Primary: Jurnal Pendidikan Guru Sekolah Dasar
journal homepage: https://primary.ejournal.unri.ac.id/
contents lists available at SINTA and DOAJ

Vol. 15 No. 3 (2026): June
http://dx.doi.org/10.33578/jpfkip.v15i3.607-622

Developing a multicultural education-based assessment instrument for Indonesian language learning in fourth-grade elementary schools

Ani Satul Inajati [1*], Ngatmini [1], Asrofah [1]

[1] Universitas PGRI Semarang, Semarang, Indonesia, 50126

Article info

Keywords: assessment instrument, elementary school, Indonesian language learning, multicultural education.

Abstract

This study developed a multicultural education-based assessment instrument for fourth-grade Indonesian language learning that integrates reading-comprehension assessment with indicators of tolerance, empathy, cooperation, and cultural appreciation. A research and development design adapted from Gall et al. was implemented in eight stages: needs analysis, information gathering, product design, expert validation, revision, limited trial, product refinement, and field implementation. The study involved three experts, five Grade IV teachers, and 108 students across the development stages; the final implementation and student response testing involved 77 students from three elementary schools in Semarang. The final product consists of an assessment blueprint, reading-comprehension tasks, and a multicultural attitude rubric. Expert validation yielded an overall mean of 92%, indicating high validity. Teacher practicality testing reached 96%, indicating that the instrument was feasible and easy to use. Student responses yielded a mean score of 92.3%, interpreted as high acceptability and perceived usefulness rather than evidence of causal learning gains. The instrument provides teachers with a curriculum-aligned assessment tool for linking Indonesian language learning with multicultural character development, while further comparative and longitudinal testing is recommended.

* Corresponding Author.
E-mail address: anisatulinajati@gmail.com (Ani Satul Inajati)
DOI: http://dx.doi.org/10.33578/jpfkip.v15i3.607-622
Received 13 December 2025; Received in revised form 24 May 2026; Accepted 25 May 2026
Available online 30 June 2026
e-ISSN 2598-5949 | p-ISSN 2303-1514 © The Authors.

1. Introduction

Primary education plays a decisive role in developing literacy, character, and social competence in culturally diverse classrooms. Assessment in this context cannot be reduced to checking whether students remember information from a text. It must also help teachers notice how students interpret social situations, respect differences, cooperate with peers, and express empathy through language. The Merdeka Curriculum places assessment within a cycle of planning,

Primary: Jurnal Pendidikan Guru Sekolah Dasar is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

learning, feedback, and improvement, while also requiring schools to strengthen the Pancasila Student Profile, including global diversity, cooperation, and social responsibility (Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi [Kemdikbudristek], 2024). This orientation is consistent with international scholarship showing that culturally responsive teaching and assessment become meaningful when students' cultural backgrounds are treated as learning resources rather than as peripheral classroom attributes (Abacioglu et al., 2020; Alhanachi et al., 2021; Aronson & Laughter, 2016; Civitillo et al., 2018; Gay, 2018; Ladson-Billings, 1995).

Evidence from different countries strengthens this argument. In the United States, Byrd (2016), Bonner et al. (2018), Min et al. (2022), and Lawson et al. (2024) reported that culturally relevant practices, teacher agency, and valid CRT-related measures are closely connected with students' engagement and with teachers' capacity to teach diverse learners. In the Netherlands and Germany, Alhanachi et al. (2021), Kehl et al. (2024), Schachner et al. (2016), and Ialuna et al. (2024, 2025) showed that diversity climate, teacher-student relationships, and critical awareness shape students' belonging and adjustment. Similar concerns appear in Asian and transnational contexts. Kong et al. (2022) examined rural English teachers' engagement with culturally responsive teaching in China; Guan et al. (2023) found that culturally responsive music teaching strengthened students' ethnic identity in a multi-ethnic school; Nguyen and Huynh (2023) described how Vietnamese teachers developed agency when teaching ethnic minority students; Hu et al. (2021) emphasized field experience in preparing pre-service teachers; and Al Alaleeli (2022), Acquah and Szelei (2020), Chuang et al. (2020), Copeland Solas and Kamalodeen (2022), and Meihami (2023) demonstrated that culturally responsive pedagogy needs explicit modelling, contextual adaptation, and valid instruments across teacher education, technology-supported learning, science, and language classrooms.

Assessment research also supports a broader view of classroom measurement. Andrade and Brookhart (2020) positioned classroom assessment as co-regulation of learning, while Panadero et al. (2018) and Lipnevich and Panadero (2021) argued that assessment and feedback become useful when they help learners judge, revise, and improve their work. Ramadhan et al. (2020) showed that teacher-made instruments require careful evidence of validity and reliability. These findings are important for Indonesian language learning because reading comprehension tasks can simultaneously measure textual understanding and reveal students' value-based responses to characters, conflicts, cultural practices, and social interaction. A good instrument must therefore connect cognitive indicators with observable multicultural indicators in a way that teachers can actually use.

Recent Indonesian studies have further shown that the issue is urgent in elementary education. Kurniaman et al. (2025) found that learning and assessment studies in Indonesian elementary schools remain fragmented and often have a local impact. Sari and Kumalasani (2025), Nainggolan (2025), Purnamasari et al. (2025), and Anggraini and Ramadan (2023) demonstrated the importance of local culture, culturally responsive learning, collaborative projects, and values of global diversity in elementary classrooms. Nora et al. (2023), Putra et al. (2024), and Hulkin and Shaleh (2024) highlighted the need to integrate character values and affective assessment into learning tools. Studies by Zufriady et al. (2025), Rofiqi et al. (2025), Wibowo et al. (2021), Guslinda et al. (2025), Muhajang et al. (2025), Nuraida et al. (2025), and Nurhaliza et al. (2026) add related evidence on multicultural instruments, culturally responsive teaching trends, Melayu culture-based materials, differentiated assessment, diversity-oriented media, and character learning. Collectively, these studies confirm that Indonesian elementary education needs instruments that are valid, practical, culturally grounded, and aligned with classroom realities.

Primary: Jurnal Pendidikan Guru Sekolah Dasar is licensed under
a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.
[e-**ISSN 2598-5949** | p-**ISSN 2303-1514**]

Preliminary observations and teacher interviews in the three participating schools revealed the same problem at the classroom level. Teachers mostly used publisher-made worksheets and teacher-made cognitive tests. These tools helped students answer comprehension questions, yet they rarely provided explicit criteria for tolerance, empathy, cooperation, or appreciation of cultural diversity. Teachers also reported difficulty translating multicultural values into observable indicators. This local evidence indicates that the research problem is not abstract. It appears in daily assessment routines, in the selection of reading materials, and in the absence of rubrics that help teachers fairly assess students' socio-cultural responses.

The scholarly gap addressed in this study is the limited availability of empirically developed assessment instruments that integrate indicators of reading comprehension and multicultural attitudes into a single classroom-ready tool. Previous work has discussed culturally responsive pedagogy, character education, locally culture-based learning, and affective assessment, but few studies have developed an Indonesian-language assessment instrument that teachers can use to assess comprehension and multicultural values simultaneously. This study, therefore, positions multicultural values not as decorative moral messages but as assessable learning indicators observable in students' reading responses, classroom interactions, and rubric-based performance.

Accordingly, this study aims to develop and evaluate a multicultural education-based assessment instrument for Grade IV Indonesian language learning. The study examines expert validity, teacher practicality, and students' responses during field implementation. Its novelty lies in integrating reading tasks, multicultural texts, a scoring blueprint, and rubric descriptors for tolerance, empathy, cooperation, and cultural appreciation. The contribution is modest but important: it offers teachers a practical way to assess literacy and multicultural character within the same Indonesian language learning process.

2. Method

This study employed a research and development (R&D) design adapted from Gall et al. (2003) to produce and evaluate a multicultural education-based assessment instrument for Grade IV Indonesian language learning. The development process was organized into eight stages: (1) needs analysis, (2) information gathering, (3) preliminary product design, (4) expert validation, (5) first revision, (6) limited trial, (7) final product revision, and (8) field implementation. The research was conducted at SD Negeri Pedurungan Kidul 03, SD Negeri Pedurungan Kidul 02, and SD Muhammadiyah 08 Semarang. Participants consisted of three expert validators, five Grade IV teachers, and 108 students across the development stages. To clarify the sample-size difference noted in the results, the 108 students represented all participants across the trial process, whereas the final implementation and student response testing involved 77 students from the three schools. Figure 1 illustrates the eight-stage development process used to transform the needs analysis into a validated, classroom-tested assessment instrument.

Primary: Jurnal Pendidikan Guru Sekolah Dasar is licensed under
a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.
[e-**ISSN 2598-5949** | p-**ISSN 2303-1514**]

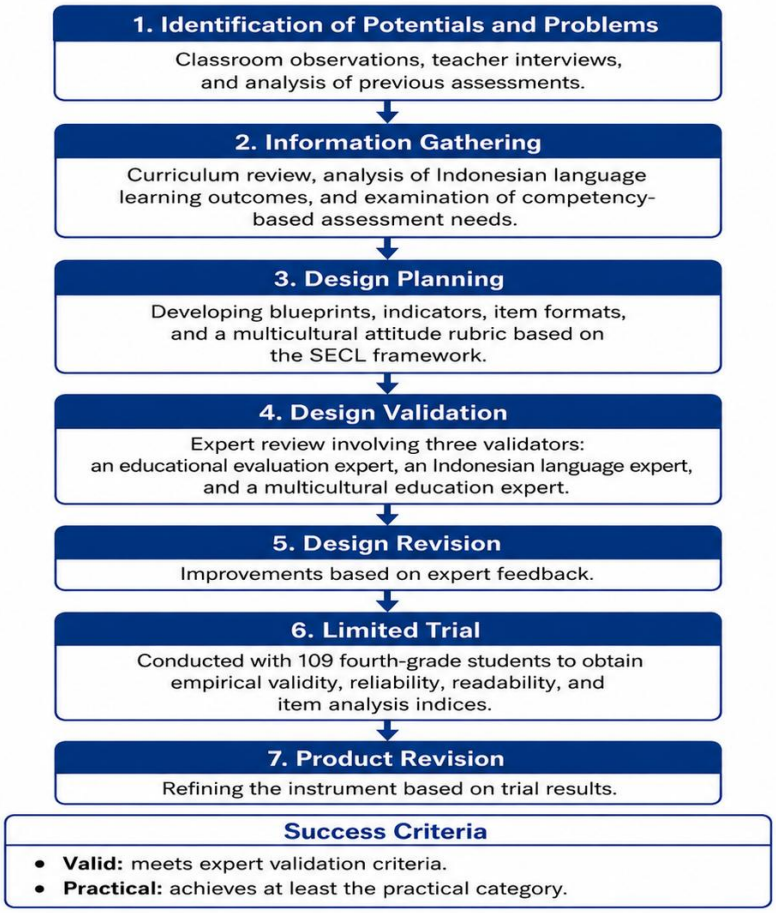

Figure 1. Research flowchart

The flowchart shows that product development moved systematically from identifying classroom problems to field implementation, so each revision was grounded in evidence rather than the Author's assumptions. The steps taken for each stage are explained as follows:

Needs analysis was conducted through classroom observation, teacher interviews, and document review of previously used assessment tools. The analysis focused on the types of assessment used, the extent to which multicultural values were represented, and teachers' needs for a practical rubric. The information-gathering stage involved reviewing Grade IV Indonesian language learning outcomes, Merdeka Curriculum assessment guidance, multicultural education theory, and classroom assessment principles (Banks, 2020; Brookhart, 2013; Kemdikbudristek, 2024).

The product design stage produced three main components: an assessment blueprint that aligns learning outcomes with reading-comprehension indicators; reading-comprehension tasks using multicultural texts and contexts; and a multicultural attitude rubric. The rubric operationalized tolerance, empathy, cooperation, and cultural appreciation into observable descriptors, enabling teachers to score students consistently during learning activities.

Expert validation involved one content expert, one language expert, and one educational evaluation expert. The validators assessed content accuracy, item construction, language quality, and integration of multicultural values. Validation scores were calculated by dividing the obtained score by the maximum score and multiplying by 100%, consistent with expert-judgment principles

Primary: Jurnal Pendidikan Guru Sekolah Dasar is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

in content validation (Aiken, 1985). The product was considered valid when the score reached the predetermined validity category. Qualitative comments from the experts were used to revise wording, indicators, and rubric descriptors.

The limited trial generated evidence on item clarity, readability, empirical item quality, difficulty, discrimination, and the rubric's usability. Reliability and item analysis were interpreted using classical test principles (Cronbach, 1951; Ebel & Frisbie, 1991). Item and rubric revisions were made when wording was ambiguous, when the difficulty was judged to be too high or too low for Grade IV students, or when indicators did not directly reflect the intended multicultural construct. This procedure was intended to reduce construct bias and improve the consistency of classroom use.

The field implementation evaluated the practicality of the teacher and students' responses to the final product. Practicality was measured using teacher questionnaires and interpreted using descriptive percentage criteria (Agustina, 2016). Student response was used to describe acceptability and perceived usefulness. Because this study did not employ a comparison group or a longitudinal pretest-posttest design, student responses are reported as perceived effectiveness rather than as causal evidence of improved learning outcomes. All quantitative data were analyzed descriptively using percentage scores, and qualitative comments were used to interpret revision needs.

3. Results

This study produced a multicultural education-based assessment instrument through eight R&D stages. The results are presented in the following order: product development, validation, revision, practicality, and student response evidence.

3.1 Identification of potentials and problems

The needs analysis showed that teachers predominantly used cognitive reading-comprehension questions and publisher-made worksheets. These tools supported basic comprehension checks but did not provide criteria for observing tolerance, empathy, cooperation, or appreciation of cultural diversity during Indonesian language learning. Teachers also needed a concise rubric that could translate multicultural values into observable classroom behaviors. This finding became the basis for developing a contextual, culturally grounded assessment instrument aligned with the Merdeka Curriculum.

3.2 Information gathering

Information was collected through a review of multicultural assessment theories, an analysis of authentic assessment approaches within the Merdeka Curriculum, an examination of Indonesian Language learning outcomes, and a review of assessment instruments used in schools. Findings reinforce the need for assessments to reflect students' cultural contexts, promote social interaction, and embed values of tolerance within the learning process.

3.3 Design planning

In this stage, three product components were developed. The components were an assessment blueprint, reading-comprehension indicators, multicultural attitude indicators (tolerance, empathy, cooperation, and cultural appreciation), reading tasks, and a rubric for scoring multicultural attitudes.

Primary: Jurnal Pendidikan Guru Sekolah Dasar is licensed under
a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.
[e-**ISSN 2598-5949** | p-**ISSN 2303-1514**]

The design aligned with the learning outcomes, learning objectives, and the learning sequence in the Merdeka Curriculum, enabling the instrument to be used in regular Grade IV Indonesian language lessons.

3.4 Design validation

The instrument was validated by three experts (content expert, language expert, and educational evaluation expert). The validation results show that all assessed components met the high validity criterion. The overall mean was 92%, with the highest score in the integration of multicultural values (94%). These results indicate that the instrument was conceptually appropriate for pilot testing after minor revision. Table 1 presents the expert validation results for content accuracy, item construction, language quality, multicultural integration, and layout clarity.

Table 1. Expert validation summary

Component Evaluated	Validator 1	Validator 2	Validator 3	Average (%)	Category
Content Accuracy	90%	92%	93%	92%	Highly Valid
Item Construction	88%	91%	90%	90%	Highly Valid
Language Quality	89%	93%	92%	91%	Highly Valid
Integration of Multicultural Values	93%	95%	94%	94%	Highly Valid
Overall Mean				**92%**	**Highly Valid**

3.5 Design revision

Minor revisions were carried out based on expert feedback, including: Simplifying item wording, improving punctuation and grammar, refining indicators of empathy and tolerance to enhance measurability. Revisions followed the product validity criteria presented below. Table 2 defines the validity categories used to interpret the expert validation percentages and to determine whether the product was ready for trial implementation. The criteria in Table 2 helped the researchers make a transparent decision about product feasibility before moving from expert review to classroom trials.

Table 2. Validity criteria

Interval	Category
$80 < P \leq 100$	Highly Valid
$60 < P \leq 80$	Valid
$40 < P \leq 60$	Moderately Valid
$20 < P \leq 40$	Low Validity
$0 < P \leq 20$	Invalid

3.6 Limited trial

The limited trial was conducted with 108 fourth-grade students to evaluate: empirical validity, reliability, readability, level of difficulty, and item discrimination. The limited trial indicated that the instrument was understandable for Grade IV students and usable for classroom implementation after revision. Items and rubric descriptors that were unclear or showed inappropriate difficulty were revised before the field implementation stage.

Primary: Jurnal Pendidikan Guru Sekolah Dasar is licensed under
a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

3.7 Product revision

Following the limited trial, several items were revised because their difficulty level was not appropriate for the target learners. The rubric descriptors were also simplified to improve scoring consistency for classroom teachers.

3.8 Field implementation (operational trial)

Field testing was conducted at three elementary schools: Pedurungan Kidul 03 Public Elementary School, Pedurungan Kidul 02 Public Elementary School, and Muhammadiyah 08 Elementary School, Semarang. The implementation phase involved teacher practicality testing and student response testing to describe acceptability and perceived usefulness. Table 3 summarizes teachers' practicality ratings for the display, content, language, and multicultural components of the assessment instrument.

Table 3. Practicality test results

No.	Assessment Aspect	Obtained Score	Maximum Score	Percentage	Practicality Level
	Interface Display				
1	The interface design of the multicultural-based assessment instrument is attractive.	38	40	95%	Very Practical
	Relevance of the Assessment Instrument				
2	Alignment of the assessment instrument with the Merdeka Curriculum	40	40	100%	Very Practical
3	Alignment of the assessment instrument with learning outcomes	40	40	100%	Very Practical
4	Alignment of the assessment instrument with learning objectives	40	40	100%	Very Practical
5	The presentation of the assessment instrument is systematic and logical according to the students' comprehension levels	40	40	100%	Very Practical
6	Relevance of the questions in the multicultural-based assessment instrument	39	40	98%	Very Practical
	Language				
7	Appropriateness of the language with students' level of thinking	39	40	98%	Very Practical
8	All information in the evaluation tool is clearly readable	36	40	90%	Very Practical
9	The language used is easy to understand	38	40	95%	Very Practical
10	Sentences used are effective and unambiguous	37	40	93%	Very Practical
11	The use of communicative language	39	40	98%	Very Practical
12	Spelling used in the evaluation tool follows the Indonesian Dictionary (KBBI)	38	40	95%	Very Practical
13	Clarity in delivering information in the evaluation tool	36	40	90%	Very Practical
14	The use of punctuation follows proper Indonesian language rules	40	40	100%	Very Practical
	Presentation				
15	The visual illustrations in the assessment instrument are attractive	36	40	90%	Very Practical

Primary: Jurnal Pendidikan Guru Sekolah Dasar is licensed under
a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.
[e-**ISSN 2598-5949** | p-**ISSN 2303-1514**]

No.	Assessment Aspect	Obtained Score	Maximum Score	Percentage	Practicality Level
	Multicultural Education				
16	The assessment instrument demonstrates mutual respect	37	40	93%	Very Practical
17	The assessment instrument reflects aspects of cultural diversity	39	40	98%	Very Practical
18	The assessment instrument shows love for the homeland	40	40	100%	Very Practical
19	The evaluation questions demonstrate tolerance toward differences	40	40	100%	Very Practical
	Total	**732**	**760**	**96%**	**Very Practical**

The practicality test conducted with Grade IV teachers from three elementary schools showed that the multicultural education-based assessment instrument achieved a very high level of practicality. Using descriptive percentage analysis and practicality criteria adapted from Agustina (2016), the total practicality score reached 96%, indicating that teachers considered the instrument clear, relevant, and feasible for classroom use. The highest-rated aspects were curriculum alignment, learning-objective alignment, systematic presentation, and tolerance-related items. The language indicators also received high ratings, showing that the instructions and item wording were understandable for elementary students.

Some teachers suggested minor visual improvements and clearer examples of rubric use. These comments were incorporated into the final product revision, allowing the instrument to be implemented without major modification. The multicultural component received consistently high ratings, particularly indicators related to mutual respect, cultural diversity, love of homeland, and tolerance. These results support the instrument's relevance for integrating character-oriented assessment into Indonesian language learning.

Overall, the practicality evidence indicates that teachers can use the instrument as a classroom-ready assessment tool. The instrument combines pedagogical alignment, language clarity, visual support, and multicultural values in a format that is manageable for Grade IV classrooms.

3.9 Student response and perceived effectiveness results

Student response testing was conducted with 77 students from three elementary schools to describe the acceptability and perceived usefulness of the instrument. The questionnaire evaluated clarity of language, ease of reading, relevance of the materials, perceived learning benefits, and alignment with multicultural education values. The results are interpreted as student response evidence rather than as causal evidence of learning gains because no comparison group, effect-size calculation, or long-term measurement was used.

Table 4 reports students' responses from the three participating elementary schools and shows how learners perceived the clarity, usefulness, and multicultural relevance of the instrument. Table 4 shows that the three schools achieved very high student response scores, ranging from 92% to 93%, with an average of 92.3%. SD Negeri Pedurungan Kidul 03 recorded the highest score (93%), followed by SD Negeri Pedurungan Kidul 02 (92%) and SD Muhammadiyah 08 Semarang (92%). These results indicate strong student acceptance of the instrument in Grade IV Indonesian language classrooms.

Primary: Jurnal Pendidikan Guru Sekolah Dasar is licensed under
a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

Table 4. Student response results

No	School Name	Total Score	Maximum Score	Percentage	Student Response Category
1	SD Negeri Pedurungan Kidul 03	1.046	1.120	93%	Very High
2	SD Negeri Pedurungan Kidul 02	1.031	1.120	92%	Very High
3	SD Muhammadiyah 08 Semarang	1.028	1.120	92%	Very High
Average		**1.035**	**1.120**	**92.3%**	**Very High**

The response pattern was consistent across public and private school contexts. The most prominent aspects were clarity of instructions, relevance of reading materials, perceived learning benefits, and sensitivity to multicultural values. Minor differences among schools may reflect variations in classroom context and student characteristics.

Accordingly, the student response results support the feasibility and perceived usefulness of the assessment instrument. Claims regarding improvement in reading comprehension or multicultural behavior should be interpreted cautiously and require stronger outcome evidence in future studies, such as pretest-posttest designs, comparison groups, effect sizes, or longer implementation periods.

In summary, the R&D process produced an assessment instrument consisting of a blueprint, reading tasks, multicultural indicators, and a rubric. Expert validation showed a high level of validity (overall mean = 92%), teacher practicality reached 96%, and student response reached 92.3%. These findings support the instrument's validity, practicality, and perceived usefulness for Grade IV Indonesian language learning, while stronger evidence is needed before making causal claims about learning improvement.

4. Discussion

The findings show that the multicultural education-based assessment instrument for Grade IV Indonesian language learning met the criteria of expert validity, teacher practicality, and high student acceptance. This result supports the idea that assessment should be part of learning, not merely an administrative device for assigning scores. Andrade and Brookhart (2020) described classroom assessment as a co-regulatory process involving teachers, students, materials, and context. In the present study, this principle is evident in how the instrument links reading tasks, multicultural texts, and observable rubric indicators. Students do not only answer questions. They also demonstrate their understanding of fairness, difference, cooperation, and empathy through their responses to Indonesian language materials.

The alignment among learning outcomes, reading comprehension indicators, and multicultural rubric descriptors can explain the high expert validation score. The validators judged the product positively because the instrument did not separate literacy from character. It embedded multicultural values into reading activities, consistent with the asset-based orientation advocated by Aronson and Laughter (2016), Civitillo et al. (2018), and Smith et al. (2022). This alignment also responds to Lawson et al. (2024), who stressed the need for clearer measures of culturally responsive practice. In assessment terms, the product moves from broad moral vocabulary to observable performance indicators. That is the central strength of the instrument.

Primary: Jurnal Pendidikan Guru Sekolah Dasar is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

The instrument also reflects international evidence that culturally responsive education requires teacher competence, critical reflection, and contextual sensitivity. Abacioglu et al. (2020) found that multicultural attitudes and perspective-taking support culturally responsive teaching. Alhanachi et al. (2021) showed that professional learning communities can strengthen teachers' culturally responsive competences. Hu et al. (2021) and Acquah and Szelei (2020) emphasized the importance of modeling and field-based experience in preparing teachers, while Kehl et al. (2024) warned that cultural responsiveness can become superficial when teachers overlook structural bias and microaggressions. These studies help explain why the present instrument includes not only reading questions but also rubric indicators that guide teachers in observing concrete classroom behavior. A rubric can discipline attention. It makes values visible.

Teacher practicality results indicate that the instrument is usable because its instructions, item format, and rubric descriptors are accessible for classroom implementation. This finding is meaningful because a theoretically strong instrument will have little effect if teachers find it too complex. Ramadhan et al. (2020) showed that teacher-made test instruments often need stronger validity and reliability evidence; Panadero et al. (2018) and Lipnevich and Panadero (2021) added that assessments must generate feedback that students and teachers can act upon. In this study, teachers rated the product as practical because it uses familiar Indonesian-language learning activities while providing a clearer structure for observing multicultural values. The instrument, therefore, extends existing classroom practice rather than replacing it with an unfamiliar procedure.

Student responses were also high. This result suggests that the reading materials, instructions, and task format were understandable and relevant for Grade IV learners. The finding is consistent with Byrd (2016), who showed that students' perceptions matter in evaluating culturally relevant teaching, and with Ialuna et al. (2024, 2025), who connected culturally responsive teaching, classroom climate, teacher-student relationships, and students' adjustment. The present evidence should still be read carefully. Student response scores indicate acceptance and perceived usefulness; they do not prove long-term changes in multicultural behavior. The current data are stronger for feasibility than for causal effectiveness.

The Indonesian evidence strengthens the contextual interpretation of these findings. Kurniaman et al. (2025) showed that reading, learning, and assessment practices in Indonesian elementary schools still need a stronger connection to broader educational use. Sari and Kumalasani (2025), Nainggolan (2025), and Guslinda et al. (2025) demonstrated that local culture can make elementary learning more meaningful. Purnamasari et al. (2025) linked culturally responsive collaborative projects to social awareness, and Anggraini and Ramadan (2023) connected elementary education to global diversity values in the Pancasila Student Profile. These findings support the present product because the instrument uses Indonesian language learning as a space for both literacy development and cultural meaning-making.

Several local studies also clarify why the instrument's affective and multicultural dimensions are necessary. Nora et al. (2023) and Putra et al. (2024) emphasized character values in elementary learning, while Hulkin and Shaleh (2024) discussed character-based affective assessment in modern learning. Zufriady et al. (2025) developed a multicultural instrument for arts, culture, and crafts learning; Rofiqi et al. (2025) mapped global trends in culturally responsive teaching; Wibowo et al. (2021) showed that multicultural-based Indonesian teaching materials could improve critical thinking; Muhajang et al. (2025) proposed differentiated assessment for inclusive teaching; Nuraida et al. (2025) designed media for learning diversity; and Nurhaliza et al. (2026) highlighted the role of technology in character learning. The present study complements these works by focusing specifically on assessment in Grade IV Indonesian language learning.

Primary: Jurnal Pendidikan Guru Sekolah Dasar is licensed under
a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.
[e-**ISSN 2598-5949** | p-**ISSN 2303-1514**]

The product's novelty lies in its integration. Reading comprehension and multicultural attitude assessment are brought into a single instrument. This integration helps teachers connect students' cognitive understanding of texts with their ability to interpret social differences respectfully. It also supports the Merdeka Curriculum's emphasis on formative, authentic, and character-oriented assessment (Kemdikbudristek, 2024). Rather than treating multicultural attitudes as general moral statements, the rubric turns them into observable classroom evidence: listening to peers, responding respectfully, recognizing cultural differences, and cooperating during learning tasks.

Nevertheless, the claims must remain proportional to the evidence. The study involved three elementary schools in one city and relied on expert validation, teacher practicality, and student response data. The results do not yet establish causal effects on reading achievement or long-term multicultural behavior. Stronger evidence would require a comparison group, repeated implementation, more complete item analysis, and longitudinal observation. Future studies should also examine whether the instrument performs equally well across rural schools, urban multicultural classrooms, schools with diverse linguistic backgrounds, and digital learning environments. Such expansion would make the instrument more robust and more useful for Indonesian elementary education.

Practically, teachers may use the instrument before, during, or after reading lessons. The reading tasks can assess comprehension, while the rubric can guide observation of empathy, cooperation, tolerance, and cultural appreciation. Assessment results can inform feedback, remediation, enrichment, and selection of texts that better reflect students' cultural worlds. Overall, the study contributes a validated and practical assessment model that links literacy learning with multicultural character development in elementary education. Its value lies not in claiming a complete solution, but in offering a workable instrument that helps teachers see what conventional reading tests often miss.

5. Conclusion and implications

This study concludes that the multicultural education-based assessment instrument developed for Grade IV Indonesian language learning is valid, practical, and positively accepted by students. The final product consists of an assessment blueprint, reading-comprehension tasks, multicultural indicators, and a rubric for assessing tolerance, empathy, cooperation, and cultural appreciation. Expert validation produced an overall mean of 92%, teacher practicality reached 96%, and student response reached 92.3%. These findings indicate that the instrument is feasible for classroom use in the tested elementary school contexts.

The implication is that teachers can use the instrument to connect Indonesian language assessment with multicultural character development, especially when teaching texts that involve social interaction and cultural diversity. Teacher training is recommended so that rubric scoring, feedback, and follow-up instruction are applied consistently. Broader testing across regions and school characteristics, and longer instructional periods, are needed to evaluate learning gains and behavioral development more rigorously.

6. Limitations

This study has several limitations. It involved only three schools in one district, so generalization to other educational contexts should be made cautiously. The field implementation was conducted over a limited period and did not use a comparison group, effect-size analysis, or long-term follow-up. Therefore, student response results should not be interpreted as definitive evidence of learning

Primary: Jurnal Pendidikan Guru Sekolah Dasar is licensed under
a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.
[e-**ISSN 2598-5949** | p-**ISSN 2303-1514**]

improvement or durable multicultural behavior change. In addition, differences in teacher scoring skill, school context, and classroom conditions were not examined in depth. Future studies should report more complete reliability and item analysis indices and test the instrument in broader, more diverse contexts.

Credit authorship contribution statement

Ani Satul Inajati contributed to the conceptualization, methodology, formal analysis, data curation, data collection, validation, and preparation of the original manuscript draft. **Dr. Ngatmini, M.Pd,** provided supervision, resources, methodological guidance, validation, project administration, and contributed to the review and editing of the manuscript. **Dr. Asrofah, M.Pd,** contributed through supervision, methodological guidance, validation, funding acquisition, formal analysis, and critical review of the manuscript. All authors have read and approved the final version of this article.

Declaration of competing interest

The authors declare that they have no financial interests or personal relationships that could have influenced the work reported in this paper.

Data Availability Statement

The data that support the findings of this study are available from the corresponding Author upon reasonable request. Because the data involves elementary school participants, the dataset is not publicly shared to protect participant privacy and confidentiality.

Ethical Declaration

This research was conducted in accordance with ethical standards for educational research involving human participants. Permission was obtained from the participating schools, and teachers, parents or guardians, and students were informed about the study objectives, procedures, voluntary participation, and the right to withdraw without consequence. Students participated with age-appropriate assent under school supervision. No personal identities were reported; all data were anonymized, stored securely, and used only for academic purposes. The authors affirm that the study upholds honesty, transparency, confidentiality, and academic integrity.

Acknowledgement

The authors would like to thank the school principals, teachers, and expert validators who participated and provided valuable support during this study. Appreciation is also extended to the Postgraduate Program of Universitas PGRI Semarang for the academic guidance and research facilities provided. This research was self-funded and received no external financial support.

Declaration of AI Statement

Portions of this manuscript were edited with the assistance of ChatGPT (OpenAI) for language refinement, clarity, and formatting consistency. The authors reviewed, verified, and approved all revised content and take full responsibility for the final manuscript.

References

Abacioglu, C. S., Volman, M., & Fischer, A. H. (2020). Teachers' multicultural attitudes and perspective-taking abilities as factors in culturally responsive teaching. British Journal of Educational Psychology, 90(3), 736-752. https://doi.org/10.1111/bjep.12328

Primary: Jurnal Pendidikan Guru Sekolah Dasar is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

Acquah, E. O., & Szelei, N. (2020). The potential of modeling culturally responsive teaching: Pre-service teachers' learning experiences. Teaching in Higher Education, 25(2), 157-173. https://doi.org/10.1080/13562517.2018.1547275

Agustina, D. (2016). Analisis kepraktisan instrumen penilaian pembelajaran di sekolah dasar. Alfabeta.

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

Al Alaleeli, S. (2022). Empowering female pre-service teachers through culturally responsive teaching: Shared and divergent views of Arab and expatriate teacher educators in the United Arab Emirates. Education 3-13, 50(5), 592-611. https://doi.org/10.1080/03004279.2021.1876133

Alhanachi, S., de Meijer, L. A. L., & Severiens, S. E. (2021). Improving culturally responsive teaching through professional learning communities: A qualitative study in Dutch pre-vocational schools. International Journal of Educational Research, 105, Article 101698. https://doi.org/10.1016/j.ijer.2020.101698

Andrade, H. L., & Brookhart, S. M. (2020). Classroom assessment as the co-regulation of learning. Assessment in Education: Principles, Policy & Practice, 27(4), 350-372. https://doi.org/10.1080/0969594X.2019.1571992

Anggraini, T. A., & Ramadan, Z. H. (2023). Implementation of 21st-century learning in cultivating the value of global diversity at Elementary School 109, Pekanbaru. Primary: Jurnal Pendidikan Guru Sekolah Dasar, 12(5), 1395-1406. https://doi.org/10.33578/jpfkip.v12i5.9921

Aronson, B., & Laughter, J. (2016). The theory and practice of culturally relevant education: A synthesis of research across content areas. Review of Educational Research, 86(1), 163-206. https://doi.org/10.3102/0034654315582066

Banks, J. A. (2020). Multicultural education: Issues and perspectives (10th ed.). Wiley.

Bonner, P. J., Warren, S. R., & Jiang, Y. H. (2018). Voices from urban classrooms: Teachers' perceptions on instructing diverse students and using culturally responsive teaching. Education and Urban Society, 50(8), 697-726. https://doi.org/10.1177/0013124517713820

Brookhart, S. M. (2013). How to create and use rubrics for formative assessment and grading. ASCD.

Byrd, C. M. (2016). Does culturally relevant teaching work? An examination from student perspectives. SAGE Open, 6(3), 1-10. https://doi.org/10.1177/2158244016660744

Chuang, H.-H., Shih, C.-L., & Cheng, M.-M. (2020). Teachers' perceptions of culturally responsive teaching in technology-supported learning environments. British Journal of Educational Technology, 51(6), 2442-2460. https://doi.org/10.1111/bjet.12921

Civitillo, S., Juang, L., & Schachner, M. K. (2018). Challenging beliefs about cultural diversity in education: A synthesis and critical review of training with pre-service teachers. Educational Research Review, 24, 67-83. https://doi.org/10.1016/j.edurev.2018.01.003

Copeland Solas, E., & Kamalodeen, V. (2022). Culturally relevant pedagogies (CRP) and culturally responsive teaching (CRT) in science education: Black success stories in Ontario. Canadian Journal of Science, Mathematics and Technology Education, 22(4), 796-817. https://doi.org/10.1007/s42330-022-00236-z

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. Psychometrika, 16(3), 297-334. https://doi.org/10.1007/BF02310555

Ebel, R. L., & Frisbie, D. A. (1991). Essentials of educational measurement (5th ed.). Prentice Hall.

Gall, M. D., Gall, J. P., & Borg, W. R. (2003). Educational research: An introduction (7th ed.). Allyn & Bacon.

Primary: Jurnal Pendidikan Guru Sekolah Dasar is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

Gay, G. (2018). Culturally responsive teaching: Theory, research, and practice (3rd ed.). Teachers College Press.

Guan, T., Luo, N., & Matsunobu, K. (2023). Nurturing student ethnic identity through culturally responsive music teaching in China. International Journal of Music Education, 41(4), 598-615. https://doi.org/10.1177/02557614221132550

Guslinda, Munjiatun, Sari, I. K., & Alifarani, A. (2025). Development of teaching materials based on Melayu culture IPAS learning in elementary schools. Jurnal PAJAR (Pendidikan dan Pengajaran), 9(2), 10116. https://doi.org/10.33578/pjr.v9i2.10116

Hammond, Z. (2015). Culturally responsive teaching and the brain: Promoting authentic engagement and rigor among culturally and linguistically diverse students. Corwin.

Hu, X., Xu, Z., Neshyba, M., Geng, Z., & Turner, R. (2021). A multi-dimensional model: Implications for preparing pre-service teachers for culturally responsive teaching. Asia-Pacific Journal of Teacher Education, 49(3), 282-299. https://doi.org/10.1080/1359866X.2020.1753169

Hulkin, M., & Shaleh. (2024). Character-based affective assessment concept in modern elementary school learning. Jurnal PAJAR (Pendidikan dan Pengajaran), 8(6), 9805. https://doi.org/10.33578/pjr.v8i6.9805

Ialuna, F., Civitillo, S., & Jugert, P. (2024). Culturally responsive teaching, teacher-student relationship, and school belongingness: A multi-informant study in ethnically diverse classrooms. Learning, Culture and Social Interaction, 47, Article 100839. https://doi.org/10.1016/j.lcsi.2024.100839

Ialuna, F., Civitillo, S., Schachner, M. K., & Jugert, P. (2025). Culturally responsive teaching self-efficacy and cultural diversity climate are positively associated with the academic and psychological adjustment of immigrant and nonimmigrant students. Cultural Diversity and Ethnic Minority Psychology, 31(4), 683-698. https://doi.org/10.1037/cdp0000697

Kehl, J., Krachum Ott, P., Schachner, M., & Civitillo, S. (2024). Culturally responsive teaching in question: A multiple case study examining the complexity and interplay of teacher practices, beliefs, and microaggressions in Germany. Teaching and Teacher Education, 152, Article 104772. https://doi.org/10.1016/j.tate.2024.104772

Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi. (2024). Panduan pembelajaran dan asesmen: Pendidikan anak usia dini, pendidikan dasar, dan pendidikan menengah (Edisi revisi tahun 2024). Badan Standar, Kurikulum, dan Asesmen Pendidikan. https://pskp.kemendikdasmen.go.id/file/download/1751862432_file.pdf

Kong, D., Zou, M., & Chen, J. (2022). English as a foreign language teacher engagement with culturally responsive teaching in rural schools: Insights from China. Frontiers in Psychology, 13, Article 990363. https://doi.org/10.3389/fpsyg.2022.990363

Kurniaman, O., Hadriana, Islami, N., Putra, Z. H., Radini, & Al Fajar, B. (2025). Trends in learning and assessment in Indonesian elementary schools: Opportunities and challenges. Primary: Jurnal Pendidikan Guru Sekolah Dasar, 14(3), 374-388. https://doi.org/10.33578/jpfkip.v14i3.p374-388

Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. American Educational Research Journal, 32(3), 465-491. https://doi.org/10.3102/00028312032003465

Lawson, T. K., Knox, J., Romero, E., Molina Palacios, A., & Fallon, L. M. (2024). Check yourself! Exploring current culturally responsive teaching assessment measures. Psychology in the Schools, 61(6), 2649-2667. https://doi.org/10.1002/pits.23189

Lipnevich, A. A., & Panadero, E. (2021). A review of feedback models and theories: Descriptions, definitions, and conclusions. Frontiers in Education, 6, Article 720195. https://doi.org/10.3389/feduc.2021.720195

Primary: Jurnal Pendidikan Guru Sekolah Dasar is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

Meihami, H. (2023). Culturally responsive teaching in TEFL programs: Exploring EFL teachers' experiences. English Teaching & Learning, 47(1), 119-144. https://doi.org/10.1007/s42321-022-00108-7

Min, M., Lee, H., Hodge, C., & Croxton, N. (2022). What empowers teachers to become social justice-oriented change agents? Influential factors on teacher agency toward culturally responsive teaching. Education and Urban Society, 54(5), 560-584. https://doi.org/10.1177/00131245211027511

Muhajang, T., Sukmanasa, E., Budiana, S., & Gani, R. A. (2025). Developing a differentiated assessment model for teaching at the Right Level: Improving the quality of inclusive teaching in elementary school. Jurnal PAJAR (Pendidikan dan Pengajaran), 9(3), 115. https://doi.org/10.33578/pjr.v9i3.115

Nainggolan, M. (2025). Integrating Toba Batak culture in elementary school learning: An effort to strengthen the local identity of the nation's children. Primary: Jurnal Pendidikan Guru Sekolah Dasar, 14(5), 812-825. https://doi.org/10.33578/jpfkip.v14i5.p812-825

Nguyen, C. D., & Huynh, T.-N. (2023). Teacher agency in culturally responsive teaching: Learning to teach ethnic minority students in the Central Highlands of Vietnam. Educational Review, 75(4), 719-743. https://doi.org/10.1080/00131911.2021.1974346

Nora, Y., Jamaris, & Marsidin, S. (2023). The validity of character value social science learning media for fifth-grade elementary students. Primary: Jurnal Pendidikan Guru Sekolah Dasar, 12(3), 818-828. https://doi.org/10.33578/jpfkip.v12i3.9904

Nuraida, Setiawan, D., & Yusnadi. (2025). Development of interactive learning media based on Smart App Creator on the material of building self-identity in diversity to improve the learning outcomes of fourth-grade elementary school students. Jurnal PAJAR (Pendidikan dan Pengajaran), 9(6), 354. https://doi.org/10.33578/pjr.v9i6.354

Nurhaliza, F., Putra, Z. H., Hermita, N., & Deprizon. (2026). 21st-century character learning: The role of media and technology in primary schools during the transition period. Jurnal PAJAR (Pendidikan dan Pengajaran), 10(1), 388. https://doi.org/10.33578/pjr.v10i1.388

Panadero, E., Andrade, H., & Brookhart, S. M. (2018). Fusing self-regulated learning and formative assessment: A roadmap of where we are, how we got here, and where we are going. The Australian Educational Researcher, 45(1), 13-31. https://doi.org/10.1007/s13384-018-0258-y

Paris, D., & Alim, H. S. (Eds.). (2017). Culturally sustaining pedagogies: Teaching and learning for justice in a changing world. Teachers College Press.

Purnamasari, I., Prasetyo, S. A., Rofian, & Cholifah, N. (2025). Collaborative project model based on a culturally responsive teaching approach to strengthening social awareness. Primary: Jurnal Pendidikan Guru Sekolah Dasar, 14(4), 539-553. https://doi.org/10.33578/jpfkip.v14i4.p539-553

Putra, E. D., Maftuh, B., & Supriatna, M. (2024). Cultivating elok budi character value in elementary school. Primary: Jurnal Pendidikan Guru Sekolah Dasar, 13(6), 234-241. https://doi.org/10.33578/jpfkip-v13i6.p234-241

Ramadhan, S., Sumiharsono, R., Mardapi, D., & Prasetyo, Z. K. (2020). The quality of test instruments constructed by teachers in Bima Regency, Indonesia: Document analysis. International Journal of Instruction, 13(2), 507-518. https://doi.org/10.29333/iji.2020.13235a

Rofiqi, A., Suciptaningsih, O. A., & Mardhatillah. (2025). Global trends in culturally responsive teaching: A systematic literature review. Jurnal PAJAR (Pendidikan dan Pengajaran), 9(6), 342. https://doi.org/10.33578/pjr.v9i6.342

Primary: Jurnal Pendidikan Guru Sekolah Dasar is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

Sari, A. P., & Kumalasani, M. P. (2025). Integrating local culture into IPAS learning: A culturally responsive teaching approach in elementary school. Primary: Jurnal Pendidikan Guru Sekolah Dasar, 14(6), 941-956. https://doi.org/10.33578/jpfkip.v14i6.941-956

Schachner, M. K., Noack, P., Van de Vijver, F. J. R., & Eckstein, K. (2016). Cultural diversity climate and psychological adjustment at school: Equality and inclusion versus cultural pluralism. Child Development, 87(4), 1175-1191. https://doi.org/10.1111/cdev.12536

Shepard, L. A. (2000). The role of assessment in a learning culture. Educational Researcher, 29(7), 4-14. https://doi.org/10.3102/0013189X029007004

Smith, T., Avraamidou, L., & Adams, J. D. (2022). Culturally relevant/responsive and sustaining pedagogies in science education: Theoretical perspectives and curriculum implications. Cultural Studies of Science Education, 17, 637-660. https://doi.org/10.1007/s11422-021-10082-4

Wibowo, S. A., Murtono, & Utaminingsih, S. (2021). The effectiveness of multicultural-based Indonesian teaching materials in improving critical thinking skills. Jurnal PAJAR (Pendidikan dan Pengajaran), 5(2), 8287. https://doi.org/10.33578/pjr.v5i2.8287

Zufriady, Al Fajar, B., Sari, I. K., Rizka, M., Rahman, F., Radini, Desmita, I., & Balqis, A. (2025). Development of multicultural instruments for elementary school teachers in arts, culture, and crafts learning: Instrument validation analysis. Jurnal PAJAR (Pendidikan dan Pengajaran), 9(5), 605-616. https://doi.org/10.33578/pjr.v9i5.318

Primary: Jurnal Pendidikan Guru Sekolah Dasar is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.
[e-**ISSN 2598-5949** | p-**ISSN 2303-1514**]