



EVALUATING ONLINE AUTHENTIC READING ASSESSMENT

¹Hendro Ardiansyah*, ²Izazol Idris, ³Muh. Khairul Wajedi Imami

^{1,2,3}Universiti Pendidikan Sultan Idris (Educational Measurement and Evaluation, Faculty of Human Development), Malaysia

ABSTRACT

The aim of the current research was to evaluate the online authentic reading assessment at the University of Muhammadiyah Malang, Indonesia. A total of 170 students who were participants in the assessment program participated in the current research. They were asked to rate and evaluate the assessment program based on five criteria: (1) planning; (2) teaching and learning; (3) student involvement; (4) coursework; and (5) course description aspects. This research employed quantitative approach. Descriptive analysis in Statistical Product and Service Solutions (SPSS) version 27 was used to analyse the data. The data were collected through an instrument adopted from the evaluation report for the instructors of the UMM. The instruments were valid and reliable to be used based on their content, face validity, and Cronbach alpha results. The results of the current research indicated that participants rated the program in a good category in all aspects of the criteria given, namely in the aspects of planning, teaching and learning, student involvement, coursework, and course description. Therefore, the online authentic reading assessment at the UMM was conducted sufficiently. The implications of the study suggest that future research should explore the integration of performance-based assessment with other approaches, such as competency-based assessment, situation-based assessment, or problem-based assessment.

Keywords: *Authentic assessment, reading, program evaluation.*

ABSTRAK

Tujuan penelitian ini yaitu mengevaluasi penilaian autentik pada subjek *reading* secara daring di Universitas Muhammadiyah Malang. Sebanyak 170 mahasiswa menjadi partisipan dalam program asesmen penelitian ini. Para partisipan diminta untuk menilai dan mengevaluasi program asesmen berdasarkan lima kriteria, yaitu: (1) perencanaan; (2) pengajaran dan pembelajaran; (3) keterlibatan siswa; (4) tugas subjek; dan (5) penjelasan subjek. Penelitian ini menggunakan pendekatan kuantitatif. Analisis deskriptif SPSS 27 digunakan untuk menganalisis data. Data dihimpun melalui sebuah instrumen yang diadopsi dari *Evaluation Report* bagi intruktur di UMM. Instrumen ini dinilai valid dan reliabel digunakan berdasarkan validitas konten, validitas muka, dan hasil perhitungan *Cronbach alpha*. Hasil dari penelitian ini mengindikasikan bahwa para partisipan telah menilai program tersebut dengan baik dalam segala aspek kriteria, seperti perencanaan, pengajaran dan pembelajaran, keterlibatan siswa, tugas subjek, dan penjelasan subjek. Dengan demikian, penilaian autentik dalam talian pada subjek *reading* dilakukan secara cukup. Implikasi penelitian ini menyarankan kepada peneliti selanjutnya untuk mengeksplorasi integrasi dari penilaian berdasarkan performa dengan beberapa pendekatan lain seperti, penilaian berdasarkan kompetensi, penilaian berdasarkan situasi, atau penilaian berdasarkan masalah.

Kata Kunci: *Penilaian autentik, reading, evaluasi program*

E-ISSN: 2621-9158
P-ISSN: 2356-0401

*Correspondence:
endoardiansyah@gmail.com

Submitted: October 20, 2024
Approved: December 5, 2024
Published: December 5, 2024

Citation: Ardiansyah, H., Idris, I., Imami, M.K.W. (2024). Evaluating online authentic reading assessment. *Celtic: A Journal of Culture, English Language Teaching, Literature and Linguistics*, 11(2), 276-292. Doi: 10.22219/celtic.v11.i2.37798

INTRODUCTION

Assessment is currently seen as important as teaching and learning since it evaluates the students' progress and examines their accomplishments (Kinash et al., 2018). Assessment aims to see how well students' performance as stated in the intended learning outcomes (Biggs et al., 2011). Meanwhile, according to Permendikbud (2013), assessment refers to the process of collecting information by utilizing various methods and approaches, such as summative and formative assessment, in order to acquire evidence and judgment regarding education product. It was considered a salient aspect of education and was able to significantly affect the students achievement (Andersson & Palm, 2018). In this case, authentic assessment is already well known among higher education and could influence students learning quality (Sokhanvar et al., 2021). It was also stated to be able to help students encounter their life pragmatic issues that they might experience (Wiewiora & Kowalkiewicz, 2019). Therefore, the practice of authentic assessment is already familiar among higher institutions, including in Indonesia. It is because in Indonesia, teaching and learning as well as assessment require the teachers to relate the class material to the authentic issues and competencies needed in a real-life situation (Sutarto & Jaedun, 2018). Other than that, the students were encouraged to have interpersonal skills, attitudes, knowledge, and psychomotor abilities that cannot be assessed through traditional assessment alone (Zaim et al., 2020). Hence, the practice of authentic assessment in Indonesia is common. Authentic assessment creates a correspondence between what is assessed at the university and what the students need in the real world (Villarroel et al., 2020). Martika and Zaim (2021) stated that authentic assessment defines evaluating knowledge and skills of pupils in the frame as close as possible to the real life setting. It is a process-oriented assessment examining the student's competency and performance in relation to real-life situations (Pohan & Ikawati, 2022). Meaning that the assessment evaluates the subject's proficiency based on a real-life situation (Sutadji et al., 2021). Therefore, it can support the personal development of students in terms of confidence and autonomous practice in real-world situations (Raymond et al., 2013; Serrano et al., 2018).

Furthermore, with the effect of the pandemic and the shift of the teaching and learning process to an online platform, authentic assessments were also conducted online. According to (Ibrahim et al., 2023), in the current digital era, the use of online assessments, including authentic assessments, has become more and more popular. The pandemic forced the practice of education to shift to online mode, including the assessment practice (Salirawati, 2021). Therefore, the use of online assessment has become familiar among educational practitioners. It stated that it would give teachers more flexibility in assessing the students (Biggs et al., 2011). Which is a great concern among university teachers related to assessment utilizing digital technology, including in Indonesia (Ibarra-Sáiz et al., 2020). However, challenges are also faced by teachers in conducting authentic online assessments. As found by Wahyuni et al. (2021), the teachers perceive assessment practice conducted online as a quite thing to do in terms of data collection and implementation. Maria and Qamaria (2021) reported that several aspects of authentic assessment, namely observing, controlling, and assisting students, became challenges for teachers in conducting online authentic assessment.

According to (Saputra et al., 2019), the implementation of authentic assessment in Indonesia has had a lot of benefits; however, there are still many issues to be investigated. In this case, it is necessary to conduct an evaluation regarding the implementation of this assessment in order to gauge and reveal the important aspects of its implementation. It is important since this assessment is a contextual approach that is associated with a real and quite complex situation (Susani, 2018). Other than that, authentic assessment requires several aspects, namely teacher competency in carrying out the assessment process, educational institutions supporting the assessment implementation, and research assessment literacy that should be known by the stakeholders (Pohan & Ikawati, 2022). It was emphasized by (Sutadji et al., 2021), who explain the characteristics of authentic assessment. First, assessments must be realistic, performance-based, and cognitively complex to make them authentic. Second, students must defend their answers and products and learn to work together. Third, it must be formative. In regard to its complexity, the evaluation of authentic assessment is necessary in order to control and enhance the performance of the authentic assessment program.

The current research aims to evaluate the authentic assessment online. This is important since the evaluation in an authentic online assessment can give a brief overview of its implementation. Other than that, according to (Ashford-Rowe et al., 2014), authentic assessment has its own standards, such as being in the form of a performance or product (outcome), ensuring the transfer of knowledge, adopting metacognition as a component, and ensuring accuracy in assessment performance. Doing evaluation in this case to examine whether the program evaluation conducted follows or achieves the basic standard. It is because evaluation in education is the process of determining the value or characteristic of an object, subject, or phenomenon (Adom et al., 2020). Therefore, doing evaluation in this aspect is to understand the value of the authentic online assessment conducted. Furthermore, research related to online authentic assessment evaluation is scarce in Indonesia. Current research is only related to authentic assessment and has yet to be conducted in an online context. For example, Ningsih and Wahyumiani (2020) evaluate the implementation of authentic assessment in junior high school in Yogyakarta. Hanifah and Irambona (2019) evaluate and describe the authentic assessment in junior high school in Kulon Progo, Indonesia. Additionally, Wangid et al. (2017) evaluate the implementation of authentic assessment in elementary schools in Yogyakarta. Reflecting these previous studies, there are missing point regarding the scope of evaluating the implementation of online authentic reading assessment in the university level since previous studies more discussed the topic in junior high school level. None of these studies evaluated the implementation of the online authentic assessment. Therefore, the current research aims to evaluate the implementation of the online authentic reading assessment in the University of Muhammadiyah Malang, Indonesia. Furthermore, this research is expected to contribute to the university stakeholders to improve the implementation of authentic reading assessment during online classroom in the university level in Indonesia.

METHOD

Participants and study location

The current research aims to evaluate the online authentic reading assessment conducted at the University of Muhammadiyah Malang. A quantitative approach through descriptive analysis was conducted to evaluate the program. In this case, the participant's perspective was considered. The participants were asked to rate the program in several aspects, namely the program planning, the program teaching and learning aspect, the students involvement in the program aspect, the program coursework aspect, and the program course description aspect. A total of 170 students are involved in the current research. They were participants in the program and had finished it.

Instrument

A social survey refers to a technique for collecting statistics information regarding the attributes, attitudes or behaviours of a population by providing standard question items for all its members (Buckingham, A; Saunders, 2004). The data was collected through a survey after the students finished the program. The survey was adopted from the evaluation report for instructor in the University of Muhammadiyah Malang and consists of 28 items. 5 items in program planning, 6 items in teaching and learning, 6 items in student involvement in the program, 5 items in program coursework, and 6 items in the program course description. The questionnaire items were evaluated for validity and reliability through content and face validity and Cronbach alpha to assess reliability.

In the current research, three experts from the University Pendidikan Sultan Idris Malaysia evaluated the questionnaires in order to confirm their content validity. The result indicated that the majority of experts agreed with the constructs and items utilized in the questionnaire (80%–100%). However, the researcher removed 'item 3' since two of the three experts stated that the item was not essential. Finally, there were 27 of the total 28 questionnaire items to be used in getting data. The suggestion for changes was successfully conducted before the pilot study process.

In terms of reliability, the questionnaires were tested in a pilot study involving 30 students. In this case, the questionnaires were distributed online and checked for reliability through Cronbach alpha. The result of the pilot study can be seen in Table 1.

Table 1. Reliability Coefficient of Questionnaire

Dimension	Number of items	Cronbach alpha value
Planning	4	0.688
Teaching and learning	6	0.642
Students involvement	6	0.679
Coursework	5	0.715
Course description	6	0.646

Table 1 displays the results of the reliability test. The Cronbach alpha for the planning dimension was 0.688, the Cronbach alpha for the teaching and learning dimension was 0.642, the Cronbach alpha for the student involvement dimension was 0.679, the Cronbach alpha for the coursework dimension was 0.715, and the Cronbach alpha for the course description dimension was 0.646. According to

(Creswell & Creswell, 2018), an instrument is said to be reliable if the coefficient value is greater than 0.6, which means that the questionnaire can be performed. In the current analysis, the value of each dimension in the instrument was higher than 0.60. Therefore, it can be stated that the questionnaires are reliable and can be used in the actual research.

Analysis

The current research examines the students perspective on authentic online reading assessment at the University of Muhammadiyah Malang. Students perspectives were obtained through a quantitative approach. The data was analysed through Statistical Product and Service Solutions (SPSS) version 27. A feature in SPSS was utilized to conduct the descriptive analysis. In this case, the percentage, total number of participants, and frequency were reported.

The objective of the study was to identify and to confirm the reliability of the scale applied in the questionnaire as purposed by (Ary, Donald; Jacobs, Lucy Cheser; Sorensen, Chris; Razavieh, 2010). The students were asked to rate the program based on a five-point scale consisting of strongly disagree (SD), disagree (D), neutral (N), agree (A), and strongly agree (SA). The percentage of students who valued the options 'Strongly Agree' and 'Agree' indicated that students showed positive attitudes with the calculation of frequency and percentage for high agreement. Otherwise, the students' percentages who rated the options 'Strongly Disagree' and 'Disagree' indicated that students showed negative attitudes with a calculation of frequency and percentage for high disagreement. Lastly, the percentage of students who valued 'Neutral' is shown separately and applied to illustrate the above two extremes.

Ethical consideration

This research concerns the research ethic, which obtained research ethic permission from the institute of graduate study at the University Pendidikan Sultan Idris Malaysia. Other than that, this research obtained permission from the director of the Language Center at the University of Muhammadiyah Malang.

DISCUSSION

The current research aims to evaluate the implementation of online authentic reading assessments at the University of Muhammadiyah Malang. The program was evaluated based on the students, who join the program, perspective. Therefore, descriptive analysis was employed to analyse the data. Several aspects of the program were evaluated, namely the planning, teaching and learning, student involvement, coursework, and course description aspects. The results are as follows:

Planning aspect

In the planning aspect, there are five criteria concerned: whether the course outline was explained at the beginning of the learning, whether the learning material was listed in the course outline, whether the learning material was uploaded to the learning and management system, and whether the learning topic included the whole course content. The result of the planning aspect can be seen in Table 2.

Table 2. Planning

No	Questionnaire Items	Scales	Students' Responses	
			Frequency	Percentage (%)
1.	<i>Course outline of this course was explained at the beginning of teaching and learning sessions.</i>	Strongly Agree	51	30
		Agree	48	48.2
		Neutral	22	12.9
		Disagree	9	5.3
		Strongly Disagree	6	3.5
Total			170	
Mean			3.9	
2.	<i>The learning materials listed in the course outline.</i>	Strongly Agree	46	27.1
		Agree	73	42.9
		Neutral	27	15.9
		Disagree	15	8.8
		Strongly Disagree	9	5.3
Total			170	
Mean			3.7	
3.	<i>Learning materials are uploaded or provided link to Learning Management System (LMS).</i>	Strongly Agree	45	26.5
		Agree	74	43.5
		Neutral	41	24.1
		Disagree	10	5.9
		Strongly Disagree	0	0
Total			170	
Mean			3.9	
4.	<i>Learning topics include the whole course content.</i>	Strongly Agree	45	26.5
		Agree	61	35.9
		Neutral	48	28.2
		Disagree	14	8.2
		Strongly Disagree	2	1.2
Total			170	
Mean			3.7	
Total Mean			3.8	

Based on table 2, most of students (48.2% of students) agree that the course outline of the program was explained at the beginning of the teaching and learning session; the students (42.9%) agree that the learning materials were listed in the course outline; the students (43.5%) agree that the learning materials are uploaded or provided a link to the learning management system (LMS); and the students (35.9%) agree that the learning topics include the whole course content. Based on the criteria given, the average student agreed that the planning section of the reading course was managed accordingly.

Teaching and learning aspect

The second aspect evaluated was the teaching and learning aspect of the program. In order to evaluate, several criteria were given, namely whether the teaching and learning session applied a variety of teaching and learning approaches, whether the teaching and learning session required students to engage in critical and creative thinking, whether the teaching and learning session helped students understand the course content, whether the teaching and learning session helped students relate theories and applications in everyday life, whether flexibility in

delivering the teaching and learning session encouraged students to learn the course, and whether the students were encouraged to search for information from a variety of resources or not. The results of the descriptive analysis regarding the perspective of the program participants is reported in Table 3.

Table 3. Teaching and Learning

No	Questionnaire Items	Scales	Students' Responses	
			Frequency	Percentage (%)
1.	<i>Teaching and learning (TnL) sessions applied a variety of TnL approaches.</i>	Strongly Agree	45	26.5
		Agree	60	35.3
		Neutral	49	28.8
		Disagree	13	7.6
		Strongly Disagree	3	1.8
		Total		170
	Mean		3.7	
2.	<i>Teaching and learning sessions require me to involve in critical and creative thinking.</i>	Strongly Agree	38	22.4
		Agree	64	37.6
		Neutral	46	27.1
		Disagree	19	11.2
		Strongly Disagree	3	1.8
		Total		170
	Mean		3.6	
3.	<i>Teaching and learning sessions help me to understand my course content.</i>	Strongly Agree	50	29.4
		Agree	60	35.3
		Neutral	38	22.4
		Disagree	20	11.8
		Strongly Disagree	2	1.2
		Total		170
	Mean		3.8	
4.	<i>Teaching and learning sessions help me to relate theories and their applications in everyday life.</i>	Strongly Agree	46	27.1
		Agree	72	42.4
		Neutral	37	21.8
		Disagree	13	7.6
		Strongly Disagree	2	1.2
		Total		170
	Mean		3.8	
5.	<i>Flexibility in delivering Teaching and learning sessions encourage me to learn the course.</i>	Strongly Agree	40	23.5
		Agree	66	38.8
		Neutral	42	24.7
		Disagree	19	11.2
		Strongly Disagree	3	1.8
		Total		170
	Mean		3.7	
6.	<i>I was encouraged to search information from a variety of resources.</i>	Strongly Agree	36	21.2
		Agree	65	38.2
		Neutral	40	23.5
		Disagree	26	15.3
		Strongly Disagree	3	1.8
		Total		170
	Mean		3.6	
Total Mean			3.7	

Based on table 3, most of students (35.3%) were agree that the teaching and learning session in the program applied variety of teaching and learning approach, the students (37.6%) were agree that the teaching and learning session in the program require the students to involve the critical and creative thinking, the students (35.3%) were agree that the teaching and learning session help the students to understand the course content, the students (42.4%) were agree that the teaching and learning session help students to relate theories and the application in everyday life, the students (38.8%) were agree that the flexibility in delivering teaching and learning session in the program encourage students to learning the course, and the students (38.2%) were agree that the students was encouraged to search information from variety of resources. From the six criteria given for the teaching and learning in the program, most students agreed. Therefore, it can be stated that the teaching and learning implemented in the program were good enough.

Student's involvement in the program

The third aspect evaluated was related to the involvement of students in the program. In this case, several criteria were given, namely: whether the students were given the opportunity to ask about the learning in the inline classroom; whether the instructor asked questions to encourage the students involvement in learning; whether the students were given the opportunity to discuss with their friends; whether the activities conducted in the teaching and learning session encouraged the students participation; whether the students were given the opportunity to give an opinion during the teaching and learning session; and whether the students were given sufficient time to practice self-learning through the learning process. The result of the students' perspective regarding their involvement in the program based on the criteria given can be seen in Table 4.

Table 4. Students' Involvement

No	Questionnaire Items	Scales	Students' Responses	
			Frequency	Percentage (%)
1.	<i>I was given the opportunity to ask about learning through online classroom.</i>	Strongly Agree	41	24.1
		Agree	65	38.2
		Neutral	45	26.5
		Disagree	15	8.8
		Strongly Disagree	4	2.4
Total			170	
Mean			3.7	
2.	<i>Instructor asked questions to encourage my involvement in learning.</i>	Strongly Agree	40	23.5
		Agree	72	42.4
		Neutral	38	22.4
		Disagree	18	10.6
		Strongly Disagree	2	1.2
Total			170	
Mean			3.7	
3.	<i>I was given the opportunity to discuss with my friends.</i>	Strongly Agree	46	27.1
		Agree	68	40
		Neutral	40	23.5
		Disagree	13	7.6

	Strongly Disagree	3	1.8
	Total	170	
	Mean	3.8	
4.	<i>The activities conducted in Teaching and learning sessions encourage my participation.</i>		
	Strongly Agree	42	24.7
	Agree	69	40.6
	Neutral	40	23.5
	Disagree	15	8.8
	Strongly Disagree	4	2.4
	Total	170	
	Mean	3.7	
5.	<i>I was given the opportunity to give opinion during Teaching and learning sessions for this semester.</i>		
	Strongly Agree	36	21.2
	Agree	71	41.8
	Neutral	51	30
	Disagree	12	7.1
	Strongly Disagree	0	0
	Total	170	
	Mean	3.7	
6.	<i>I was given sufficient time to practice self-learning through the learning process.</i>		
	Strongly Agree	38	22.4
	Agree	66	38.8
	Neutral	44	25.9
	Disagree	20	11.8
	Strongly Disagree	2	1.2
	Total	170	
	Mean	3.6	
Total Mean		3.7	

Based on table 4, most of students (38.2%) were agree that they were given opportunity to ask about the learning through online classroom, the students (42.4%) were agree that the instructor asked question to encourage the students involvement in the learning, the students (40%) were agree that the they were given opportunity to discuss with friends, the students (40.6%) were agree that the activities conducted in the teaching and learning session encourage students participation, the students (41.8%) were agree that they were given opportunity to give opinion during the teaching and learning session, the students (38.8%) were agree that they were given sufficient time to practice self-learning through the learning process. From six criteria evaluating the students' involvement in the program, most students agreed. Therefore, it can be stated that the program involved students in its activities.

Program coursework aspect

The fourth aspect evaluated was the course aspect of the program. In this case, five criteria were given: whether the coursework given is related to the course content or not; whether the coursework given is suitable for students learning time or not; whether the coursework given helps students relate the course content to the real world or not; whether the coursework conducted is guided by the instructor periodically or not; and whether the coursework is evaluated based on rubrics as informed earlier or not. The student's perspective on the program coursework based on the criteria given can be seen in Table 5.

Table 5. Program Coursework

No	Questionnaire Items	Scales	Students' Responses	
			Frequency	Percentage (%)
1.	<i>Coursework given is related to the course content.</i>	Strongly Agree	53	31.2
		Agree	64	37.6
		Neutral	31	18.2
		Disagree	21	12.4
		Strongly Disagree	1	0.6
		Total		170
	Mean		3.8	
2.	<i>Coursework given is suitable to Student Learning Time (SLT).</i>	Strongly Agree	35	20.6
		Agree	78	45.9
		Neutral	41	24.1
		Disagree	10	5.9
		Strongly Disagree	6	3.5
		Total		170
	Mean		3.7	
3.	<i>Coursework given helps me to relate the course content to the real world.</i>	Strongly Agree	34	20
		Agree	82	48.2
		Neutral	42	24.7
		Disagree	9	5.3
		Strongly Disagree	3	1.8
		Total		170
	Mean		3.7	
4.	<i>Coursework conducted is guided by the instructor periodically.</i>	Strongly Agree	35	20.6
		Agree	76	44.7
		Neutral	41	24.1
		Disagree	12	7.1
		Strongly Disagree	6	3.5
		Total		170
	Mean		3.7	
5.	<i>Coursework is evaluated based on rubrics as informed earlier.</i>	Strongly Agree	30	17.6
		Agree	87	51.2
		Neutral	37	21.8
		Disagree	15	8.8
		Strongly Disagree	1	0.6
		Total		170
	Mean		3.7	
Total Mean			3.7	

Based on table 5, most of the students (37.6%) agreed that the coursework given is related to the course content; the students (45.9%) agreed that the coursework given is suitable for their learning time; the students (48.2%) agreed that the coursework given helps students relate the coursework content to the real world; the students (44.7%) agreed that the coursework conducted is guided by the instructor periodically; and the students (51.2%) agreed that the coursework is evaluated based on rubrics as informed. From the five criteria for evaluating the coursework in the program, most students agree. Therefore, it can be stated that the coursework in the program was suitable for the students.

Course description aspect

The fifth aspect evaluated was the course description. In this case, there are several criteria given, namely whether the course content depicts the body of knowledge or not, whether the students learning time corresponds to the credit hours or not, whether the course content is relevant to their program or not, whether the course is important to the program or not, whether the course offered in accordance with the program or not, and whether the students were satisfied with the course or not. The student's perspective on the program course based on the criteria given can be seen in Table 6.

Table 6. Course Description

No	Questionnaire Items	Scales	Students' Responses	
			Frequency	Percentage (%)
1.	<i>Course content depicts the body of knowledge.</i>	Strongly Agree	37	21.8
		Agree	69	40.6
		Neutral	44	25.9
		Disagree	15	8.8
		Strongly Disagree	5	2.9
		Total		170
	Mean		3.6	
2.	<i>The student learning time corresponds to the credit hours.</i>	Strongly Agree	41	24.1
		Agree	67	39.4
		Neutral	50	29.4
		Disagree	11	6.5
		Strongly Disagree	1	0.6
		Total		170
	Mean		3.8	
3.	<i>Course content is relevant to my programme.</i>	Strongly Agree	44	25.9
		Agree	55	32.4
		Neutral	51	30
		Disagree	19	11.2
		Strongly Disagree	1	0.6
		Total		170
	Mean		3.7	
4.	<i>This course is very important for my programme.</i>	Strongly Agree	49	28.8
		Agree	65	38.2
		Neutral	43	25.3
		Disagree	12	7.1
		Strongly Disagree	1	0.6
		Total		170
	Mean		3.8	
5.	<i>This course is offered according to programme structure.</i>	Strongly Agree	47	27.6
		Agree	71	41.8
		Neutral	43	25.3
		Disagree	7	4.1
		Strongly Disagree	2	1.2
		Total		170
	Mean		3.9	
6.	<i>Overall, I am satisfied with this course.</i>	Strongly Agree	56	32.9
		Agree	75	44.1
		Neutral	25	14.7
		Disagree	14	8.2
		Strongly Disagree	0	0
		Total		170
	Mean		3.9	

Total	170
Mean	4
Total Mean	3.8

Based on table 6, most of the students (40.6%) agreed that the course content depicted the body of knowledge, the students (39.4%) agreed that the students learning time corresponded to the credit hours, the students (32.4%) agreed that the course content was relevant to the students program, the students (38.2%) agreed that the course was important for the program, the students (41.8%) agreed that the course was offered in accordance with the program structure, and the students (44.1%) agreed that they were satisfied with the course. From the five criteria for evaluating the course in the program, most students agree. Therefore, it can be stated that the course in the program was sufficient.

DISCUSSION

The current research aims to evaluate the implementation of online authentic reading assessments at the University of Muhammadiyah Malang, Indonesia. The participants of the authentic online assessment were asked to rate the program based on five criteria: the planning aspect, the teaching and learning aspect, the student involvement aspect, the coursework aspect, and the course description aspect. The results indicated that the majority of participants rated the implementation of online authentic reading assessments as good. In the planning aspect, the program was stated and outlined at the beginning of the program so the participants could have an overview of its implementation. The complete materials used throughout the program were listed and uploaded in the learning management system, making it easier for participants to access and download the materials. In the aspect of teaching and learning, the program stated to apply a variety of teaching and learning approaches that trigger students creativity and critical thinking. The teaching and learning were also implemented with flexibility, allowing participants to search for information, which can help students relate theories learned to their everyday lives.

In the aspect of student involvement in the program, the program stated to give participants much space to interact with the instructor by asking questions through the online classroom and discussing with classmates. The instructor also tries to encourage participant involvement by giving sufficient time for participants to express their opinions and practicing self-learning. That encourages the engagement of the participants throughout the program implementation. In terms of the coursework aspect, the program stated that it would be able to give an overall overview regarding the course content. The coursework also considers students conditions, such as setting aside sufficient time for students, which is set up by the instructor periodically. Therefore, following the coursework given in the program, students were said to be able to relate the content in the program to their real-life situation. In terms of course description, the program arranged the description in a way that can help participants understand the program flow easily; it also arranged it based on the program's aim and purpose. Therefore, the participants agree that the program is relevant, corresponds with the credit hours, and the course offered is important for the program.

The current research proves the sufficiency of an online authentic reading assessment. This is significant since much of the literature regarding authentic assessment was disarticulated with little or no references to the significance of philosophical discussion of the assessment process (McArthur, 2023). The result of the current research can depict the literature on authentic assessment with empirical data gained through the program participant's overview. Other than that, the participants perspective on the current research proves the contribution of the current research to the body of literature regarding authentic assessment. It is because Ajjawi et al. (2020) reported that less is known about the participants perspectives on the authentic assessment program. Ensuring the program alignment and authenticity based on the staff perspective is familiar, but evaluating the authentic assessment program based on the participant perspective is scarce. So, the current research helps to evaluate the main purpose of assessment evaluation, such as whether the student or the program participants perceive that there is alignment between the program implementation and what actually should be assessed in the authentic assessment program (Ajjawi et al., 2020). It is important since the assessment program should exactly align with the material already taught during the teaching and learning process (Hopwood et al., 2020).

The significance of the online authentic reading assessment conducted at the University of Muhammadiyah Malang proved through this research can give an overview of other similar programs. It can be an example in terms of the implementation and the evaluation criteria used. Other authentic assessments or online authentic assessments can adopt the program principle in terms of its implementation and the assessment indicators. It is because, according to Zaim et al. (2020), the average teacher or educational practitioner in Indonesia heavily relies on and is accustomed to the traditional assessment method, namely pencil and paper assessment. It was emphasized by Adnan et al. (2019), who found that even Indonesian teachers have sufficient knowledge of authentic assessment but do not have adequate competency in conducting authentic assessment. Research conducted in West Java, Indonesia, by Saputra et al. (2019) reported that the teachers faced the usual implementation difficulties in authentic assessment, which included concerns relating to the students, time and effort requirements, validity and reliability problems, resource management, the transformation of evidence, and subjectivity. Pardimin (2013) reported that Indonesian teachers still have difficulties developing the authentic assessment instrument. Lastly, Ekoati et al., (2022) also stated that although the teachers have long years of service, it does not mean the teachers understand and promote the authentic assessment appropriately. Therefore, the current research is significant in order to provide empirical examples related to authentic assessment implementation and evaluation.

CONCLUSION

The aim of the current research was to evaluate the online authentic reading assessment at the University of Muhammadiyah Malang, Indonesia. The current research considers the program participants to evaluate the program. Therefore, the program participants were asked to rate and evaluate the assessment program based on five criteria: the planning aspect, the teaching and learning aspect, the student involvement aspect, the coursework aspect, and the course description

aspect. Descriptive analysis in Statistical Product and Service Solutions (SPSS) version 27 was used to analyse the data. The result of the current research indicated that the participants rated the program in a good category in all aspects of the criteria given, namely in the aspects of planning, teaching and learning, student involvement, coursework, and course description. Therefore, the online authentic reading assessment at the University of Muhammadiyah Malang was conducted sufficiently.

Aside from its limitation of conducting the research in the Language Center of University of Muhammadiyah Malang, the researcher considered that this research has benefits for university lecturers as learning prototype to improve the quality of online authentic assessment for reading course. Based on the results of the study, the future researchers are expected to conduct the study about online authentic assessment with various English courses, large sample, and different object of study in Indonesia. Furthermore, the researchers are suggested to perform a study regarding the effectiveness of authentic assessment in reading course during online classroom. The findings may be compared to this research.

REFERENCES

- Adnan, Suwandi, S., Nurkamto, J., & Setiawan, B. (2019). Teacher competence in authentic and integrative assessment in Indonesian language learning. *International Journal of Instruction*, 12(1), 701–716.
<https://doi.org/10.29333/iji.2019.12145a>
- Adom, D., Mensah, J. A., & Dake, D. A. (2020). Test, measurement, and evaluation: Understanding and use of the concepts in education. *International Journal of Evaluation and Research in Education*, 9(1), 109–119.
<https://doi.org/10.11591/ijere.v9i1.20457>
- Ajjawi, R., Tai, J., Huu Nghia, T. Le, Boud, D., Johnson, L., & Patrick, C. J. (2020). Aligning assessment with the needs of work-integrated learning: the challenges of authentic assessment in a complex context. *Assessment and Evaluation in Higher Education*, 45(2), 304–316.
<https://doi.org/10.1080/02602938.2019.1639613>
- Andersson, C., & Palm, T. (2018). Reasons for teachers' successful development of a formative assessment practice through professional development—a motivation perspective. *Assessment in Education: Principles, Policy and Practice*, 25(6), 1–22. <https://doi.org/10.1080/0969594X.2018.1430685>
- Ary, Donald; Jacobs, Lucy Cheser; Sorensen, Chris; Razavieh, A. (2010). *Introduction to Research in Education*. 6.
- Ashford-Rowe, K., Herrington, J., & Brown, C. (2014). Establishing the critical elements that determine authentic assessment. *Assessment and Evaluation in Higher Education*, 39(2), 205–222.
<https://doi.org/10.1080/02602938.2013.819566>
- Biggs, John, Tang, & Catherine. (2011). *Teaching For Quality Learning At University* (Vol. 2011).
http://books.google.se/books/about/Teaching_for_Quality_Learning_at_Univers.html?id=XhjRBrDAESkC&pgis=1
- Buckingham, A; Saunders, P. (2004). Survey Methods Workbook. *Polity Press*, 1, 1–27.

- Creswell, W. J., & Creswell, J. D. (2018). Research Design: Qualitative, Quantitative and Mixed Methods Approaches. In *Journal of Chemical Information and Modeling* (Vol. 53, Issue 9). file:///C:/Users/Harrison/Downloads/John W. Creswell & J. David Creswell - Research Design_ Qualitative, Quantitative, and Mixed Methods Approaches (2018).pdf%0Afile:///C:/Users/Harrison/AppData/Local/Mendeley Ltd./Mendeley Desktop/Downloaded/Creswell, Cr
- ekoati, E. S., Rakhman, F., Zulaeha, I., & Subyantoro. (2022). The Necessity of the Indonesian Language Authentic Assessment Model Development Based on the Teacher Interview Results. *Proceedings of the 6th International Conference on Science, Education and Technology (ISET 2020)*, 574(Iset 2020), 94–98. <https://doi.org/10.2991/assehr.k.211125.020>
- Hanifah, M., & Irambona, A. (2019). Authentic assessment: Evaluation and its application in science learning. *Psychology, Evaluation, and Technology in Educational Research*, 1(2), 81. <https://doi.org/10.33292/petier.v1i2.4>
- Hopwood, C. J., Harrison, A. L., Amole, M., Girard, J. M., Wright, A. G. C., Thomas, K. M., Sadler, P., Ansell, E. B., Chaplin, T. M., Morey, L. C., Crowley, M. J., Emily Durbin, C., & Kashy, D. A. (2020). Properties of the Continuous Assessment of Interpersonal Dynamics Across Sex, Level of Familiarity, and Interpersonal Conflict. *Assessment*, 27(1), 40–56. <https://doi.org/10.1177/1073191118798916>
- Ibarra-Sáiz, M. S., Rodríguez-Gómez, G., Boud, D., Rotsaert, T., Brown, S., Salinas-Salazar, M. L., & Rodríguez-Gómez, H. M. (2020). The future of assessment in Higher Education. *RELIEVE - Revista Electronica de Investigacion y Evaluacion Educativa*, 26(1), 1–6. <https://doi.org/10.7203/relieve.26.1.17323>
- Ibrahim, M. M., Malik, M., & Avianti, R. A. (2023). Lecturers' perceptions of authentic assessment in times of COVID-19 pandemic: a case of Indonesian universities. *Journal of Applied Research in Higher Education*, 15(5), 1514–1529. <https://doi.org/10.1108/JARHE-02-2022-0041>
- Kinash, S., McGillivray, L., & Crane, L. (2018). Do university students, alumni, educators and employers link assessment and graduate employability? *Higher Education Research and Development*, 37(2), 301–315. <https://doi.org/10.1080/07294360.2017.1370439>
- Maria Ulfa M, S., & Suci Qamaria, R. (2021). The Death of Authentic Language Learning Assessment in Indonesia During The Global Pandemic. *PTK: Jurnal Tindakan Kelas*, 1(2), 81–93. <https://doi.org/10.53624/ptk.v1i2.21>
- Martika, S. E., & Zaim, M. (2021). The Implementation of Authentic Assessment on Reading Comprehension of the Tenth Grade of SMAN 8 Padang. *Journal of English Language Teaching*, 10(10), 129–137. <https://doi.org/10.24036/jelt.v10i1.111679>
- McArthur, J. (2023). Rethinking authentic assessment: work, well-being, and society. *Higher Education*, 85(1), 85–101. <https://doi.org/10.1007/s10734-022-00822-y>
- Ningsih, R. S., & Wahyumiani, N. (2020). Evaluating the implementation of authentic assessments in junior high school English lesson. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 24(1). <https://doi.org/10.21831/pep.v24i1.28037>
- Pardimin. (2013). Analisis of the Indonesia Mathematics Teachers Ability in

- Applying Authentic Assessment. *Integration of Climate Protection and Cultural Heritage: Aspects in Policy and Development Plans. Free and Hanseatic City of Hamburg*, 26(4), 1–37.
- Permendikbud. (2013). Peraturan Menteri Pendidikan dan Kebudayaan Indonesia Nomor 65. *Kementerian Pendidikan Dan Kebudayaan*, 53(9), 1689–1699.
- Pohan, J. E., & Ikawati, E. (2022). Authentic assessment of communication skills high school student in Indonesia. *Journal of Positive School Psychology*, 6(2), 596–605. <http://journalppw.com>
- Raymond, J. E., Homer, C. S. E., Smith, R., & Gray, J. E. (2013). Learning through authentic assessment: An evaluation of a new development in the undergraduate midwifery curriculum. *Nurse Education in Practice*, 13(5), 471–476. <https://doi.org/10.1016/j.nepr.2012.10.006>
- Salirawati, D. (2021). Authentic Assessment in The Pandemic Period. *Journal of The Indonesian Society of Integrated Chemistry*, 13(1), 21–31. <https://doi.org/10.22437/jisic.v13i1.11716>
- Saputra, E. R., Hamied, F. A., & Suherdi, D. (2019). *A Case of Authentic Assessment in Indonesian Secondary EFL Classroom Context: Teachers' struggle*. 254(Conaplin 2018), 198–201. <https://doi.org/10.2991/conaplin-18.2019.40>
- Serrano, M. M., O'Brien, M., Roberts, K., & Whyte, D. (2018). Critical Pedagogy and assessment in higher education: The ideal of 'authenticity' in learning. *Active Learning in Higher Education*, 19(1), 9–21. <https://doi.org/10.1177/1469787417723244>
- Sokhanvar, Z., Salehi, K., & Sokhanvar, F. (2021). Advantages of authentic assessment for improving the learning experience and employability skills of higher education students: A systematic literature review. *Studies in Educational Evaluation*, 70(March), 101030. <https://doi.org/10.1016/j.stueduc.2021.101030>
- Susani, R. G. (2018). the Implementation of Authentic Assessment in Extensive Reading. *International Journal of Education*, 11(1), 87. <https://doi.org/10.17509/ije.v11i1.9113>
- Sutadji, E., Susilo, H., Wibawa, A. P., Jabari, N. A. M., & Rohmad, S. N. (2021). Adaptation strategy of authentic assessment in online learning during the covid-19 pandemic. *Journal of Physics: Conference Series*, 1810(1). <https://doi.org/10.1088/1742-6596/1810/1/012059>
- Sutarto, H. P., & Jaedun, M. P. D. (2018). Authentic assessment competence of building construction teachers in Indonesian vocational schools. *Journal of Technical Education and Training*, 10(1), 91–108. <https://doi.org/10.30880/jtet.2018.10.01.008>
- Villarroel, V., Boud, D., Bloxham, S., Bruna, D., & Bruna, C. (2020). Using principles of authentic assessment to redesign written examinations and tests. *Innovations in Education and Teaching International*, 57(1), 38–49. <https://doi.org/10.1080/14703297.2018.1564882>
- Wahyuni, L. G. E., Dewi, N. L. P. E. S., & Paramartha, A. A. G. (2021). Authentic assessment practice teachers' perceived knowledge. *Proceedings of the 2nd International Conference on Technology and Educational Science (ICTES 2020)*, 540(Ictes 2020), 316–323.
- Wangid, M. N., Mustadi, A., Senen, A., & Herianingtyas, N. L. R. (2017). The

- evaluation of authentic assessment implementation of Curriculum 2013 in Elementary School. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 21(1), 104–115.
<https://doi.org/10.21831/pep.v21i1.15779>
- Wiewiora, A., & Kowalkiewicz, A. (2019). The role of authentic assessment in developing authentic leadership identity and competencies. *Assessment and Evaluation in Higher Education*, 44(3), 415–430.
<https://doi.org/10.1080/02602938.2018.1516730>
- Zaim, M., Refnaldi, & Arsyad, S. (2020). Authentic assessment for speaking skills: Problem and solution for english secondary school teachers in Indonesia. *International Journal of Instruction*, 13(3), 587–604.
<https://doi.org/10.29333/iji.2020.13340a>