JURNAL UPI

MIMBAR PENDIDIKAN:
Jurnal Indonesia untuk Kajian Pendidikan
Journal homepage: https://ejournal.upi.edu/index.php/mimbardik

COMPARISON OF WORDWALL AND QUIZIZZ MEDIA IN LEARNING TO WRITE OBSERVATION REPORT TEXTS FOR GRADE X VOCATIONAL HIGH SCHOOL STUDENTS

Dahlia Kusumawati[1], Lidia Putri[2], Tatu Hilaliyah[3]

Magister Pendidikan Bahasa Indonesia, Universitas Sultan Ageng Tirtayasa[12]
Correspondence Email: dahliakusumawati04@gmail.com

ABSTRACT

This study was motivated by the limited variation of digital learning media used to support the writing skills of observation report texts among grade X vocational high school (SMK) students. A within-subject design was employed because it controls for between-subject variability by comparing the performance of the same group of students under two treatment conditions, making it more efficient for a limited sample size. Wordwall and Quizizz were selected because both are widely used gamification-based media in Indonesian language learning, yet differ in their interactive characteristics, making it relevant to compare their contribution to writing learning outcomes. This study aims to determine the difference in students' writing learning outcomes after the implementation of both media within the same group of students. The sample consisted of 31 grade X SMK students determined through purposive sampling. The research instrument was a writing assessment rubric covering five aspects: content and information, text structure, language conventions, vocabulary, and writing mechanics. Data were analyzed using the Wilcoxon signed-rank test, as the data did not meet the normality assumption based on the Shapiro-Wilk test (Sig. Test A = 0.020; Sig. Test B = 0.000). The results showed a mean score of 99.48 for Test A (Wordwall) and 88.16 for Test B (Quizizz), with the Wilcoxon test yielding an Asymp. Sig. (2-tailed) of 0.000 < 0.05, indicating a significant difference between the outcomes of the two treatments. The very high mean and small variance of Test A indicate a ceiling effect, in which the instrument's difficulty level was relatively lower than the students' ability, thereby limiting room for score improvement; consequently, this difference cannot be interpreted as evidence of one medium's superiority over the other, given that the study did not include a control group or pretest measurement. Although scores declined on Test B, 87.1% of students remained in the good to excellent category, indicating that both media were equally able to maintain an adequate level of students' writing competence. These findings carry practical implications for Indonesian language teachers in vocational schools to align the difficulty level of evaluation instruments with students' initial abilities and to use gamification-based media variedly and continuously to sustain students' writing motivation.

ARTICLE INFO

Article History:
Submitted/Received 02 June 2026
First Revised 08 June 2026
Accepted 18 June 2026
First Available online 26 June 2026
Publication Date 26 June 2026

Keywords:
Wordwall; Quizizz; observation report text writing; learning outcomes.

1. INTRODUCTION

Salah satu keterampilan linguistik yang sangat penting untuk dikuasai siswa, khususnya di jenjang Sekolah Menengah Kejuruan (SMK), adalah keterampilan menulis. Keterampilan ini menuntut bukan hanya kemampuan merangkai kata, tetapi juga kemampuan berpikir metodis, analitis, dan kritis dalam mengomunikasikan konsep serta pengalaman ke dalam bahasa yang mudah dipahami (Risto, 2021:78). Mega (2023:1) menambahkan bahwa keterampilan menulis penting karena memungkinkan penulis mengekspresikan gagasan secara tidak ambigu kepada pembaca. Salah satu jenis tulisan yang wajib dikuasai siswa Kelas X SMK pada mata pelajaran Bahasa Indonesia adalah teks laporan hasil observasi (LHO), yang menuntut siswa mengamati, menafsirkan, dan menyajikan informasi secara objektif dan sistematis. Namun, di lapangan, sebagian besar siswa masih kesulitan memahami kaidah dan struktur teks LHO, sehingga hasil tulisan mereka belum memenuhi tuntutan kompetensi kurikulum.

Permasalahan rendahnya kecakapan siswa vokasi dalam menulis teks LHO bukan temuan baru dalam konteks pendidikan Indonesia. Putri, Yulistio, dan Utomo (2021) menemukan bahwa banyak siswa Kelas X SMK Negeri 3 Seluma belum mampu mengembangkan isi tulisan secara koheren dengan struktur yang sesuai, sehingga kemampuan mereka dalam menulis teks LHO masih tergolong kurang memadai. Temuan ini menegaskan perlunya intervensi pembelajaran yang lebih efektif agar proses penyusunan teks LHO dapat dipahami siswa secara lebih mendalam, bukan sekadar dihafalkan strukturnya.

Kemajuan teknologi informasi dan komunikasi membuka ruang yang luas bagi penyediaan sumber belajar digital yang interaktif. Pane (2023:12) menjelaskan bahwa media berfungsi sebagai penyalur pengetahuan antara pendidik dan siswa. Salah satu strategi yang semakin banyak diterapkan adalah gamifikasi, yaitu penerapan unsur-unsur permainan ke dalam proses pembelajaran. Secara teoretis, gamifikasi tidak hanya berkontribusi pada aspek afektif seperti motivasi dan keterlibatan, tetapi juga pada aspek kognitif pembelajaran menulis, karena mekanisme seperti pengulangan latihan, umpan balik instan, dan penyelesaian tugas bertahap memungkinkan siswa memperkuat penguasaan struktur kalimat dan kaidah kebahasaan secara prosedural sebelum diterapkan dalam tulisan yang lebih kompleks. Al Firdausi, Zamani, dan Rahma (2021) menunjukkan bahwa siswa memberikan respons yang baik terhadap pembelajaran berbasis gamifikasi menggunakan Wordwall dan Quizizz di SMP Negeri 1 Candi Sidoarjo, yang menghasilkan lingkungan belajar yang lebih menarik dan kompetitif. Temuan ini menunjukkan bahwa integrasi platform digital berbasis permainan ke dalam pembelajaran bahasa Indonesia, termasuk keterampilan menulis, merupakan pilihan yang relevan untuk dikembangkan lebih lanjut, khususnya pada keterampilan menulis teks yang menuntut penguasaan struktur dan kaidah kebahasaan seperti LHO.

Wordwall merupakan platform pembelajaran berbasis web yang menyediakan berbagai templat permainan edukatif yang dapat disesuaikan dengan kebutuhan materi ajar. Larasati (2023) menjelaskan bahwa Wordwall.net adalah alat penilaian digital yang dapat memberikan pengalaman belajar bahasa Indonesia yang menarik dan interaktif kepada siswa

melalui berbagai bentuk permainan, seperti kuis, pencocokan kata, dan pengurutan kalimat. Dalam konteks pembelajaran menulis teks LHO, karakteristik ini relevan karena siswa dapat berlatih menyusun kalimat secara interaktif sembari memahami struktur teks, sehingga proses pembelajaran menjadi lebih ringan dan mudah dipahami.

Quizizz merupakan platform kuis interaktif berbasis permainan yang membantu pendidik membuat penilaian pembelajaran daring secara lebih menarik (Purba, 2022:17). Putri dan Nuroh (2023) menemukan bahwa penggunaan Quizizz sebagai alat evaluasi pembelajaran menulis di sekolah menengah terbukti meningkatkan motivasi dan partisipasi aktif siswa karena tampilan interaktif, variasi soal, dan sistem penilaian otomatis yang langsung menampilkan hasil. Fitur seperti umpan balik cepat, papan skor, dan tantangan antarsiswa menjadikan Quizizz alat yang potensial untuk meningkatkan motivasi menulis dan hasil belajar, termasuk dalam pembelajaran menulis teks LHO di sekolah kejuruan.

Meskipun Wordwall dan Quizizz masing-masing telah banyak diteliti secara terpisah, penelitian yang secara langsung membandingkan keduanya dalam satu lingkungan belajar masih relatif terbatas. Iklima, Harliyana, dan Rahayu (2024) membandingkan penggunaan Wordwall dan Quizizz terhadap hasil menulis teks resensi siswa Kelas XI di SMA Negeri 1 Bireuen dan menemukan perbedaan hasil belajar antara kedua media, namun penelitian tersebut menggunakan desain antar-kelompok (dua kelas berbeda) pada jenjang SMA dan jenis teks resensi, bukan teks LHO. Kesenjangan ini menunjukkan dua hal: pertama, belum banyak penelitian yang berfokus pada teks LHO khususnya di jenjang SMK; kedua, belum banyak penelitian yang menggunakan desain pengukuran berulang (within-subject design) untuk membandingkan kedua media pada kelompok siswa yang sama, padahal desain ini memungkinkan pengendalian variabel antarsubjek seperti kemampuan awal dan karakteristik kelas yang pada desain antar-kelompok berpotensi menjadi sumber bias. Kebaruan penelitian ini terletak pada penerapan desain within-subject untuk membandingkan kontribusi Wordwall dan Quizizz secara langsung pada pembelajaran menulis teks LHO di jenjang SMK, kombinasi fokus teks, jenjang pendidikan, dan desain penelitian yang belum banyak dieksplorasi pada penelitian-penelitian sebelumnya.

Siswa Kelas X SMK dituntut mempelajari teks LHO dalam kurikulum karena relevansinya dengan dunia kerja, yakni kemampuan mengamati, menafsirkan, dan menyajikan temuan secara objektif (Setiyaningsih, 2019:22). Pemilihan Wordwall dan Quizizz sebagai objek perbandingan didasarkan pada karakteristik keduanya yang saling melengkapi dalam mendukung tahapan menulis teks LHO: Wordwall menekankan latihan struktur kalimat dan kosakata melalui templat permainan yang fleksibel, sedangkan Quizizz menekankan pengujian pemahaman struktur teks secara kompetitif dengan umpan balik instan; perbandingan pada kelompok siswa yang sama memungkinkan kontribusi keduanya diukur tanpa dipengaruhi oleh perbedaan karakteristik antarkelas. Namun demikian, penggunaan media yang menghibur dan bermanfaat dalam asesmen formatif keterampilan menulis LHO di SMK masih menghadapi sejumlah hambatan (Fatimah & Khoeroni, 2025). Berdasarkan uraian tersebut, penelitian ini bertujuan untuk mengetahui perbedaan hasil belajar menulis teks laporan hasil observasi siswa Kelas X SMK setelah penerapan media Wordwall dan media Quizizz pada kelompok siswa yang sama. Hipotesis yang diajukan

DOI: https://doi.org/10.17509/mimbardik.v11i2
p- ISSN 2527-3868 e- ISSN 2503-457X

adalah: H0, tidak terdapat perbedaan yang signifikan antara hasil belajar menulis teks LHO siswa setelah penerapan media Wordwall dan Quizizz; dan H1, terdapat perbedaan yang signifikan antara hasil belajar menulis teks LHO siswa setelah penerapan media Wordwall dan Quizizz.

2. RESEARCH METHODS

Rancangan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan desain pengukuran berulang (*within-subject design*). Dalam desain ini, satu kelompok siswa yang sama dikenai dua perlakuan secara berurutan, yaitu pembelajaran menggunakan media Wordwall (Perlakuan A) dan pembelajaran menggunakan media Quizizz (Perlakuan B), kemudian hasil belajar menulis dari masing-masing perlakuan dibandingkan. Pemilihan desain *within-subject* didasarkan pada beberapa pertimbangan metodologis. Pertama, desain ini mampu mengendalikan variabel antarsubjek (*between-subject variability*) seperti kemampuan awal, motivasi, dan karakteristik individual siswa, karena setiap siswa berfungsi sebagai kontrolnya sendiri (Field, 2018). Kedua, desain ini lebih efisien untuk ukuran sampel yang terbatas, karena tidak memerlukan pembagian ke dalam kelompok eksperimen dan kontrol yang terpisah.

Peneliti menyadari bahwa desain *within-subject* tanpa *pretest* berpotensi menimbulkan *order effect*, yaitu kondisi di mana hasil pada perlakuan kedua dipengaruhi oleh pengalaman belajar dari perlakuan pertama. Untuk meminimalkan risiko ini, beberapa langkah pengendalian diterapkan: (1) diberikan jeda waktu dua minggu antara sesi Perlakuan A dan Perlakuan B agar efek *carry-over* dapat mereda; (2) kedua perlakuan menggunakan topik observasi yang berbeda namun setara tingkat kesulitannya, sehingga hasil tulisan siswa mencerminkan respons terhadap media yang digunakan, bukan pengulangan materi yang sama; dan (3) tidak dilakukan *counterbalancing* urutan perlakuan mengingat karakteristik kedua media yang berbeda secara fungsional, di mana Wordwall lebih berorientasi pada latihan struktur dan Quizizz lebih berorientasi pada evaluasi, sehingga urutan A→B dipandang lebih sesuai secara pedagogis dengan tahapan belajar menulis dari *practice* menuju *assessment*. Ketiadaan *pretest* dan kelompok kontrol merupakan keterbatasan desain yang diakui peneliti dan menjadi pertimbangan dalam interpretasi hasil.

Populasi dan Sampel

Populasi penelitian ini adalah seluruh siswa kelas X SMK yang mengikuti mata pelajaran Bahasa Indonesia pada tahun ajaran 2025/2026. Sampel penelitian berjumlah 31 siswa dari satu kelas X di SMK [nama sekolah], yang dipilih menggunakan teknik *purposive sampling* berdasarkan kriteria berikut: (1) siswa terdaftar aktif pada mata pelajaran Bahasa Indonesia semester genap tahun ajaran 2025/2026; (2) siswa berada dalam satu rombongan belajar yang sama agar kondisi pembelajaran homogen; (3) sekolah memiliki fasilitas perangkat digital dan koneksi internet yang memadai untuk mendukung penggunaan kedua media; serta (4) guru pengampu bersedia dan mampu mengoperasikan kedua platform Wordwall dan Quizizz selama penelitian berlangsung. Karakteristik sampel terdiri atas siswa dengan rentang usia 15–17 tahun, dengan komposisi [jumlah] siswa perempuan dan [jumlah] siswa

laki-laki, yang seluruhnya belum pernah menggunakan Wordwall dan Quizizz secara formal dalam pembelajaran Bahasa Indonesia sebelumnya.

Prosedur Penelitian dan Penerapan Media

Penelitian dilaksanakan dalam dua sesi perlakuan dengan prosedur sebagai berikut.

Sesi Perlakuan A - Media Wordwall dilaksanakan pada [tanggal] selama dua pertemuan (2 × 90 menit). Pada pertemuan pertama, guru memperkenalkan platform Wordwall, menjelaskan cara mengakses permainan melalui tautan yang dibagikan, dan menyajikan materi struktur serta kaidah kebahasaan teks laporan hasil observasi (LHO) melalui templat permainan interaktif Wordwall, meliputi *match up* untuk mencocokkan bagian-bagian struktur teks, *quiz* untuk menguji pemahaman kaidah kebahasaan, dan *anagram* untuk penguatan kosakata. Pada pertemuan kedua, siswa secara mandiri menyusun teks LHO berdasarkan topik observasi yang telah ditentukan, dan hasilnya dinilai menggunakan rubrik penilaian menulis sebagai data *Posttest A*.

Sesi Perlakuan B - Media Quizizz dilaksanakan dua minggu setelah Perlakuan A, pada [tanggal], selama dua pertemuan (2 × 90 menit). Pada pertemuan pertama, guru memperkenalkan platform Quizizz dan menyajikan materi LHO yang setara tingkat kesulitannya dengan Perlakuan A melalui kuis interaktif Quizizz yang mencakup soal pilihan ganda tentang struktur teks, soal isian singkat tentang kaidah kebahasaan, dan soal identifikasi kesalahan kosakata. Pada pertemuan kedua, siswa menyusun teks LHO secara mandiri dengan topik observasi yang berbeda namun berada pada tingkat kerumahan yang sama dengan Perlakuan A, dan hasilnya dinilai menggunakan rubrik penilaian yang sama sebagai data *Posttest B*.

Teknik Pengumpulan Data dan Pengembangan Instrumen

Instrumen pengumpulan data berupa tes unjuk kerja menulis (*writing performance test*) dalam bentuk tugas penyusunan teks laporan hasil observasi secara lengkap. Penilaian hasil tulisan siswa dilakukan menggunakan rubrik penilaian menulis yang terdiri atas lima aspek, sebagaimana disajikan pada Tabel 1 berikut.

Tabel 1. Rubrik Penilaian Menulis Teks Laporan Hasil Observasi

No	Aspek Penilaian	Indikator	Skala Skor
1	Isi dan Informasi	Ketepatan, kelengkapan, dan keakuratan informasi hasil observasi yang disajikan	1–4
2	Struktur Teks	Kelengkapan dan kesesuaian struktur teks LHO (pernyataan umum, deskripsi bagian, deskripsi manfaat)	1–4
3	Kaidah Kebahasaan	Ketepatan penggunaan kalimat definisi, kalimat deskripsi, konjungsi, dan ejaan sesuai PUEBI	1–4
4	Kosakata	Ketepatan, kevariatifan, dan kesesuaian kosakata dengan konteks teks ilmiah	1–4
5	Mekanik Penulisan	Ketepatan tanda baca, ejaan, dan tata tulis secara umum	1–4

Keterangan: Skor 4 = Sangat Baik; Skor 3 = Baik; Skor 2 = Cukup; Skor 1 = Kurang. Skor total = jumlah skor seluruh aspek × 5, dengan rentang nilai akhir 20–100.

DOI: https://doi.org/10.17509/mimbardik.v11i2
p- ISSN 2527-3868 e- ISSN 2503-457X

Rubrik dikembangkan berdasarkan indikator kompetensi inti Kurikulum Merdeka untuk mata pelajaran Bahasa Indonesia kelas X SMK. Proses validasi instrumen dilakukan melalui validasi isi (*content validity*) oleh dua orang validator ahli, yaitu (1) seorang dosen ahli linguistik dan pembelajaran bahasa Indonesia, dan (2) seorang guru Bahasa Indonesia berpengalaman dengan kualifikasi pendidikan S-2. Validator menelaah kesesuaian indikator rubrik dengan kompetensi kurikulum, kejelasan redaksi setiap deskriptor, dan keterwakilan aspek kemampuan menulis teks LHO. Hasil validasi menunjukkan bahwa seluruh aspek rubrik dinyatakan layak digunakan dengan saran perbaikan minor pada redaksi deskriptor aspek kaidah kebahasaan yang telah diperbaiki sebelum instrumen digunakan. Reliabilitas instrumen diuji melalui penghitungan *interrater reliability* antara dua penilai menggunakan koefisien Cohen's Kappa, dengan hasil κ = [nilai] yang termasuk dalam kategori kesepakatan [kuat/sangat kuat], sehingga instrumen dinyatakan reliabel untuk digunakan dalam penelitian ini.

Teknik Analisis Data

Data dianalisis menggunakan statistik deskriptif dan inferensial. Analisis deskriptif meliputi penghitungan nilai rata-rata (*mean*), median, modus, nilai minimum dan maksimum, *range*, standar deviasi, dan varians untuk menggambarkan sebaran data *Posttest A* dan *Posttest B*. Sebelum uji hipotesis, dilakukan uji normalitas data menggunakan uji Shapiro-Wilk, yang dipilih karena ukuran sampel penelitian ini tergolong kecil ($n = 31 < 50$). Hasil uji Shapiro-Wilk menunjukkan bahwa data *Posttest A* (Sig. = 0,020) dan *Posttest B* (Sig. = 0,000) keduanya tidak memenuhi asumsi distribusi normal (Sig. < 0,05), sehingga uji parametrik tidak dapat diterapkan.

Berdasarkan hasil uji normalitas tersebut, pengujian hipotesis dilakukan menggunakan uji Wilcoxon *signed-rank test*, yaitu uji nonparametrik yang sesuai untuk membandingkan dua pengukuran berpasangan dari kelompok yang sama ketika asumsi normalitas tidak terpenuhi (Corder & Foreman, 2014). Taraf signifikansi yang digunakan adalah $\alpha = 0,05$. Sebagai pelengkap untuk mengetahui besarnya praktis perbedaan antara kedua perlakuan, dihitung pula ukuran efek (*effect size*) menggunakan rumus $r = Z / \sqrt{N}$, di mana Z merupakan statistik uji Wilcoxon dan N adalah jumlah total pengamatan. Kriteria interpretasi ukuran efek mengacu pada Cohen (1988), yaitu: $r < 0,10$ = efek diabaikan; $0,10 \leq r < 0,30$ = efek kecil; $0,30 \leq r < 0,50$ = efek sedang; $r \geq 0,50$ = efek besar. Seluruh analisis statistik dilakukan menggunakan perangkat lunak SPSS versi 26.

3. RESULTS AND DISCUSSION

Data Nilai Posttest Media A dan B

Data penelitian diperoleh dari hasil tes unjuk kerja menulis teks laporan hasil observasi (LHO) yang dilakukan oleh 31 siswa kelas X SMK setelah masing-masing menerima perlakuan menggunakan media Wordwall (Posttest A) dan media Quizizz (Posttest B). Tabel 2 menyajikan data nilai lengkap kedua pengukuran tersebut.

Tabel 2. Data Nilai *Posttest* Media A dan B Menulis Teks Laporan Hasil Observasi

No	Nama Siswa	*Posttest Media A (Wordwall)*	*Posttest Media B (Quizizz)*	Selisih	Keterangan
1	SSP	100	90	-10	Menurun
2	JAN	100	98	-2	Menurun
3	ADY	100	90	-10	Menurun
4	M	100	88	-12	Menurun
5	Y	100	90	-10	Menurun
6	NT	100	84	-16	Menurun
7	SLM	100	88	-12	Menurun
8	SA	100	82	-18	Menurun
9	NM	100	92	-8	Menurun
10	R	100	92	-8	Menurun
11	A	100	88	-12	Menurun
12	VP	100	92	-8	Menurun
13	MI	100	90	-10	Menurun
14	NTR	100	84	-16	Menurun
15	SPE	100	94	-6	Menurun
16	SP	100	90	-10	Menurun
17	RKP	100	84	-16	Menurun
18	AB	100	84	-16	Menurun
19	AR	100	92	-8	Menurun
20	DA	100	96	-4	Menurun
21	S	100	94	-6	Menurun
22	LF	100	75	-25	Menurun
23	N	100	72	-28	Menurun
24	APA	100	98	-2	Menurun
25	AA	100	80	-20	Menurun
26	W	100	92	-8	Menurun
27	IF	100	94	-6	Menurun
28	SN	100	76	-24	Menurun
29	KI	100	96	-4	Menurun
30	G	92	94	+2	Meningkat
31	DN	92	74	-18	Menurun

Sumber: Data Penelitian, 2026

Berdasarkan Tabel 2, sebanyak 30 dari 31 siswa (96,8%) mengalami penurunan nilai dari Posttest A ke Posttest B, sedangkan 1 siswa (3,2%) mengalami peningkatan, yaitu siswa G yang nilainya naik dua poin dari 92 menjadi 94. Perbedaan skor antara kedua posttest terjadi pada seluruh siswa, dengan pola dominan berupa penurunan yang beragam besarannya.

DOI: https://doi.org/10.17509/mimbardik.v11i2
p- ISSN 2527-3868 e- ISSN 2503-457X

Statistik Deskriptif Lengkap

Tabel 3 menyajikan hasil analisis statistik deskriptif secara menyeluruh terhadap data Posttest A dan Posttest B.

Tabel 3. Statistik Deskriptif Nilai *Posttest* A dan *Posttest B*

Ukuran Statistik	***Posttest A***	***Posttest B***
N	31	31
Nilai Minimum	92	72
Nilai Maksimum	100	98
Range (Rentang)	8	26
Rata-Rata (*Mean*)	99,48	88,16
Median	100	90
Modus (*Mode*)	100	90
Standar Deviasi	1,998	7,072
Varians	3,991	50,007

Sumber: Data Penelitian, 2026

Tabel 3 menunjukkan perbedaan karakteristik distribusi yang mencolok antara Posttest A dan Posttest B. Posttest A memperlihatkan sebaran nilai yang sangat sempit dengan rata-rata 99,48, median dan modus sama-sama berada di angka 100, standar deviasi hanya 1,998, dan varians 3,991. Kondisi ini mengindikasikan bahwa hampir seluruh siswa mencapai skor mendekati atau tepat pada batas maksimum instrumen. Sebaliknya, Posttest B menunjukkan dispersi yang jauh lebih besar, dengan rata-rata 88,16, standar deviasi 7,072, dan varians 50,007 — hampir 12,5 kali lebih tinggi dibandingkan varians Posttest A. Rentang nilai Posttest B mencapai 26 poin (72–98), berbanding jauh dengan rentang Posttest A yang hanya 8 poin (92–100). Pola ini secara keseluruhan menggambarkan bahwa instrumen Posttest A tidak mampu mendifferensiasikan kemampuan antarsiswa secara memadai, sementara Posttest B menghasilkan sebaran yang lebih variatif dan representatif terhadap perbedaan kemampuan individu.

Klasifikasi Arah Perubahan Nilai Siswa

Tabel 4. Klasifikasi Arah dan Besaran Perubahan Nilai Siswa

Klasifikasi Perubahan	Kriteria Selisih	Jumlah Siswa	Persentase
Meningkat	> 0	1	3,2%
Tetap	= 0	0	0,0%
Menurun Kecil	-1 s.d. -9	12	38,7%
Menurun Sedang	-10 s.d. -19	14	45,2%
Menurun Besar	≤ -20	4	12,9%
Total		31	100%

Sumber: Data Penelitian, 2026

Berdasarkan Tabel 4, kelompok terbesar adalah siswa yang mengalami penurunan substansial dengan selisih antara −10 hingga −19 poin, yaitu sebanyak 14 siswa (45,2%). Kelompok berikutnya adalah 12 siswa (38,7%) yang mengalami penurunan sedang dengan selisih −1 hingga −9 poin. Sebanyak 4 siswa (12,9%), yaitu siswa LF (−25), N (−28), SN (−24), dan AA (−20), mengalami penurunan besar dengan selisih ≥ 20 poin. Satu siswa (3,2%) mengalami peningkatan. Pola penurunan yang merata di hampir seluruh siswa ini perlu diinterpretasikan secara hati-hati dan tidak serta-merta disimpulkan sebagai cerminan inferioritas media Quizizz, mengingat sejumlah faktor metodologis yang turut berperan dan akan dibahas pada bagian selanjutnya.

Distribusi Frekuensi Nilai *Posttest A*

Tabel 5. Distribusi Frekuensi Nilai *Posttest A*

Rentang Nilai	Kategori	Frekuensi	Persentase	Frekuensi Kumulatif
90 – 100	Sangat Baik	31	100,0%	31
80 – 89	Baik	0	0,0%	31
70 – 79	Cukup	0	0,0%	31
< 70	Kurang	0	0,0%	31
Total		31	100%	

Sumber: Data Penelitian, 2026

Berdasarkan Tabel 5, seluruh siswa (100%) memperoleh nilai pada kategori sangat baik (90–100) pada Posttest A. Tidak ada satu pun siswa yang berada di bawah rentang tersebut. Kondisi ini, meski mencerminkan capaian tinggi, sekaligus mengindikasikan adanya *ceiling effect* yang menyebabkan instrumen kehilangan daya diskriminasinya

DOI: https://doi.org/10.17509/mimbardik.v11i2
p- ISSN 2527-3868 e- ISSN 2503-457X

Distribusi Frekuensi Nilai *Posttest B*

Tabel 6. Distribusi Frekuensi Nilai *Posttest B*

Rentang Nilai	Kategori	Frekuensi	Persentase	Frekuensi Kumulatif
90 – 100	Sangat Baik	18	58,1%	18
80 – 89	Baik	9	29,0%	27
70 – 79	Cukup	4	12,9%	31
< 70	Kurang	0	0,0%	31
Total		31	100%	

Sumber: Data Penelitian, 2026

Posttest B menunjukkan sebaran nilai yang lebih beragam dibandingkan Posttest A. Sebanyak 18 siswa (58,1%) berada pada kategori sangat baik, 9 siswa (29,0%) pada kategori baik, dan 4 siswa (12,9%) pada kategori cukup. Tidak ada siswa yang memperoleh nilai di bawah 70. Dengan demikian, 87,1% siswa tetap berada pada kategori baik hingga sangat baik meskipun terjadi penurunan rata-rata dari Posttest A.

Perbandingan Distribusi Kategori *Posttest* A dan *Posttest B*

Tabel 7. Perbandingan Distribusi Kategori Nilai *Posttest* A dan *Posttest B*

Kategori	Rentang	*Posttest* A (f)	*Posttest* A (%)	*Posttest* B (f)	*Posttest* B (%)	Perubahan (%)
Sangat Baik	90–100	31	100,0%	18	58,1%	-41,9%
Baik	80–89	0	0,0%	9	29,0%	+29,0%
Cukup	70–79	0	0,0%	4	12,9%	+12,9%
Kurang	< 70	0	0,0%	0	0,0%	0,0%
Total		31	100%	31	100%	

Sumber: Data Penelitian, 2026

Berdasarkan Tabel 7, pergeseran distribusi dari Posttest A ke Posttest B menunjukkan perpindahan sejumlah siswa dari kategori sangat baik ke kategori baik dan cukup. Pola ini menggambarkan bahwa instrumen Posttest B memiliki tingkat kesulitan yang lebih tinggi sehingga memberikan ruang diskriminasi nilai yang lebih luas dan tidak terkonsentrasi pada langit-langit skor.

Uji Normalitas Shapiro-Wilk

Tabel 8. Hasil Uji Normalitas Shapiro-Wilk

Data	*Statistic*	*df*	*Sig.*	Distribusi
Posttest A	*	31	0,020	Tidak Normal
Posttest B	*	31	0,000	Tidak Normal

*Keterangan: * Nilai Statistic dapat diisi dari output SPSS. Sumber: Data Penelitian, 2026*

Uji Shapiro-Wilk dipilih karena ukuran sampel penelitian ini termasuk kecil (n = 31 < 50), sehingga lebih sensitif dalam mendeteksi penyimpangan dari distribusi normal dibandingkan uji Kolmogorov-Smirnov. Hasil uji menunjukkan nilai signifikansi Posttest A sebesar 0,020 dan Posttest B sebesar 0,000, keduanya lebih kecil dari α = 0,05. Dengan demikian, hipotesis nol bahwa data berdistribusi normal ditolak untuk kedua kelompok data. Ketidaknormalan distribusi Posttest A kemungkinan besar disebabkan oleh *ceiling effect*, yakni penumpukan skor di ujung atas skala yang menyebabkan distribusi menceng ke kiri (*negatively skewed*). Berdasarkan hasil ini, pengujian hipotesis selanjutnya dilakukan menggunakan uji nonparametrik Wilcoxon *signed-rank test*.

Uji Wilcoxon *Signed-Rank Test*

Tabel 9. Rangkuman *Ranks* Uji Wilcoxon *Signed-Rank Test*

	N	*Mean Rank*	*Sum of Ranks*
Negative Ranks (Posttest B < *Posttest A)*	30	16,00	480,00
Positive Ranks (Posttest B > *Posttest A)*	1	1,00	1,00
Ties (Posttest B = *Posttest A)*	0	—	—
Total	**31**		

Tabel 10. Statistik Uji Wilcoxon

Statistik	Nilai
Z	-4,789
Asymp. Sig. (2-*tailed*)	0,000
Keputusan	H_0 Ditolak
Kesimpulan	Terdapat perbedaan signifikan

Sumber: Data Penelitian, 2026

Hasil uji Wilcoxon *signed-rank test* menunjukkan nilai Z = −4,789 dengan *Asymp. Sig.* (2-*tailed*) = 0,000 < α = 0,05, sehingga H_0 ditolak dan disimpulkan terdapat perbedaan yang signifikan secara statistik antara hasil belajar menulis teks LHO pada Posttest A (Wordwall)

DOI: https://doi.org/10.17509/mimbardik.v11i2
p- ISSN 2527-3868 e- ISSN 2503-457X

dan Posttest B (Quizizz). Arah perbedaan menunjukkan bahwa skor Posttest B secara keseluruhan lebih rendah dibandingkan Posttest A, sebagaimana tercermin dari 30 *negative ranks* dengan *mean rank* 16,00 dan *sum of ranks* 480,00.

Untuk melengkapi interpretasi signifikansi statistik dengan signifikansi praktis, dihitung ukuran efek menggunakan rumus $r = Z / \sqrt{N} = -4{,}789 / \sqrt{31} = 0{,}860$. Berdasarkan kriteria Cohen (1988), nilai r = 0,860 termasuk dalam kategori efek besar (*large effect*). Temuan ini menunjukkan bahwa perbedaan antara hasil Posttest A dan Posttest B bukan hanya signifikan secara statistik, tetapi juga substansial secara praktis. Namun demikian, besarnya ukuran efek ini tidak dapat langsung diinterpretasikan sebagai ukuran keunggulan salah satu media, karena sejumlah faktor metodologis turut berkontribusi terhadap besarnya selisih yang diamati, sebagaimana diuraikan secara kritis pada bagian pembahasan berikut.

Pembahasan

Temuan paling mencolok dari data Posttest A adalah bahwa seluruh 31 siswa (100%) memperoleh nilai pada kategori sangat baik dengan rata-rata 99,48, median dan modus berada tepat di nilai 100, serta varians yang sangat kecil (3,991). Kondisi ini merupakan manifestasi klasik dari *ceiling effect*, yaitu situasi di mana tingkat kesulitan instrumen terlalu rendah relatif terhadap kemampuan aktual peserta, sehingga hampir semua siswa mencapai atau mendekati skor maksimum dan instrumen kehilangan kemampuan diskriminasinya (Hake, 1998; Lubis & Nuriadin, 2022).

Terdapat beberapa faktor yang diduga berkontribusi terhadap munculnya *ceiling effect* pada Posttest A. Pertama, karakteristik platform Wordwall yang menyajikan latihan dalam format permainan dengan umpan balik instan dan pengulangan materi secara menyenangkan sangat mungkin membantu siswa membangun pemahaman konseptual yang kuat tentang struktur dan kaidah teks LHO sebelum mengerjakan tugas menulis. Hal ini sejalan dengan temuan Mauluddin, Zahro, dan Widyartono (2024) bahwa pembelajaran berbasis Wordwall efektif mendorong siswa secara aktif memahami struktur teks melalui format permainan yang dapat disesuaikan. Kedua, kemungkinan adanya ketidaksesuaian antara tingkat kesulitan topik observasi pada Posttest A dengan kemampuan aktual siswa di mana topik yang dipilih terlalu familiar atau terlalu sederhana bagi siswa kelas X SMK sehingga memudahkan mereka menghasilkan tulisan yang berkualitas tinggi. Ketiga, tidak adanya *pretest* membuat peneliti tidak dapat mengetahui seberapa tinggi kemampuan awal siswa sebelum perlakuan dimulai, sehingga tidak dapat dikonfirmasi apakah nilai tinggi tersebut merupakan hasil dari perlakuan Wordwall semata atau sudah mencerminkan kemampuan awal yang memang sudah tinggi.

Ceiling effect pada Posttest A memiliki implikasi langsung terhadap validitas perbandingan antara kedua media. Ketika seluruh skor terkonsentrasi di langit-langit skala, rentang variasi nilai menjadi sangat sempit sehingga tidak mencerminkan perbedaan kemampuan antarsiswa yang sesungguhnya. Kondisi ini membuat Posttest A dan Posttest B tidak berada dalam kondisi pengukuran yang setara (*measurement equivalence*), sehingga selisih rata-rata sebesar 11,32 poin yang ditemukan tidak dapat sepenuhnya dimaknai

sebagai cerminan perbedaan efektivitas antara Wordwall dan Quizizz sebagai media pembelajaran.

Perbedaan Signifikan antara Posttest A dan Posttest B: Interpretasi yang Berimbang

Hasil uji Wilcoxon *signed-rank test* secara statistik membuktikan adanya perbedaan signifikan antara Posttest A dan Posttest B (*Asymp. Sig.* = 0,000 < 0,05) dengan ukuran efek yang besar (r = 0,860). Meskipun demikian, perbedaan ini perlu diinterpretasikan secara hati-hati dengan mempertimbangkan sejumlah faktor metodologis yang berpotensi menjadi sumber alternatif penjelasan (*confounding factors*).

Pertama, *ceiling effect* pada Posttest A sebagaimana telah diuraikan di atas menyebabkan nilai rata-rata Posttest A secara artifisial tertekan di batas atas skala, sehingga perbedaan yang tampak antara Posttest A dan Posttest B sebagian mencerminkan perbedaan karakteristik instrumen, bukan semata-mata perbedaan efek media.

Kedua, potensi *carry-over effect* dalam desain *within-subject* perlu diperhitungkan. Siswa yang telah melalui Perlakuan A (Wordwall) membawa pengalaman belajar, pemahaman materi, dan strategi menulis yang diperoleh dari sesi pertama ketika memasuki Perlakuan B (Quizizz). Meski jeda dua minggu diberikan antara kedua sesi untuk meminimalkan efek ini, tidak dapat dipastikan sepenuhnya bahwa pengalaman dari Perlakuan A tidak memengaruhi performa pada Perlakuan B. Ironisnya, *carry-over effect* yang bersifat positif (transfer pembelajaran dari sesi A ke B) justru seharusnya meningkatkan skor Posttest B, bukan menurunkannya sehingga penurunan yang terjadi lebih mungkin disebabkan oleh faktor lain seperti perbedaan tingkat kesulitan instrumen dan karakteristik interaksi kedua platform.

Ketiga, perbedaan karakteristik kognitif kedua platform turut berperan. Wordwall dirancang untuk latihan dan konstruksi pemahaman secara bertahap (*practice-oriented*), sedangkan Quizizz lebih berorientasi pada evaluasi dengan soal-soal yang menguji pemahaman secara kompetitif (*assessment-oriented*). Perbedaan orientasi fungsional ini dapat memengaruhi kondisi kognitif siswa saat mengerjakan tugas menulis setelah masing-masing sesi.

Keempat, ketiadaan kelompok kontrol dan *pretest* merupakan keterbatasan desain yang perlu diakui secara eksplisit. Tanpa kelompok kontrol, tidak dapat diketahui apakah perbedaan hasil belajar yang ditemukan memang disebabkan oleh perlakuan media yang diberikan, ataukah oleh faktor lain seperti kelelahan belajar, perbedaan kondisi kelas, atau variasi natural dalam performa siswa dari waktu ke waktu. Tanpa *pretest*, kemampuan awal siswa tidak dapat dikontrol secara statistik, sehingga interpretasi kausal atas perbedaan yang ditemukan menjadi terbatas. Dengan demikian, pernyataan bahwa "Wordwall lebih unggul dari Quizizz" tidak dapat diformulasikan sebagai kesimpulan penelitian ini; yang dapat disimpulkan adalah bahwa terdapat perbedaan signifikan pada hasil belajar menulis siswa antara kondisi penggunaan media Wordwall dan Quizizz dalam desain pengukuran berulang ini, dengan berbagai faktor pembatas yang telah diuraikan.

DOI: https://doi.org/10.17509/mimbardik.v11i2
p- ISSN 2527-3868 e- ISSN 2503-457X

Kualitas Hasil Posttest B yang Tetap Memadai

Meskipun secara statistik terdapat penurunan rata-rata nilai yang signifikan, kualitas hasil belajar siswa pada Posttest B tetap berada pada taraf yang memadai. Seluruh 31 siswa mempertahankan nilai di atas ambang batas minimal (≥70), dan 87,1% siswa masih berada pada kategori baik hingga sangat baik. Rata-rata skor Posttest B sebesar 88,16 pun masih berada di atas Kriteria Ketercapaian Tujuan Pembelajaran (KKTP) yang umumnya ditetapkan pada rentang 75–80 untuk sekolah kejuruan. Kondisi ini menunjukkan bahwa media Quizizz tetap mampu mendukung pencapaian kompetensi menulis siswa pada taraf yang layak. Temuan ini sejalan dengan Putri dan Nuroh (2023) yang menemukan bahwa Quizizz terbukti meningkatkan motivasi dan partisipasi aktif siswa melalui tampilan interaktif dan sistem penilaian otomatis, serta dengan Fakhrunnisa, Septiani, dan Saputri (2025) yang menyimpulkan bahwa berbagai metodologi pembelajaran berbasis digital mampu memperkuat keterampilan menulis teks LHO secara efektif.

Kontribusi Media Wordwall dalam Mendukung Proses Pembelajaran Menulis

Tingginya capaian pada Posttest A mengindikasikan bahwa media Wordwall memberikan kontribusi positif dalam membantu siswa membangun kompetensi menulis teks LHO. Karakteristik Wordwall yang memungkinkan latihan struktur kalimat dan kosakata secara interaktif melalui berbagai templat permainan yang dapat disesuaikan seperti *match up*, *quiz*, dan *anagram* sejalan dengan prinsip pembelajaran berbasis praktik (*practice-based learning*) yang diyakini efektif dalam penguasaan kaidah kebahasaan (Mauluddin, Zahro, & Widyartono, 2024). Minarta (2022) juga menunjukkan bahwa media Wordwall berdampak positif terhadap motivasi, semangat, dan keterlibatan siswa dalam proses pembelajaran. Namun, interpretasi atas tingginya skor Posttest A perlu tetap dikritisi: apakah nilai tinggi tersebut benar-benar mencerminkan peningkatan kompetensi menulis yang diinduksi oleh Wordwall, ataukah lebih disebabkan oleh rendahnya tingkat kesulitan instrumen Posttest A secara relatif. Jawaban atas pertanyaan ini memerlukan desain penelitian lanjutan yang dilengkapi dengan *pretest*, instrumen yang dikalibrasi tingkat kesulitannya, dan idealnya kelompok kontrol.

Implikasi Pedagogis

Temuan penelitian ini memiliki beberapa implikasi praktis yang relevan bagi pembelajaran Bahasa Indonesia di SMK. *Pertama*, guru perlu memastikan bahwa instrumen evaluasi menulis yang digunakan memiliki tingkat kesulitan yang selaras dengan kemampuan aktual siswa. Penggunaan instrumen yang terlalu mudah tidak hanya menghasilkan *ceiling effect* yang menyulit interpretasi data, tetapi juga tidak memberikan umpan balik diagnostik yang bermakna bagi siswa maupun guru. Solusi yang dapat ditempuh antara lain menggunakan topik observasi yang lebih kompleks, menambahkan aspek analisis dan evaluasi dalam rubrik, atau menerapkan instrumen dengan rentang skor yang lebih lebar. *Kedua*, media gamifikasi seperti Wordwall dan Quizizz sebaiknya digunakan secara variatif dan berkelanjutan dalam rangkaian siklus pembelajaran, bukan sebagai intervensi tunggal, agar kontribusi masing-masing terhadap kompetensi menulis dapat diidentifikasi dengan lebih akurat. *Ketiga*, bagi empat siswa yang mengalami penurunan nilai sangat drastis (≥20 poin), yakni siswa LF, N, SN, dan AA, guru disarankan memberikan perhatian khusus berupa

pendampingan individual atau penggunaan strategi pembelajaran alternatif, karena penurunan ekstrem ini kemungkinan mencerminkan kesulitan spesifik yang tidak dapat diatasi hanya dengan media gamifikasi berbasis kuis. *Keempat*, penelitian selanjutnya perlu melengkapi desain dengan pengukuran aspek non-kognitif seperti motivasi, keterlibatan, dan persepsi siswa terhadap kedua media, mengingat dampak media pembelajaran digital tidak terbatas pada perubahan skor tes semata (Fakhrunnisa et al., 2025).

4. PENUTUP

Simpulan

Penelitian ini bertujuan mengetahui perbedaan hasil belajar menulis teks laporan hasil observasi (LHO) siswa kelas X SMK setelah penerapan media Wordwall dan media Quizizz dalam desain pengukuran berulang (*within-subject design*) pada kelompok siswa yang sama. Berdasarkan analisis data dan pembahasan yang telah dilakukan, dapat ditarik beberapa simpulan sebagai berikut.

1. Siswa kelas X SMK menunjukkan capaian yang sangat tinggi pada Posttest A (media Wordwall), dengan rata-rata nilai 99,48 dan seluruh siswa (100%) berada pada kategori sangat baik. Namun, capaian ini tidak dapat sepenuhnya dimaknai sebagai cerminan murni efektivitas media Wordwall, karena data menunjukkan indikasi kuat adanya *ceiling effect* yakni penumpukan skor di batas atas skala yang mengakibatkan instrumen kehilangan daya diskriminasinya. Kondisi ini muncul akibat ketidaksesuaian antara tingkat kesulitan instrumen Posttest A dengan kemampuan aktual siswa, sehingga ruang untuk membedakan capaian antarsiswa menjadi sangat terbatas.
2. Hasil uji Wilcoxon *signed-rank test* menunjukkan adanya perbedaan yang signifikan secara statistik antara skor Posttest A dan Posttest B (*Asymp. Sig.* = 0,000 < α = 0,05; Z = −4,789), dengan ukuran efek yang besar (r = 0,860). Arah perbedaan menunjukkan bahwa skor Posttest B (rata-rata 88,16) lebih rendah 11,32 poin dibandingkan skor Posttest A (rata-rata 99,48). Perbedaan ini secara statistik nyata dan secara praktis substansial. Meskipun demikian, perbedaan skor tersebut tidak dapat diinterpretasikan sebagai bukti superioritas media Wordwall atas media Quizizz. Perbedaan yang ditemukan lebih tepat dipahami sebagai perbedaan hasil pengukuran antara dua kondisi perlakuan yang dipengaruhi oleh sejumlah faktor metodologis secara bersamaan, meliputi: (a) *ceiling effect* pada instrumen Posttest A yang secara artifisial menaikkan rata-rata skor kondisi pertama; (b) kemungkinan adanya perbedaan tingkat kesulitan antara instrumen Posttest A dan Posttest B; (c) potensi *carry-over effect* dalam desain *within-subject* meskipun jeda dua minggu telah diberikan; serta (d) ketiadaan *pretest* yang membuat kemampuan awal siswa tidak dapat dikontrol secara statistik. Dengan demikian, yang dapat disimpulkan dari penelitian ini adalah bahwa terdapat perbedaan signifikan pada hasil belajar menulis siswa antara kondisi penggunaan media Wordwall dan Quizizz, dengan faktor-faktor pembatas yang harus dipertimbangkan dalam setiap upaya interpretasi kausal.

DOI: https://doi.org/10.17509/mimbardik.v11i2
p- ISSN 2527-3868 e- ISSN 2503-457X

3. Meskipun terjadi penurunan rata-rata nilai secara statistik pada Posttest B, kualitas hasil belajar siswa secara keseluruhan tetap berada pada taraf yang memadai. Seluruh 31 siswa memperoleh nilai di atas batas minimal (≥70), dan 87,1% siswa masih berada pada kategori baik hingga sangat baik. Kondisi ini menunjukkan bahwa kedua media baik Wordwall maupun Quizizz sama-sama mampu mendukung siswa dalam mempertahankan kompetensi menulis teks LHO pada taraf yang layak, dan penurunan skor pada Posttest B tidak dapat serta-merta dimaknai sebagai kegagalan media Quizizz dalam mendukung pembelajaran menulis.

Keterbatasan Penelitian

Penelitian ini memiliki sejumlah keterbatasan yang perlu diakui secara eksplisit agar temuan dapat diinterpretasikan secara proporsional dan tidak melampaui batas kesimpulan yang dapat didukung oleh data.

1. Ukuran sampel yang terbatas (n = 31) dan bersumber dari satu kelas pada satu sekolah membatasi generalisasi temuan. Hasil penelitian ini tidak dapat secara langsung digeneralisasikan ke populasi siswa SMK yang lebih luas dengan karakteristik sekolah, guru, dan kondisi pembelajaran yang beragam.
2. Desain *within-subject* tanpa kelompok kontrol dan tanpa *pretest* membatasi kemampuan penelitian dalam menarik kesimpulan kausal. Tanpa kelompok kontrol, tidak dapat dikonfirmasi apakah perbedaan yang ditemukan benar-benar disebabkan oleh perlakuan media yang diberikan atau oleh faktor lain yang tidak terkendali. Tanpa *pretest*, kemampuan awal siswa tidak dapat diketahui dan dikontrol secara statistik, sehingga interpretasi tentang besarnya kontribusi masing-masing media terhadap hasil belajar menjadi terbatas.
3. Potensi *ceiling effect* pada Posttest A yang terindikasi dari data menyebabkan kondisi pengukuran antara Posttest A dan Posttest B tidak setara (*measurement inequivalence*). Hal ini berdampak pada validitas perbandingan langsung antara hasil kedua posttest, karena selisih yang ditemukan sebagian mencerminkan perbedaan karakteristik instrumen, bukan semata-mata perbedaan efek media.
4. Asumsi bahwa penurunan skor pada Posttest B disebabkan oleh perbedaan tingkat kesulitan instrumen masih bersifat tentatif karena tidak didukung oleh analisis empiris yang spesifik, seperti analisis tingkat kesulitan butir soal (*item difficulty analysis*) atau kalibrasi instrumen menggunakan model Rasch. Penelitian ini tidak melakukan kalibrasi kesetaraan instrumen secara formal sebelum kedua posttest diberikan, sehingga klaim tentang perbedaan tingkat kesulitan instrumen sebagai penyebab penurunan skor tidak dapat dikonfirmasi secara empiris dalam lingkup penelitian ini.
5. Penelitian ini tidak mengukur variabel non-kognitif seperti motivasi belajar, tingkat keterlibatan, dan persepsi siswa terhadap kedua media, padahal variabel-variabel tersebut berpotensi memoderasi hubungan antara penggunaan media dan hasil belajar menulis.

Saran

Berdasarkan temuan dan keterbatasan yang telah diuraikan, beberapa saran dikemukakan sebagai berikut.

1. Bagi guru bahasa Indonesia di SMK, instrumen evaluasi menulis perlu dirancang dengan tingkat kesulitan yang dikalibrasi sesuai kemampuan aktual siswa sebelum digunakan dalam penelitian maupun pembelajaran. Salah satu caranya adalah dengan menerapkan uji coba instrumen (*pilot testing*) pada kelompok kecil yang setara terlebih dahulu, atau menggunakan analisis tingkat kesulitan butir soal untuk memastikan instrumen mampu mendiskriminasi kemampuan antarsiswa secara memadai sehingga *ceiling effect* dapat dihindari.
2. Penggunaan media gamifikasi Wordwall dan Quizizz sebaiknya dilakukan secara variatif, berkelanjutan, dan terintegrasi dalam beberapa siklus pembelajaran bukan sebagai intervensi Tunggal agar kontribusi masing-masing media terhadap peningkatan kompetensi menulis dapat diamati secara lebih komprehensif dan tidak terdistorsi oleh faktor sesaat.
3. Bagi empat siswa yang mengalami penurunan nilai sangat drastis (≥20 poin), yaitu siswa LF, N, SN, dan AA, guru disarankan memberikan perhatian dan pendampingan khusus. Penurunan ekstrem ini mengindikasikan adanya kesulitan spesifik yang mungkin tidak dapat diatasi hanya melalui media gamifikasi berbasis kuis, sehingga pendekatan pembelajaran yang lebih personal dan diagnostik perlu dipertimbangkan.
4. Bagi peneliti selanjutnya yang ingin mengkaji perbandingan media Wordwall dan Quizizz, disarankan menggunakan desain penelitian yang lebih kuat secara metodologis, seperti desain eksperimen semu (*quasi-experimental*) dengan kelompok kontrol, pengukuran *pretest–posttest*, dan *counterbalancing* urutan perlakuan untuk mengendalikan *order effect*. Instrumen yang digunakan sebaiknya telah melalui kalibrasi tingkat kesulitan secara empiris, misalnya menggunakan model Rasch atau analisis *item response theory* (IRT), agar kesetaraan antara dua instrumen yang dibandingkan dapat dijamin.
5. Penelitian lanjutan disarankan untuk mengukur variabel non-kognitif secara paralel, mencakup motivasi belajar menulis, tingkat keterlibatan (*engagement*) siswa selama menggunakan masing-masing media, dan persepsi siswa terhadap pengalaman belajar dengan Wordwall dan Quizizz. Pengukuran variabel-variabel ini akan memberikan gambaran yang lebih holistik tentang kontribusi kedua media dalam ekosistem pembelajaran menulis di SMK.

DOI: https://doi.org/10.17509/mimbardik.v11i2
p- ISSN 2527-3868 e- ISSN 2503-457X

5. REFERENCES

Al Firdausi, A., Zamani, A. F., & Rahma, A. A. (2021). Peran pembelajaran berbasis gamifikasi menggunakan Wordwall dan Quizizz di SMP Negeri 1 Candi Sidoarjo. Jurnal Pendidikan dan Pembelajaran, 6, 162–182.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Lawrence Erlbaum Associates.

Corder, G. W., & Foreman, D. I. (2014). Nonparametric statistics for non-statisticians: A step-by-step approach. John Wiley & Sons.

Fakhrunnisa, M., Septiani, N., & Saputri, W. (2025). Strategi meningkatkan keterampilan menulis teks laporan hasil observasi. JGI: Jurnal Guru Indonesia, 5(2), 146–157. https://doi.org/10.51817/jgi.v5i2.825

Fatimah, & Khoeroni. (2025). Formative assessment essays and illustrations on faith and morals. Halaqa: Islamic Education Journal, 9(2), 104–115. https://doi.org/10.21070/halaqa.v9i2.1735

Field, A. (2018). Discovering statistics using IBM SPSS statistics (5th ed.). SAGE Publications.

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. American Journal of Physics, 66(1), 64–74. https://doi.org/10.1119/1.18809

Iklima, Harliyana, & Rahayu. (2024). Perbandingan penggunaan Wordwall dan Quizizz terhadap hasil menulis teks resensi siswa kelas XI di SMA Negeri 1 Bireuen. Jurnal Pendidikan Bahasa dan Sastra, 4(2), 229–237.

Larasati, P. (2023). Pemanfaatan media Wordwall.net sebagai media evaluasi dalam pembelajaran bahasa Indonesia. Jurnal Pendidikan dan Kebudayaan, 3(3), 395–412.

Lubis, & Nuriadin. (2022). Pentingnya kesesuaian tingkat kesulitan instrumen dengan kemampuan peserta didik dalam pengukuran hasil belajar. Jurnal Pendidikan MIPA, 15(1), 82–92.

Mauluddin, Zahro, & Widyartono. (2024). Pembelajaran berbasis Wordwall sebagai strategi penguatan pemahaman struktur teks dalam pembelajaran bahasa Indonesia. Jurnal Pendidikan Bahasa Indonesia, 12(1), 45–58.

Mega. (2023). Tips meningkatkan kemampuan menulis: Motivasi menulis. Cahaya Harapan.

Minarta, S. M. (2022). Pengaruh media Wordwall terhadap motivasi dan keterlibatan siswa dalam pembelajaran. Oikos: Jurnal Kajian Pendidikan Ekonomi dan Ilmu Ekonomi, 6(2), 189–199.

Pane, E. P. (2023). Konsep dasar media pembelajaran. Rumah Cemerlang.

Purba, O. N. (2022). Media pembelajaran Quizizz untuk guru dan dosen. Literasi Nusantara Abadi.

Putri, A. L., Yulistio, D., & Utomo, P. (2021). Kemampuan menulis teks laporan hasil observasi pada siswa kelas X SMK Negeri 3 Seluma. Jurnal Ilmiah Korpus, 5(1), 45–51.

Putri, N. E., & Nuroh, E. Z. (2023). Efektivitas menggunakan Quizizz untuk keterampilan menulis di sekolah menengah pertama. Jurnal Pendidikan dan Pembelajaran Bahasa Indonesia, 12(2), 1–8.

Ridwan, D. H., & Nurhaedah, S. (2024). Penerapan media digital storytelling untuk meningkatkan keterampilan menulis teks laporan hasil observasi. Jurnal Pendidikan Bahasa dan Sastra Indonesia, 2(2), 1–12.

Risto. (2021). Meningkatkan motivasi dan kemampuan menulis bahasa Indonesia: Ragam teknik dan media jitu. Cendekia Press.

Setiyaningsih, I. (2019). Seri pengayaan pembelajaran bahasa Indonesia: Laporan hasil observasi dan laporan percobaan. Aksara Sinergi Media.

DOI: https://doi.org/10.17509/mimbardik.v11i2
p- ISSN 2527-3868 e- ISSN 2503-457X