, Teaching, and Assessing. In Longman New York.

Ansari, K. (2020). Arah pembelajaran bahasa dan sastra indonesia pada era revolusi industri 4.0. Pustaka Diksi.

Barthes, R., & Duisit, L. (1975). An introduction to the structural analysis of narrative. New Literary History, 6(2), 237–272.

Blohm, S., Versace, S., Methner, S., Wagner, V., Schlesewsky, M., & Menninghaus, W. (2022). Reading poetry and prose: Eye movements and acoustic evidence. Discourse Processes, 59(3), 159–183.

Chaudhuri, S., Pickering, A., Dooley, M., & Bhattacharya, J. (2024). Beyond the words: Exploring individual differences in the evaluation of poetic creativity. Plos One, 19(10), e0307298.

Fatmawati, F., Haryati, G., Rustan, N. A., & Istiqomah, A. (2024). Profil Kemampuan Literasi Bahasa Indonesia Siswa Kelas Xii Sma Muhammadiyah Maumere. Lingua Rima: Jurnal Pendidikan Bahasa Dan Sastra Indonesia, 13(1).


ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v29i1.83343

Fayers, P. (2004). Item Response Theory for Psychologists. Quality of Life Research. https://doi.org/10.1023/b:qure.0000021503.45367.f2

Gasong, D. (2019). Apresiasi Sastra Indonesia.

Griffin, P. (1991). Literacy Assessment: Merging Teaching, Learning, and Assessment.

Hakim, H. R., & Utami, S. F. (2024). Analisis Struktural Cerpen "Assyarru Bis-Syarri" Karya Abdul Fattah Shobri & Ali Umar. Enggang: Jurnal Pendidikan, Bahasa, Sastra, Seni, Dan Budaya, 4(2), 172–182.

Hambleton, R. K., Hariharan, S., & Rogers, H. J. (1991). Fundamentals of Item Response Theory. SAGE, 2. https://doi.org/10.2307/2075521

Hambleton, R., & Patsula, L. (1998). Adapting tests for use in multiple languages and cultures. Social Indicators Research. https://doi.org/10.1023/A:1006941729637

Hermawan, D. (2019). Pemanfaatan hasil analisis novel seruni karya almas sufeeya sebagai bahan ajar sastra di sma. METAMORFOSIS| Jurnal Bahasa, Sastra Indonesia Dan Pengajarannya, 12(1), 11–20.

Ifianti, T., & Fitriani, E. (2022). Drama dalam Pembelajaran Bahasa Inggris: Tantangan Dan Harapan. ALFABETA: Jurnal Bahasa, Sastra, Dan Pembelajarannya, 5(1), 10–18.

Kidwell, T. (2017). Understanding Culture: A Literature Review Comparing Three Cultural Pedagogies. Register Journal, 10(2). https://doi.org/10.18326/rgt.v10i2.221-233

Lianawati, W. S. (2019). Menyelami keindahan sastra Indonesia. Bhuana Ilmu Populer.

Mansyur, U. (2016). Inovasi pembelajaran bahasa indonesia melalui pendekatan proses. RETORIKA: Jurnal Bahasa, Sastra, Dan Pengajarannya, 9(2), 256786.

Marantika, J. E. R. (2014). Drama dalam pembelajaran bahasa dan sastra. Tahuri, 11(2), 92–102.

Mardapi, D. (2012). Pengukuran Penilaian & Evaluasi Pendidikan. In Nuha Medika. Nuha Media.

Mellyzar, M., Rochintaniawati, D., Riandi, R., Fakhrah, F., & Virijai, F. (2025). The Merdeka Curriculum and P5: A Review of Teachers' Understanding and Students' Experiences. Jurnal Wahana Pendidikan, 12(1), 123–140.

Mirnawati, M. (2015). Tinjauan terhadap Problematika Pembelajaran Sastra Indonesia pada Pendidikan Formal. Aksara, 16(1).

Moussa, M. A., & Amer, A. N. E. S. (2024). The Academic Motivation Scale: Evaluation Evidence of Intrinsic, Extrinsic, and Amotivation in Faculty of Education Students. International Journal of Psychology and Educational Studies, 11(4), 283–294.

Mustadi, A., Amelia, R., Budiarti, W. N., Anggraini, D., Amalia, E., & Susandi, A. (2021). Strategi pembelajaran keterampilan berbahasa dan bersastra yang efektif di Sekolah Dasar. UNY Press.

MW, der Linden, W. J. Van, & Hambleton, R. K. (1997). Handbook of Modern Item Response Theory. Journal of the American Statistical Association. https://doi.org/10.2307/2965612

Nitko, A. J., & Brookhart, S. M. (2011). Educational Assessment of Students. Human Movement Science. https://doi.org/10.1016/j.humov.2005.02.001


ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v29i1.83343

Oakleaf, M. (2009). Writing information literacy assessment plans: A guide to best practice. Communications in Information Literacy. https://doi.org/10.15760/comminfolit.2010.3.2.73

Priyandani, Y., Suliani, N. N. W., & Mustofa, A. (2015). Pembelajaran menulis teks cerita pendek pada kelas VII SMPN 13 Bandar Lampung. Jurnal Kata: Bahasa, Sastra, Dan Pembelajarannya, 3(4).

Purba, A. (2010). Pengantar Ilmu Sastra. USUpress.

Rachmah, S., & Huda, M. (2021). Realisasi Pembelajaran dalam Jaringan pada Pelajaran Bahasa Indonesia di Sekolah Menengah Pertama. Jurnal Pendidikan Bahasa Dan Sastra Indonesia Undiksha, 11(3), 338–345.

Reise, S. P., & Waller, N. G. (2009). Item Response Theory and Clinical Measurement. Annual Review of Clinical Psychology. https://doi.org/10.1146/annurev.clinpsy.032408.153553

Retnawati, H. (2016). Proving Content Validity of Self-Regulated Learning Scale (The comparison of Aiken index and Expanded Gregory Index). REiD (Research and Evaluation in Education), 2(2), 155–164. https://doi.org/10.21831/reid.v2i2.11029

Retnawati, H. (2017). Learning Trajectory of Item Response Theory Course using Multiple Softwares. Olympiads in Informatics 11, 123–142. https://doi.org/10.15388/ioi.2017.10

Sari, R. F., & Lubis, F. (2017). Kemampuan Menulis Naskah Drama Siswa Kelas VIII MTs Nurul Khairiyah Sei Tuan Tahun Pembelajaran 2016/2017. Asas: Jurnal Sastra, 6(2), 195972.

Siswanto, W. (2008). Pengantar teori sastra. Grasindo.

Stiggins, R. J., & Chappuis, J. (2012). An introduction to student involved assessment for learning. Pearson Assessment Training Institute.

Sunaryo, H., Zuriah, N., & Kusniarti, T. (2017). Model adaptasi nilai kearifan lokal dalam pembelajaran sastra berkarakter. SOSIOHUMANIKA, 10(1), 107–116.

Widhiarso, W. (2013). Aplikasi Model Rasch Campuran dalam Mengevaluasi Pengukuran Harga Diri. Jurnal Penelitian Dan Evaluasi Pendidikan. https://doi.org/10.21831/PEP.V17I1.1367

Windiatmoko, D. U. (2024, June 21). *Menerawang program sastra masuk kurikulum*. Balai Pengembangan dan Pembinaan Bahasa.

Woods, A., & Baker, R. (1985). Item response theory. Language Testing. https://doi.org/10.1177/026553228500200202

Wuryani, W. (2013). Pesona karya sastra dalam pembelajaran bahasa dan budaya indonesia. Semantik, 2(2), 87–101.


ISSN (print) 2685-7111 | ISSN (online) 2338-6061