



Invention: Journal Research and Education Studies
Volume 7 Nomor 1 March 2026

The Invention: Journal Research and Education Studies is published three (3) times a year

(March, July and November)

Focus : Education Management, Education Policy, Education Technology, Education Psychology, Curriculum Development, Learning Strategies, Islamic Education, Elementary Education

LINK : <https://pusdikra-publishing.com/index.php/jres>

Students' Perceptions of Performance-Based Authentic Assessment in Speaking Recount Text at SMA Swasta Bandung Medan

Neni Afrida Sari Harahap¹, Yudhistira Aidil Fitrah²

^{1,2} Universitas Negeri Medan, Indonesia

ABSTRACT

This study aims to describe the implementation of performance-based authentic assessment in speaking recount text activities and analyze students' perceptions of this method. The research was conducted at SMA Swasta Bandung Medan, utilizing 10 student respondents as the primary data source. A quantitative descriptive design was applied, employing a survey method for data collection. Data were gathered through an 18-item Likert scale questionnaire distributed via Google Form. The collected data were analyzed utilizing descriptive statistics to calculate mean scores and percentages. The findings establish that storytelling of personal experiences is the dominant type of assessment utilized (100% frequency). Furthermore, students hold a 'Very Positive' overall perception of this assessment method, evidenced by a total mean score of 4.02 out of 5.00. The highest-rated indicator was the authenticity of the assessment (mean 4.17). The implications of this study confirm that performance-based authentic assessment is positively perceived and effectively supports student learning. However, the study also highlights crucial pedagogical and psychological implications, specifically the necessity for educators to actively manage student performance anxiety, mitigate classroom noise, and allocate adequate preparation time.

Authentic Assessment, Performance-Based Assessment, Speaking Skill, Recount Text, Student Perception.

Kata Kunci

Corresponding Author:

yudhistiraaf.2242121004@mhs.unimed.ac.id

INTRODUCTION

Assessment constitutes a fundamental and indispensable component within the broader educational ecosystem. It functions not merely as an instrument for measuring student achievement at the end of a module, but also serves as an ongoing pedagogical guide intended to enhance the overall quality of instruction. Lorrie A. Shepard (2000) posits that assessment should be seamlessly integrated with instruction, serving as a critical mechanism to support learning rather than functioning exclusively for summative grading purposes. This significant paradigm shift repositions assessment from a purely

evaluative tool to a meaningful, dynamic scaffold for holistic learner development.

In the specific realm of English language pedagogy, assessment assumes an even greater degree of complexity, particularly when addressing productive skills such as speaking. Oral proficiency extends significantly beyond simple grammatical accuracy; it must encompass conversational fluency, interactive communicative competence, self-confidence, and the critical capacity to convey meaning authentically within real-world communicative contexts. H. Douglas Brown (2004) articulates heavily that speaking assessment must necessarily incorporate actual language performance. He notes that communicative competence fundamentally eludes adequate measurement through conventional paper-and-pencil testing methodologies, necessitating a more direct observational approach.

Among the various textual genres systematically employed in speaking instruction, recount text occupies a highly significant position, explicitly requiring learners to orally reconstruct past experiences or chronological events. This particular genre holds immense pedagogical relevance as it draws directly upon students' personal repositories of experience. Whether students are recounting memorable holidays, significant life events, or routine daily activities, the genre fosters deep, meaningful engagement with the target language. However, the direct assessment of students' oral performance in recount text presentation presents considerable logistical and evaluative challenges for educators. Teachers require appropriate, well-structured assessment frameworks capable of evaluating learners' ability to retell events with appropriate fluency, exact pronunciation accuracy, and coherent textual structure.

Performance-based authentic assessment emerges as a particularly suitable and effective approach for evaluating speaking competence, especially within the context of recount text. Grant Wiggins (1998) defines authentic assessment as an evaluation methodology that necessitates students' active engagement with real-world tasks, thereby demonstrating a meaningful application of acquired knowledge and linguistic skills. Within the specific context of recount text delivery, students are tasked with narrating authentic personal experiences, thereby rendering the assessment process inherently meaningful and contextually relevant to their lived realities. Arthur Hughes (2003) further corroborates this perspective, asserting unequivocally that performance tasks yield much more valid evidence of learners' genuine linguistic capabilities compared to traditional indirect testing methods.

At SMA Swasta Bandung Medan, performance-based authentic assessment has been systematically and continually implemented in speaking activities, including those specifically involving recount text materials. Despite this ongoing implementation, a critical problem remains: students' perceptual responses to such assessment practices remain inadequately understood. The psychological construct of student perception assumes critical importance in the educational environment. It fundamentally influences learners' motivational orientation, their intrinsic self-confidence, and their participatory engagement in public speaking activities. A positive perceptual disposition toward assessment practices tends to organically enhance learners' confidence and encourage their active participation. Conversely, negative perceptions may precipitate severe performance anxiety and subsequent psychological withdrawal from communicative tasks.

Consequently, this specific investigation is directed toward an in-depth exploration of students' perceptions concerning the implementation of performance-based authentic assessment in speaking recount text activities at SMA Swasta Bandung Medan. To guide this inquiry, the present study addresses two primary research questions: (1) What types of performance-based authentic assessment are implemented in speaking recount text activities at SMA Swasta Bandung Medan?, and (2) How is the process of implementing performance-based authentic assessment in speaking recount text activities at SMA Swasta Bandung Medan?. In alignment with these questions, the study aims to comprehensively describe the implementation of this assessment framework, analyze students' perceptions deeply, and identify the effect of this assessment on students' confidence and participation. Through these interconnected objectives, the research intends to provide a holistic understanding of how authentic assessment practices shape both the instructional process and learner outcomes in oral recount text production.

The anticipated findings of this research are expected to contribute meaningfully to multiple educational stakeholders. For English language educators at SMA Swasta Bandung Medan and similar contexts, it offers valuable insights into how students perceive these practices, thereby informing the design of more pedagogically sound speaking assessments. For students, it provides an opportunity for their voices to be represented, potentially fostering greater awareness, increased confidence, and active participation in their communicative development.

RESEARCH METHOD

This study methodically utilizes a quantitative descriptive design framework. The primary purpose of this specific research architecture is to objectively measure and comprehensively describe students' internal perceptions of performance-based authentic assessment in speaking recount text activities at SMA Swasta Bandung Medan. A quantitative approach was deemed most appropriate because this study strictly aims to collect numerical, standardized data that can be analyzed statistically to describe students' perceptions accurately. According to Creswell (2014), quantitative research is an established approach for testing objective theories by critically examining the relationship among specific variables. In this methodology, variables are measured using carefully structured instruments so that numbered data can be subsequently analyzed using rigorous statistical procedures. In this study, the sole variable measured is the students' perception of performance-based authentic assessment in speaking recount text. The descriptive nature of the design allows the researcher to document and describe the current status of students' perceptions without manipulating any variables or environmental conditions.

The empirical data utilized in this study are exclusively primary data obtained directly from the active respondents. The designated source of data is students formally enrolled at SMA Swasta Bandung Medan who have previously and specifically experienced performance-based authentic assessment in speaking recount text activities. The respondents were selected utilizing a purposive sampling technique. The strict criteria for selecting respondents dictated that they must be students who have participated in speaking activities utilizing recount text, have been directly assessed using performance-based authentic assessment, and have explicitly stated their willingness to participate in the research. While the initial methodological framework targeted 30 respondents as adequate for quantitative descriptive research, the actual data collection phase gathered responses from 10 eligible students.

Data collection was executed through the distribution of highly structured questionnaires to the students. The applied technique was a formal survey method whereby questionnaire sheets were presented to the respondents. The sequential steps of data collection proceeded as follows: First, the researcher formally requested permission from the school administration and the respective English teacher to conduct the study. Second, the researcher clearly explained the exact purpose and scope of the research to the participating students. Third, the researcher distributed the digital questionnaire sheets to the

students utilizing the Google Form platform to ensure accessibility. Fourth, students were instructed to complete the questionnaire sheets honestly based on their personal lived experiences. Students were provided sufficient time to complete the task. The primary instrument utilized was the aforementioned questionnaire sheets. The questionnaire grid was meticulously developed based on the theoretical indicators proposed by Gerritsen-van Leeuwenkamp, Joosten-ten Brinke, and Kester (2018). It consists of six distinct indicators mapped to a total of 18 specific statements. Indicator 1: Effects of assessment on learning covers items 1, 2 (positive) and 3 (negative). Indicator 2: Fairness of assessment covers items 4, 5 (positive) and 6 (negative). Indicator 3: Conditions of assessment covers items 7, 8 (positive) and 9 (negative). Indicator 4: Interpretation of assessment results covers items 10, 11 (positive) and 12 (negative). Indicator 5: Authenticity of assessment covers items 13, 14 (positive) and 15 (negative). Indicator 6: Credibility of assessment covers items 16, 17 (positive) and 18 (negative). In total, the instrument contained 12 positive statements and 6 negative statements. Notably, no neutral statements were included in the base design of the questionnaire to force directional feedback. For positive statements, the scoring logic assigned 5 for SA down to 1 for SD, while negative statements were strictly reverse-scored. Prior to deployment, content validity was verified by consulting with the academic supervisor and English language assessment experts to ensure item relevance and clarity.

The raw data were meticulously analyzed utilizing descriptive statistics. According to Sugiyono (2013), descriptive statistics are purposefully used to analyze data by describing the collected figures exactly as they are without attempting to make broader population generalizations. The analytical steps were highly sequential: First, all responses were tabulated into a master table. Second, each response was converted to a numerical score based on the Likert scoring system. Third, the researcher calculated the mean score for each item and indicator (Total Score / Number of Respondents). Fourth, mean scores were assigned to predefined interpretation categories: 4.01-5.00 (Very Positive), 3.01-4.00 (Positive), 2.01-3.00 (Neutral), 1.01-2.00 (Negative), and 0.00-1.00 (Very Negative). Fifth, response percentages were calculated to show frequency distributions.

RESULT AND DISCUSSION

The empirical data collected in this study meticulously addresses the two primary research questions formulated in the initial chapter: (1) the specific types of performance-based authentic assessment implemented in speaking recount text activities, and (2) the systematic process of implementing this

assessment, which is comprehensively evaluated through the students' perceptual responses to the experience. The empirical findings are derived from the responses generated by 10 students from SMA Swasta Bandung Medan who strictly met the purposive sampling criteria. The demographic profile of these active respondents reveals that a majority were female, constituting 60% of the sample (6 students), while male respondents made up the remaining 40% (4 students). Academically, the overwhelming majority of the respondents (90%, or 9 students) were enrolled in Grade 10, with only a single respondent (10%) originating from Grade 11. In terms of age demographics, the largest cohort consisted of 16-year-old students (70%, 7 students), followed by 15-year-old students (20%, 2 students), and one 17-year-old student (10%). This age distribution is highly typical for senior high school students navigating Grades 10 and 11.

The first focal point of the analysis investigates the specific types of performance-based authentic assessments physically implemented by the English educators. An astounding 100% of the respondents (all 10 students) confirmed that they were explicitly required to perform storytelling by orally narrating their own genuine, personal experiences in front of the classroom. These authentic narratives encompassed past vacations, highly memorable life events, or routine daily activities. In addition to the primary storytelling component, 80% of the respondents (8 students) reported engaging in formal Oral Presentations, wherein they presented their structured recount texts orally to both the teacher and their peers. Furthermore, 40% of the respondents (4 students) noted that an interactive Question and Answer segment was deliberately integrated into the evaluation; in this phase, students had to respond extemporaneously to questions posed by either the teacher or peers immediately after delivering their story. These specific findings perfectly align with the core definition of authentic assessment, which strictly demands the execution of real-world tasks that demonstrate a meaningful, practical application of linguistic knowledge. The authenticity of this targeted approach was highly praised by the student body. As respondent Zainal Rizky explicitly stated, "I like that I can tell my real experience". Another respondent, Clarissa Amelia, further elaborated on this pedagogical value, stating, "I find the most value in the authenticity of the task".

The operational process of implementing this assessment was meticulously analyzed through the students' questionnaire responses. The execution of this pedagogical strategy was observed to occur in three distinct, sequential stages. To fully understand the efficacy of this process, the implementation stages are directly evaluated through the lens of the students'

perceptions. Prior to any student performance, the teacher actively and explicitly explained the analytical assessment criteria to the class. Students felt strongly that the teacher dispensed impeccably fair assessments (mean score 4.4) and effectively explained the evaluation criteria before the assessment commenced (mean score 4.0). The students were fully aware that they would be rigorously evaluated on pronunciation, fluency, grammar, vocabulary, and recount text organization. Respondent Julian Sinaga reinforced the necessity of this preparatory transparency, noting, "Teachers should explain the rubric clearly such as pronunciation, fluency, grammar, vocabulary, and organization". During the actual assessment phase, students stood before the class to perform their oral recount texts. This dynamic stage was evaluated by the students through the Conditions of Assessment and Authenticity of Assessment indicators. The time duration provided for this speaking assessment was widely considered sufficient (mean score 4.0, 'Positive' category). Additionally, the general classroom atmosphere was perceived to be supportive of student confidence, yielding a mean score of 4.0 ('Positive'). The implementation of this process also yielded 'Positive' Effects on Learning (mean score 4.00), as students felt the performance directly helped them understand their linguistic weaknesses (mean score 4.3) and increased their motivation to learn (mean score 4.1). However, the implementation process was not without environmental and psychological challenges. The quantitative data flagged a 'Neutral' mean score of 3.0 regarding distractions caused directly by ambient classroom noise, strongly suggesting that the teacher's noise management protocol during peer performances requires immediate improvement. Psychologically, the most frequently mentioned drawback during this active process was severe performance anxiety. Clarissa Amelia articulated that the "primary drawback is performance anxiety" and that the "pressure of speaking publicly can lead to cognitive blocks, such as forgetting vocabulary". Following the individual performances, the teacher concluded the process by providing final grades alongside targeted, constructive feedback. This final stage was highly rated through the Interpretation of Assessment Results and Credibility of Assessment indicators. Students exhibited a remarkably high level of comprehension regarding the grades they received (mean score 4.1, 'Very Positive'). Critically, they found the teacher's post-performance feedback exceptionally helpful for their ongoing linguistic improvement (mean score 4.3, 'Very Positive'). Respondent Esa Nabila highlighted this specific impact, stating, "The teacher's feedback helps me improve my speaking recount text". The process was deemed highly credible (mean score 4.03, 'Very Positive'), with students heavily trusting the teacher's final assessment results (mean score 4.1).

Despite the overwhelmingly positive reception to the grading and feedback process, students consistently voiced a deep desire for systemic adjustments regarding preparation logistics. Almost all students urgently requested more practice time prior to the assessment phase. Zainal Rizky noted, "Students need more practice time before the assessment and the teacher should give examples and feedback". Muhammad Abiyyu echoed this logistical concern, requesting "more in-class rehearsal time" and "more demonstrations and clear feedback before the actual grading starts". Furthermore, students recognized severe time constraints inherent in large classes; Intan Nurlaini noted it "takes a lot of time to assess each student". Ultimately, when utilizing the students' perceptions to evaluate the overarching process, the implementation of this performance-based authentic assessment was categorized as "Very Positive", achieving a cumulative mean score of 4.02 across all metrics. This conclusively demonstrates that while there are minor logistical challenges regarding rehearsal time and classroom noise, the pedagogical process from initial preparation to final post-assessment feedback is highly effective, credible, and profoundly meaningful to the learners.

The comprehensive findings of this study provide robust verification for multiple theoretical frameworks surrounding educational assessment. Foremost, the overwhelming dominance of storytelling as a personal recount task deeply verifies Wiggins' (1998) foundational theory that true authentic assessment actively mandates students to perform real-world tasks that demonstrate meaningful knowledge application. Furthermore, the empirical data directly verifies Brown's (2004) speaking assessment framework, which insists that actual language performance tasks provide highly valuable, actionable learning opportunities and effectively assist students in pinpointing their linguistic weaknesses. The diverse perceptual responses similarly substantiate Baidoo-Anu and DeLuca's (2025) identification of factors influencing perception; the qualitative data starkly highlights how tangible experiences with performance nervousness, the availability of practice time, and the quality of targeted feedback fundamentally shape holistic student views regarding assessment. All six indicators formulated by Gerritsen-van Leeuwenkamp et al. (2018) were successfully confirmed as critical factors in measuring student perception, with task authenticity proving to be the most influential element. Regarding broad implications toward English language learning, the current study yields profound psychological and pedagogical insights. Psychologically, there is an observable dual impact on the learner: acute initial performance anxiety that is predictably followed by a significant increase in self-efficacy and confidence through repeated practice iterations.

This crucially implies that teachers must actively manage cognitive anxiety by normalizing speaking mistakes and meticulously designing assessment sequences that allow for progressive skill development rather than relying on one-time, high-stakes performances. Pedagogically, the primary implication is that educators should persist in utilizing highly authentic tasks, as they substantially enhance intrinsic motivation among learners by making language use relevant to their daily lives. However, to fully optimize this communicative learning experience, educators are mandated to explicitly explain grading criteria beforehand, deliberately allocate sufficient in-class rehearsal time, and aggressively manage classroom noise conditions to minimize debilitating environmental performance distractions.

CONCLUSION

The implementation of performance-based authentic assessment in speaking recount text activities at SMA Swasta Bandung Medan fundamentally and practically relies on the interactive storytelling of personal experiences. This specific pedagogical method aligns seamlessly with the core academic principles of authentic assessment, as it mandates students to perform actual, real-world communicative tasks that are deeply and contextually relevant to their lived realities, such as recounting memorable vacations or detailing routine daily activities. The instructional process itself is systematically and transparently executed by the educators; it begins with a clear analytical explanation of criteria during preparation, moves to structured public implementation, and concludes with the crucial provision of highly beneficial, actionable feedback post-assessment.

The students involved in this study consistently hold a demonstrably positive perception of this precise assessment paradigm. This highly favorable outlook is driven primarily by the high authenticity of the tasks, their direct relevance to daily real-world communication, and their tangible effectiveness in identifying linguistic weaknesses and progressively enhancing overall speaking fluency. While the overall reception is quantifiably very strong, specific logistical challenges remain pertinent hurdles. Performance-induced anxiety, debilitating auditory distractions resulting from unmanaged classroom noise, severe time constraints inherent in large classroom settings, and an acute student demand for extended rehearsal time remain critical issues that directly compromise the qualitative conditions of the assessment environment. It is strongly advised to deliberately allocate adequate in-class rehearsal time, formally allowing students to practice speaking extensively in pairs or small collaborative groups before subjecting them to individual, graded

performances. This systemic approach is necessary to organically reduce performance anxiety. Furthermore, teachers must establish and rigidly enforce clear expectations for audience behavior to immediately mitigate distracting ambient classroom noise during peer assessments. Teachers should also continue providing highly specific, actionable feedback that highlights strengths and explicitly offers guidance for future development.

Students are deeply encouraged to reframe their views on performance-based assessments, seeing them as vital, low-risk opportunities for linguistic improvement rather than stressful, high-stakes evaluations. They should proactively utilize regular practice opportunities both at home and in class. Crucially, students must actively develop psychological coping strategies, such as deep breathing exercises and positive self-talk, to manage public speaking nervousness effectively. Future researchers investigating this domain are recommended to substantially expand the demographic sample size across multiple educational institutions and diverse grade levels to increase the statistical generalizability of the findings. They should deploy comprehensive mixed-methods approaches, intentionally combining quantitative surveys with deep-dive structured interviews and rigorous classroom observations, to glean a more nuanced understanding of students' and teachers' authentic assessment experiences.

ACKNOWLEDGEMENT

The authors would like to express their sincere gratitude to the English teacher and students of SMA Swasta Bandung Medan for their participation and cooperation in this study. Special thanks are also extended to the lecturer for guidance, support, and valuable feedback throughout the research process. In addition, the authors appreciate the support from all parties who have contributed to the completion of this study.

REFERENCES

- Baidoo-Anu, D., & DeLuca, C. (2025). Do students perceive assessment differently? Exploring the diverse ways students conceive assessment and its impact on their assessment experiences and engagement. *Cogent Education*, 12(1), 1-18.
- Brown, G. T. L., & Hirschfeld, G. H. F. (2008). Students' conceptions of assessment: Links to outcomes. *Assessment in Education: Principles, Policy & Practice*, 15(1), 3-17.
- Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. Pearson Education.

- Byman Frisé, L. (2024). Eyes on the ball: Teachers' rating processes when assessing a national L2 speaking test through the lens of their scoring document. *Studies in Language Assessment*, 13(1), 1-36.
- Cal State LA. (n.d.). Week 9: Authentic assessment. <https://web.calstatela.edu//faculty/jshindi/teaching/Week%209%20Authentic%20Assessment.htm>
- Chen, S. (2025). Getting started with a manageable and tailored performance assessment. McREL Institute.
- Creswell, J. W. (2014). Research design: Qualitative, quantitative, and mixed methods approaches (4th ed.). SAGE Publications.
- Fulcher, G. (2015). Assessing second language speaking. *Language Teaching*, 48(3), 1-36.
- Gerritsen-van Leeuwenkamp, K. J., Joosten-ten Brinke, D., & Kester, L. (2018). Developing questionnaires to measure students' expectations and perceptions of assessment quality. *Cogent Education*, 5(1), 1-16.
- Harris, D. P. (1990). Testing English as a second language. McGraw-Hill.
- Hughes, A. (2003). *Testing for language teachers (2nd ed.)*. Cambridge University Press.
- Hyland, K. (2004). *Genre and second language writing*. University of Michigan Press.
- Knapp, P., & Watkins, M. (2005). Genre, text, grammar: Technologies for teaching and assessing writing. UNSW Press.
- Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4-14.
- Sugiyono. (2013). Metode penelitian pendidikan: Pendekatan kuantitatif, kualitatif, dan R&D. Alfabeta.
- Wiggins, G. (1998). Educative assessment: Designing assessments to inform and improve student performance. Jossey-Bass.