



## Pengembangan Instrumen Penilaian Teks Deskripsi Berbasis HOTS Berbantuan Aplikasi Quizizz di Kelas VII SMP Amalia Medan

Titin Mariani\*, Surya M. Hutagalung, & Abdurahman Adisaputera

Magister Pendidikan Bahasa dan Sastra Indonesia, Universitas Negeri Medan, Indonesia

### ABSTRACT

This study develops a Higher Order Thinking Skills (HOTS)-based assessment instrument for descriptive texts using Quizizz for seventh-grade students at SMP Amalia Medan. The research employs Sugiyono's (2015) R&D method, implementing seven stages: problem identification, data collection, product design, validation, revision, testing, and final revision. Validation by three experts yielded a 94% validity rate, confirming strong instrument reliability. Of 30 developed questions, 25 were valid ( $r > 0.3291$ ) with a reliability coefficient of 0.78 ( $> 0.60$ ). The difficulty level was "Moderate", and two questions with low discriminative power were removed. Most distractors were effective, with over 5% of students selecting alternative answers. Trial results indicated students' HOTS levels were categorized as "Good", demonstrating the instrument's effectiveness in evaluating higher-order thinking in descriptive text comprehension. This study highlights the potential of Quizizz as an engaging digital tool for HOTS-based assessments.

### ARTICLE HISTORY

Submitted 04 07 2024  
Revised 12 03 2025  
Accepted 17 03 2025  
Published 20 03 2025

### KEYWORDS

HOTS assessment; descriptive text; Quizizz; test reliability; digital learning evaluation.

### \*CORRESPONDANCE AUTHOR

[titinmariani85@gmail.com](mailto:titinmariani85@gmail.com)

DOI: <https://doi.org/10.30743/bahastra.v9i2.9543>

### PENDAHULUAN

*Higher Order Thinking Skills* (HOTS) merupakan salah satu kompetensi esensial dalam pendidikan abad ke-21, termasuk dalam pembelajaran bahasa Indonesia. Namun, instrumen evaluasi yang digunakan di sekolah saat ini belum sepenuhnya berbasis HOTS, terutama dalam penilaian pemahaman teks deskripsi. Sebagian besar guru masih menggunakan metode penilaian konvensional yang berfokus pada keterampilan berpikir tingkat rendah, sehingga belum sepenuhnya mendorong siswa untuk mengembangkan kemampuan berpikir kritis dan analitis. Beberapa penelitian sebelumnya telah mengembangkan instrumen evaluasi berbasis HOTS dalam berbagai mata pelajaran, tetapi masih terbatas dalam penerapannya menggunakan Quizizz sebagai alat penilaian pada teks deskripsi. Oleh karena itu, penelitian ini bertujuan untuk mengembangkan instrumen penilaian HOTS berbantuan Quizizz, sehingga dapat membantu guru dalam melakukan evaluasi yang lebih efektif dan interaktif.

Belajar merupakan proses interaksi individu dengan lingkungannya yang dapat berlangsung di berbagai situasi. Menurut Bakar (2013), peran guru tidak hanya sebatas menyampaikan informasi, tetapi juga membimbing proses pembelajaran guna meningkatkan kemampuan berpikir siswa (Bakar, 2013). Dalam konteks ini, Suratman et al. (2020) menekankan bahwa guru memiliki peran penting dalam meningkatkan kualitas pendidikan (Suratman et al., 2020). Dewi (2015) menambahkan bahwa beberapa keterampilan utama yang diperlukan di abad ke-21 meliputi pemecahan masalah, berpikir kritis, pengambilan keputusan, komunikasi, kerja sama, literasi digital, serta tanggung jawab personal dan sosial (Dewi, 2015). Senada dengan itu, Retnawati et al. (2018) menyatakan bahwa kemampuan berpikir kritis dalam menyelesaikan masalah merupakan aspek krusial dalam pendidikan modern (Retnawati et al., 2018).

Pada era digital saat ini, institusi pendidikan menghadapi tantangan untuk menghasilkan siswa yang memiliki keterampilan berpikir kritis, inovatif, kolaboratif, dan komunikatif, sebagaimana dikonsepsikan dalam HOTS (Uswatun & Herina, 2019). HOTS bukan sekadar kemampuan menghafal atau menyampaikan kembali informasi, tetapi menuntun siswa untuk mengolah, menganalisis, serta menarik kesimpulan dari data yang diperoleh (Widana, 2017). Rajendran (dalam Nugroho, 2018) menjelaskan bahwa HOTS mendorong siswa untuk mengevaluasi informasi secara kritis dan menghasilkan kesimpulan yang logis. Selain itu, menurut Underbakke et al. (dalam Sani, 2019), HOTS mencakup pemikiran strategis, yakni kemampuan menggunakan informasi untuk menyelesaikan masalah, mengevaluasi argumen, serta merancang solusi. Kemendikbud (2017) merinci karakteristik soal HOTS sebagai berikut: (1) Mengukur kemampuan berpikir tingkat tinggi, mencakup pemecahan masalah, berpikir kritis, berpikir kreatif, penalaran, serta pengambilan keputusan; (2) Berbasis permasalahan kontekstual, yang memungkinkan siswa menerapkan konsep

dalam situasi nyata; serta (3) Menggunakan bentuk soal beragam, sebagaimana yang digunakan dalam PISA untuk mengukur pemahaman yang lebih mendalam.

Dalam kurikulum Sekolah Menengah Pertama (SMP), materi teks deskripsi bertujuan untuk mengajarkan siswa mengidentifikasi informasi serta memahami isi teks berdasarkan objek tertentu, seperti tempat wisata, bangunan bersejarah, atau suasana sebuah acara. Menurut Alwi et al. (2015), identifikasi merupakan proses menentukan identitas suatu objek, sementara deskripsi merupakan pemaparan secara terperinci. Saddhono & Slamet (2015) menegaskan bahwa deskripsi memungkinkan pembaca membayangkan secara imajinatif objek yang digambarkan oleh penulis. Hal ini sejalan dengan pandangan Keraf (2017), yang menyatakan bahwa deskripsi merupakan teknik penulisan yang berfokus pada perincian objek untuk membangun pengalaman sensorik bagi pembaca. Dengan demikian, pemahaman terhadap teks deskripsi menuntut siswa tidak hanya memahami isi teks secara literal tetapi juga menafsirkan makna yang lebih dalam.

Dalam konteks pembelajaran, guru memiliki peran dalam mengevaluasi sejauh mana siswa memahami teks deskripsi. Evaluasi merupakan proses pemberian nilai terhadap data hasil pengukuran guna mengetahui ketercapaian tujuan pembelajaran (Arikunto, 2015). Pengukuran yang akurat memerlukan instrumen yang valid dan reliabel, sehingga mampu menghasilkan data yang representatif. Namun, di sekolah-sekolah, banyak guru masih menggunakan instrumen penilaian secara konvensional tanpa memperhatikan validitas dan reliabilitasnya. Guru bahasa Indonesia cenderung menggunakan soal berbasis cetak tanpa memanfaatkan teknologi digital dalam proses evaluasi. Minimnya pemanfaatan aplikasi digital dalam penilaian kemungkinan disebabkan oleh kurangnya pengalaman guru dalam menggunakan alat bantu elektronik. Oleh karena itu, penelitian ini mendorong penggunaan Quizizz sebagai platform evaluasi berbasis digital yang memungkinkan guru merancang instrumen penilaian HOTS secara lebih efektif dan menarik.

Penelitian ini berfokus pada pengembangan instrumen tes berbantuan Quizizz dengan tingkat kesulitan HOTS untuk materi teks deskripsi pada kelas VII di SMP Amalia Medan. Implementasi Quizizz sebagai alat evaluasi merupakan inovasi dalam penilaian hasil belajar siswa pada level HOTS, serta sebagai langkah strategis dalam meningkatkan kualitas soal berbasis HOTS. Hasil penelitian ini diharapkan dapat menjadi referensi bagi guru dalam merancang instrumen tes HOTS yang lebih baik di masa mendatang. Selain itu, penelitian ini juga dapat menjadi acuan dalam pengembangan instrumen serupa untuk mata pelajaran lain di tingkat SMP.

Sebagai sebuah platform berbasis web, Quizizz memungkinkan guru merancang soal dengan tingkat kesulitan yang dapat disesuaikan dengan kompetensi siswa. Menurut Jauhariya (2021), Quizizz lebih efektif dibandingkan penilaian konvensional karena memiliki tampilan yang menarik, fitur waktu yang terkontrol, serta fleksibilitas dalam pelaksanaan evaluasi di berbagai kondisi. Hal ini sejalan dengan penelitian yang dilakukan oleh Shifatun Nisa dan Triesninda Pahlefi (2021), yang mengembangkan instrumen evaluasi berbasis HOTS dengan Quizizz pada mata pelajaran kearsipan di SMK. Selain itu, penelitian Hamidah dan Wulandari (2021) juga menunjukkan bahwa Quizizz efektif dalam meningkatkan keterampilan berpikir tingkat tinggi pada mata pelajaran Otomatisasi Tata Kelola Humas dan Keprotokolan di SMK Ipems Surabaya. Studi lain oleh Dhini et al. (2019) mengembangkan instrumen penilaian berbasis HOTS pada teks deskripsi di SMP, namun tanpa integrasi aplikasi digital. Perbedaan utama penelitian ini dengan studi sebelumnya terletak pada penerapan Quizizz dalam pengembangan instrumen evaluasi HOTS pada teks deskripsi di SMP Amalia Medan.

Agar Quizizz semakin kompetitif sebagai alat evaluasi pembelajaran yang mendukung peningkatan kualitas pendidikan di Indonesia, pengembangan instrumen tes berbantuan aplikasi ini perlu dilakukan secara berkelanjutan, khususnya dalam konteks pembelajaran teks deskripsi di SMP Amalia Medan. Dengan inovasi ini, diharapkan guru dapat lebih terbantu dalam menilai keterampilan berpikir tingkat tinggi siswa secara lebih akurat dan efisien.

## **METODE**

Metode penelitian yang digunakan dalam penelitian ini adalah penelitian dan pengembangan atau (R&D) (Sugiyono (2015). Menurutnya, penelitian pengembangan merupakan metode yang digunakan untuk menghasilkan suatu produk, di mana efektivitas dan kelayakan produk tersebut harus diuji terlebih dahulu sebelum diterapkan secara luas. Model pengembangan dalam penelitian ini terdiri dari 10 tahap, namun penelitian ini hanya menerapkan 7 tahap utama, yaitu analisis potensi dan masalah, pengumpulan data, desain produk, validasi desain, revisi desain, uji coba produk, dan revisi produk tahap 1. Pembatasan ini dilakukan karena penelitian hanya dilaksanakan di SMP Amalia

Medan dan dikembangkan khusus pada Tujuan Pembelajaran (TP) 3.1 dan TP 4.1, sehingga produk yang dihasilkan tidak didistribusikan secara luas. Penelitian ini bertujuan untuk mengembangkan instrumen penilaian berbasis HOTS (*Higher Order Thinking Skills*) dalam pembelajaran teks deskripsi untuk siswa kelas VII SMP, serta mengevaluasi kesesuaian dan kelayakan instrumen yang dikembangkan.

Subjek penelitian ini adalah 34 siswa kelas VII SMP Amalia Medan. Penelitian ini menggunakan data kualitatif dan kuantitatif, yang dianalisis dengan melibatkan ahli materi serta praktisi atau guru Bahasa Indonesia. Sebanyak 30 soal yang telah disusun menjadi sumber analisis data kuantitatif. Tahap pertama dalam analisis butir soal adalah uji validitas, di mana suatu butir soal dikategorikan valid apabila  $r$  tabel (0,3291) lebih kecil dari  $r$  hitung yang terdapat dalam tabel statistik butir soal (Emy, 2019; Maulidia & Pahlevi, 2020; Najihah et al., 2018).

**Tabel 1. Kriteria interpretasi validitas item**

Angka	Kriteria Interpretasi
0,80 – 1,00	Sangat tinggi
0,60 – 0,80	Tinggi
0,40 – 0,60	Cukup
0,20 – 0,40	Rendah
0,00 – 0-20	Sangat rendah

Sumber: (Arikunto, 2018)

Analisis soal kedua adalah reliabilitas; jika suatu pemeriksaan dapat menghasilkan hasil yang konsisten, maka pemeriksaan tersebut dapat dianggap reliable (Arikunto, 2018).

**Tabel 2. Kriteria Interpretasi Reliabilitas**

Angka	Kriteria Interpretasi
0,80 – 1,00	Sangat tinggi
0,60 – 0,79	Tinggi
0,40 – 0,59	Cukup
0,20 – 0,39	Rendah
0,00 – 0,19	Sangat rendah

Sumber: (Sunarti & Selly, 2014)

Analisis daya pembeda soal juga digunakan untuk membedakan siswa yang memiliki kemampuan rendah dari siswa yang memiliki kemampuan tinggi (Arifin, 2009; Najihah et al., 2018).

**Tabel 3. Kriteria Daya Pembeda Soal**

Angka	Kriteria Interpretasi
0,40 – ke atas	Sangat baik, soal diterima
0,30 – 0,39	Baik, diterima dengan revisi
0,20 – 0,29	Cukup, dengan perbaikan
0,19 – ke bawah	Kurang, soal dibuang

Sumber: (Basuki & Hariyanto, 2016)

Analisis butir soal selanjutnya adalah, tingkat kesukaran soal. Soal dengan kriteria tidak terlalu mudah dan tidak terlalu sulit merupakan ciri-ciri soal yang baik (Najihah et al., 2018, p. 23).

**Tabel 4. Kriteria Tingkat Kesukaran**

Angka	Kriteria Interpretasi
0,90 – ke atas	Sangat mudah, soal ditolak
0,71 – 0,89	Baik, diterima dengan revisi
0,31 – 0,70	Cukup, dengan perbaikan
0,21 – 30	Kurang, soal dibuang
0,20 – ke bawah	Sangat sukar, soal ditolak

Sumber: (Basuki & Hariyanto, 2016)

Analisis terakhir adalah analisis butir soal pengecoh. Pengecoh dapat dikatakan efektif apabila dipilih oleh minimal 5% dari jumlah keseluruhan siswa (Basuki & Hariyanto, 2016).

**Tabel 5. Kriteria Soal Pengecoh**

Angka	Kriteria Interpretasi
76% - 125%	Sangat baik
51% - 75% atau 126% - 150%	Baik
26% - 50% atau 151% - 175%	Cukup baik
0% - 25% atau 176% - 200%	Jelek
<200%	Sangat jelek

Sumber: (Arifin, 2009)

Kemudian Analisis data kuantitatif selanjutnya yaitu kemampuan tingkat berpikir kritis siswa, hal ini dapat diketahui setelah siswa selesai mengerjakan butir soal berbasis HOTS.

**Tabel 6. Kriteria Kemampuan Berpikir Tingkat Tinggi**

Angka	Kriteria Interpretasi
81-100	Sangat baik
61-80	Baik
41-60	Cukup
21-40	Kurang
<20	Sangat kurang

Sumber: (Purbaningrum, 2017).

## HASIL DAN PEMBAHASAN

### Pengembangan Instrumen Penilaian Berbasis HOTS pada Teks Deskripsi di Kelas VII SMP Swasta Amalia

Penelitian ini menggunakan metode penelitian dan pengembangan (Research and Development/R&D) yang mengacu pada model yang dikemukakan oleh Sugiyono (2015). Model pengembangan ini terdiri dari sepuluh tahapan, namun penelitian ini hanya menerapkan tujuh tahapan, yaitu (1) analisis potensi dan masalah, (2) pengumpulan data, (3) desain produk, (4) validasi desain, (5) revisi desain, (6) uji coba produk, dan (7) revisi produk tahap 1. Penelitian ini dilakukan di SMP Swasta Amalia Medan dan difokuskan pada Tujuan Pembelajaran (TP) 3.1 dan 4.1, sehingga instrumen yang dikembangkan tidak didistribusikan secara luas. Tujuan utama penelitian ini adalah mengembangkan instrumen penilaian berbasis HOTS (Higher Order Thinking Skills) pada teks deskripsi di kelas VII SMP, serta mengevaluasi kelayakan dan kesesuaian instrumen tersebut.

Pada tahap analisis potensi dan masalah, peneliti melakukan observasi dan wawancara dengan guru Bahasa Indonesia. Hasil observasi menunjukkan bahwa sekolah ini masih menggunakan instrumen penilaian konvensional, yang lebih menekankan pada kemampuan memahami dan mengingat dibandingkan dengan analisis dan evaluasi. Akibatnya, siswa kurang mampu menyelesaikan soal berbasis HOTS yang memerlukan keterampilan berpikir kritis. Selain itu, guru masih mengandalkan pengajaran secara manual tanpa memanfaatkan media digital, karena fasilitas pembelajaran berbasis teknologi di sekolah masih terbatas.

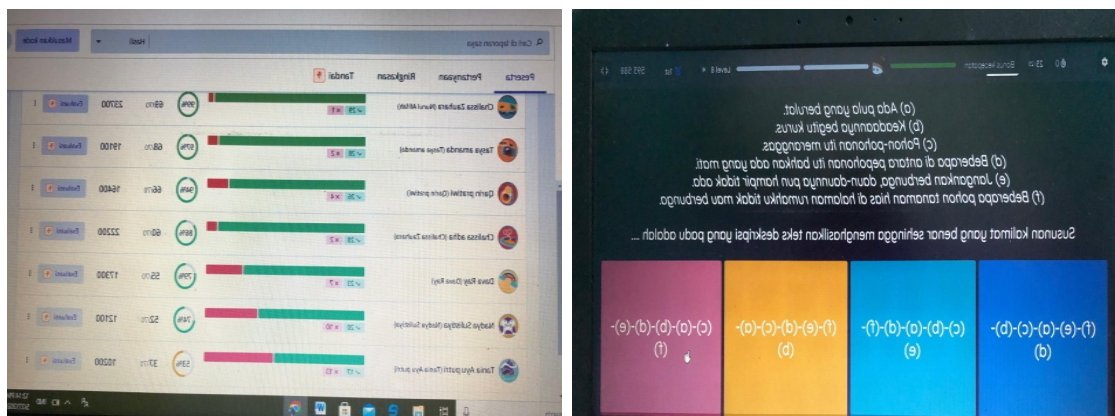
Pada tahap pengumpulan data, peneliti melakukan wawancara tidak terstruktur dengan guru Bahasa Indonesia. Hasil wawancara menunjukkan bahwa guru menggunakan buku pegangan, LKS, dan kumpulan soal latihan sebagai sumber utama dalam menguji pemahaman siswa terhadap materi teks deskripsi. Akibatnya, siswa belum terbiasa menghadapi soal yang berbasis analisis dan evaluasi. Selain itu, proses penilaian masih dilakukan secara konvensional melalui tes berbasis kertas, yang kurang efektif dalam mendorong siswa untuk berpikir secara mendalam dan sistematis.

Berdasarkan temuan tersebut, peneliti merancang desain produk berupa instrumen penilaian berbasis HOTS dalam bentuk soal pilihan ganda (*multiple choice*). Pemilihan format ini sejalan dengan penelitian yang dilakukan oleh Widiyawati et al. (2019), yang menyatakan bahwa soal pilihan ganda lebih objektif dan mudah dalam pengolahan data.

Sebanyak 30 butir soal dikembangkan, dan setelah dilakukan analisis lebih lanjut, 25 soal terbaik dipilih untuk digunakan dalam uji coba.

Pada tahap validasi desain, tiga ahli validator menilai kualitas instrumen penilaian. Validator terdiri dari guru Bahasa Indonesia di SMP Swasta Amalia sebagai validator materi dan bahasa, serta dosen dari Universitas Negeri Medan sebagai validator evaluasi. Hasil validasi menunjukkan bahwa instrumen memperoleh nilai 98%, yang termasuk dalam kategori "sangat kuat", sehingga dianggap layak untuk diuji coba lebih lanjut.

Selanjutnya, pada tahap revisi desain, instrumen penilaian HOTS diperbaiki berdasarkan masukan dari validator. Perbaikan meliputi penyesuaian ejaan, ketepatan tanda baca, penggunaan huruf kapital, serta pemilihan diksi yang lebih tepat. Setelah perbaikan dilakukan, instrumen kemudian dikonversikan ke dalam aplikasi Quizizz menggunakan menu "Create", sehingga dapat diakses secara digital oleh siswa.



Gambar 1. Bentuk Instrumen Penilaian di Aplikasi Quizizz

Pada tahap uji coba produk, instrumen diuji pada 34 siswa kelas VII SMP Swasta Amalia. Aplikasi Quizizz digunakan sebagai media evaluasi digital yang dapat diakses melalui komputer, laptop, atau ponsel siswa. Hasil pekerjaan siswa dapat dianalisis melalui fitur "Laporan", yang memungkinkan guru dan peneliti untuk mengunduh hasil dalam bentuk file Excel.

Berdasarkan hasil uji coba, ditemukan bahwa 25 soal dinyatakan valid, sementara 5 soal tidak valid. Soal yang memenuhi kriteria validitas mencakup nomor 1, 2, 3, 4, 6, 7, 8, 9, 11, 12, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 25, 26, 27, 29, dan 30, sedangkan soal yang tidak valid adalah nomor 5, 10, 13, 24, dan 28. Hal ini sesuai dengan hasil penelitian Arifin & Retnawati (2017), Desilva et al. (2020), Najihah et al. (2018), serta Widiyawati et al. (2019), yang menyatakan bahwa pengembangan instrumen HOTS berbasis *multiple choice test* memerlukan seleksi ketat terhadap soal-soal yang valid dan layak digunakan.

**Kelayakan Instrumen HOTS Berbantuan Aplikasi Quizizz pada Tujuan Pembelajaran (T.P.) 3.1 dan 4.1**

Berdasarkan analisis data kuantitatif dan kualitatif, penelitian ini menghasilkan evaluasi terhadap kelayakan instrumen penilaian berbasis *Higher Order Thinking Skills* (HOTS) untuk teks deskripsi pada siswa kelas VII SMP Swasta Amalia Medan. Produk akhir penelitian berupa instrumen penilaian berbasis HOTS yang diintegrasikan dalam aplikasi Quizizz. Setelah dilakukan perbaikan terhadap ejaan, ketepatan penggunaan tanda baca, huruf kapital, serta pemilihan diksi yang tepat, instrumen penilaian ini dinyatakan layak digunakan berdasarkan hasil validasi dari ahli materi dan praktisi (guru Bahasa Indonesia). Hasil validasi instrumen oleh tiga validator menunjukkan bahwa nilai rata-rata yang diberikan adalah 94%, yang masuk dalam kategori "sangat kuat", sebagaimana disajikan dalam tabel berikut:

Tabel 7. Hasil Akhir Penilaian Validator

Validator	Penilaian Validator	Kriteria
Ahli Materi	87%	Sangat kuat
Praktisi (Ahli Bahasa)	99%	Sangat kuat
Ahli Evaluasi	97%	Sangat kuat
<b>Rata-Rata</b>	<b>94%</b>	<b>Sangat kuat</b>

Sumber: Data diolah oleh peneliti

Kelayakan instrumen HOTS berbantuan aplikasi Quizizz pada tujuan pembelajaran (T.P.) 3.1 dan 4.1 dievaluasi berdasarkan perhitungan data kuantitatif dan kualitatif. Produk akhir penelitian ini berupa instrumen penilaian berbasis HOTS dalam teks deskripsi untuk siswa kelas VII SMP Swasta Amalia Medan yang diintegrasikan dengan aplikasi Quizizz. Setelah dilakukan perbaikan terhadap ejaan, ketepatan penggunaan tanda baca, huruf kapital, dan pemilihan diksi, instrumen ini dinyatakan layak digunakan berdasarkan hasil validasi dari ahli materi dan praktisi (guru bahasa Indonesia). Hasil validasi instrumen oleh tiga validator menunjukkan bahwa nilai rata-rata yang diberikan adalah 94%, yang masuk dalam kategori "sangat kuat". Validasi dilakukan oleh tiga ahli dengan rincian: ahli materi memberikan nilai 87%, ahli bahasa (praktisi) memberikan nilai 99%, dan ahli evaluasi memberikan nilai 97%. Dengan demikian, instrumen ini dianggap sangat layak untuk digunakan sebagai alat evaluasi dalam pembelajaran berbasis HOTS.

Hasil analisis butir soal mencakup validitas, reliabilitas, tingkat kesulitan, daya pembeda, serta efektivitas pengecoh (distraktor). Validitas butir soal diuji menggunakan korelasi point biserial. Menurut Budiman dan Jailani (2014), soal pilihan ganda dianggap valid jika  $r$  hitung lebih besar dari  $r$  tabel (0,3291). Hasil analisis menunjukkan bahwa 25 soal memiliki nilai lebih dari 0,3291 sehingga dikategorikan valid, sementara 5 soal tidak valid, yakni soal nomor 5 ( $r = 0,232 < 0,3291$ ), soal nomor 10 ( $r = 0,275 < 0,3291$ ), soal nomor 13 ( $r = 0,213 < 0,3291$ ), soal nomor 24 ( $r = 0,134 < 0,3291$ ), dan soal nomor 28 ( $r = 0,089 < 0,3291$ ). Oleh karena itu, lima soal ini dihapus atau direvisi. Reliabilitas soal diuji menggunakan koefisien Alpha Cronbach. Menurut Akhsan et al. (2020), soal dianggap reliabel jika skornya di atas 0,60. Hasil uji reliabilitas menunjukkan nilai 0,78, yang masuk dalam kategori "tinggi", sehingga instrumen ini dapat diandalkan sebagai alat evaluasi HOTS.

Tingkat kesulitan soal dianalisis menggunakan Tabel Prop Correct. Wantoro et al. (2019) menyatakan bahwa soal dikatakan semakin sulit jika indeks tingkat kesukaran lebih kecil. Berdasarkan hasil uji coba, diperoleh kategori tingkat kesulitan soal sebagai berikut: kategori "Sulit" terdapat pada soal nomor 11 (0,26) dan soal nomor 16 (0,27); kategori "Mudah" terdapat pada soal nomor 3, 6, dan 28 (nilai antara 0,73 dan 0,91); sementara kategori "Sedang" mencakup 25 soal dengan nilai antara 0,31 dan 0,70. Mayoritas soal berada dalam kategori "Sedang", sehingga dapat digunakan untuk mengukur pemahaman siswa secara proporsional. Selanjutnya, daya pembeda soal diuji menggunakan Tabel Biserial Point. Wardany et al. (2017) menyatakan bahwa daya pembeda bertujuan untuk mengevaluasi perbedaan antara siswa upper group dan lower group. Dari hasil uji coba, ditemukan bahwa 25 soal memenuhi kriteria "Sangat Baik", satu soal memenuhi kriteria "Baik" (Nomor 10,  $r = 0,378$ ), dua soal memenuhi kriteria "Cukup" (Nomor 5,  $r = 0,281$  dan Nomor 13,  $r = 0,211$ ), serta dua soal memenuhi kriteria "Kurang" (Nomor 24,  $r = 0,057$  dan Nomor 28,  $r = 0,090$ ). Dua soal dengan daya pembeda "Kurang" dihapus karena tidak memenuhi standar validitas.

Efektivitas pengecoh (distraktor) diuji menggunakan Tabel Prop Endorsing. Menurut Arifin dan Retnawati (2017), pengecoh dikatakan efektif jika dipilih oleh minimal 5% siswa. Hasil analisis menunjukkan bahwa sebagian besar pengecoh berfungsi dengan baik karena lebih dari 5% siswa memilih opsi pengecoh. Namun, terdapat dua soal yang distraktornya tidak berfungsi optimal, yaitu soal nomor 10, di mana opsi B dan D tidak dipilih oleh 5% dari total siswa, dan soal nomor 13, di mana opsi A dan C tidak dipilih oleh 5% dari total siswa. Karena kedua soal ini tidak memenuhi kriteria validitas pengecoh, maka dihapus dari instrumen akhir. Dengan demikian, berdasarkan hasil validasi dan uji kelayakan, instrumen penilaian berbasis HOTS dalam teks deskripsi kelas VII SMP berbantuan aplikasi Quizizz telah dinyatakan valid, reliabel, dan layak digunakan. Hasil penelitian ini diharapkan dapat meningkatkan efektivitas evaluasi pembelajaran serta mendorong pengembangan instrumen HOTS berbasis digital di masa depan.

### Analisis Kemampuan Berpikir Tingkat Tinggi Siswa Dilihat dari Hasil Pengerjaan Soal Berbasis HOTS

Analisis kemampuan berpikir tingkat tinggi (*Higher Order Thinking Skills/HOTS*) siswa dapat diukur setelah mereka menyelesaikan soal yang telah disusun berdasarkan prinsip HOTS. Hasil pengujian terhadap siswa disajikan dalam tabel berikut, yang menunjukkan distribusi nilai berdasarkan kategori penilaian:

**Table 8. Perolehan Nilai Siswa**

Jumah Siswa	Nilai siswa	Kriteria
5	82,8 – 89,3	Sangat Baik
14	64,7 – 72,8	Baik
11	45,5 – 56,4	Cukup
3	24,7 - 40	Kurang
1	21,50	Sangat Kurang

Sumber: Data diolah oleh peneliti

Berdasarkan hasil analisis, siswa yang masuk dalam kategori “kurang” dan “sangat kurang” menunjukkan beberapa faktor penyebab utama, seperti rendahnya pemahaman terhadap materi, kurangnya kemampuan penalaran, serta lemahnya keterampilan pemecahan masalah. Hal ini sejalan dengan temuan Budiman dan Jailani (2014) serta Widiyawati et al. (2019), yang menyatakan bahwa kesulitan siswa dalam berpikir kritis sering kali disebabkan oleh kurangnya latihan dalam mengerjakan soal-soal dengan tingkat kesulitan tinggi. Oleh karena itu, diperlukan peran aktif guru dalam memberikan model pembelajaran yang sesuai serta menyusun instrumen evaluasi berbasis HOTS untuk melatih kemampuan berpikir tingkat tinggi siswa (Arifin & Retnawati, 2017). Pengembangan soal berbasis HOTS yang terintegrasi dalam sistem pembelajaran akan membantu siswa mengembangkan keterampilan analitis, evaluatif, dan kreatif dalam menyelesaikan permasalahan akademik maupun kehidupan nyata.

## SIMPULAN

Berdasarkan hasil penelitian, pengembangan instrumen penilaian berbasis HOTS menggunakan aplikasi Quizizz telah terbukti valid dan layak digunakan dalam mengukur kemampuan berpikir tingkat tinggi siswa pada materi Teks Deskripsi kelas VII SMP Swasta Amalia Medan. Dari 30 soal yang dikembangkan dalam ranah C4, C5, dan C6, sebanyak 25 soal dinyatakan valid berdasarkan hasil analisis validitas dengan  $r$  hitung lebih besar dari  $r$  tabel (0,3291) dan reliabilitas sebesar  $0,78 > 0,60$ . Hasil validasi dari tiga ahli menunjukkan tingkat validasi sebesar 94%, yang masuk dalam kategori "sangat kuat." Selain itu, tingkat kesulitan soal berada dalam kategori "sedang," dengan dua soal yang memiliki daya pembeda "kurang" sehingga perlu dibuang. Hasil uji coba juga menunjukkan bahwa pengecoh soal berfungsi dengan baik, karena lebih dari 5% siswa memilih opsi jawaban yang disediakan. Secara keseluruhan, instrumen penilaian ini menunjukkan bahwa siswa memiliki kemampuan berpikir tingkat tinggi dengan kriteria "Baik."

Untuk penelitian selanjutnya, disarankan agar pengembangan instrumen penilaian berbasis HOTS dilakukan dengan menerapkan model penelitian pengembangan yang lebih komprehensif, seperti model Sugiyono yang terdiri dari sepuluh tahapan lengkap. Peneliti selanjutnya juga perlu memastikan bahwa setiap instrumen yang dikembangkan telah melalui uji kelayakan secara menyeluruh guna memperoleh kualitas soal yang optimal. Selain itu, penggunaan teknologi digital dalam pembelajaran perlu terus diintegrasikan, dengan memanfaatkan aplikasi tambahan selain Quizizz yang dapat mendukung efektivitas evaluasi pembelajaran dan meningkatkan motivasi belajar siswa.

## REFERENSI

- Abu, B. N. (2013). *Kurikulum ke arah penghasilan kemahiran berfikiran kritis, kreatif dan inovatif*.
- Alwi, H., Dardjowidjojo, S., Lapoliwa, H., & Moeliono, A. M. (2015). *Kamus besar bahasa Indonesia* (Edisi ke-4). Jakarta: PT Gramedia Pustaka Utama.
- Arifin, Z. (2009). *Evaluasi pembelajaran: Prinsip, teknik, dan prosedur*. Bandung: PT Remaja Rosdakarya.
- Arifin, Z., & Retnawati, H. (2017). Pengembangan instrumen pengukur higher order thinking skills matematika siswa SMA kelas X. *Jurnal Pendidikan Matematika*, 12(1), 98–108. <https://doi.org/10.21831/pg.v12i1.14058>
- Arikunto, S. (2015). *Dasar-dasar evaluasi pendidikan* (Edisi ke-5). Jakarta: Bumi Aksara.
- Arikunto, S. (2018). *Dasar-dasar evaluasi pendidikan*. Jakarta: Bumi Aksara.
- Basuki, I., & Hariyanto. (2016). *Asesmen pembelajaran*. Bandung: Remaja Rosdakarya.
- Budiman, A., & Jailani. (2014). Pengembangan instrumen asesmen higher order thinking skill (HOTS) pada mata pelajaran matematika SMP kelas VIII semester 1. *Jurnal Riset Pendidikan Matematika*, 1(2), 139–151. <https://journal.uny.ac.id/index.php/jrpm/article/view/2671>
- Desilva, D., Sakti, I., & Medriati, R. (2020). Pengembangan instrumen penilaian hasil belajar fisika berorientasi HOTS (higher order thinking skills) pada materi elastisitas dan hukum Hooke. *Jurnal Kumparan Fisika*, 3(1), 41–50. <https://doi.org/10.33369/jkf.3.1.41-50>
- Dewi, F. (2015). Proyek buku digital: Upaya peningkatan keterampilan abad 21 calon guru sekolah dasar melalui model pembelajaran berbasis proyek. *Jurnal UPI Edu*, 9(2). <https://doi.org/10.17509/md.v9i2.3248>
- Dhini, M. S., Sunarti, I., & Riadi, B. (2019). Pengembangan instrumen penilaian berbasis HOTS (higher order thinking skill) teks deskripsi SMP kelas VII. *Jurnal Kata (Bahasa, Sastra, dan Pembelajarannya)*, 1(1), 1–6.
- Emy, T. (2019). Berpikir HOTS pada metode pembelajaran problem based learning IPS. *Jurnal Penelitian dan Pendidikan IPS*, 13(2), 1–6. <http://ejournal.unikama.ac.id/index.php/JPPi>
- Hamidah, M., & Wulandari, S. (2021). Pengembangan instrumen penilaian berbasis HOTS menggunakan aplikasi "Quizizz". *Efisiensi: Kajian Ilmu Administrasi*, 18(1), 105–124. <https://doi.org/10.21831/efisiensi.v18i1.36997>

- Kementerian Pendidikan dan Kebudayaan. (2017). *Modul penyusunan soal higher order thinking skill (HOTS)*. Jakarta: Direktorat Pembinaan SMA Ditjen Pendidikan Dasar dan Menengah.
- Keraf, G. (2018). *Eksposisi dan deskripsi*. Ende: Nusa Indah.
- Maulidia, F., & Pahlevi, T. (2020). Pengembangan instrumen penilaian tes soal pilihan ganda berbasis HOTS pada mata pelajaran administrasi umum jurusan OTKP SMK Negeri 1 Lamongan. *Jurnal Pendidikan Administrasi Perkantoran*, 8(1), 136–145. <https://journal.unesa.ac.id/index.php/jpap/article/view/8114>
- Najihah, A. R., Serevina, V., & Sari, D. K. (2018). The development of high order thinking skills (HOTS) assessment instrument for temperature and heat learning. *Jurnal Penelitian dan Pengembangan Pendidikan Fisika*, 4(1), 19–26. <https://doi.org/10.21009/1.04103>
- Nisa, S., & Pahlevi, T. (2021). Pengembangan instrumen penilaian HOTS berbantuan Quizizz pada mata pelajaran kearsipan SMK. *Edukatif: Jurnal Ilmu Pendidikan*, 3(5), 2146–2159. <https://doi.org/10.31004/edukatif.v3i5.756>
- Nugroho, A. (2018). *HOTS kemampuan berpikir tingkat tinggi: Konsep, pembelajaran, penilaian dan soal-soal*. Grasindo.
- Purbaningrum, K. A. (2017). Kemampuan berpikir tingkat tinggi siswa SMP dalam pemecahan masalah matematika ditinjau dari gaya belajar. *Jurnal Penelitian dan Pembelajaran Matematika*, 10(2), 40–49. <https://doi.org/10.30870/jppm.v10i2.2029>
- Putri, A. (2022). *Pengembangan e-instrument test berbantuan Quizizz untuk mengukur HOTS pada materi redoks di kelas X SMA* (Skripsi, Universitas Jambi).
- Putri, N. A., & Jauhariyah, M. N. R. (2021). Penilaian higher order thinking skills (HOTS) menggunakan Quizizz pada materi usaha dan energi. *Kappa Journal*, 5(1), 88–101.
- Retnawati, H., Djidu, H., Kartianom, K., Apino, E., & Anazifa, R. D. (2018). Teachers' knowledge about higher order thinking skills and its learning strategy. *Problems of Education in the 21st Century*, 76(2), 215–230. <https://doi.org/10.33225/pec/18.76.215>
- Saddhono, K., & Slamet, S. (2015). *Pembelajaran keterampilan berbahasa Indonesia: Teori dan aplikasi* (Edisi ke-2). Yogyakarta: Graha Ilmu.
- Sani, R. A. (2019). *Pembelajaran berbasis HOTS (higher order thinking skill)*. Tangerang: Tira Smart.
- Sugiyono. (2015). *Metode penelitian kuantitatif, kualitatif, dan R&D*. Bandung: Alfabeta.
- Suratman, B., Wulandari, S. S., Nugraha, J., & Narmaditya, B. S. (2020). Does teacher certification promote work motivation and teacher performance? A lesson from Indonesia. *International Journal of Educational Development*, 77, 102–108.
- Widana, I. W. (2017). *Modul penyusunan soal higher order thinking skill (HOTS)*. Direktorat Jenderal Pendidikan Dasar dan Menengah. [http://repo.ikipgribali.ac.id/id/eprint/651/1/MODUL%20PENYUSUNAN%20SOAL%20HOTS\\_Dit%20PSMA%202017.pdf](http://repo.ikipgribali.ac.id/id/eprint/651/1/MODUL%20PENYUSUNAN%20SOAL%20HOTS_Dit%20PSMA%202017.pdf)
- Widiyawati, Y., Nurwahidah, I., & Sari, D. S. (2019). Pengembangan instrumen integrated science test tipe pilihan ganda beralasan untuk mengukur HOTS peserta didik. *Jurnal Saintifika*, 21(2), 1–14. <https://doi.org/10.19184/jst.v21i2.13394>