

THE EFFECT OF USING QUIZIZZ AND WORDWALL MEDIA ON COGNITIVE LEARNING OUTCOMES OF HEALTH MATERIALS

Naufal Izha Manggalih¹, Bertika Kusuma Prastiwi¹, Pandu Kresnapati¹

¹Universitas PGRI Semarang, Physical Education Health and Recreation, Semarang, Indonesia

*Coressponding Author. Email: noval11845@gmail.com

*Email Author: noval11845@gmail.com, bertikakusuma@gmail.com,
pandukresnapati@upgris.ac.id

Received: April 2026, Revised: May 2026, Accepted: June 2026

Abstract

Conventional instructional methods remain inadequate for delivering health education on sensitive topics such as Sexually Transmitted Diseases (STDs), particularly in vocational high school settings where epidemiological urgency coexists with deep sociocultural barriers that inhibit open classroom discussion. This study compared the effectiveness of Quizizz and Wordwall platforms on the cognitive learning outcomes of grade XI students at SMK Negeri 1 Kersana. A quasi-experimental approach with a *non-equivalent pretest-posttest group design* was applied, positioning both classes as active experimental groups: Experimental Group I used Quizizz ($n=36$) and Experimental Group II used Wordwall ($n = 33$), totaling 69 students selected through purposive sampling during the Ramadan academic period. The 25-item multiple-choice instrument was validated through expert judgment, producing an Aiken's V index of 0.91 and a Cohen's Kappa reliability coefficient of 0.603 ($p = 0.003$), both confirming strong instrument quality. Data were analyzed using normalized N-gain scores and the Mann-Whitney U non-parametric test. Results showed that the Quizizz group achieved a mean N-Gain of 0.7556 (75.56%), categorized as high effectiveness, while the Wordwall group obtained 0.4968 (49.68%), categorized as moderate. A statistically significant difference was confirmed ($p = 0.000$), with an effect size coefficient $r = 0.503$ (*large effect*), indicating that the performance gap between groups reflects a substantive instructional difference rather than statistical chance. The competitive architecture of Quizizz, particularly its leaderboard system and real-time corrective feedback, produced stronger cognitive gains than the template-based exploration model of Wordwall. Quizizz is recommended as the primary instructional tool for PJOK teachers delivering reproductive health content in vocational settings, especially under conditions that restrict physical activity.

Keywords: *Gamification; Health Education; sexually transmitted diseases; Quizizz; Wordwall*

INTRODUCTION

Physical Education (PE) is an important part of the elementary school curriculum because Physical Education, Sports, and Health subjects at the vocational high school level carry out responsibilities that go beyond physical coaching alone. The curriculum explicitly targets the formation of comprehensive health literacy as the foundation of students' long-term quality of life.

However, the reality in the field shows that conditions that are far from ideal for PJOK learning achievements, especially in materials with a high cognitive dimension, are still often not optimal. Wibowo et al. (2016) emphasized that a uniform instructional approach that relies on one-way transmission consistently results in superficial mastery of concepts. This pattern becomes even more problematic when dealing with the subject of sexually transmitted diseases (STDs), a topic that bears a double burden: high epidemiological urgency on the one hand and strong sociocultural barriers on the other.

The urgency of mastering STD material among adolescents is not just a curricular demand but a response to real health realities. The World Health Organization (2025) records that more than one million new cases of curable STDs occur every day globally, the majority of which are asymptomatic so that transmission takes place unknowingly. At the national level, the Ministry of Health of the Republic of Indonesia (2025) recorded 356,638 individuals identified as ODHIV out of an estimated total of 564,000, with 37% coming from key population groups. The data of Johnston et al. (2021) detail this condition further: HIV-positive diagnoses reach 30% in the group of young men who have sex with men (LSL) aged 15–24 years in urban Indonesia. This age range coincides directly with the age of vocational school students, making reproductive health education not just an option but a must.

The biggest challenge actually arises from the pedagogical side. The topic of sexuality in Indonesia often clashes with cultural norms and religious values, creating a closed classroom atmosphere full of psychological barriers. Sukmawati et al. (2023) found that the reproductive health literacy of Indonesian adolescents is still in the moderate category with significant gaps between genders, while Todesco et al. (2023) identified social stigma as a structural barrier that limits adolescents' access to trusted reproductive information. Initial observations at SMK Negeri 1 Kersana corroborate this pattern: students' average scores on reproductive health materials consistently do not reach the KKM, indicating that the prevailing lecture method has not been able to penetrate these affective barriers while building meaningful cognitive understanding.

This is where the real need lies: not just a medium to replace lectures, but an approach that is simultaneously able to reduce psychological barriers and increase cognitive engagement in one integrated mechanism. Gamification is here as an answer to this need. Ariani (2020) explained that gamification is a learning strategy that adopts elements of games into the context of non-game learning to encourage active involvement and motivation of students. The effectiveness of this approach has been empirically confirmed in the context of Indonesian education. Nurjannah et al. (2021) prove that the application of gamification significantly increases students' motivation to learn through challenge mechanisms, points, and reward systems. Winatha et al. (2020) reinforce



these findings by showing a significant difference in motivation and learning achievement between classes that implement game-based learning and classes that do not, with a significance value of $0.001 < 0.05$. The relevance of gamification for PMS materials in particular has been proven by Haruna et al. (2021) who through a quasi-experimental design showed that gamification-based interventions in adolescent sexual health education resulted in an average posttest score of 79.23 compared to 51.93 in the non-treatment group, proving that the mechanism of play not only makes learning more interesting, but actively breaks down the psychological barriers that prevent adolescents from processing information reproduction openly.

Two platforms that operationalize gamification principles with different architectures are Quizizz and Wordwall. Quizizz structures the learning experience through leaderboard-based competitions and instant answer feedback that transforms the evaluation process into a positive-pressure activity (Wang & Tahir, 2020). Wordwall offers content exploration through a variety of game templates that can be calibrated according to the characteristics of the material (Auliya et al., 2021). Both have the potential to create a more fluid and less intimidating learning space for sensitive topics like STDs. Nevertheless, these different mechanisms are likely to produce unequal cognitive impacts: Quizizz with its external competition has the potential to be more potent at engagement at the level of understanding (C2) and applying (C3), while Wordwall with its visual exploration may be more effective for reinforcement at the level of remembering (C1). This architectural difference makes the direct comparison between the two of them of high scientific value, but it has never been specifically studied in the context of reproductive health education in vocational schools.

Previous studies on both platforms, while quite abundant, left some gaps that have not yet been reached. Citra & Rosy (2020), Hidayati & Aslam (2021), and Saputra et al. (2024) proved the effectiveness of Quizizz, while Muntaza et al. (2022) and Minarta & Pamungkas (2022) confirmed the benefits of Wordwall, but the entire study evaluated each platform separately against the no-treatment condition. The only study that directly compared the two, namely Iklima et al. (2023), focused on productive writing skills and not cognitive understanding of health materials, so the findings cannot be generalized in the context of this study.

The novelty of this research rests on three things that have not been reached simultaneously by previous studies. First, this study is a direct comparative study between Quizizz and Wordwall that places the two as two active experimental groups, rather than comparing one of them to the passive group without treatment. Second, the domain studied is the cognitive learning outcomes of PMS materials in vocational schools, a combination that is specifically not yet available in the existing literature. Third, this research was carried out in a typical ecological context, namely the



month of Ramadan, where the limitation of physical activity naturally strengthens the relevance of theoretical digital learning media as the main instrument, not just a complement. The contribution of this research to the development of digital pedagogy is to produce evidence-based recommendations on which gamification platform is most effective for PJOK teachers to use in learning reproductive health materials in a vocational environment.

Based on the description above, this study aims to: (1) analyze the influence of the use of Quizizz and Wordwall media on students' cognitive learning outcomes on PMS materials; and (2) identify platforms that demonstrate higher levels of effectiveness in improving students' cognitive understanding. The working hypothesis (H_a) proposed is that there is a significant difference in influence between the use of Quizizz and Wordwall on the cognitive learning outcomes of grade XI students of SMK Negeri 1 Kersana on PMS material, with a hypothesis of zero (H_0), which states that the two media do not produce a significant difference in impact.

METHODOLOGY

This study uses a quantitative approach with a quasi-experimental design. Sugiyono (2019) suggested that the quasi-experimental approach was chosen when it was not possible for the researcher to exercise full control over external variables, as is common in formal education settings that use naturally formed classes. The specific design applied in this study is a *non-equivalent pretest-posttest group design*, which is a research pattern that allows comparison of achievement trajectories between two groups without prior random assignment.

One important thing that needs to be emphasized is the position of the two groups in this study. Although the design used was classified as a *non-equivalent control group design*, the two groups in this study both received different active treatments, so they were positioned as experimental groups rather than as experimental and control pairs in the conventional sense. Experimental Group I received learning using the Quizizz platform, while Experimental Group II received learning using the Wordwall platform. This assertion is methodologically important because the study was designed to compare the effectiveness of two treatments that were both active, not to measure the difference between treatment and no treatment. The research was carried out at SMK Negeri 1 Kersana, Brebes, in the even semester of the 2025/2026 school year, coinciding with the Ramadan fasting period which pedagogically became the basis for consideration for the selection of digital-based theoretical materials as a substitute for field activities, so that the physical condition of fasting students was maintained.

The research population includes all students of grade XI of SMK Negeri 1 Kersana which totals 549 students, spread across 16 study groups across 8 expertise programs, namely Fashion (2



classes, 69 students), Light Vehicle Engineering (4 classes, 140 students), Welding Engineering (2 classes, 68 students), Accounting (3 classes, 105 students), Visual Communication Design (3 classes, 105 students), and Culinary (2 classes, 62 students). The sampling technique used is *purposive sampling*. Arikunto (2018) explained that this technique is based on certain considerations and goals that are in accordance with the characteristics desired by the researcher. Two classes from the Fashion Skills program, namely XI Fashion A ($n = 36$) and XI Fashion B ($n = 33$), were selected as a sample with a total of 69 students on the basis of two main considerations. First, the proportion of female students who were dominant in both classes had direct theoretical relevance to reproductive health risks in STD materials, given that a number of sexually transmitted infections clinically had a more significant impact on women. Second, both classes are taught by the same PJOK teacher so that the potential for instructional bias due to differences in teaching styles can be suppressed to a minimum.

It needs to be explicitly acknowledged that the use of *purposive sampling* involving only one expertise program from a single school carries methodological consequences that cannot be ignored. This decision has the potential to cause selection bias because the characteristics of fashion planning students, both in terms of gender composition, socioeconomic background, and learning orientation, do not represent the overall population of vocational school students, which is much more diverse. Therefore, the generalization of the findings of this study should be done carefully and limited to contexts that demographically have similarities to the sample studied. Follow-up research with a wider range of samples, covering a wide range of different skill programs and schools, is needed to produce more transferable conclusions.

The independent variables in this study are the learning platforms set for each group, namely Quizizz for Experimental Group I and Wordwall for Experiment Group II, while the bound variable is the students' cognitive learning outcomes on PMS material measured through pretest and posttest instruments.

The evaluation instrument used consists of 25 multiple-choice questions that include three cognitive levels based on the revised Bloom taxonomy, namely remembering (C1), understanding (C2), and applying (C3), as classified by Anderson & Krathwohl (2001, in Gunawan & Palupi, 2016). The initial collection of question items amounted to 30 items and was then reduced to 25 items after going through an *expert judgment* process carried out by health education lecturers and PJOK teachers as experienced education practitioners in their fields.

The quality of the instrument is guaranteed through two complementary validation procedures. First, content *validity* is assessed through a structured expert panel study. Arikunto (2018) explained that the validity of the content is related to the extent to which the question items



that are prepared as a whole reflect the domain of material that students should master according to the learning objectives that have been set so that a test can be declared valid if it is able to measure what is intended to be measured accurately and representatively.

Table 1. Instrument Validation Results

Uji Instrumen	Nilai	Keterangan
Aiken's V (rata-rata)	0,91	Sangat Valid ($> 0,60$)
Cohen's kappa	0,603	<i>Substantial Agreement</i>
Sig. (p)	0,003	Signifikan ($p < 0,05$)

The quantitative results of instrument validation were obtained through two mutually reinforcing statistical indices. The content validity index of Aiken's V is calculated based on the assessment of the expert panel and results in an average value of 0.91. This value substantially exceeds the minimum required threshold of 0.60, so all 25 confirmed questions have a very strong content conformity with the applicable curriculum standards. The reliability of the instrument was evaluated using an *inter-rater reliability* approach, which was operationalized through the Cohen's Kappa coefficient. Two independent evaluators, namely PJOK teachers and health education lecturers, assessed the same question items using standardized assessment guidelines. The resulting Cohen's Kappa coefficient is 0.603 with a significance value of $p = 0.003$ ($p < 0.05$), which is included in the category of *substantial agreement*. Buka et al. (2025) explained that *inter-rater reliability* is the degree of consistency of assessment given by two or more assessors to the same object or participant using uniform criteria, and this is very crucial for subjective types of assessments. Arikunto (2018) emphasized that reliable instruments provide consistent results even when used at different times or by different assessors, a characteristic that greatly determines the credibility of research data. The combination of Aiken's V value of 0.91 and Cohen's Kappa coefficient of 0.603 provides quantitative assurance that the test instrument used is objective, free from the risk of personal bias, and has a very high feasibility to measure students' cognitive learning outcomes validly and reliably.

The data collection procedure is carried out in three stages of the meeting, as illustrated in Image 1. The first meeting was used for the implementation of a pretest to map the students' initial abilities before the intervention. At the second meeting, the two groups received substantially equivalent treatment but different platforms: the two classes first listened to educational videos and PowerPoint presentations, followed by formative quiz sessions using Quizizz with the "Shuffled Multiple Choice" feature (experimental class I) and the Wordwall template "Open the Box" (experimental class II), with 10 questions each. In both sessions, the appointment of students who



answered was carried out through digital randomization using the Wheel of Spin application to maintain even participation and reduce student awkwardness in discussing sensitive topics. The third meeting was used for the implementation of the posttest to measure the final achievement of students.

Data analysis was carried out in two stages. First, cognitive improvement was measured using a normalized N-Gain score that categorized the effectiveness of the intervention into three levels: low (< 0.30), moderate ($0.30-0.70$), and high (> 0.70) (Hake, 1998, in Safitri & Supeni Edie, 2020). Second, comparative hypothesis testing was performed using the Mann-Whitney U non-parametric test through SPSS software, considering that the Shapiro-Wilk normality test showed an abnormal data distribution in the Quizizz group (Sugiyono, 2019).

Given the comparative hypothesis testing using Mann-Whitney U non-parametric statistics, the practical significance of the media intervention was calculated by estimating the effect size of the r coefficient. The value of the coefficient r is obtained by converting the statistical value Z from the computerized output divided by the square root of the total sample (N). The criteria for interpreting the strength of the effect refer to the *Cohen* standard, which is an index of 0.10 for the small category, 0.30 for the medium category, and 0.50 for the large category.

RESULTS

Statistical Prerequisite Analysis

Before the hypothesis test is carried out, two prerequisite tests are taken to ensure the accuracy of the analysis techniques to be used, namely the normality test and the variance homogeneity test.

Table 2. Normality test results (Shapiro-Wilk)

Tests of Normality							
	Kelas	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Pretest	Quizizz	.144	36	.056	.927	36	.020
	Wordwall	.119	33	.200*	.959	33	.249
Posttest	Quizizz	.213	36	.000	.829	36	.000
	Wordwall	.138	33	.110	.962	33	.302

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

The normality test using Shapiro-Wilk was chosen because it was more sensitive in small to medium-sized samples compared to Kolmogorov-Smirnov (Ghozali, 2021). The results showed



that the Quizizz group's data were not normally distributed, either on the pretest ($p = 0.020$) and the postes ($p = 0.000$), while the Wordwall group met the normality assumption at both stages ($p = 0.249$ and $p = 0.302$). The abnormal distribution in the Quizizz group does not indicate an instrument failure but rather reflects the occurrence of a ceiling effect: the high enthusiasm of students for the Quizizz competitive mechanism encourages the majority of them to achieve scores close to the maximum score simultaneously, so that the distribution curve slopes to the right.

Table 3 Results of the Variance Homogeneity Test (Levene's Test)

Test of Homogeneity of Variances				
	Levene Statistic	df1	df2	Sig.
Pretest	2.310	1	67	.133
Posttest	4.113	1	67	.047

In the aspect of homogeneity of variance in table 3, Ghazali (2021) states that the Levene's Test is the most robust procedure for testing the similarity of variance between two or more groups, as it does not require a normal data distribution. The pretest data showed a significance value of 0.133 ($p > 0.05$), which confirmed that both classes departed from an equivalent condition of variance before the intervention was administered. This initial equivalence is an important foundation to ensure that the differences in achievement that appear at the end of the study are purely the effect of the differences in the media tested, not the result of the initial ability gap between classes. However, after the intervention, the homogeneity of variance was no longer met ($p = 0.047 < 0.05$), indicating that the two platforms had qualitatively different instructional impacts. Quizizz tends to concentrate students' grades in the high range, while Wordwall produces a more varied distribution. This condition methodologically strengthens the justification for using the Mann-Whitney U non-parametric test as a hypothesis test instrument in this study.

Learning Effectiveness Test (N-Gain)

Table 4. Learning Effectiveness Statistics (N-Gain Score) Quizizz

Descriptive Statistics					
	N	Minimum	Maximum	Mean	Std. Deviation
NGain_Score	36	.18	1.00	.7556	.20821
NGain_Persen	36	18.18	100.00	75.5599	20.82095
Valid N (listwise)	36				



Table 5. Learning Effectiveness Statistics (N-Gain Score) Wordwall

Descriptive Statistics					
	N	Minimum	Maximum	Mean	Std. Deviation
NGain_Score	33	-.29	1.00	.4968	.25734
NGain_Persen	33	-28.57	100.00	49.6754	25.73365
Valid N (listwise)	33				

Table 6. Percentage Increase

Indikator Statistik	Kelompok Quizizz	Kelompok Wordwall
Rata-rata (Mean) N-Gain	0,7556	0,4968
Persentase Efektivitas	75,56%	49,68%
Kategori N-Gain	Tinggi	Sedang
Tingkat Efektivitas	Efektif	Cukup Efektif

Cognitive improvement was measured using a normalized N-Gain score that classified the level of effectiveness of the intervention into three categories: low (< 0.30), moderate ($0.30-0.70$), and high (> 0.70) (Hake, 1998, in Safitri & Supeni Edie, 2020). The calculation of the N-Gain Score in the two groups (Tables 4, 5, and 6) produced quite contrasting differences. The Quizizz group achieved an average N-Gain of 0.7556 (75.56%) which was in the high category, while the Wordwall group obtained an N-Gain of 0.4968 (49.68%) in the medium category.

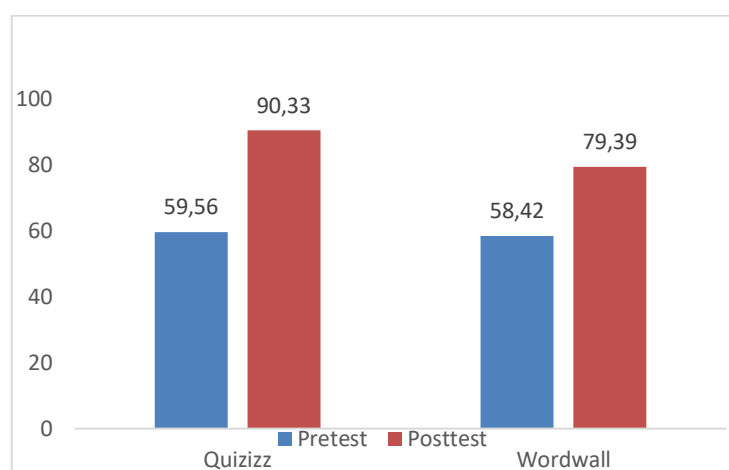


Image 1 Comparison of Average Pretest and Posttest Scores

This pattern of cognitive improvement can be observed more intuitively through Image 2 which maps the average trajectory of the score from the initial condition to the end of the

intervention. In the pretest stage, the two groups departed from almost close proximity: 59.56 for Quizizz and 58.42 for Wordwall, an initial balance that indicated the absence of an innate cognitive advantage between the two. Entering the posttest stage, the gap between the two widened significantly: the Quizizz group jumped to 90.33 while the Wordwall group only reached 79.39. The difference of 10.94 points in this final achievement confirms that the apparent difference is not just a random variation, but a concrete reflection of the difference in instructional thrust between the two platforms. While both have a positive impact on learning outcomes, the acceleration of understanding shown by each platform differs markedly. These findings were then statistically tested to prove whether the difference was significant or just a coincidence.

Hypothesis Test (Mann-Whitney U) and Effect Size

Comparative hypothesis testing was performed using the Mann-Whitney U test via SPSS, chosen because the assumption of normality was not met in the Quizizz group. Ghazali (2021) stated that this test is the most appropriate non-parametric alternative to compare two independent groups when the data distribution is abnormal because the test is rank-based so that it is not affected by the slope of the distribution.

Table 7. Hypothesis Test Results (Mann-Whitney U)

Test Statistics ^a	
	Value
Mann-Whitney U	249.000
Wilcoxon W	810.000
Z	-4.178
Asymp. Sig. (2-tailed)	.000
a. Grouping Variable: Class	

Based on the results of the statistical test in Table 7, an Asymp value was obtained. Sig. (2-tailed) is 0.000. Ghazali, (2021) explains that the decision to reject H_0 in the Mann-Whitney U test is taken when the significance value obtained is smaller than the set error level ($\alpha = 0.05$). Since the significance value of 0.000 is far below that threshold, H_0 is rejected and H_a is accepted. Thus, it can be concluded that there is a significant difference in the influence between the use of Quizizz and Wordwall on students' cognitive learning outcomes in PMS materials.

In addition to showing statistically significant differences, an in-depth analysis of the strength of the intervention was emphasized through the results of the effect size test. Based on the conversion index of the statistical value Z in the Mann-Whitney U test through the equation



function:

$$r = \frac{Z}{\sqrt{N}}$$

The final R-coefficient value of 0.503 was obtained. Referring to Cohen's standardization parameters, the acquisition of these numbers indicates that the use of interactive digital media has a contribution to the influence in the Large Category (*Large Effect*) on improving the cognitive understanding of students' PMS materials. These empirical findings confirm that the learning outcomes gap between experiment group I and experiment group II was purely driven by the effectiveness of different game mechanics' architecture differences on both platforms, rather than a statistical anomaly that occurred by chance in the field.

Table 8. Average Rating Comparison (Mean Rank)

		Ranks		
	Kelas	N	mean rank	Sum of Ranks
Nilai	Postest_Quizizz	36	44.58	1605.00
	Postest_Wordwall	33	24.55	810.00
	Total	69		

The analysis of the average rating in Table 8 reinforces this conclusion, where the Quizizz group recorded a mean rank of 44.58, substantially outperforming the Wordwall group which only obtained a mean rank of 24.55. This wide range of rankings is further confirmation that Quizizz's prowess is consistent and does not happen by chance.

DISCUSSION

The overall results of the study prove that gamification-based media can encourage meaningful changes in the quality of learning in materials that contain socioculturally sensitive content. The topic of STDs, which often causes psychological discomfort in adolescents (Sukmawati et al., 2023), has proven to be approachable more openly when packaged through competitive and fun digital rides.

Quizizz's advantage in producing a high-category N-Gain (0.7556) is rooted in three design mechanisms that work synergism. First, the leaderboard structure creates a positive competitive pressure that encourages student vigilance on an ongoing basis throughout the session. Second, the randomization of the question sequence inhibits the mechanical memorization strategy and forces each student to process the question independently from the beginning. Third, instant feedback after each response allows for real-time correction of understanding without relying on teacher



evaluation, a process that directly supports cognitive reinforcement at the level of understanding (C2) and applying (C3) based on the revised Bloom taxonomy (Anderson & Krathwohl, 2001, in Gunawan & Palupi, 2016)). This pattern is consistent with the findings of Citra & Rosy (2020), Saputra et al. (2024), and Hidayati & Aslam (2021) who associate immediate feedback mechanisms with more structured cognitive development.

A wordwall with a moderate N-Gain (0.4968) still makes a meaningful positive contribution. The Open the Box template combined with the Wheel of Spin mechanism presents an element of anticipation that maintains students' focus during learning. Muntaza et al. (2022) suggest that Wordwall is more accommodating for students with visual learning styles because the exploration of diverse interfaces helps to internalize material more intuitively, while Minarta & Pamungkas (2022) confirm the superiority of Wordwall over conventional methods. The difference in achievement between the two platforms reflects fundamental differences in cognitive orientation: Quizizz encourages active engagement through competition and immediate correction, while Wordwall relies more on exploration that doesn't always result in in-depth processing at the C2 and C3 levels.

However, the observed differences in achievement cannot be fully attributed to the intrinsic effectiveness of the platform alone. A number of external factors have the potential to moderate results immeasurably. First, different digital literacy skills between students can affect the smooth navigation of the platform independently of mastering the material, so that students who are more familiar with online quiz applications have the potential to respond to Quizizz more efficiently. Second, learning motivation that was not directly measured in this study may play an important role in moderation: the Quizizz competition mechanism may increase extrinsic motivation in competitively oriented students, but may instead trigger counterproductive anxiety in other students. Ariani (2020) reminds that gamification is basically more effective at encouraging short-term engagement and does not necessarily result in consistent deep internalization. Third, the variation in the quality of internet connections between students in one room has the potential to affect the smooth use of real-time-based Quizizz more than Wordwall which is relatively more tolerant of network latency. The absence of data on these variables limits the ability of the study to draw stronger causal conclusions.

The context of implementation during the month of Ramadan adds a relevant ecological dimension. The physical limitations due to fasting, which usually hinder the components of PJOK practice, have been successfully converted into opportunities to strengthen theoretical material. Digital gamification has been shown to maintain productive intellectual engagement under physiological stress, with game dynamics shifting students' focus from physical fatigue to cognitive



challenge. It should be noted that physiological conditions during fasting have the potential to weaken cognitive processing capacity in general, and this impact may not be uniform between students, so individual variations in physical endurance may contribute to the observed score variability.

The Wheel of Spin mechanism as a participant randomization tool also makes its own contribution. Random selection eliminates social asymmetry in participation and distributes the responsibility of involvement evenly across the class, while reducing the psychological burden on students when asked to discuss reproductive topics in front of peers. Haruna et al. (2021) prove that the success of adolescent sexual health education is not only determined by the depth of content, but also by the accuracy of participation management mechanisms, a principle that is strengthened and expanded by the findings of this study in the context of Indonesian vocational education..

CONCLUSION

This research resulted in two mutually reinforcing conclusions. First, there was a statistically significant difference in cognitive learning outcomes between Experimental Group I (Quizizz) and Experimental Group II (Wordwall) on PMS material, confirmed by the Mann-Whitney U test with a value of $p = 0.000$ ($p < 0.05$). Second, Quizizz has proven to be more effective instructionally than Wordwall, as reflected in the high category N-Gain of 0.7556 (75.56%) compared to N-Gain while Wordwall is 0.4968 (49.68%), strengthened by the value of effect size coefficient $r = 0.503$ which is included in the Large Effect category. This advantage comes from the synergy between leaderboard-based competitions, question randomization, and instant corrective feedback that strengthens cognitive engagement at the C2 and C3 levels.

Based on these findings, Quizizz is recommended as the main instructional instrument for PJOK teachers in learning reproductive health materials at vocational schools, especially in conditions that limit physical activity. Some limitations of the research need to be stated openly. The sample only included two classes from one specialty program (Fashion Dress) in one school, with a homogeneous gender composition, so it did not represent the diversity of the SMK student population more broadly. The use of purposive sampling chosen on the basis of theoretical and practical considerations limits the statistical generalization of these findings; The results of the study are only valid in contexts that demographically have similarities to the sample studied. In addition, moderating variables such as digital proficiency, intrinsic learning motivation, and technology access were not directly measured, so the relative contribution of each factor to the difference in achievement between the two platforms could not be properly isolated. Measurements using only short-term pretest-postes schemes also could not confirm whether the resulting cognitive



gains were long-lasting or were a temporary effect of involvement during the intervention.

Based on these limitations, further research is recommended to: (1) expand the scope of the sample by involving various expertise programs and more than one school to improve the generalizability of the findings; (2) integrate measurements of moderating variables such as digital proficiency and learning motivation as covariates in the analysis so that the contribution of each factor can be mapped more accurately; (3) apply a longitudinal design with a follow-up session several weeks after the intervention to assess the durability of long-term cognitive retention; (4) test the potential combination or rotation between Quizizz and Wordwall in a single learning unit to identify possible synergistic effects between competitive mechanisms and visual exploration; and (5) replicating similar studies on other PJOK health materials to test whether Quizizz's comparative advantages are consistent across topics or specific to socioculturally sensitive materials.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest in this research.

REFERENCE

- Ariani, D. (2020). Gamifikasi untuk Pembelajaran. *Jurnal Pembelajaran Inovatif*, 03(02), 144–149. <https://doi.org/10.21009/JPI.032.09>
- Arikunto, S. (2018). *Dasar-Dasar Evaluasi Pendidikan, Edisi 3* (R. Damayanti (ed.); 3rd ed.). PT Bumi Aksara.
- Auliya, A., Latipah, N., & Suhi. (2021). The Development Of Based Evaluation Instruments Wordwall For Science Courses Of Junior High School Class Vii. *Tarbiyah: Jurnal Ilmiah Kependidikan*, 10(2), 73–83. <https://jurnal.uin-antasari.ac.id/index.php/jtijk/article/view/4566>
- Buka, N., Jayamin, M. A., Kholis, N., Rasyid, N. A., Pascasarjana, P., & Alauddin, U. (2025). Peran Validitas dan Rehabilitas Dalam Evaluasi Pembelajaran. *Socius: Jurnal Penelitian Ilmu-Ilmu Sosial*, 02(June), 224–227. <https://doi.org/10.5281/zenodo.15557188>
- Citra, C. A., & Rosy, B. (2020). Keefektifan Penggunaan Media Pembelajaran Berbasis Game Edukasi Quizizz Terhadap Hasil Belajar Teknologi Perkantoran Siswa Kelas X SMK Ketintang Surabaya. *Jurnal Pendidikan Administrasi Perkantoran (JPAP)*, 8, 261–272. <https://doi.org/10.26740/jpap.v8n2.p261-272>
- Ghozali, I. (2021). *Aplikasi Analisis Multivariate Dengan Program IBM SPSS 26 Edisi 10* (10th ed.). Badan Penerbit Universitas Diponegoro.
- Gunawan, I., & Palupi, A. R. (2016). Revisi Taksonomi Bloom Ranah Kognitif: Kerangka Landasan Untuk Pembelajaran, Pengajaran, Dan Asesmen. *Jurnal Pendidikan Dasar Dan Pembelajaran*, 1, 98–117. <https://doi.org/10.25273/pe.v2i02.50>
- Haruna, H., Okoye, K., Zainuddin, Z., Hu, X., & Chu, S. (2021). Gamifying Sexual Education for



Adolescents in a Low-Tech Setting: Quasi-Experimental Design Study Corresponding Author : *JMIR Serious Games*, 9. <https://doi.org/10.2196/19614>

- Hidayati, I. D., & Aslam. (2021). Efektivitas Media Pembelajaran Aplikasi Quizizz Secara Daring Terhadap Perkembangan Kognitif Siswa. *Jurnal Pedagogi Dan Pembelajaran*, 4(2), 251–257. <https://doi.org/10.23887/jp2.v4i2.37038>
- Iklima, Harliyana, I., & Rahayu, R. (2023). Perbandingan Penggunaan Wordwall Dengan Quizizz Terhadap Hasil Menulis Resensi Kelas Xi Di Sma Negeri 1 Bireuen. *Kande : Jurnal Ilmiah Pendidikan Bahasa Dan Sastra Indonesia*, 4(2), 229–237. <https://doi.org/10.29103/jk.v4i2.13446>
- Johnston, L. G., Soe, P., Setiawati, A., Artha, W., Tarinanda, C., & Putri, A. (2021). Alarmingly High HIV Prevalence Among Adolescent and Young Men Who have Sex with Men (MSM) in Urban Indonesia. *AIDS and Behavior*, 25(11), 3687–3694. <https://doi.org/10.1007/s10461-021-03347-0>
- Kementerian Kesehatan Republik Indonesia. (2025). *Berani Tes, Berani Lindungi Diri, Kemenkes Targetkan Eliminasi HIV dan IMS Tahun 2030*. <https://kemkes.go.id/id/berani-tes-berani-lindungi-diri-kemenkes-targetkan-eliminasi-hiv-dan-ims-tahun-2030>
- Minarta, S. M., & Pamungkas, H. P. (2022). Efektivitas Media Wordwall Untuk Meningkatkan Hasil Belajar Ekonomi Siswa Man 1 Lamongan. *Oikos: Jurnal Kajian Pendidikan Ekonomi Dan Ilmu Ekonomi*, VI, 189–199. <https://doi.org/10.23969/oikos.v6i2.5628>
- Muntaza, F., Dhiyaa, A., & Haqq, U. (2022). The Effectiveness of Wordwall Application Learning Media on Students ' Cognitive Development at SMK Al Firdaus Banyuwangi. *Permadani: Journal of Teaching and Educational Management*, 1(2), 40–48. <https://permadani.amiin.or.id/index.php/permadani/article/view/13>
- Nurjannah, Kaswar, A. B., & Kasim, E. W. (2021). Efektifitas gamifikasi dalam pembelajaran matematika. *JURNAL MathEdu (Mathematic Education Journal)*, 4(2), 189–193. <https://doi.org/10.37081/mathedu.v4i2.2492>
- Safitri, F., & Supeni Edie, S. (2020). Penerapan Model Pembelajaran Problem Based Learning Menggunakan Strategi Pair Checks untuk Meningkatkan Hasil Belajar Siswa di MAN 2 Boyolali. *Unnes Physics Education Journal*, 9(1), 96–105. <https://journal.unnes.ac.id/sju/upej/article/view/38286>
- Saputra, K. W., Amalia, F. P. M. P. I. Q., & Rahman, K. (2024). Pengaruh Penggunaan Media Pembelajaran Interaktif Quizizz Terhadap Hasil Belajar Siswa Kelas 10 Program Keahlian Dkv Smk Negeri 10 Malang. *Jurnal Teknologi Informasi Dan Ilmu Komputer (JTIK)*, 11(6), 1177–1184. <https://doi.org/10.25126/jtiik.2023118108>
- Sugiyono. (2019). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D* (M. Dr.Ir.Sutopo.S.Pd (ed.); 2 Cetakan). Alfabeta.
- Sukmawati, I., Afdal, A., Fikri, M., Hariko, R., Zikra, Z., Suri, G. D., & Syapitri, D. (2023). Improving Research , and Practice to Reproductive Health Understanding : An Indonesian Perspective for Sexual Education. *Universal Journal of Public Health*, 11(4), 441–447. <https://doi.org/10.13189/ujph.2023.110408>



- Todesco, M., Breman, J., Haryanto, N. N., Kok, G., & Massar, K. (2023). Effect evaluation of a comprehensive sexuality education intervention based on socio-emotional learning among adolescents in Jakarta, Indonesia. *Frontiers in Public Health*, 11. <https://doi.org/10.3389/fpubh.2023.1254717>
- Wang, A. I., & Tahir, R. (2020). Computers & Education The effect of using Kahoot ! for learning – A literature review. *Computers & Education*, 149(May 2019), 103818. <https://doi.org/10.1016/j.compedu.2020.103818>
- Wibowo, N., Saptosari, N., & Wibowo, N. (2016). Upaya Peningkatan Keaktifan Siswa Melalui Pembelajaran Berdasarkan Gaya Belajar Di Smk Negeri. *Jurnal Electronics, Informatics, and Vocational Education (ELINVO)*, 1(2), 128–139. <https://doi.org/10.21831/elinvo.v1i2.10621>
- Winatha, K. R., Setiawan, I. M. D., & STMIK STIKOM Indonesia. (2020). Pengaruh Game-Based Learning Terhadap Motivasi dan Prestasi Belajar. *Scholaria: Jurnal Pendidikan Dan Kebudayaan*, 10(3), 198–206. <https://doi.org/10.24246/j.js.2020.v10.i3.p198-206>
- World Health Organization. (2025). *Sexually transmitted infections (STIs) Fact Sheets*. <https://www.who.int/news-room/fact-sheets/detail/sexually-transmitted-infections-%28stis%29>

