

**PENGEMBANGAN EVALUASI BERBASIS HOTS DENGAN MENGGUNAKAN QUIZIZZ UNTUK MENINGKATKAN PROBLEM SOLVING SKILL DAN SELF EFFICACY SISWA**

**Eka Firmansyah<sup>1)\*</sup>, Fina Samaro Ma'mun<sup>1)</sup>, Gitta Maharani<sup>2)</sup>**

<sup>1</sup>Magister Pendidikan Matematika, Universitas Pasundan, Indonesia

<sup>2</sup>SMK Bina Karya 1 Karawang, Indonesia

[eka\\_firmansyah@unpas.ac.id](mailto:eka_firmansyah@unpas.ac.id)

Informasi Artikel	Abstrak
<p>Submitted: Jan 2, 2026 Revised: April 3, 2026 Accepted: May 20, 2026</p> <p><b>Kata Kunci</b> HOTS; Komunikasi Pengembangan Evaluasi; Problem Solving Skill; Quizizz; Self Efficacy.</p>	<p><b>Tujuan:</b> Fokus riset diarahkan pada perancangan sekaligus pengujian kelayakan suatu perangkat penilaian matematika berorientasi Higher Order Thinking Skills (HOTS) yang dioperasikan melalui platform Quizizz. Selain menghasilkan produk evaluasi, kegiatan ini turut menelaah perubahan kemampuan <i>problem solving</i> serta <i>self-efficacy</i> peserta didik setelah penggunaan instrumen tersebut.</p> <p><b>Metode:</b> Pendekatan yang diterapkan berbentuk R&amp;D dengan kerangka kerja ADDIE. Pelaksanaannya melibatkan peserta didik kelas X TMI 4 dan X TMI 2 di SMK Bina Karya 1 Karawang. Perangkat pengumpulan data mencakup formulir penilaian pakar (bidang materi, media, serta konstruksi instrumen), seperangkat tes <i>problem solving</i>, dan kuesioner <i>self-efficacy</i>. Pengolahan data peningkatan kemampuan menggunakan uji Wilcoxon disertai perhitungan N-gain, sedangkan pengaruh perlakuan dihitung melalui indeks effect size.</p> <p><b>Hasil:</b> Validasi memperoleh skor rerata 4,17 untuk aspek materi, 3,77 pada aspek media, dan 3,93 pada aspek instrumen, sehingga produk dinyatakan dapat digunakan. Hasil pengujian Wilcoxon menunjukkan nilai Asymp. Sig di bawah 0,05 yang menandakan adanya perubahan signifikan pada <i>problem solving skill</i>. Pada variabel <i>self-efficacy</i>, perolehan N-gain berada pada kategori sedang dengan rincian magnitude 0,36, strength 0,39, serta generality 0,33. Nilai effect size terhadap <i>problem solving</i> tercatat 0,62 dalam kategori sedang, sedangkan pada <i>self-efficacy</i> termasuk kategori tinggi.</p> <p><b>Simpulan:</b> Keluaran riset berupa instrumen evaluasi matematika berbasis HOTS terintegrasi dengan Quizizz yang telah melewati tahap penelaahan ahli dan uji terbatas. Perangkat tersebut tersedia sebagai opsi penilaian yang mendukung penguatan keterampilan pemecahan masalah sekaligus memperbesar tingkat keyakinan diri peserta didik saat menghadapi tugas matematika.</p>
<p><b>Keywords</b> Evaluation Development Communication; Problem-Solving Skill; Quizizz; Self-Efficacy.</p>	<p style="text-align: center;"><b>Abstract</b></p> <p><b>Purpose:</b> The research focus is directed toward the design simultaneously with feasibility testing of a mathematics assessment apparatus oriented toward Higher Order Thinking Skills (HOTS) that is operated through the Quizizz platform. Beyond producing an evaluation product, this activity also examines the transformation of problem solving ability as well as self-efficacy of learners following the utilization of said instrument.</p> <p><b>Method:</b> The approach applied takes the form of R&amp;D with ADDIE working framework. Its implementation involves learners of class X TMI 4 and X TMI 2 at SMK Bina Karya 1 Karawang. The data collection apparatus encompasses expert assessment instruments (material domain, media, as well as instrument construction), a set of problem solving tests, and self-efficacy questionnaires. Data processing of ability enhancement utilizes Wilcoxon test accompanied by N-gain calculation, whereas treatment influence is calculated through effect size index.</p> <p><b>Results:</b> Validation obtained a mean score of 4.17 for the material aspect, 3.77 on the media aspect, and 3.93 on the instrument aspect, such that the product is declared usable. Wilcoxon testing results demonstrate Asymp. Sig value below 0.05 which signifies the existence of significant transformation in problem solving</p>

---

skill. On the self-efficacy variable, N-gain acquisition resides within the moderate category with magnitude specification of 0.36, strength 0.39, as well as generality 0.33. Effect size value toward problem solving is recorded at 0.62 within the moderate category, whereas on self-efficacy it is included in the high category.

**Conclusion:** Research output takes the form of HOTS-based mathematics evaluation instrument integrated with Quizizz that has passed through expert scrutinization stage and limited testing. Said apparatus is available as an assessment option that supports the strengthening of problem-solving skills simultaneously with increasing the level of self-confidence of learners when confronting mathematics tasks.

---

## PENDAHULUAN

Penilaian dalam pembelajaran matematika perlu mengakomodasi kemampuan berpikir tingkat tinggi, yang mencakup penalaran kritis, kreativitas, komunikasi efektif, dan kolaborasi. Oleh karena itu, penilaian tidak hanya berorientasi pada hasil akhir, tetapi juga diarahkan untuk mengevaluasi proses serta operasi kognitif kompleks peserta didik. Evaluasi matematika abad ke-21 tidak lagi relevan jika hanya berfokus pada aspek ingatan dan penguasaan prosedur rutin, melainkan harus secara eksplisit diarahkan pada pengembangan dan pengukuran Higher Order Thinking Skills (HOTS) (Boral dkk, 2025; Jailani dkk, 2023). Evaluasi pembelajaran menempati posisi sentral dalam keseluruhan sistem instruksional. Fungsinya melampaui pencatatan capaian kompetensi; instrumen evaluasi juga bertindak sebagai pemicu aktivitas mental tingkat lanjut. Pada ranah matematika, hasil evaluasi menjadi dasar refleksi empiris bagi guru guna menyusun penyesuaian strategi yang selaras dengan orientasi HOTS (Ansya dkk., 2024; Kusumaningtyas dkk., 2023; Saleem dkk., 2025). Kenyataan di lapangan memperlihatkan dominasi soal prosedural dan repetitif, sehingga eksplorasi kemampuan analitis tingkat tinggi belum memperoleh ruang memadai.

Kebijakan nasional melalui program Asesmen Kompetensi Minimum (AKM) menempatkan urgensi penguatan berpikir kritis serta pemecahan masalah. Soal HOTS menuntut aktivitas mental berupa analisis, evaluasi, dan kreasi, serta kreasi yang dikaitkan dengan konteks kehidupan nyata (Rahmawatingrum dkk., 2019; Putri dkk., 2024). Dalam konteks ini, evaluasi tidak hanya berfungsi mengukur capaian kompetensi, tetapi juga berperan sebagai pemicu berkembangnya kemampuan berpikir tingkat tinggi peserta didik. Secara konseptual, soal HOTS berada pada level C4, C5, dan C6 dalam Taksonomi Bloom Revisi (Muslika, 2025; Widana, 2017; Manalu & Harahap, 2025; Masrifah dkk., 2025). Selain itu, efektivitas evaluasi berbasis HOTS turut dipengaruhi oleh media dan platform pembelajaran yang digunakan. Platform digital yang interaktif, menyediakan umpan balik langsung, serta memfasilitasi diskusi antarpeserta didik cenderung lebih mendukung pengembangan HOTS dibandingkan tes konvensional (Zhan dkk., 2023; Yuniarti dkk., 2024; Huang dkk., 2022; Lu dkk., 2025; Dermawan dkk., 2025; Setyaningrum, 2024; Putri & Aznam, 2019).

Ekspansi teknologi digital menghadirkan alternatif mekanisme evaluasi berbasis daring melalui Quizizz yang sangat relevan untuk mendukung penilaian HOTS. Tidak sekadar mengandalkan fitur gamifikasi umum seperti timer dan leaderboard untuk memicu kecepatan berpikir, Quizizz juga menyediakan tipe soal terbuka (open-ended), fitur mengurutkan (reorder), serta integrasi stimulus multimedia. Fitur-fitur ini secara strategis memaksa peserta didik melakukan analisis mendalam, mengevaluasi informasi, hingga mengonstruksi argumen secara kritis melalui sistem umpan balik yang instan. Sejumlah studi melaporkan peningkatan

motivasi dan partisipasi melalui pemanfaatan aplikasi tersebut (Muchuweni & Kariyana, 2025; Aritonang, 2025; Astuti dkk., 2024; Fadda dkk., 2021). Kendati demikian, implementasinya sering terbatas pada pertanyaan tingkat rendah. Di sisi lain, dimensi afektif berupa *self-efficacy* memiliki kontribusi signifikan terhadap keberhasilan belajar. Keyakinan individu atas kapasitas dirinya memengaruhi persistensi ketika menghadapi persoalan kompleks (Yang dkk., 2024; Liu dkk., 2024; Holenstein dkk., 2021; Shone dkk., 2023; Živković, 2023). Paparan tugas berorientasi HOTS dinilai dapat memperkuat dimensi tersebut (Sasson & Tifferet, 2024).

Penelitian sebelumnya telah mengulas evaluasi HOTS maupun penggunaan Quizizz dalam matematika (Huei dkk., 2021; Pitoyo dkk., 2020; Yunus & Hua, 2021). Sebagian besar kajian tersebut memiliki keterbatasan karena hanya berfokus pada media atau satu variabel kognitif tertentu; misalnya, laporan Handayani dkk. (2024) yang hanya berfokus pada peningkatan berpikir kritis, serta studi Zhang & Crawford (2023) dan Hassan dkk. (2024) yang sebatas menyoroti kenaikan keterlibatan serta pengalaman belajar melalui platform digital. Kebaruan penelitian ini hadir untuk mengatasi celah tersebut dengan mengintegrasikan pengembangan produk sistem evaluasi HOTS yang dirakit secara terstruktur dalam Quizizz, serta menguji dampaknya terhadap problem solving skill (Amalina & Vidákovich, 2023; Olivares dkk., 2020; Ukobizaba dkk., 2021) dan self-efficacy (Sagitarini dkk., 2023) secara bersamaan guna mengisi celah empiris yang belum tersentuh pada penelitian-penelitian terdahulu. Berdasarkan uraian tersebut, tujuan penelitian ini adalah untuk mengembangkan evaluasi berbasis HOTS dengan menggunakan Quizizz serta menganalisis efektivitasnya dalam meningkatkan *problem solving skill* dan *self efficacy* siswa.

## **METODE PENELITIAN**

### ***Desain Penelitian***

Rangkaian pengembangan mengikuti tahapan ADDIE, dimulai dari analisis kebutuhan peserta didik dan karakteristik materi barisan serta deret aritmetika. Fase desain memuat penyusunan kisi-kisi dan perancangan butir HOTS beserta konfigurasi tampilan pada Quizizz. Fase pengembangan melibatkan telaah oleh pakar materi, media, dan instrumen. Uji coba terbatas dilaksanakan pada kelompok siswa untuk memperoleh data empiris. Tahap akhir berupa evaluasi produk guna menilai tingkat kelayakan serta pengaruhnya terhadap variabel riset.

### ***Subjek***

Penelitian dilaksanakan di SMK Bina Karya 1 Karawang dengan melibatkan 64 siswa sebagai partisipan total. Teknik pengambilan sampel yang digunakan adalah purposive sampling, di mana kelas X TMI 4 dan X TMI 2 dipilih sebagai subjek uji coba produk evaluasi berbasis HOTS yang dikembangkan. Pemilihan kedua kelas ini didasarkan pada rekomendasi guru mata pelajaran serta karakteristik kemampuan akademik siswa yang cenderung homogen.

### ***Instrumen***

Perangkat pengumpulan data mencakup dua kategori utama. Pertama, instrumen tes berupa 5 butir soal uraian problem solving berbasis HOTS. Instrumen ini disusun berdasarkan kisi-kisi indikator yang meliputi: identifikasi masalah, penyusunan rencana, pelaksanaan strategi, serta verifikasi hasil. Sebelum digunakan, instrumen tes telah melalui uji validitas ahli (*expert judgment*) dan dinyatakan valid secara empiris melalui uji coba lapangan dengan nilai reliabilitas sebesar 0,81 (kategori tinggi). Kedua, instrumen non-tes berbentuk kuesioner self-efficacy menggunakan skala Likert (1–5) yang terdiri atas 30 butir item pernyataan. Kuesioner

ini memuat dimensi magnitude (contoh indikator: keyakinan menghadapi tingkat kesulitan soal), strength (contoh indikator: daya tahan menyelesaikan tugas sulit), dan generality (contoh indikator: penerapan keyakinan pada berbagai situasi belajar).

### **Analisis Data**

Pengolahan data ditempuh melalui dua jalur yang berbeda karakter. Jalur pertama bersifat pemaparan numerik untuk membaca hasil telaah pakar serta distribusi skor *self-efficacy* peserta didik. Jalur kedua bergerak pada ranah pengujian hipotesis dengan prosedur nonparametrik Wilcoxon guna mendeteksi perubahan *problem solving skill* dan *self-efficacy* sebelum serta sesudah perlakuan. Tambahan analisis berupa kalkulasi effect size dilakukan untuk memetakan derajat pengaruh instrumen HOTS yang diimplementasikan melalui Quizizz.

## **HASIL DAN PEMBAHASAN**

### **Hasil Penelitian**

Temuan empiris memperlihatkan bahwa perangkat evaluasi HOTS yang dioperasikan lewat Quizizz memenuhi kriteria penggunaan dalam pembelajaran matematika. Skor rerata dari pakar materi tercatat 4,17, pakar media 3,77, serta pakar instrumen 3,93. Seluruh nilai tersebut berada pada rentang kategori layak menurut pedoman penilaian yang diterapkan.

**Tabel 1.** Hasil Analisis Validasi Ahli terhadap Instrumen/Evaluasi

No.	Nama Validator	Konstruksi	HOTS	Bahasa	Rata-rata Total
1	Validator 1	3,86	4,00	4,00	3,95
2	Validator 2	3,71	4,75	3,67	4,04
3	Validator 3	3,86	4,25	4,67	4,26
Rata-rata per Aspek		3,81	4,33	4,11	4,08

Kelayakan ini menunjukkan bahwa produk telah memenuhi aspek substansi materi, konstruksi soal, bahasa, serta tampilan media.

Pada aspek *problem solving skill*, terjadi peningkatan skor dari pretes ke postes. Rata-rata pretes sebesar 25,4 meningkat menjadi 41,30 pada postes.

**Tabel 2.** Hasil Analisis Efektivitas *Problem Solving Skill*

<b>KELAS EKSPERIMEN</b>	
Rata-rata PRETES	25,4
Rata-rata POSTES	41,30
Standar Deviasi	25,71
<b>EFFECT SIZE</b>	<b>0,62</b>

Indeks effect size yang diperoleh sebesar 0,62 menempati klasifikasi sedang. Angka tersebut mengindikasikan adanya daya dorong yang terukur terhadap peningkatan kemampuan pemecahan masalah, meskipun tidak mencapai level kuat atau ekstrem.

Pada variabel *self-efficacy*, seluruh dimensi mengalami peningkatan. Dimensi *magnitude* meningkat dari 56 menjadi 72 dengan *effect size* 2,93, dimensi *strength* dari 59 menjadi 75 dengan *effect size* 1,43, dan dimensi *generality* dari 58 menjadi 72 dengan *effect size* 1,65. Secara keseluruhan, efektivitas terhadap *self-efficacy* berada pada kategori tinggi.

**Tabel 3.** Hasil Analisis Efektivitas *Self Efficacy*

<b>Indikator Self Efficacy</b>	<b>Pretes</b>	<b>Postes</b>	<b>Standar Deviasi</b>	<b>Effect Size</b>
Magnitude	56	72	5.46	2.93
Strength	59	75	11.16	1.43
Generality	58	72	8.47	1.65

### ***Pembahasan***

Penentuan kelayakan tidak semata-mata bersandar pada angka validasi pakar. Struktur butir soal menunjukkan keterpenuhan indikator HOTS, komposisi pertanyaan memenuhi kaidah konstruksi, redaksi kalimat terkontrol, dan tampilan media tersusun sistematis. Konfigurasi ini mengharuskan peserta didik melakukan penguraian informasi, penilaian alternatif solusi, serta penyelesaian persoalan non-rutin dengan tahapan berpikir yang bertingkat. Hasil tersebut memperluas bukti yang dikemukakan Aryana dkk. (2025) mengenai kontribusi instrumen HOTS terhadap penguatan proses kognitif tingkat tinggi.

Data peningkatan *problem solving skill* menampilkan perubahan yang konsisten setelah peserta didik berinteraksi dengan soal berkarakter HOTS. Pola soal yang tidak prosedural memaksa pemilihan strategi secara sadar dan pemeriksaan ulang langkah penyelesaian. Situasi ini membentuk latihan mental berulang pada ranah analitis. Hasil tersebut memiliki koherensi dengan simpulan Irianto dkk. (2021) serta Jailani dkk. (2023) yang menempatkan evaluasi HOTS sebagai faktor signifikan dalam pengembangan kemampuan pemecahan masalah matematika.

Pada dimensi self-efficacy, tingkat efektivitas intervensi tergolong tinggi. Peningkatan yang signifikan ini khususnya terlihat pada dimensi strength dan generality. Secara teoritis, kenaikan kedua dimensi ini dipicu oleh keberhasilan siswa dalam menyelesaikan tantangan soal-soal berbasis HOTS di platform Quizizz. Mengacu pada teori efikasi diri Albert Bandura, keberhasilan tersebut menstimulasi terbentuknya mastery experience (pengalaman menguasai sesuatu), yang merupakan sumber terkuat dalam membangun efikasi diri seseorang. Ketika siswa berulang kali berhasil memecahkan masalah kompleks, strength (keteguhan keyakinan mereka untuk tidak mudah menyerah) akan menguat, dan secara simultan menumbuhkan generality (keyakinan bahwa mereka juga mampu menerapkan kemampuan analisis tersebut pada berbagai situasi atau topik matematika lainnya). Pemanfaatan Quizizz menghadirkan respons langsung terhadap jawaban dan mekanisme interaksi yang aktif. Kondisi tersebut berimplikasi pada peningkatan keyakinan individu dalam menghadapi tugas matematis. Bukti ini memiliki kesesuaian dengan laporan Sofyan dkk (2025) terkait pengaruh platform digital interaktif terhadap penguatan kepercayaan diri akademik.

Keunggulan penelitian ini dibandingkan penelitian terdahulu terletak pada pengintegrasian pengembangan instrumen evaluasi HOTS dengan pengujian efektivitasnya terhadap dua aspek secara simultan, yaitu aspek kognitif (*problem solving skill*) dan afektif (*self-efficacy*), sehingga memberikan kontribusi yang lebih komprehensif terhadap kajian evaluasi pembelajaran matematika berbasis digital.

### ***Implikasi***

Secara teoritis, penelitian ini memperkuat konsep bahwa evaluasi berbasis HOTS tidak hanya mengembangkan kemampuan berpikir tingkat tinggi, tetapi juga berkontribusi terhadap peningkatan keyakinan diri siswa dalam belajar matematika. Dalam praktik kelas, pendidik memperoleh alternatif evaluasi digital yang tidak monoton melalui Quizizz. Fungsi evaluasi bergeser dari sekadar pencatatan hasil menjadi perangkat stimulasi berpikir terstruktur sekaligus sarana pembentukan sikap positif terhadap matematika. Pada tataran kebijakan pendidikan, hasil ini menyediakan dasar argumentatif bagi perumusan sistem evaluasi berbasis teknologi yang kompatibel dengan karakter pembelajaran abad ke-21 dan orientasi kompetensi berpikir tingkat tinggi.

### ***Keterbatasan dan Rekomendasi Penelitian Lanjutan***

Pelaksanaan penelitian memuat sejumlah batasan. Lingkup uji coba terfokus pada satu institusi dengan jumlah kelas terbatas sehingga cakupan generalisasi belum luas. Rancangan penelitian tidak melibatkan kelompok kontrol, sehingga komparasi efektivitas antarpendekatan belum sepenuhnya terpetakan. Materi kajian juga terbatas pada topik barisan dan deret aritmetika. Pengembangan selanjutnya direkomendasikan menggunakan desain eksperimen dengan keberadaan kelompok kontrol untuk memperoleh perbandingan yang lebih kuat. Ekspansi konteks dapat diarahkan pada jenjang pendidikan berbeda dan variasi materi lain. Variabel tambahan seperti motivasi belajar, keterlibatan siswa, maupun kemampuan berpikir kritis dapat dimasukkan guna memperkaya struktur analisis.

### **SIMPULAN**

Penelitian ini berhasil menyintesis pengembangan instrumen evaluasi HOTS matematika berbasis Quizizz yang valid dan praktis dari aspek materi, konstruksi, dan bahasa. Lebih dari sekadar alat uji digital, platform ini terbukti menjadi solusi integratif yang mampu menstimulasi aspek kognitif dan afektif siswa secara simultan. Dampak nyata berupa peningkatan problem solving skill dan *self-efficacy* menunjukkan bahwa lingkungan evaluasi berbasis digital yang interaktif dapat mereformasi paradigma penilaian konvensional. Instrumen ini tidak hanya berfungsi sebagai alat rekam capaian belajar, melainkan sebagai media akselerasi kemampuan berpikir kritis yang transformatif dalam pembelajaran abad ke-21. Sebagai rekomendasi, peneliti selanjutnya harus menguji stabilitas instrumen ini pada sampel yang lebih luas dengan latar belakang akademis yang bervariasi. Selain itu, studi lanjutan perlu dilakukan untuk memantau apakah efek positif intervensi digital ini bersifat konsisten dalam jangka panjang, serta mengeksplorasi integrasi fitur Quizizz lainnya yang berpotensi memengaruhi variabel psikologis dan motivasi belajar siswa pada topik matematika yang berbeda.

### **DAFTAR PUSTAKA**

- Amalina, I., & Vidákovich, T. (2023). Cognitive and socioeconomic factors that influence the mathematical problem-solving skills of students. *Heliyon*, 9(9), 1-14. <https://doi.org/10.1016/j.heliyon.2023.e19539>.
- Ansya, Y., Alfianita, A., Syahkira, H., & S. (2024). Optimizing mathematics learning in fifth grades: the critical role of evaluation in improving student achievement and character. *Progres Pendidikan*, 5(3), 302-311. <https://doi.org/10.29303/prospek.v5i3.1120>

- Aritonang, A., Warsodirejo, P., & Tanty, H. (2025). Peningkatan motivasi dan hasil belajar matematika siswa kelas x sma negeri 2 medan melalui model problem-based learning berbantuan teknologi Quizizz. *Journal on Education*, 7(2), 10178-10185. <https://doi.org/10.31004/joe.v7i2.8023>.
- Aryana, S., Sobari, T., , W., Fauziya, D., & Mustika, R. (2025). The effectiveness of 21st century competency assessment model and hots in improving writing skills: an experimental research. *Journal of Language Teaching and Research*, 16(4), 1115-1128. <https://doi.org/10.17507/jltr.1604.07>.
- Astuti, D., Huda, C., Darminingsih, A., & Setyawati, R. (2024). Peningkatan aktivitas belajar siswa melalui model problem based learning berbantuan media Quizizz pada kelas v sdn mlatiharjo 02 semarang. *Journal on Education*, 7(1), 3441-3447. <https://doi.org/10.31004/joe.v7i1.6791>.
- Boral, K., Mandal, S., Kumbhakar, M., & Dutta, S. (2025). Analysis of higher order thinking skills of intended mathematics learning outcomes in india. *MIER Journal of Educational Studies Trends and Practices*, 146-176. <https://doi.org/10.52634/mier/2025/v15/i1/2856>.
- Dermawan, D., Wuryandan, W., Herwin, H., Ikhsan, M., Giwangsa, S., Murrone, F., Agusman, D., & Fadli, R. (2025). The Development of HOTS Assessment via Mobile Web Based Quizziz App. *TEM Journal*, 14(1). <https://doi.org/10.18421/tem141-12>.
- Fadda, D., Pellegrini, M., Vivianet, G., & Callegher, C. (2021). Effects of digital games on student motivation in mathematics: A meta-analysis in K-12. *J. Comput. Assist. Learn*, 38, 304-325. <https://doi.org/10.1111/jcal.12618>. *J. Comput. Assist. Learn*
- Handayani, N., Aw, S., Zamroni, Z., Imanita, M., Setiawan, J., & Fadli, M. (2024). Development of higher order thinking skill assessment instruments in social studies learning. *International Journal of Evaluation and Research inn Education (IJERE)*, 13(2), 923-933. <https://doi.org/10.11591/ijere.v13i2.26448>.
- Hassan, N., Rahman, M., & Sumintono, B. (2024). Enhancing integration of technology in authentic assessment for education: a structured review. *Journal of Advanced Research in Applied Sciences and Engineering Technology*, 2(2025), 58-78. <https://doi.org/10.37934/araset.54.2.5878>.
- Hernanz, V., Latorre-Coscolluela, C., Suárez, C., & Lanchares-Sancho, E. (2024). Revitalising learning in three university contexts: unleashing the power of the Quizizz app to increase *self-efficacy*, intrinsic motivation, satisfaction and performance. *Education and Information Technologies*, 29, 23917 - 23935. <https://doi.org/10.1007/s10639-024-12779-9>.
- Holenstein, M., Bruckmaier, G., & Grob, A. (2021). How do self-efficacy and self-concept impact mathematical achievement? The case of mathematical modelling. *The British Journal of Educational Psychology*, 92. <https://doi.org/10.1111/bjep.12443>.

- Huang, Y., Silitonga, L., Murti, A., & Wu, T. (2022). Learner engagement in a business simulation game: impact on higher-order thinking skills. *Journal of Educationnal Computing Research*, 61, 96 - 126. <https://doi.org/10.1177/07356331221106918>.
- Huei, L., Yunus, M., & Hashim, H. (2021). Strategy to improve english vocabulary achievement during covid-19 epidemic. Does Quizizz help?. *Journal of Education and e-Learning Research*, 8(2), 135-142. <https://doi.org/10.20448/journal.509.2021.82.135.142>
- Irianto, B., Saleh, M., Nurhaidah, N., & Taufiq, T. (2021). Exploring students' learning strategies and self-regulated learning in solving mathematical higher-order thinking problems. *European Journal of Educational Research*, 10(2), 743-756. <https://doi.org/10.12973/eu-jer.10.2.743>.
- Jailani, J., Retnawati, H., Rafi, I., Mahmudi, A., Arliani, E., Zulnaidi, H., Hamid, H., & Prayitno, H. (2023). A phenomenological study of challenges that prospective mathematics teachers face in developing mathematical problems that require higher-order thinking skills. *Eurasia Journal of Mathematics, Science and Technology Education*, 19(10), em2339. <https://doi.org/10.29333/ejmste/13631>.
- Kusumaningtyas, D., Manyunuh, M., Kurniasari, E., Awaln, A., Rahmaniati, R., & Febriyanti, A. (2023). Enhancing learning outcomes: A study on the development of higher order thinking skills based evaluation instruments for work and energy in high school physics. *Indonesian Journal on Learning and Advanced Education (IJOLAE)*, 14-31. <https://doi.org/10.23917/ijolae.v6i1.23125>.
- Liu, R., Jong, C., & Fan, M. (2024). Reciprocal relationship between *self-efficacy* and achievement in mathematics among high school students. *Large-scale Assessments in Education*, 12, 1-16. <https://doi.org/10.1186/s40536-024-00201-2>.
- Lu, Z., Chiu, M., Wang, S., Mao, W., & Lei, H. (2025). Effects of augmented reality on students' higher-order thinking: A meta-analysis. *Journal of Educational Computing Research*, 63, 728 - 747. <https://doi.org/10.1177/07356331241309644>.
- Manalu, S., & Harahap, N. (2025). Higher order thinking skills in reading questions of english textbook for 10th grade high school students. *Invention: Journal Research and Education Studies*, 6 (1), 139-154. <https://doi.org/10.51178/invention.v6i1.2422>.
- Masrifah, K., Ahsanuddin, M., Mahliatussikah, H., & Ismail, Z. (2025). Evaluating higher-order thinking skills representation in the 2022 Quadra Arabic language textbooks for Indonesian Madrasah Aliyah. *Al-Lisan: Jurnal Bahasa (e-Journal)*, 10(2), 255-269. <https://doi.org/10.30603/al.v10i2.6703>.
- Muchuwani, T., Jojo, Z., & Kariyana, I. (2025). Enhancing Mathematics Instruction through Quizizz: A Systematic Literature Review. *Available at SSRN 5585230*.
- Musliha, S., Sudana, D., & Wirza, Y. (2021). The analysis of higher order thinking skills (HOTS) in the test questions constructed by english teachers. *Proceedings of the Fifth*

*International Conference on Language, Literature, Culture, and Education (ICOLLITE 2021)*. <https://doi.org/10.2991/assehr.k.211119.095>.

- Olivares, D., Lupiáñez, J., & Segovia, I. (2020). Roles and characteristics of *problem solving* in the mathematics curriculum: A review. *International Journal of Mathematical Education in Science and Technology*, 52, 1079 - 1096. <https://doi.org/10.1080/0020739x.2020.1738579>.
- Pham, A. (2022). University students attitudes towards the application of Quizizz in learning english as a foreign language. *Int. J. Emerg. Technol. Learn*, 17, 278-290. <https://doi.org/10.3991/ijet.v17i19.32235>.
- Pitoyo, M., Sumardi, S., & Asib, A. (2020). Gamification-based assessment: The washback effect of Quizizz on students learning in higher education. *International Journal of Language Education*, 4(1), 1-10. <https://doi.org/10.26858/ijole.v4i2.8188>.
- Putri, A., & Aznam, N. (2019). The effectiveness of science learning media using focusky software on junior high school students higher order thinking skills. *Scientiae Educatia: Jurnal Pendidikan Sains*, 8(1), 12-22. <https://doi.org/10.24235/sc.educatia.v8i1.3886>.
- Putri, F., Marmoah, S., & Supianto, S. (2024). Analyzing students mathematical problem solving skills through hots-based questions at the elementary school level. *Social, Humanities, and Educational Studies (SHES): Conference Series* (Vol. 7, No. 1, pp. 85-95). <https://doi.org/10.20961/shes.v7i1.84280>.
- Rahmawatiningrum, A., Kusmayadi, T., & Fitriana, L. (2019). Student's ability in solving higher order thinking skills (HOTS) mathematics problem based on learning achievement. *Journal of Physics: Conference Series*, 1318. <https://doi.org/10.1088/1742-6596/1318/1/012090>.
- Sagitarini, N. M. D., Candiasa, I. M., & Pujawan, I. G. N. (2023). Pengaruh ketahananmalangan, regulasi diri dan efikasi diri terhadap prestasi belajar matematika. *Jurnal Penelitian Dan Evaluasi Pendidikan Indonesia*, 13(1), 27-43. <https://doi.org/10.23887/jpepi.v13i1.1874>
- Sasson, I., & Tifferet, S. (2024). Enhancing undergraduate metacognitive awareness and self-efficacy: effective instructional practices for research question formulation. *European Journal of Education*, 60(1), e12888. <https://doi.org/10.1111/ejed.12888>.
- Saleem, R., Fatima, K., & Shehbaz, I. (2025). The Role of Classroom Assessment in Enhancing Effective Learning: A Critical Analysis. *Journal of Arts and Linguistics Studies*, 3(1), 1333-1346.
- Setyaningrum, W., Pastoriko, F., Fabian, K., & Ying, C. (2024). The effect of scaffolding-based digital instructional media on higher-order thinking skills. *Journal on Mathematics Education*, 15(4), 1077-1094. <https://doi.org/10.22342/jme.v15i4.pp1077-1094>.
- Shone, E., Weldemeskel, F., & Worku, B. (2023). The role of students mathematics perception and self-efficacy toward their mathematics achievement. *Psychology in the Schools*, 61(1), 103-122. <https://doi.org/10.1002/pits.23033>.

- Sofyan, A., Nissa, I. C., Judijanto, L., Nugraha, M. S., & Santi, D. (2025). The Impact of Digital Formative Assessment on Student Motivation and Learning Outcomes in the Technological Era. *Journal of Hunan University Natural Sciences*, 52(7). <https://doi.org/10.55463/issn.1674-2974.52.7.9>.
- Ukobizaba, F., Nizeyimana, G., & Mukuka, A. (2021). Assessment strategies for enhancing students' mathematical problem-solving skills: A review of literature. *Eurasia Journal of Mathematics, Science and Technology Education*, 17. <https://doi.org/10.29333/ejmste/9728>.
- Widana, I. W. (2017). Higher order thinking skills assessment (HOTS). *JISAE: Journal of Indonesian Student Assessment and Evaluation*, 3(1), 32-44. <https://doi.org/10.21009/jisae.031.04>.
- Yang, Y., Maeda, Y., & Gentry, M. (2024). The relationship between mathematics *self-efficacy* and mathematics achievement: Multilevel analysis with NAEP 2019. *Large-scale Assessment in Education*, 12, 1-29. <https://doi.org/10.1186/s40536-024-00204-z>.
- Yuniarti, N., Rahmawati, Y., Anwar, M., Al Hakim, V., Hidayat, H., Hariyanto, D., Husna, A., & Wang, J. (2024). Augmented reality-based higher order thinking skills learning media: Enhancing learning performance through self-regulated learning, digital literacy, and critical thinking skills in vocational teacher education. *European Journal of Education*, 59(4), e12725. <https://doi.org/10.1111/ejed.12725>.
- Yunus, C., & Hua, T. (2021). Exploring a gamified learning tool in the esl classroom: The case of Quizizz. *Journal of Education and e-Learning Research*, 8(1), 103-108. <https://doi.org/10.20448/journal.509.2021.81.103.108>.
- Zainudin, Z., & Zulkiply, N. (2023). Gamification in learning: students motivation and cognitive engagement in learning business using Quizizz. *Asian Journal of University Education (AJUE)*, 19(4), 823-833. <https://doi.org/10.24191/ajue.v19i4.24928>
- Zhan, Y., Yan, Z., Wan, Z., Wang, X., Zeng, Y., Yang, M., & Yang, L. (2023). Effects of online peer assessment on higher-order thinking: A meta-analysis. *Br. K. Educ. Technol*, 54, 817-835. <https://doi.org/10.1111/bjet.13310>.
- Zhang, Z., & Crawford, J. (2023). EFL learners motivation in a gamified formative assessment: The case of Quizizz. *Education and Information Technologies*, 29, 6217 - 6239. <https://doi.org/10.1007/s10639-023-12034-7>.
- Živković, M., Pellizzoni, S., Doz, E., Cuder, A., Mammarella, I., & Passolunghi, M. (2023). Math *self-efficacy* or anxiety? The role of emotional and motivational contribution in math performance. *Social Psychology of Education*, 1-23. <https://doi.org/10.1007/s11218-023-09760-8>.