JPPIPA 12(10) (2026)

Jurnal Penelitian Pendidikan IPA

http://jppipa.unram.ac.id/index.php/jppipa/index

Development of Think Pair Share Learning Model Assessment in English Subjects on the Oral Communication Ability of Students to Support Sustainable Development Goal (SDG) 4: Quality Education

Meria Fitriwati[1], Alwen Bentri[1], Aban Hidayati[1], Nofri Hendri[1], Rayendra[1*]

[1] Educational Technology, Postgraduate School, Universitas Negeri Padang, Padang, Indonesia.

Received: July 12, 2026
Revised: September 25, 2026
Accepted: September 30, 2026
Published: October 31, 2026

Corresponding Author:
Rayendra
rayendra@fip.unp.ac.id

DOI: 10.29303/jppipa.v12i10.15990

Open Access

© 2026 The Authors. This open access article is distributed under a (CC-BY License)

Abstract: This study aimed to develop a Think Pair Share-Based Speaking Assessment Model (TPS-SAM) to improve the oral communication skills of Grade X Madrasah Aliyah students while supporting the implementation of Sustainable Development Goal (SDG) 4 on Quality Education. The study employed a Research and Development (R&D) approach using the ADDIE model, consisting of Analysis, Design, Development, Implementation, and Evaluation. The product comprised a teacher guide, speaking assessment rubric, observation sheet, student task sheet, scoring template, and student reflection sheet, all integrated into the Think Pair Share learning syntax. Product validity was evaluated using Aiken's V, while reliability was examined using Cronbach's Alpha and Intraclass Correlation Coefficient (ICC). Practicality was assessed through teacher and student response questionnaires, whereas effectiveness was determined using a one-group pretest–posttest design, normalized gain (N-Gain), paired sample t-test, and Cohen's d. The results indicated that the developed model was highly valid (Aiken's V = 0.92), highly reliable (Cronbach's Alpha = 0.93; ICC = 0.91), and highly practical based on teacher (95.5%) and student (95.0%) responses. Students' oral communication skills improved significantly, with the mean score increasing from 61.43 to 84.67, an N-Gain of 0.60 (moderate), a significant paired sample t-test ($p < 0.001$), and a large effect size (Cohen's d = 1.82). These findings indicate that TPS-SAM is a valid, practical, and effective assessment model for enhancing students' oral communication skills and supporting authentic assessment practices in English language learning.

Keywords: Authentic assessment; Oral communication skills; SDGs 4; Speaking assessment; Think pair share

Introduction

Oral communication skills are one of the main competencies that students must have in the 21st century (Anas & Sapri, 2022; Indarta et al., 2022). The mastery of speaking skills in English not only serves as a means of conveying information, but also as an indicator of the ability to think critically, collaborate, solve problems, and participate in global communication (Wiartis, 2021; Wijayanto, 2023). Therefore, English learning in schools is no longer oriented towards mastering the linguistic aspect alone, but is directed at the development of communicative competencies that enable students to use language effectively in various real-life contexts (Setiawan, 2021; Qamariah, 2023).

The urgency of improving oral communication skills is also in line with the Sustainable Development Goals (SDGs) agenda, especially Goal 4 (Quality Education) which emphasizes the importance of providing quality, inclusive, and sustainable education

How to Cite:
Fitriwati, M., Bentri, A., Hidayati, A., Hendri, N., & Rayendra, R. (2026). Development of Think Pair Share Learning Model Assessment in English Subjects on the Oral Communication Ability of Students to Support Sustainable Development Goal (SDG) 4: Quality Education. *Jurnal Penelitian Pendidikan IPA*, *12*(10), 19-27. https://doi.org/10.29303/jppipa.v12i10.15990

for all students. Target 4.1 emphasizes that every student must obtain meaningful learning outcomes, while Target 4.7 encourages the development of 21st century competencies through learning innovations that are able to improve critical thinking, communication, collaboration, and creativity skills. In this context, the assessment system has a strategic role because it not only functions to measure learning outcomes, but also becomes an instrument to improve the quality of the learning process through providing continuous feedback.

Speaking skills assessment is still one of the main challenges in learning English. In contrast to reading or writing skills which are relatively easy to measure through written tests, speaking skills are complex abilities because they involve aspects of fluency, accuracy, pronunciation, vocabulary mastery, interaction, and confidence in communicating simultaneously (Fiktorius, 2020; Hanafi et al., 2021; Zaim, 2018). This complexity causes many teachers to have difficulty in compiling assessment instruments that are able to measure speaking performance objectively, validly, and reliably. These problems are still found in English learning in Indonesia, especially at the Madrasah Aliyah level. Although the Independent Curriculum has emphasized the importance of implementing authentic assessment and student-centered learning, implementation in the field shows that the assessment process is still dominated by written tests. Teachers tend to provide assessments based on the final learning outcomes, while the communication process that takes place during learning activities has not been systematically documented. This condition causes students' speaking skills to not receive proportional attention compared to other language skills (Gunawan et al., 2022; Linggasari & Rochaendi, 2022).

The results of the preliminary study conducted by the researcher also showed a similar condition. Based on learning observations, interviews with English teachers, and analysis of assessment documents, it is known that the assessment of speaking skills has not been carried out in a structured manner. Teachers generally do not use a standard assessment rubric, the feedback given is still general, and the documentation of the results of speaking assessments has not been carried out systematically. As a result, teachers have difficulty in monitoring the development of students' oral communication skills over time, while students also do not get clear information about aspects that need improvement in their speaking skills.

Various previous studies have proven that the use of authentic assessments can improve the quality of speaking skills assessment because learners are assessed based on real performance in authentic communication situations (Arsita & Fathoni, 2022; Ismail et al., 2023; Linggasari & Rochaendi, 2022; Rosyada, 2024). On the other hand, the Think Pair Share (TPS) cooperative learning model developed by Lyman provides opportunities for students to think independently, discuss with partners, and then convey the results of the discussion to a larger group (share). This model has been proven to improve students' participation, confidence, and the quality of oral communication because it provides a collaborative learning environment and provides opportunities for gradual speaking practice (Benjelloun, 2021; Darmawan & Lestariningsih, 2023; Maulani et al., 2020).

However, the results of the literature review show that most previous studies have focused attention on the effectiveness of the Think Pair Share model in improving learning outcomes or speaking skills, while the development of assessment instruments that specifically integrate the Think Pair Share syntax into the assessment process is still very limited. Research on authentic assessments generally only addresses the use of speaking assessment rubrics without linking them to the stages of cooperative learning, while research on Think Pair Share focuses more on learning strategies than on assessment systems that support the process (Apriliani et al., 2021; Wulandari, 2023). Thus, there is still a research gap, namely the lack of an assessment model based on Think Pair Share that is able to simultaneously assess process aspects (think and pair) and outcome aspects (share) in English learning.

Based on these gaps, this study offers novelty in the form of the development of a Think Pair Share-Based Speaking Assessment Model that integrates the principles of authentic assessment, performance assessment, and formative assessment into each stage of the Think Pair Share model. The assessment model developed not only produces validated speaking assessment rubrics, but also provides observation sheets, usage guides, assessment recapitulation formats, and student reflection sheets as a unit of assessment toolkit that teachers can use in evaluating oral communication skills objectively, systematically, and continuously.

Therefore, the purpose of this study is to develop a valid, practical, and effective learning assessment model based on Think Pair Share to improve the oral communication skills of class X students of Madrasah Aliyah. The results of this research are expected to contribute to the development of an authentic assessment system in English language learning while supporting the implementation of Sustainable Development Goal (SDG) 4, especially in an effort to improve the quality of education through learning innovations and assessments oriented to 21st century competency development.

Method

This research is research and development (Research and Development; R&D) which aims to develop the Think Pair Share-Based Speaking Assessment Model (TPS-SAM) as an authentic assessment model to improve students' oral communication skills in English subjects at Madrasah Aliyah. The development model used is ADDIE (Analysis, Design, Development, Implementation, and Evaluation) developed by Branch. The ADDIE model was chosen because it has systematic, flexible, and development-oriented stages for educational products that meet the aspects of validity, practicality, and effectiveness (Branch, 2009; Sugiyono, 2018). Through these five stages, the resulting product is expected to be able to become an assessment instrument that is in accordance with the characteristics of communication-based English learning.

The Analysis stage is carried out to identify the needs of teachers and students for speaking skill assessment (Aspriyani & Suzana, 2020). Needs analysis was carried out through observation of the English learning process, semi-structured interviews with teachers, analysis of learning tools and assessment instruments used in schools, analysis of student characteristics, and studies of the Learning Outcomes (CP) of the Independent Curriculum (Triana et al., 2024). The results of the analysis are the basis for formulating product specifications, assessment indicators, and characteristics of the assessment model to be developed.

The Design stage is focused on designing an assessment model based on Think Pair Share. At this stage, a product blueprint was prepared consisting of a user guide (Teacher Guide), Speaking Assessment Rubric, learning observation sheet, student assignment sheet, assessment recapitulation format, and student reflection sheet. The development of rubrics refers to the indicators of speaking ability in the Common European Framework of Reference for Languages (CEFR) which includes fluency, accuracy, pronunciation, vocabulary, and interaction. All of these indicators are integrated into the Think-Pair-Share learning syntax, so that the assessment not only focuses on the final outcome of speaking, but also assesses the communication process that takes place at each stage of learning.

Furthermore, in the Development stage, the initial prototype of the assessment system was developed and validated by three validators consisting of learning evaluation experts, English education experts, and educational technology experts. The validation process was carried out on aspects of content feasibility, construct feasibility, language feasibility, and presentation feasibility using a five-level Likert scale assessment instrument (Annajati & Wikarya, 2023; Eriansyah & Baadilla, 2023). Inputs, suggestions, and recommendations from validators are used as the basis for product revision until an assessment model that meets the valid criteria to be implemented in learning is obtained.

The Implementation stage was carried out through a limited trial of two English teachers and thirty students of class X of Madrasah Aliyah Negeri 3 Pesisir Selatan who were selected using purposive sampling techniques. The trial aims to determine the level of practicality of using the assessment model by teachers and students, as well as test the effectiveness of the model in improving students' oral communication skills. The effectiveness of the product was measured using a one-group pretest-posttest design, namely by comparing students' speaking skills before and after the implementation of the Think Pair Share-Based Speaking Assessment Model.

The last stage is Evaluation, which is a thorough evaluation of the entire development process based on the results of expert validation, practicality tests, and effectiveness tests. Evaluation is carried out to identify the advantages and weaknesses of the product so that final revisions can be carried out before the assessment model is declared suitable for use in English learning at Madrasah Aliyah. This stage produces the final product in the form of a Think Pair Share-Based Speaking Assessment Model (TPS-SAM) that has met the aspects of validity, practicality, and effectiveness.

The research instruments used include expert validation sheets, observation sheets on learning implementation, teacher response questionnaires, student response questionnaires, interview guidelines, and Think Pair Share-based speaking skills assessment rubrics. The assessment rubric was developed based on five indicators of oral communication ability, namely fluency, accuracy, pronunciation, vocabulary, and interaction. All instruments are prepared based on theoretical studies and the results of needs analysis so that they are able to measure the quality of the products developed comprehensively.

The research data was analyzed quantitatively and qualitatively. The validity of the instrument's content is analyzed using Aiken's V, the instrument is declared valid if the value of Aiken's V ≥ 0.80.

$$V = \frac{\sum s}{n\,(c-1)} \tag{1}$$

The reliability of the questionnaire was analyzed using Cronbach's Alpha with a coefficient value of α≥ 0.70 as a reliability criterion.

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum_{i=1}^{k} \sigma_i^2}{\sigma_t^2}\right) \tag{2}$$

To ensure the consistency of assessment on the speaking skills rubric, the Intraclass Correlation Coefficient (ICC) is used as a measure of inter-rater reliability.

The practicality of the model is analyzed based on the average percentage of teacher and student response questionnaire scores.

$$\overline{X} = \frac{\sum X}{N} \quad (3)$$

It is then interpreted into the categories of very practical, practical, quite practical, less practical, and impractical.

The effectiveness of the assessment model was analyzed the effectiveness of the assessment model was analyzed using Normalized Gain (N-Gain) to determine the improvement of students' oral communication skills after product implementation.

$$g = \frac{S_{post} - S_{pre}}{S_{max} - S_{pre}} \quad (4)$$

Furthermore, the difference in pretest and posttest scores was analyzed using a paired sample t-test at a significance level of 0.05.

$$t = \frac{\overline{D}}{S_D/\sqrt{n}} \quad (5)$$

To determine the magnitude of the influence of the developed model on the improvement of oral communication skills, this study also calculated the effect size using Cohen's d.

$$d = \frac{\overline{X}_{post} - \overline{X}_{post}}{S_p} \quad (6)$$

With the standard deviation combine is calculated using the formula:

$$S_p = \sqrt{\frac{S_{pre}^2 + S_{post}^2}{2}} \quad (7)$$

Meanwhile, data from observations, interviews, and field notes were analyzed using thematic analysis techniques through the stages of data reduction, data presentation, and conclusion drawn. The integration of quantitative and qualitative analysis was carried out to obtain a comprehensive picture of the quality of the Think Pair Share-Based Speaking Assessment Model developed.

Result and Discussion

Needs Analysis Results

In the analysis stage, the need for the development of the Think Pair Share-Based Speaking Assessment Model (TPS-SAM) was identified through learning observations, interviews with English teachers, analysis of learning documents, and studies of the Learning Outcomes of the Independent Curriculum. The results of the analysis show that the process of assessing speaking skills in English learning at Madrasah Aliyah is still oriented towards assessing learning outcomes through written tests. Assessment of students' speaking performance has not been carried out systematically because teachers do not have instruments that are able to measure the communication process at each stage of learning.

Classroom observation shows that teachers provide opportunities for students to discuss and present learning results, but these activities are not accompanied by a standard assessment rubric. Assessments are more based on teachers' general observations so that assessment results tend to be subjective. In addition, documentation of the development of students' speaking skills has not been carried out on an ongoing basis so that teachers have difficulty in monitoring the development of students' oral communication skills.

The results of the interviews also show that teachers need an assessment model that is practical, easy to use, and able to integrate learning activities and assessment processes in one learning unit. Based on these findings, it is concluded that it is necessary to develop an assessment model based on Think Pair Share that is able to assess students' oral communication skills authentically, objectively, and sustainably.

Product Design

The design stage is carried out based on the results of the needs analysis that has been obtained in the previous stage. The product developed is called the Think Pair Share-Based Speaking Assessment Model (TPS-SAM). The model is designed as an authentic assessment tool that integrates the Think Pair Share learning syntax with the speaking skills assessment process.

The product consists of six main components, namely Teacher Guide, Speaking Assessment Rubric, Observation Sheet, Student Task Sheet, Scoring Template, and Student Reflection Sheet. The six components are designed in an integrated manner so that they are able to support the implementation of assessments starting from the learning planning stage, the implementation of assessments, providing feedback, to documentation of the development of students' oral communication skills.

The assessment rubric was developed based on the indicators of speaking ability that refer to the Common European Framework of Reference (CEFR), namely fluency, accuracy, pronunciation, vocabulary, and interaction. Each indicator is equipped with a performance description so that teachers can provide assessments more objectively and consistently.

Product Validation Results (Development)

The development stage resulted in an initial prototype of the Think Pair Share-Based Speaking Assessment Model which was then validated by three validators consisting of learning evaluation experts, English education experts, and educational technology experts.

Validation was carried out on four main aspects, namely content feasibility, constructability, language feasibility, and presentation feasibility. A summary of the validation results is presented in Table 1.

Table 1. Product validation results

Aspects	Aiken's V	Category
Content eligibility	0.92	Very valid
Construct eligibility	0.90	Very valid
Language qualifications	0.95	Very valid
Presentation eligibility	0.91	Very valid
Average	0.92	Very valid

Based on the validation results, an Aiken's V value of 0.92 was obtained so that the product met the valid category. so that the Think Pair Share-Based Speaking Assessment Model meets the category of being very valid. The validator provided several inputs in the form of improving the editorial of the indicators, adding implementation examples to the teacher's guide, and simplifying the format of the observation sheet. All inputs have been used as a basis for product revision before field trials.

Results of Reality Test

Once the product is valid, the reliability test of the instrument is carried out using Cronbach's Alpha and Intraclass Correlation Coefficient (ICC).

Table 2. Reliability test results

Instrument	Value	Category
Cronbach's Alpha	0.93	Very reliable
ICC (Average Measures)	0.91	Very good

The results of the analysis showed that Cronbach's Alpha value was 0.93, which indicates that the internal consistency of the instrument is in the very high category. The Intraclass Correlation Coefficient (ICC) value of 0.91 indicates an excellent level of agreement between assessors so that the rubric can be used consistently by teachers.

Product Practicality Test Results (Implementation)

The implementation stage is carried out through a limited trial of teachers and students to determine the level of practicality of using the assessment model.

Table 3. Teacher practicality

Aspect	Percentage (%)	Category
Ease of use	96	Very practical
Indicator clarity	94	Very practical
Ease of assessment	95	Very practical
Ease of documentation	97	Very practical
Average	95.5	Very practical

Table 4. Practicality of students

Aspect	Percentage (%)	Category
Ease of understanding tasks	94	Very practical
Clarity of feedback clarity	95	Very practical
Ease of participating in polling station activities	96	Very practical
Learning motivation	95	Very practical
Average	95	Very practical

The results of the analysis showed that the level of practicality according to teachers was 95.5%, while according to students it was 95.0%. Based on practicality criteria, the assessment model is included in the category of very practical.

The teacher stated that the assessment model is easy to use because the assessment indicators have been systematically compiled according to the Think Pair Share syntax. In addition, the existence of a grade recapitulation format and observation sheets makes it easier for teachers to document the development of students' oral communication skills. Students also responded positively because they received clearer feedback on their speaking skills.

Product Effectiveness Test Results

The effectiveness of the Think Pair Share-Based Speaking Assessment Model was analyzed based on the results of the pretest and posttest of students' oral communication skills.

Table 5. Pretest and posttest results

Test	Mean	SD
Pretest	61.43	7.84
Posttest	84.67	6.91

The results of the analysis showed that there was an increase in the average oral communication skills of students after using the developed assessment model. To determine the level of improvement in this ability, analysis was carried out using Normalized Gain (N-Gain).

Table 6. N-Gain results

Mean Pretest	Mean Posttest	N-Gain	Category
Pretest	61.43	7.84	Medium

An N-Gain value of 0.60 indicates that the improvement in students' oral communication skills is

in the medium category, which indicates an increase in ability after the use of TPS-SAM.

Next, a paired sample t-test was carried out to determine the significance of the increase.

Table 7. Paired sample t-test results

Variable	t	df	Sig.(2-tailed)
Pretest-posttest	-17.84	29	0.000

The test results showed a value of Sig. = 0.000 < 0.05, so there was a significant difference between the students' oral communication skills before and after using the Think Pair Share-Based Speaking Assessment Model. The magnitude of the influence of the assessment model was analyzed using Cohen's d.

Table 8. Effect size results

Cohoen's d	Interpretation
1.82	The effect is huge

Cohen's d value of 1.82 indicates that the application of the Think Pair Share-Based Speaking Assessment Model has a huge influence on improving students' oral communication skills.

Discussion

This study aims to develop the Think Pair Share-Based Speaking Assessment Model (TPS-SAM) as an authentic assessment model to improve the oral communication skills of class X students of Madrasah Aliyah. The results of the study show that the developed model has met the three main characteristics of research and development products, namely valid, practical, and effective. The findings indicate that TPS-SAM is not only suitable for use as an instrument for assessing speaking skills, but also able to support a learning process oriented towards the development of 21st century communication competencies as emphasized in Sustainable Development Goal (SDG) 4 on Quality Education.

Validity of Think Pair Share-Based Speaking Assessment Model

The validation results showed that TPS-SAM obtained an average Aiken's V score of 0.92, which is in the very valid category. The score shows that all components of the model, including the user guide, assessment rubric, observation sheet, assignment sheet, recapitulation format, and reflection sheet are in accordance with the learning objectives, student characteristics, and indicators of oral communication skills developed based on the Common European Framework of Reference (CEFR).

The high level of validity indicates that the developed assessment model has met the validity of the content and the validity of the construct. This happens because the development of the instrument begins through a comprehensive needs analysis, then is designed based on the Think-Pair-Share learning syntax, so that each assessment indicator represents the learning activities that students actually do during the learning process. Thus, the assessment not only measures the final outcome of speaking, but also accommodates individual thinking processes (think), interaction in pairs, and the ability to convey ideas to the group (share).

These findings are in line with opinion of Nurhaifa et al. (2020) which states that the assessment of speaking skills will have a high level of validity if the assessment indicators really represent the communication competence to be measured. In addition, Hasanah (2022) It also emphasizes that effective speaking assessments must include aspects of fluency, accuracy, pronunciation, vocabulary, and interaction, all of which have been integrated into the TPS-SAM rubric.

Reliability of Assessment Instruments

In addition to having a high level of validity, the developed instruments also show excellent reliability. Cronbach's Alpha value of 0.93 indicates that each indicator in the rubric has high internal consistency, while an Intraclass Correlation Coefficient (ICC) value of 0.91 indicates excellent agreement between raters.

Reliability is one of the main requirements for an assessment instrument because the assessment results must provide consistent information even if used by different assessors. In this study, the high ICC score showed that the description of the indicators at each rubric level had been clearly formulated so as to minimize the subjectivity of teachers when assessing speaking skills.

The findings support the theory of O'Malley and Pierce who stated that authentic assessment requires a detailed rubric so that the assessment of student performance takes place objectively and consistently (El-Koumy, 2004). Thus, TPS-SAM is able to reduce the variation in assessment that has often occurred in speech skill assessments.

Practicality of Assessment Models

The results of the implementation showed that teachers gave a level of practicality of 95.5%, while students gave a positive response of 95.0%, which is included in the category of very practical. These findings suggest that the assessment model is easy to use in learning without significantly increasing the workload of teachers.

This practicality is influenced by the product design that unites all assessment tools in one integrated package (assessment toolkit). Teachers no longer compile rubrics, observation sheets, or recapitulation formats separately because all components are available in one assessment system. In addition, assessment

indicators that follow the Think-Pair-Share stage help teachers make systematic observations during the learning process.

For students, the existence of assessment rubrics and reflection sheets provides a clear picture of the indicators of speaking success so that they better understand the aspects that must be improved. This condition is in accordance with the concept of assessment for learning, which is an assessment that not only functions to provide value, but also becomes a means of learning through continuous feedback (Reimann et al., 2021).

Effectivity Think Pair Share-Based Speaking Assessment Model

The results of the study showed an increase in students' oral communication skills after using TPS-SAM. The average score increased from 61.43 in the pretest to 84.67 in the posttest. The N-Gain value of 0.60 showed an increase in the medium category, while the results of the paired sample t-test showed a significant difference ($p < 0.001$). The magnitude of the model's influence is also shown by Cohen's d-value of 1.82, which belongs to the category of very large effects.

The increase occurred because the assessment model developed not only serves as an evaluation tool, but also becomes part of the learning process. In the think stage, students are given the opportunity to formulate ideas independently so that they have time to prepare answers. The pair stage allows students to discuss and get feedback from study partners so that confidence increases. Furthermore, at the share stage, students convey the results of the discussion to a larger group so that they have a more authentic communication experience.

The integration of assessments at each stage provides an opportunity for teachers to provide direct feedback during the learning process. Formative feedback has been proven to help students correct pronunciation, grammar, and vocabulary usage errors before making a final presentation. Thus, the assessment process functions as an integral part of learning, not just as an evaluation activity after learning is completed.

These findings reinforce the results of the study by Hidayat et al. (2026) which states that the Think Pair Share model is able to increase students' participation and courage in speaking in English. This research also complements previous research because it not only proves the effectiveness of the Think Pair Share learning strategy, but also produces an assessment model that is specifically designed to follow the learning syntax.

Contribution to the Development of Authentic Assessments and SDG 4

The main novelty of this research lies in the development of the Think Pair Share-Based Speaking Assessment Model (TPS-SAM) which integrates authentic assessment, performance assessment, and formative assessment into all stages of Think-Pair-Share learning. In contrast to previous research which generally only developed speaking assessment rubrics or tested the effectiveness of the Think Pair Share model as a learning strategy, this study produced a complete and integrated assessment system, including teacher guides, assessment rubrics, observation sheets, grade recapitulation formats, and student reflection sheets.

This model makes a practical contribution to English teachers because it is able to produce a more objective, systematic, and sustainable assessment process. In addition, TPS-SAM also supports the implementation of the Independent Curriculum which emphasizes the importance of formative assessments and student-centered learning. From a broader perspective, the results of this study support the achievement of Sustainable Development Goal (SDG) 4, especially in improving the quality of education through the innovation of assessment systems that are able to develop communication competencies, collaboration, critical thinking, and self.

Conclusion

This research has succeeded in developing the Think Pair Share-Based Speaking Assessment Model (TPS-SAM) as an authentic assessment model for English learning in Madrasah Aliyah that meets valid, practical, and effective criteria. The validation results showed that the product had a very high level of validity (Aiken's $V = 0.92$), supported by excellent instrument reliability (Cronbach's Alpha = 0.93 and ICC = 0.91). The results of the implementation showed that the model obtained a very high level of practicality based on the responses of teachers (95.5%) and students (95.0%), so that it was easy to apply in the learning process. In addition, the use of TPS-SAM was proven to significantly improve students' oral communication skills, as shown by an increase in the average score from 61.43 to 84.67, an N-Gain value of 0.60 (medium category), a significant paired sample t-test result ($p < 0.001$), and a Cohen's d value of 1.82 which showed a very large influence. Thus, TPS-SAM can be recommended as a cooperative learning-based assessment model that is able to support the implementation of authentic assessments, improve the quality of English language learning, and contribute to the achievement of Sustainable Development Goal

(SDG) 4 through strengthening 21st century communication competencies.

Acknowledgments
I would like to thank Prof. Dr. Alwen Bentri, M.Pd., Prof. Dr. Abna Hidayati, M.Pd, Dr. Nofri Hendri, M.Pd., and Dr. Rayendra, M.Pd.

Author Contributions
Conceptualization, M.F. and A.B.; methodology, software, formal analysis, investigation, resources, data curation, writing—original draft preparation, visualization, M.F.; validation, A.H., N.H., and R.; writing—review and editing, supervision, A.B. All authors have read and agreed to the published version of the manuscript.

Funding
This research received no external funding.

Conflicts of Interest
The authors declare no conflict of interest.

References

Anas, N., & Sapri, S. (2022). Komunikasi Antara Kognitif dan Kemampuan Berbahasa. *EUNOIA (Jurnal Pendidikan Bahasa Indonesia)*, *1*(1), 1. https://doi.org/10.30821/eunoia.v1i1.997

Annajati, S., & Wikarya, Y. (2023). Pengembangan Media Pembelajaran Pembelajaran Interaktif Berbasis Google Slide pada Materi Seni Rupa Tiga Dimensi Kelas X SMA Negeri 8 Mukomuko. *Jurnal Pendidikan Tambusai*, *7*(1), 1181–1188. https://doi.org/10.31004/jptam.v7i1.5405

Apriliani, K., Hermawan, G. S., & Yeni, Y. (2021). Pelaksanaan Penilaian Autentik dalam Pembelajaran Bahasa Jepang di Kelas X SMK Negeri 1 Singaraja Berbasis Kurikulum 2013. *Jurnal Pendidikan Bahasa Jepang Undiksha*, *7*(2), 171–181. https://doi.org/10.23887/jpbj.v7i2.33276

Arsita, S. I. P., & Fathoni, A. (2022). Analisis Faktor Hambatan Guru dalam Melaksanakan Authentic Assesment di Sekolah Dasar. *Jurnal Basicedu*, *6*(4), 6605–6612. https://doi.org/10.31004/basicedu.v6i4.3319

Aspriyani, R., & Suzana, A. (2020). Pengembangan E-Modul Interaktif Materi Persamaan Lingkaran Berbasis Realistic Mathematics Education Berbantuan Geogebra. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, *9*(4), 1099. https://doi.org/10.24127/ajpm.v9i4.3123

Benjelloun, M. (2021). The Role of 'Think-Pair-Share' in Enhancing the Moroccan University Students' Speaking Skills in Online Classes. *International Journal of Second and Foreign Language Education*, *1*(1), 17–27. https://doi.org/10.33422/ijsfle.v1i1.60

Branch, R. M. (2009). *Instructional Design: The ADDIE Approach*. Springer US. https://doi.org/10.1007/978-0-387-09506-6

Darmawan, Y. R. R., & Lestariningsih, F. E. (2023). Students' Views Toward the Use of Think-Pair-Share Strategy on Students' Oral Communication Skill in an EFL Context. *Journal of Languages and Language Teaching*, *11*(4), 798. https://doi.org/10.33394/jollt.v11i4.8694

El-Koumy, A. S. A. (2004). Language Performance Assessment: Current Trends in Theory and Research. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2365667

Eriansyah, Y., & Baadilla, I. (2023). Model Discovery Learning untuk Meningkatkan Kemampuan Berpikir Kritis Siswa pada Muatan Pelajaran Bahasa Indonesia di Sekolah Dasar. *Edukasiana: Jurnal Inovasi Pendidikan*, *2*(3), 151–158. https://doi.org/10.56916/ejip.v2i3.378

Fiktorius, T. (2020). Mendongkrak Autentisitas Penilaian Bahasa Inggris Melalui Penerapan Sistem Penilaian Elasi. *Jurnal Guru Dikmen dan Diksus*, *2*(1), 38–48. https://doi.org/10.47239/jgdd.v2i1.46

Gunawan, D., Mustofa, B., & Wahyudin, D. (2022). Pengembangan Desain Pembelajaran Berbasis Verbal Linguistik Intelligence untuk Meningkatkan Keterampilan Berbahasa Peserta Didik. *Jurnal Basicedu*, *6*(2), 2979–2993. https://doi.org/10.31004/basicedu.v6i2.2541

Hanafi, N., Farmasari, S., Mahyuni, M., Amin, M., & Lestari, Y. B. (2021). Pelatihan Pengembangan Model Penilaian Otentik (Authentic Assessment) pada Pembelajaran Bahasa Inggris Sekolah Dasar bagi Guru-Guru Bahasa Inggris Sekolah Dasar di Kabupaten Lombok Barat. *Jurnal Pengabdian Magister Pendidikan IPA*, *4*(2). https://doi.org/10.29303/jpmpi.v4i2.855

Hasanah, Y. A. (2022). Implementasi Strategi Bermain Peran dengan Stik Boneka untuk Meningkatkan Kreativitas dan Kemampuan Berbicara Bahasa Inggris. *Realita: Jurnal Penelitian dan Kebudayaan Islam*, *17*(1). https://doi.org/10.30762/realita.v17i1.1384

Hidayat, D. Z., Hidayat, A., Wahyuni, M., Marta, R., & Kusuma, Y. Y. (2026). Penerapan Model Pembelajaran Think Pair Share (TPS) untuk Meningkatkan Keterampilan Berbicara Siswa Sekolah Dasar. *Jurnal Ilmiah Profesi Pendidikan*, *11*(1), 411–419. https://doi.org/10.29303/jipp.v11i1.4407

Indarta, Y., Jalinus, N., Waskito, W., Samala, A. D., Riyanda, A. R., & Adi, N. H. (2022). Relevansi Kurikulum Merdeka Belajar dengan Model Pembelajaran Abad 21 dalam Perkembangan Era Society 5.0. *EDUKATIF: Jurnal Ilmu Pendidikan*, *4*(2),

3011–3024. https://doi.org/10.31004/edukatif.v4i2.2589

Ismail, R., Erawadi, E., & Zulhammi, Z. (2023). Implementasi Penilaian Autentik pada Mata Pelajaran Al-Qur'an Hadits di MAN 1 Tapanuli Tengah. *ISLAMIKA*, *5*(3), 1293–1310. https://doi.org/10.36088/islamika.v5i3.3683

Linggasari, E., & Rochaendi, E. (2022). Indonesian Language Learning in Elementary Schools Through Life Skills Education Model. *LITERASI (Jurnal Ilmu Pendidikan)*, *13*(1), 40. https://doi.org/10.21927/literasi.2022.13(1).40-62

Maulani, M. N., Romansyah, H. K., & Atiyawati, M. I. (2020). Using Think Pair Share Teaching Strategy to Students in Teaching Speaking. *The Journal of English Literacy Education: The Teaching and Learning of English as a Foreign Language*, *6*(2), 101–106. https://doi.org/10.36706/jele.v6i2.9951

Nurhaifa, I., Hamdu, G., & Suryana, Y. (2020). Rubrik Penilaian Kinerja pada Pembelajaran STEM Berbasis Keterampilan 4C. *Indonesian Journal of Primary Education*, *4*(1), 101–110. https://doi.org/10.17509/ijpe.v4i1.24742

Qamariah, Z. (2023). Eksposur Materi Autentik di Kelas Bahasa Inggris sebagai Bahasa Asing. *Jurnal Pendidikan dan Sastra Inggris*, *3*(1), 51–62. https://doi.org/10.55606/jupensi.v3i1.1229

Reimann, N., Sambell, K., Sadler, I., & Kreber, C. (2021). Addressing the Challenges of Assessment and Feedback in Business Schools: Developing Assessment Practices Which Support Learning. In *Handbook of Teaching and Learning at Business Schools*. https://doi.org/10.4337/9781789907476.00025

Rosyada, A. (2024). The Innovation of Arabic Language Learning Assessment: Performance Assessment on Kalam Skills in MAN 3 Sleman. *Inovasi-Jurnal Diklat Keagamaan*, *18*(1), 47–59. https://doi.org/10.52048/inovasi.v18i1.478

Setiawan, I. (2021). Pengajaran Bahasa Inggris dalam Kurikulum 2013: Suatu Tinjauan dan Perspektif. *Jurnal Ilmiah Pendidikan dan Keislaman*, *1*(2), 104–112. https://doi.org/10.55883/jipkis.v1i2.12

Sugiyono, S. (2018). *Metode Penelitian Pendidikan Pendekatan Kuantitatif, Kualitatif dan R&D*. Bandung: Alfabeta.

Triana, N., Utaminingsih, S., & Khamdun. (2024). The Effectiveness of Steam-Based Interactive Module to Improving Learning Outcomes of Elementary School Students. *ICCCM Journal of Social Sciences and Humanities*, *3*(1), 37–45. https://doi.org/10.53797/icccmjssh.v3i1.6.2024

Wiartis, W. (2021). Penerapan Model Pembelajaran Project Based Learning untuk Meningkatkan Kemampuan Belajar Siswa Melalui Kegiatan Kolaborasi Mata Pelajaran Bahasa Inggris di SMP Negeri 6 Batam. *Daiwi Widya*, *7*(5), 17. https://doi.org/10.37637/dw.v7i5.673

Wijayanto, A. (2023). *Holistik Teknologi Pembelajaran dalam Bahasa dan Literasi*. Open Science Framework. https://doi.org/10.31219/osf.io/dwme7

Wulandari, R. (2023). Implementation of The Think Pair Share Learning Model to Improve Student Learning Outcomes on Algebra Function Limits Material in Class XI IIS 1. *Inovasi-Jurnal Diklat Keagamaan*, *17*(2), 194–204. https://doi.org/10.52048/inovasi.v17i2.438

Zaim, M. (2018). *Asesmen Otentik: Implementasi dan Permasalahannya dalam Pembelajaran Bahasa Inggris di Sekolah Menengah*. INA-Rxiv. https://doi.org/10.31227/osf.io/wc8ha