Randwick International of Education and Linguistics Science (RIELS) Journal
Vol. 5, No. 1, March 2024 | Page 196-210 |
https://www.randwickresearch.com/index.php/rielsj

Innovation in English Language Learning Assessment in Class VII of SMP Negeri 1 Dolok Merawan: Development of a HOTS-Based Evaluation Tool with the Help of Wordwall

DOI: *https://doi.org/10.47175/rielsj.v5i1.911*

| Nurul Aprilia Yusri[1,*] | Abdul Hasan Saragih[2] | Naeklan Simbolon[3] |

[1,2,3]*Department of Education Technology, Postgraduate, Universitas Negeri Medan, Medan, Sumatera Utara, Indonesia*

*nurulapriliay@gmail.com

ABSTRACT

An evaluation instrument or evaluation tool is a tool that meets academic requirements. The evaluation instrument still uses paper evaluation tools and does not use online learning evaluation. The aim of this research is to develop a wordwall-assisted HOTS-based English language learning evaluation tool. This type of research applies the Research and Development (R&D) research method. The subjects in this research were class VII students of SMP Negeri 1 Dolok Merawan who used the HOTS-based learning evaluation tool with the help of Wordwall, for a total of 32 students. The object of research is English-language reading material in the form of wordwall assistance. This research has been validated through stages by material experts, media experts, instructional design experts, language experts, and evaluation experts. This research has trial stages, namely initial field trials, main field trials, and operational field trials. This research also tested the effectiveness of the wordwall-assisted HOTS-based English learning evaluation tool. The research results show that students who use the HOTS-based English language learning evaluation tool assisted by Wordwall are "more effective" than those who use conventional learning methods. This is shown by the calculation results obtained: tcount = 2.28 and ttable = 2.00, where 2.28 > 2.00 for a significance level of 0.05. The effectiveness value of the wordwall-assisted HOTS-based English language learning evaluation tool is higher, namely 80.28%, than student learning outcomes using conventional methods with a value of 75.25%.

KEYWORDS

development of evaluation tools; english learning results; HOTS

INTRODUCTION

Learning success is largely determined by the selection of learning evaluation instruments. Choosing a learning evaluation tool is something that is very essential in the learning mechanism, because choosing the right learning evaluation tool will facilitate the learning process and make it easier to achieve learning goals. Setemen (2010: 43) Evaluation can be applied to identify how far students have mastered the material. Evaluation plays an essential role in carrying out learning in educational institutions, where, in its implementation, teachers must master the assessment process and learn pedagogical skills. The accuracy of data on students' competencies or barriers to learning refers to the evaluation mechanism and instruments applied. After the learning mechanism activities carried out by teachers and students, the teacher will assess as part of the evaluation to see the extent to which students have succeeded in mastering the teaching materials provided

with measurement instruments in the form of tests or non-tests, which are the aim of assessing learning gains. Thus, teachers must pay attention to the development of learning evaluation tools, namely HOTS-based learning evaluation tools, for the characteristics and needs of students in accordance with the curriculum, which demands willingness to participate and activate in learning.

According to Daryanti and Dwicahyono (2014), learning evaluation tools are a series of teaching materials that are systematized and written so that there is a learning nuance that is appropriate to the students' existence. Learning evaluation tools are divided into two types, namely print and online learning evaluation tools. Printed learning evaluation tools are a series of learning evaluation tools in the form of sheets of paper that contain learning materials, summaries, and various instructions for working on learning tasks that need to be completed by students. Meanwhile, online learning evaluation tools are in the form of audio, for example, radio. Audiovisual learning evaluation tools, for example, CAI (computer-assisted instruction), can be categorized as web-based learning evaluation tools (web-based learning materials).

Providing English lessons in class is expected to develop students' insight and skills in applying language to communicate. According to Brown (2001: 233), when studying English, students need to master four areas of language skills, which include listening, speaking, reading, and writing. Listening and reading are receptive language competencies, while speaking and writing are productive language competencies. To support this program, students must be accustomed to thinking at a higher level, or HOTS. This is because to be able to compete in a digitalized, universal world, students must have the ability to think critically and creatively. So that HOTS in English learning can be used as a provision when faced with the Industrial Revolution 4.0. If students do not have HOTS skills, they will have difficulty when faced with international challenges in the future.

From observations carried out by researchers at SMP Negeri 1 Dolok Merawan, Kec. Dolok Merawan, Kab. Serdang Bedagai, North Sumatra in December 2022, it was discovered that the school had carried out a type of assessment in accordance with the 21st century, namely by combining formative (when learning is taking place) and summative (as an instrument of accountability for whether learning is taking place) assessments, but in its implementation there were many new alignments that were aligned with the needs and context of current developments. Evaluation of learning, especially English subjects used by the school, is not yet HOTS-based. The learning evaluation used so far is still based on Lower Order Think Skills (LOTS), and there is no change from the previous year, meaning that teachers encourage students to memorize more because they are not yet fluent in making HOTS learning evaluations. The LOTS-based learning evaluation used by the teacher is still based on printed books, so it does not require students to be able to think creatively, critically, and at a high level.

The limited variety of learning evaluation tools causes students to lack interest in the evaluation process. This is because teachers have not been able to develop learning evaluation tools, especially those that use technology in the process, such as smartphones and computers. It can be seen that when students take exams at SMP Negeri 1 Dolok Merawan, they still use paper evaluation tools and do not use online learning evaluation. In this technological era, teachers should be able to use online learning evaluations. Online-based evaluation has more advantages than paper-based evaluation.

According to Redecker (2013: 79), online-based evaluation can increase the efficiency and effectiveness of tests. Apart from that, online-based evaluation can also increase the validity and reliability of test scores. Meanwhile, from the student aspect, the use of

online-based evaluation can increase student motivation in learning, concentration, and performance. According to Barth B. Riley and Adam C. Carle (2012), evaluation tools printed on paper have various disadvantages, including requiring large costs for printing test papers and using excess paper, which is now inefficient, and requiring more space for storing test question scripts and sheets. answers, it takes time to copy questions and correct them, even if you use a scanning tool. The use of paper as a learning evaluation tool at this time also influences students' interest in working on questions, even though utilizing technology and information in evaluating learning can increase students' interest in solving various problems.

Based on terminology, Echols and Shadly (1976) stated that evaluation comes from the word evaluation, which means evaluation. Thoha (2003) stated that "evaluation is a planned activity to identify the condition of an object by using tools, and the results are compared with parameters to reach conclusions.".

In general, Cross (2013) argues that "evaluation is a mechanism that determines the circumstances in which goals have been achieved." This definition explains clearly that evaluation is a mechanism for obtaining information and understanding and conveying this information to those making decisions.

This is in line with Wirawan (2011: 8), who states that evaluation aims to collect information, measure performance, and assess the usefulness of the evaluation object in relation to the parameters, objectives, or criteria of the object.

The following is a more detailed explanation of the differences between evaluation, assessment, measurement, and testing in Table 1:

Table 1. Differences in Evaluation, Assessment, Measurement and Tests

Aspects	Definition	Process	Results
Evaluation	Activities that include two components, namely measurement and assessment	Determination of decisions regarding assessment output pass/fail	Decision or justificationi
Evaluation	Make decisions according to good standards or bad	Presentation of attributes for obtaining	Interpretation measurements with qualitative characteristics
Measurement	The mechanism of determining the quantity of a thing that creates a Number	Comparing test results with specific criteria	Numbers or scores are quantitative
Test	Instrument for measuring a person's individual capabilities	Testing	Obtaining tests or sheets Work

(Source: Adi, 2016)

Based on the expert opinion above, it can be concluded that evaluation, assessment, measurement, and testing have their own roles and positions. Evaluation is broader in scope than assessment, while assessment focuses on specific components that are part of the evaluation. Evaluation and assessment have qualitative characteristics, while measurements and tests have quantitative characteristics in the form of ordinal data.

Ratnawulan and Rusdiana (2015) define learning evaluation as a mechanism for measuring and assessing various student capabilities in learning, such as insight, ethics, and skills, to create decisions regarding student capabilities.

Mehrens, W. A. and I.J. Lehman (1991) Learning evaluation is a mechanism for obtaining data and information needed to determine how far learning has taken place in order to create assessments and improvements needed to optimize the output.

Arikunto (2019) means that learning evaluation is a mechanism for collecting data to identify how far educational targets have been achieved. In line with Gronlund (1976), learning evaluation is a systematic mechanism for determining or making decisions about how various goals have been achieved by students.

Evaluation has many functions. Another use of evaluating learning, according to Sajadi (2013), is determining, reviewing the curriculum, comparing, predicting, and realizing. Meanwhile, according to Jahanian (2012), evaluating education has various functions and benefits, including diagnosing, revising the syllabus, comparing, and analyzing needs. Therefore, it can be said that learning evaluation is a mechanism or instrument applied to the learning mechanism that is useful for identifying and determining how far the student's competence has developed in understanding the material that has been given.

Wordwall is a platform that provides educational features that are not included in the application but are a website that provides educational quiz features and can be an interactive game that can be applied to teaching. The quiz can be developed to be more attractive and enjoyable with existing features. Apart from that, Wordwall can be used to fill gaps in studying.

According to Wagstaff (1999), Wordwall has various types of evaluation games that are attractive and increase students' interest, such as crossword puzzles, which can be applied to testing students' ability to remember and their astuteness; word searches, which can test students' astuteness; and match-ups, which can be applied to testing students' understanding. of teaching materials. By implementing a wordwall, students can develop critical and active reading competencies. Those who experience difficulties in identifying the right vocabulary can use wordwalls as reference material.

Wordwall is quite easy to use for students or teachers, unlike Kahoot, which requires downloading an application. The way Wordwall works is by simply sending a link to students, and they can rely on the quiz that the teacher has created. Not many people know about wordwall, but research on wordwall was conducted by Agrin (2020). In her research, Agrin (2020) stated that when presenting this game, her students felt it was attractive and were enthusiastic about playing it.

The various types of games that Wordwall offers are in the form of quizzes and crosswords. There are also other types of games available, such as Random Wheel, True or False, Missing Word, and many more. Apart from that, Wordwall also provides a duration limit feature, and the dashboard will contain the names of students who took the quiz along with the grades they got. Apart from that, on the wall, the teacher can review the length of time students need to complete the quiz, and the teacher can also review the parts where students made mistakes when answering the quiz. This can be used as material for the teacher to review with students who do not understand the lesson being presented.

Thorne & Thomas (2009) stated that HOTS is a thinking mechanism at a higher level than remembering facts or re-explaining something studied to other individuals. HOTS requires individuals to understand, formulate, and relate facts to concepts, group, manipulate, trace the truth in an ongoing event, and explore solutions to problems that occur.

HOTS is a student thinking mechanism at a higher cognitive level that is developed from various cognitive descriptions and techniques and educational taxonomies such as educational bloom taxonomy, problem-solving techniques, and assessment. HOTS

includes the capabilities of problem solving, creative and critical thinking, debating, and decision-making.

According to Widodo (2013), with HOTS, students can differentiate their ideas clearly, argue well, handle problems, construct presentations, hypothesize, and understand various complex issues clearly. HOTS will occur when individuals correlate current information with what is already in their memory and organize it to achieve goals or identify treatments for conditions that are difficult to resolve. The core objective of HOTS is to increase students' thinking capabilities at a higher level, especially with regard to the ability to think critically in obtaining various types of information, creativity in solving problems with existing insights, and making decisions in various complex conditions.

Based on Bloom's Taxonomy, which has been improved by Brookhart (2010), the cognitive domain is divided into LOTS (low-level thinking) and HOTS (high-level thinking). LOTS covers cognitive domains such as memory, understanding, and application. Meanwhile, HOTS includes study, evaluation, and creation (Anderson & Krathwohl, 2001).

Based on the problem boundaries described above, the problem formulation for this research is: (1) Is HOTS-based evaluation of English language learning assisted by wordwall suitable for application to class VII students of SMP Negeri 1 Dolok Merawan?; and (2) Is the wordwall-assisted HOTS-based English learning evaluation effective for class VII students at SMP Negeri 1 Dolok Merawan?

RESEARCH METHODS

This research was conducted to create a HOTS-based learning evaluation with the help of wordwall media for English language learning. According to Borg and Gall (Sugiyono, 2019), this research is included in the type of research and development (R&D), namely the type of research that develops a new product or complements one that is already available.

According to Sugiyono (2019), R&D is a research method applied to create special products and test their effectiveness. Simply put, it can be explained that R&D is a research method that is deliberate, systematic, aimed at, or instructed to identify, conclude, revise, develop, create, and test the effectiveness of products, forms, techniques, services, and special mechanisms that are more competent and newest. Effectiveness, efficiency, and productivity are high and meaningful. An almost similar definition is presented by Neuman (2003): this research is a mechanism applied to develop and validate various products applied to learning mechanisms.

From these various views, it can be concluded that the method applied in research is a type of research to create various products needed to educate through the development and validity of these products, the final stage of which is evaluation.

Learning development, especially HOTS-based learning evaluation with the help of wordwall media for Class VII Middle School students at SMP Negeri 1 Dolok Merawan, The characteristic of this research is that it perfects the evaluation tools that were previously available. The creation of HOTS-based learning evaluations with the help of wordwall media is expected to be carried out with careful preparation and design in a systematic manner and can be developed in line with expertise, syllabus, and RPP so that the learning evaluation tool becomes more effective and has been modified into a less complex research plan. .

This research was conducted at SMP Negeri 1, Dolok Merawan. For class VII students in Semester I of the 2022-2023 academic year. The research will be carried out in August 2023. The determination of treatment is adjusted to the existing school calendar.

Research subjects are the individuals involved in it. The subjects of this research were class VII students of SMP Negeri 1 Dolok Merawan who used the HOTS-based learning evaluation tool with the help of Wordwall, for a total of 32 students. The object of research is English-language reading material in the form of wordwall assistance.

A trial of the HOTS-based learning evaluation product with the help of a wordwall was carried out after this product was validated through discussions with two learning media expert educators and two English language material expert educators. From the trial, shortcomings could be identified, which then became material for improving the HOTS-based learning evaluation product with the help of a wordwall. . In the trial, the targets assessed were weaknesses or obstacles that emerged in further revisions (Sugiyono, 2019).

The product testing stages are: (1) First Stage: Phase I testing of the use of HOTS-based learning evaluation products with the help of a wordwall totaling 5 students. Revision I; (2) Second Stage: Phase II testing of the use of HOTS-based learning evaluation products with the help of a wordwall totaling 15 students. Revision II.; (3) Third Stage: Stage III testing of users of HOTS-based learning evaluation products with the help of Wordwall with a total of 32 students. Revision III; (4) Fourth Stage: Procurement of HOTS-based learning evaluation products with the help of a wordwall; (5) Fifth Stage: Product assessment of HOTS-based learning evaluation products with the help of a wordwall.

The instrument for collecting data in this research is a product assessment instrument that has been developed. The core instrument used in collecting data is a questionnaire sheet on reading skills aimed at: (1) students' needs (2) teacher needs, (3) material expert validators, (4) media expert validators, (5) expert validators of instructional design, (6) language expert validators, and (7) learning evaluation expert validators (8) testing tools for students. The main instrument was used to collect data in product trials.

Evaluation Tool Feasibility Questionnaire Analysis. The assessment form filled in by experts is presented in the product feasibility table so that it becomes the basis for revising each English evaluation instrument. The sheet is then studied to identify the quality of the product created by the researcher.

$$\text{Score Presentation} = \frac{\text{Number of Indicators per Category}}{\text{Number of Indicators Total Category}} \times 100\%$$

(Source: Yamasari, 2010)

From calculations applying the formula above, a figure in the form of a percentage (%) is obtained. This score classification is then transformed into a classification in the form of a percentage, then interpreted using qualitative sentences as shown in Table 2 below:

Table 2. Indicator Presentation Criteria for Products that Have Been Developed

No	Score Interval	Interpretation
1	0.00 – 2.49	Not Good / Not Worth It
2	2.50 – 3.32	Not Good / Not Appropriate
3	3.33 – 4.16	Good / Decent
4	4.17 – 5.00	Very Good / Very Decent

(Source: Sriadhi, 2018)

Presentation of student learning outcomes from experimental and control classes with assessment criteria is in Table 3 below:

Table 3. Assessment Criteria

Value	Criteria	Percentage (%)
A	Very Good	81-100%
B	Good	61-80%
C	Sufficient	41-60%
D	Poor	21-40%
E	Very Poor	0-20%

Next, the learning outcomes were tested by applying the t-test
Hypothesis:
Ho : $\mu A1 \leq \mu A2$
Ha : $\mu A1 > \mu A2$
$\mu A1$ = average learning outcomes of control class students
$\mu A2$ = average learning outcomes of experimental class students

RESULTS AND DISCUSSION

Results

Product validation aims to identify the opinions of material, media, instructional design, language and evaluation experts regarding material provisions, learning aspects and validity of content, media, teaching design, language and evaluation. The validation obtained by material experts can be reviewed in Table 4 below:

Table 4. Validation Results by Material Experts

No	Assessment Aspect	Material Expert	Criteria
1	Guide and Information	4,50	Very Good
2	Multimedia Content/Materials	4,50	Very Good
3	Evaluation	4,12	Good
Final Score Interval		**4,37**	Very Good

Media expert validation of the development of a wordwall-assisted HOTS-based English learning evaluation tool.

Table 5. Validation Results by Media Experts

No	Assessment Aspect	Media Expert		Average	Criteria
		1	2		
1	Guide and Information	4,33	4,66	4,49	Very Good
2	Program Performance	4,57	4,28	4,42	Very Good
3	Media Systematics	4,66	4,33	4,49	Very Good
4	Media Aesthetics	4,50	4,25	4,37	Very Good
5	Narration and Audio Quality	4,85	4,42	4,63	Very Good
6	Video Quality	4,50	4,50	4,50	Very Good
7	Principles of Multimedia Design	4,42	4,00	4,21	Very Good
Final Score Interval		**4,54**	**4,34**	**4,44**	**Very Good**

Validation of instructional design experts on the development of a wordwall-assisted HOTS-based English learning evaluation tool.

Table 6. Validation Results by Instructional Design Experts

No	Assessment Aspects	Design Expert		Average	Criteria
		1	2		
1	Kurikulum	5,00	5,00	5,00	Very Good
2	Metode	4,50	4,16	4,33	Very Good
3	Evaluasi	5,00	4,75	4,87	Very Good
Final Score Interval		**4,83**	**4,63**	**4,73**	**Very Good**

Linguist validation of the development of a wordwall-assisted HOTS-based English learning evaluation tool

Table 7. Validation Results by Linguistic Experts

No	Assessment Aspects	Linguist Expert	Criteria
1	Language Rules	4,80	Very Good
2	Readability	4,88	Very Good
3	Completeness of the Manual	5,00	Very Good
4	Languages	5,00	Very Good
Final Score Interval		**4,92**	**Very Good**

Validation of evaluation experts on the development of HOTS-based English language learning evaluation tools assisted by wordwall

Table 8. Validation Results by Evaluation Experts

No	Assessment Aspects	Evaluation Experts		Average	Criteria
		1	2		
1	Hint	4,66	4,00	4,33	Very Good
2	Construction	4,42	3,43	3,93	Good
3	Languages	4,66	4,00	4,33	Very Good
Final Score Interval		**4,58**	**3,81**	**4,20**	Very Good

Operational field trials were carried out on a total of 32. The results of operational field trials in the form of assessment scores on the wordwall-assisted HOTS-based English learning evaluation tool can be seen in Table 9. below:

Table 9. Operational Field Trial Assessment Scores

Aspects	Indicators	Average	Criteria
Information Guide	- Description of wordwall media is very clear	4,56	Very Good
	- Guide to using wordwall media is easy to understand	4,75	Very Good
	- Rumusan tujuan sangat jelas	4,71	Very Good
	- The formulation of objectives is very clear	4,68	Very Good
Presentation of Learning	- Material according to the topic of discussion	4,78	Very Good
	- Material supports achievement of learning objectives	4,68	Very Good
	- The material is appropriate to the thinking abilities of students (users)	4,65	Very Good
	- The material is in accordance with current developments in science and technology	4,84	Very Good
	- The material explains concepts or theories according to learning objectives	4,56	Very Good
	- Presentation of material arranged sequentially (hierarchically)	4,68	Very Good
	- The breadth of material coverage is in accordance with the learning objectives	4,71	Very Good

Aspects	Indicators	Average	Criteria
	- Use of terms and formulas is easy to understand	4,90	Very Good
	- Providing examples or illustrations that are easy to understand	4,75	Very Good
	- Material summary includes all material presented	4,68	Very Good
	- Duration of use according to the material presented	4,68	Very Good
	- Spelling and grammar are easy to understand	4,84	Very Good
	- Writing material arranged in an orderly manner (systematic)	4,78	Very Good
	- Media provides instructions for taking exercises/exam	4,84	Very Good
	- Practice/exam questions support achievement of learning objectives	4,84	Very Good
Evaluation	- The exam time corresponds to the number and level of difficulty of the questions	4,71	Very Good
	- Question items vary according to learning objectives	4,93	Very Good
	- Exam results are reviewed or returned in the media	4,87	Very Good
Media Design and Facilities	- Easy to use learning media	4,84	Very Good
	- Command buttons have accurate links (hyperlinks)	4,90	Very Good
	- Learning media can be run without damage	4,87	Very Good
	- Media provides interactive facilities for users	4,84	Very Good
	- Letters, numbers and symbols on the media are written clearly	4,93	Very Good
	- Visual images (graphics) on media are very good	4,53	Very Good
	- Audio quality is very good	4,90	Very Good
	- Video quality is very good 4.78 Very Good	4,78	Very Good
	- Animation quality is very good	4,71	Very Good
	- The coloring of the media content is very good	4,84	Very Good
Pedagogical Effects	- his media provides what students need	4,93	Very Good
	- Students are interested in using this media to learn	4,96	Very Good
	- This media makes students more enthusiastic/active in learning	4,96	Very Good
	- This media makes it easier to understand learning material	4,78	Very Good
	- This media makes it easier to increase student capabilities	4,75	Very Good
Total Score		**177,12**	Very Good
Score Interval		**4,78**	

Table 9 identifies that the assessment score interval for operational field trials on 32 students is 4.78 and has the criteria "Very Good" or "Very Eligible".

Table 10. Frequency Distribution of Student Learning Outcomes Taught Using Wordwall-Assisted HOTS-Based English Learning Evaluation Tools

No	Class Interval	Fi	Percentage
1	60-65	4	12,5%
2	66-71	2	6,25%
3	72-77	5	15,63%
4	78-83	9	28,12%
5	84-89	5	15,63%
6	90-95	7	21,87%
Total		32	100%

Based on Table 10 above, it can be seen that the value of student learning outcomes taught by applying the HOTS-based English learning evaluation tool with the help of wordwall is 11 students with a percentage of 34.38% below the average, as many as 9 students with a percentage of 28. .12% is right on average, and there are 12 students with a percentage of 37.5% above average. Furthermore, the learning gains of students who are taught using conventional methods.

The distribution of the frequency of prostest learning results for students who were taught using this method can be seen in Table 11 below:

Table 11. Frequency Distribution of Student Learning Outcomes Taught Using Conventional Methods

No	Class Interval	Fi	Percentage
1	50-56	2	6,25%
2	57-63	3	9,38%
3	64-70	5	15,63%
4	71-77	9	28,12%
5	78-84	7	21,87%
6	85-91	6	18,75%
Total		32	100%

Based on Table 11 above, it is identified that the learning achievement scores of students taught using conventional methods, namely 10 students (31.26%) are below the average, as many as 9 students with a percentage of 8.12% are right on average. , and as many as 13 students with a percentage of 40.62% above the average.

The normality test was applied to the data for both treatments, namely the experimental and control class groups. The normality test results can be reviewed in Table 12 below:

Table 12. Summary of Normality Tests for Experimental and Control Class Data

No	Class	Lcount	Ltable	Conclusion
1	Experiment	0,113	0,156	Normal
2	Control	0,083	0.156	Normal

The calculation results in Table 12 show a value of 0.083<0.156. From these data, it can be concluded that the posttest data from the experimental and control classes are distributed normally at a significance level of 0.05.

Table 13. Summary of Homogeneity Tests for Experimental and Control Class Data

Class	N	dk (n-1)	Varians (Si^2)	Fcount	Ftable	Conclusion
Experiment	32	31	84,33	1,35	1,82	Homogen
Control	32	31	114,29			

Listing the F distribution with dk = n-1 at a significance level of 0.05, we get F0.05 = 1.82. The calculation results in Table 4.18 of the experimental and control class data obtained a value of 1.35 < 1.82. Based on these data, it can be concluded that Ho is accepted, which means the data comes from a homogeneous group. So that the experimental class and control class data spread normally.

Hypothesis testing in this research was carried out by applying the t test formula. The t test was carried out to identify significant differences between learning gains using the wordwall-assisted HOTS-based English learning evaluation tool and student learning gains

in classes taught using traditional methods. From the calculation results, tcount = 2.28 and ttable = 2.00, where 2.28 > 2.00 for a significance level of 0.05. From this data, it can be concluded that learning gains using the HOTS-based English language learning evaluation tool assisted by Wordwall are higher than learning results using conventional methods at a significance level of 5%.

The effectiveness value of the wordwall-assisted HOTS-based English learning evaluation tool is obtained as follows:

$$X = \frac{\text{total score obtained}}{\text{ideal score}} \text{ x } 100\%$$

$$X = \frac{2569}{3200} \text{ x } 100 = 80{,}28\ \%$$

Meanwhile, the effectiveness of learning using conventional methods is obtained as follows:

$$X = \frac{\text{total score obtained}}{\text{ideal score}} \text{ x } 100\%$$

$$X = \frac{2408}{3200} \text{ x } 100 = 75{,}25\%$$

Thus, the learning outcomes of students who use the HOTS-based English language learning evaluation tool assisted by Wordwall are higher, with a value of 80.28%, than the learning outcomes of students using conventional methods, with a value of 75.25%. So it can be concluded that the wordwall-assisted HOTS-based English learning evaluation tool is more effectively applied to increase student learning gains.

Discussion

The development of a wordwall-assisted HOTS-based English learning evaluation tool was guided by the R&D model from Borg and Gall. This research was instructed to create a product in the form of a wordwall-assisted HOTS-based English learning evaluation tool that is applied to improve learning mechanisms or student skills.

Walker (2012) stated in her research that evaluation should be meaningful and encourage in-depth learning to develop independent and motivated thinkers while meeting assessment criteria by providing feedback at the end of the meeting.

The HOTS-based wordwall-assisted English language learning evaluation tool development product has been developed by paying attention to aspects of material, media, instructional design, language, and evaluation. This research and development was carried out to create a wordwall-assisted HOTS-based English learning evaluation tool that is suitable for use in English subjects for class VII students at SMP Negeri 1 Dolok Merawan so that it can simplify the learning mechanism and increase student learning gains.

This research and development procedure begins with a preliminary study, collecting data, creating a product, and product validation testing, which will then be revised and perfected based on suggestions and assessments from experts. The next stage is user testing to create a product that is feasible and useful in the learning mechanism.

To determine whether a product is suitable or not, a validation test is carried out. Validation of products aims at the views of various experts. The improved aspects were completed based on data analysis and testing as well as suggestions from experts and students for using HOTS-based English language learning evaluation tools assisted by

Wordwall to identify various aspects that are common in the development mechanism of a product.

The following is a summary of the results of the feasibility test assessment of the wordwall-assisted HOTS-based English learning evaluation tool by various experts and tests, which can be seen in Table 14 below:

Table 14. Average Summary of Feasibility Test Assessment Results

No	Respondent	Assessment Results	Criteria
1	Material Expert	4,44	Very Good
2	Media Expert	4,73	Very Good
3	Design Expert	4,92	Very Good
4	Linguist Expert	4,92	Very Good
5	Experts Evaluation	4,20	Very Good
6	Initial Field Trials	3,90	Good
7	Main Field Trials	4,38	Very Good
8	Operational Field Test	4,78	Very Good
Average		**4,53**	**Very Good**

Based on research data obtained, there are differences in learning outcomes taught by applying HOTS-based English language learning evaluation instruments assisted by wordwall and by applying traditional methods. namely, the average score for teaching using the HOTS-based English language learning evaluation tool assisted by Wordwall is higher than that using conventional methods. According to Arifin's opinion (2011: 12), the purpose of learning evaluation is to identify the effectiveness and efficiency of the learning network, both with regard to objectives, teaching materials, techniques, facilities, references, environment, or systems.

In accordance with existing research regarding the wordwall-assisted HOTS-based English language learning evaluation tool by Mursid et al. (2017), the research aims to identify whether the learning gains planned for learning with HOTS are higher if taught via online web and offline computers; this is higher if accompanied by a positive rather than negative innovative attitude; and in order to identify the interaction of ICT-based learning facilities with students' innovative attitudes. It is known that learning gains that involve learning plans using online web-based media are greater than offline ICT; learning gains that are accompanied by positive innovative behavior are higher than negative ones; and there is contact between ICT-based learning media and innovative attitudes towards learning acquisition.

Domun and Bahadur (2014) shows in her research that students need to assess their own performance accurately so that they can improve their learning. This study involved the design and development of a self-assessment tool based on the Revised Blooms Taxonomy Framework. The results showed substantial dissimilarity in scores between the treatment and control groups when the independent T-test was used. Group A's success percentage was higher than Group B.

Agus Ramdani et al (2020) in their research, which aims to develop learning tools, especially instruments for evaluating science learning in junior high schools, that, after being applied in the classroom, are expected to develop the skills of 21st century students. The results of this research show that the tools developed by teachers are in a complete category, the learning objectives are clear, although the stages still need to be revised, they are in accordance with various assessment principles, they have been presented in line with

the curriculum, the sentences in the questions are interactive, and they use language that is in line with the rules and does not create bias in the interpretation.

Mursid, Sitompul, and Saragih (2019) in their research entitled, which aims to identify differences between HOTS-based learning models for vocational learning design learning acquisition and differences in creative thinking abilities on learning outcomes in vocational learning design. The results of the research show that there are differences in learning outcomes for vocational learning designs that are taught using HOTS-based learning with multimedia, which are higher than learning outcomes provided using expository learning models with multimedia. There are differences in learning outcomes for vocational learning designs that have higher creative thinking capabilities. than those with low creative thinking capabilities.

Based on research into the development of a HOTS-based English language learning evaluation tool with the help of Wordwall, it can be seen that the results of the t test obtained t_{count} = 2.28 while t_{table} = 2.00, where t_{count} = 2.28 > t_{table} = 2.00. So it can be concluded that the learning gains of students who apply the HOTS-based English learning evaluation tool with the help of Wordwall are higher than the learning gains of students who apply old learning techniques. This can be seen from the average score taught using the HOTS-based English learning evaluation tool with the help of Wordwall, which is 80.28% higher than those using traditional learning methods, namely 75.25%. This data shows that the English learning evaluation tool is based on Wordwall-assisted HOTS and can increase student learning gains. The disparity in learning outcomes using HOTS-based English language learning evaluation tools assisted by wordwall using conventional methods is 5.03%.

CONCLUSION

Based on research obtained on the development of a wordwall-assisted HOTS-based English learning evaluation tool, several conclusions can be drawn, namely: (1) The product in the form of a wordwall-assisted HOTS-based English language learning evaluation tool has been proven to be a suitable product for use by every class VII student at SMP Negeri 1 Dolok Merawan. The assessments of material experts, media experts, instructional design experts, language experts, and evaluation experts show quite high scores and are classified as "very good" criteria. Initial, primary, and operational field trials also created criteria scores of "good" to "excellent." (2) From the data obtained and processed, the average value of learning gains for students who apply the HOTS-based English learning evaluation tool assisted by Wordwall shows that students who use the HOTS-based English learning evaluation tool assisted by Wordwall are "more effective" than those who apply the method of conventional learning. This is shown by the calculation results obtained: t_{count} = 2.28 and t_{table} = 2.00, where 2.28 > 2.00 for a significance level of 0.05. The effectiveness value of the HOTS-based English language learning evaluation tool assisted by Wordwall is higher, namely 80.28%, than the results of student learning using conventional methods, with a value of 75.25%.

REFERENCES

Adi, S. (2016). Evaluasi Pembelajaran di SD. Tangerang Selatan: Universitas Terbuka.

Agrin, D. F. A. (2020). Pengembangan Bahan Ajar Menulis Teks Prosedur dengan Media Berbasis Aplikasi Android G-Teks Pro (Genre Teks Prosedur) untuk Siswa Kelas VII SMPN 216 Jakarta. Skripsi pada Universitas Negeri Jakarta. http://repository.unj.ac.id/9976/

Ramdani, A., Jufri, A. W. & Jamaluddin. (2020). "Pengembangan Media Pembelajaran Berbasis Android Pada Masa Pandemi Covid-19 Untuk Meningkatkan Literasi Sains Peserta Didik." *Jurnal Kependidikan: Hasil Penelitian Dan Kajian Kepustakaan Di Bidang Pendidikan, Pengajaran Dan Pembelajaran 6*(3). https://e-journal.undikma.ac.id/index.php/jurnalkependidikan/article/view/2924

Anderson, L. W. & Krathwohl, D. R. (2001). A Taxonomy for Learning Teaching and Assessing, A Revision of Bloom's Taxonomy of Educational Objective. New York: Addison Wesley Lengman, inc.

Arifin, Z. (2016). Evaluasi Pembelajaran (Prinsip, Teknik, dan Prosedur), Cetakan Kedelapan. Jakarta: Rosda Karya.

Arikunto, S. (2019). Prosedur Penelitian. Yogyakarta: PT Reneka Cipta.

Barth, R. & Carle, A. (2012). Comparison of Two Bayesian Methods to Detect Mode Effects Between Paper-Based and Computerized Adaptive Assessments: A Preliminary Monte Carlo Study. *Journal of Department of Health System Science*, 12 (1), 124. https://bmcmedresmethodol.biomedcentral.com/articles/10.1186/1471-2288-12-124

Brookhart, S. M. (2010), How To Assess Higher-Order Thinking Skills InYour Classroom, USA: ASCD. https://www.daneshnamehicsa.ir/userfiles/files/1/7-%20How%20to%20Assess%20Higher-Order%20Thinking%20Skills%20in%20Your%20Classroom%20-%20Copy%201.pdf

Brown, H. D. (2001). Teaching by Principles an Interactive Approach to Language Pedagogy Second Edition. New York: Pearson EducationCompany.

Cross, M. (2013). Social Media Security: Leveraging Social Networking While Mitigating Risk. Amsterdam: Elsevier Science. https://www.amazon.com/Social-Media-Security-Leveraging-Networking/dp/1597499862

Daryanti & Dwicahyono, A. (2014). Pengembangan perangkat pembelajaran (silabus, rpp, phb, bahan ajar). Yogyakarta: Gava Media.

Domun, M. & Bahadur, G. K. (2014). Design and Development of a Self-Assessment Tool and Investigating its Effectiveness for E-Learning. *European Journal of Open, Distance and E-Learning. 17*(1), 1-25.

Echols, J. M., & Shadly, H. (1976). Kamus Inggris-Indonesia. Jakarta: PT Gramedia.

Gronlund, N.E. (1976). Measurement and evaluation in teaching. New York: Macmillan Publishing Company. https://onesearch.id/Author/Home?author=Norman+E.+Gronlund

Jahanian, R. (2012). Educational Evaluation: Functions and Applications in Educational Contexts. *International Journal of Academic Research in Economics and Management Sciences*, 1(2), 1-9. https://www.semanticscholar.org/paper/Educational-Evaluation%3A-Functions-and-Applications-Jahanian/4b9457558b717e5a19aae3f63539b320d8643da0

Mehrens, W. A., & Lehmann, I. J. (1991). Measurement and Evaluation in Education and Psychology. Fort Woth: Holt, Rinehart and Winston, Inc.

Mursid, R., Saragih A. H., Simbolon, N. (2017). Pengaruh Media Pembelajaran Berbasis TIK Dan Sikap Inovasi Terhadap Hasil Belajar Perencanaan Pembelajaran Dengan HOTS. *Jurnal Teknologi Informasi dan Komunikasi dalam Pendidikan*, 4 (2), 179-187. https://jurnal.unimed.ac.id/2012/index.php/teknologi/article/view/8765

Mursid, R., Sitompul, H., Saragih, A. H. (2019). The Effect of Learning Model Based on Higher Order Thinking Skills (HOTS) with Multimedia and Creative Thinking Abilities on Learning Outcomes of Vocational Learning Design. Advances in Social Science, *Education and Humanities Research*, 384 (1), 759-763. https://www.atlantis-press.com/proceedings/aisteel-19/125931943

Neuman, W. L. (2003). Social Research Methods: Qualitative and Quantitative Approach. Boston: Pearson Education Inc.

Ratnawulan, E., & Rusdiana. (2015). Evaluasi Pembelajaran DenganPendekatan Kurikulum 2013. Bandung: PT Pustaka Setia.

Redecker, C., & Johannessen, Ø. (2013). Changing Assessment: Towards a New Assessment Paradigm Using ICT. *European Journal of Education*, 48 (1), 79-96. https://www.researchgate.net/publication/256460822_Changing_Assessment_-_Towards_a_New_Assessment_Paradigm_Using_ICT

Sajadi, M. (2013). Formative Evaluation as Efficient Strategy for LearningMathematics. *Mathematics Education Trends and Research Journal*, 1(1), 13. https://www.researchgate.net/publication/274132768_Formative_Evaluation_as_Efficient_Strategy_for_Learning_Mathematics

Setemen, K. (2010). Pengembangan evaluasi pembelajaran online. *Jurnal Pendidikan Dan Pengajaran*, *43*(3), 207–214.

Sriadhi (2018). Instrumen Penilaian Multimedia Pembelajaran. Universitas NegeriMedan.

Sugiyono (2019). Metode Penelitian Kuantitatif, Kualitatif, dan R&D. Bandung: Alphabet.

Thoha, M. (2003). Perilaku Organisasi Konsep Dasar Dan Aplikasi. Jakarta: Rajawali.

Thorne, A., & Thomas, G. (2009). How to increase higher level thinking. CenterFor Development and Learning. https://www.semanticscholar.org/paper/How-to-Increase-Higher-Order-Thinking-Thomas-Thorne/e1ac293c0dd2aeb6f1d61ed17fde3547cd483bcc

Wagstaff, J. M. (1999). Teaching Reading and Writing With WordWall. U.S.A : Scholastic Inc. https://books.google.com/books/about/Teaching_Reading_and_Writing_with_Word_W.html?id=u3M5Qjoefx4C

Walker, D. M. (2012). Classroom Assessment Techniques: An Assessment andStudent Evaluation Method. *Scientific Research*. Vol.3, 903-907. https://www.scirp.org/journal/paperinformation.aspx?paperid=23555

Widodo, & Kadarwatii, S. (2013). Higher Order Thinking Berbasis Pemecahan Masalah Untuk Meningkatkan Hasil Belajar Berorientasi Pembentukan Karakter Siswa, *Jurnal Cakrawala Pendidikan*, 5 (1), 161-171. https://journal.uny.ac.id/index.php/cp/article/view/1269

Wirawan. (2011). Evaluasi Teori Model Standar Aplikasi dan Profesi, Contoh Aplikasi Evaluasi Program: Pengembangan Sumber Daya Manusia Program Nasional Pemberdayaan Masyarakat (PNPM) Mandiri Pedesaan, Kurikulum. Perpustakaan, dan Buku Tes. Jakarta: Raja Grafindo Persad.

Yamasari, Y. (2010). Pengembangan Media Pembelajaran Matematika Berbasis ICT yang Berkualitas. *Seminar Nasional Pascasarjana X-ITS*, Surabaya. https://salamsemangat.files.wordpress.com/2011/05/pengembangan-matematika-berbasis-tik.pdf