Biocaster : Jurnal Kajian Biologi
E-ISSN 2808-277X; P-ISSN 2808-3598
Volume 6, Issue 4, October 2026; Page, 2461-2480
Email: biocasterjournal@gmail.com

PENGEMBANGAN INSTRUMEN TES HOTS BERBASIS *QUIZIZZ* PADA MATERI INTERAKSI MAKHLUK HIDUP DENGAN LINGKUNGANNYA UNTUK SISWA KELAS VII

Ayuni[1]*, Vica Dian Aprelia Resti[2], & Septi Kurniasih[3]
[1,2,&3]Program Studi Pendidikan IPA, Fakultas Keguruan dan Ilmu Pendidikan, Universitas Sultan Ageng Tirtayasa, Jalan Ciwaru Raya Nomor 25, Serang, Banten 42117, Indonesia
**Email: ayuni7608@gmail.com*
Submit: 13-07-2026; Accepted: 23-07-2026; Published: 30-10-2026

ABSTRAK: Penelitian ini bertujuan untuk menghasilkan instrumen tes HOTS, mengetahui kevalidan dari pengembangan instrumen tes *higher order thinking skills* berbasis *quizizz* materi interaksi makhluk hidup dengan lingkungannya untuk siswa kelas VII. Penelitian ini menggunakan puisi sebagai media penyaji soal. Penelitian ini menggunakan metode penelitian dan pengembangan model Sugiyono yang dimodifikasi hingga tahap revisi produk. Penelitian ini divalidasi oleh validator yaitu ahli materi, ahli evaluasi, ahli media, dan guru IPA menggunakan angket. Hasil dari penelitian ini menunjukkan bahwa instrumen yang dikembangkan memperoleh kategori sangat valid dengan hasil validasi ahli materi mendapatkan persentase sebesar 85,34%; ahli evaluasi sebesar 90,84%; ahli media sebesar 81,46%; dan guru IPA mendapatkan persentase sebesar 58,24%. Masukkan yang diberikan oleh validator meliputi penyempurnaan redaksi soal penggunaan stimulus yang kontekstual, perbaikan tampilan media, serta penyusunan rubrik penilaian yang lebih rinci, revisi dilakukan berdasarkan masukan dan saran yang diberikan validator sehingga instrumen tes HOTS berbasis *quizizz* yang dikembangkan dinyatakan layak sebagai alat evaluasi pembelajaran IPA untuk mengukur kemampuan berpikir tingkat tinggi peserta didik.

Kata Kunci: Evaluasi Pembelajaran, HOTS, Instrumen Tes, IPA, *Quizizz*.

ABSTRACT: *This study aims to produce a HOTS test instrument and determine the validity of the development of a Quizizz-based higher order thinking skills test instrument on the interaction of living things with their environment for seventh-grade students. This study uses poetry as the medium for presenting questions. This study uses the Sugiyono model of research and development method modified to the product revision stage. This study was validated by validators, namely material experts, evaluation experts, media experts, and science teachers using a questionnaire. The results of this study indicate that the developed instrument is categorized as very valid with the validation results of material experts obtaining a percentage of 85.34%; evaluation experts 90.84%; media experts 81.46%; and science teachers getting a percentage of 58.24%. Input provided by the validators included refinement of the question wording using contextual stimuli, improvements to the media display, and the development of a more detailed assessment rubric. Revisions were made based on input and suggestions provided by the validators so that the developed Quizizz-based HOTS test instrument was declared suitable as a science learning evaluation tool to measure students' higher-order thinking skills.*

Keywords: *Learning Evaluation, HOTS, Test Instrument, Science, Quizizz.*

How to Cite: Ayuni, A., Resti, V. D. A., & Kurniasih, S. (2026). Pengembangan Instrumen Tes HOTS Berbasis *Quizizz* pada Materi Interaksi Makhluk Hidup dengan Lingkungannya untuk Siswa Kelas VII. *Biocaster : Jurnal Kajian Biologi, 6*(4), 2461-2480. https://doi.org/10.36312/biocaster.v6i4.1869

Biocaster : Jurnal Kajian Biologi *is Licensed Under a CC BY-SA Creative Commons Attribution-ShareAlike 4.0 International License*.

Biocaster : Jurnal Kajian Biologi
E-ISSN 2808-277X; P-ISSN 2808-3598
Volume 6, Issue 4, October 2026; Page, 2461-2480
Email: biocasterjournal@gmail.com

PENDAHULUAN

Pendidikan terus mengalami perkembangan seiring dengan perubahan global, terutama dalam bidang teknologi dan informasi. Perkembangan tersebut mendorong munculnya konsep pendidikan abad ke-21 yang menuntut peserta didik untuk memiliki berbagai keterampilan guna menghadapi tantangan dan dinamika kehidupan modern (Wero *et al.*, 2021). Salah satu keterampilan yang penting untuk dikembangkan adalah kemampuan berpikir tingkat tinggi (*Higher Order Thinking Skills* atau HOTS). Kemampuan tersebut memungkinkan peserta didik untuk menganalisis informasi, mengevaluasi suatu permasalahan, serta menghasilkan gagasan atau solusi secara kreatif. Penguasaan keterampilan berpikir tingkat tinggi menjadi bekal penting bagi peserta didik dalam mengidentifikasi, menganalisis, dan memecahkan berbagai permasalahan yang dijumpai dalam kehidupan sehari-hari maupun lingkungan sekitarnya (Ichsan *et al.*, 2019).

Pembelajaran abad ke-21 menempatkan kemampuan berpikir tingkat tinggi (*Higher Order Thinking Skills* atau HOTS) sebagai salah satu kompetensi esensial yang perlu dikembangkan melalui proses pendidikan. Kemampuan berpikir tingkat tinggi mencakup tiga tingkatan kognitif, yaitu menganalisis (C4), mengevaluasi (C5), dan mencipta (C6) sebagaimana diklasifikasikan dalam Taksonomi Bloom revisi. Ketiga tingkatan kognitif tersebut berperan penting dalam mengembangkan kemampuan berpikir kritis, logis, dan kreatif peserta didik sehingga mampu menghadapi berbagai permasalahan secara sistematis dan rasional (Ratnasari *et al.*, 2021). Penguasaan kemampuan HOTS memungkinkan peserta didik tidak hanya memahami dan mengingat informasi, tetapi juga mengintegrasikan berbagai konsep, menganalisis serta mengevaluasi suatu permasalahan, dan merumuskan alternatif solusi yang relevan. Oleh karena itu, proses pembelajaran perlu dirancang secara konstruktif dengan memberikan kesempatan kepada peserta didik untuk mengembangkan kemampuan berpikir secara mendalam, analitis, dan reflektif, sehingga orientasi pembelajaran tidak terbatas pada penguasaan materi melalui aktivitas menghafal (Fitria *et al.*, 2022).

Pengembangan kemampuan berpikir tingkat tinggi (*Higher Order Thinking Skills* atau HOTS) menjadi salah satu aspek penting dalam implementasi Kurikulum Merdeka di Indonesia. Kurikulum tersebut mendorong guru untuk mengembangkan instrumen penilaian yang tidak hanya berorientasi pada kemampuan mengingat informasi, tetapi juga mampu mengukur keterampilan berpikir tingkat tinggi peserta didik, meliputi kemampuan menganalisis, mengevaluasi, dan mencipta. Namun, sejumlah hasil penelitian menunjukkan bahwa kemampuan HOTS peserta didik masih belum berkembang secara optimal. Penelitian yang dilakukan oleh Ratnasari *et al.* (2021) menunjukkan bahwa kemampuan HOTS peserta didik pada indikator menganalisis dan mengevaluasi berada pada kategori sedang, sedangkan kemampuan mencipta masih tergolong rendah. Temuan tersebut mengindikasikan bahwa peserta didik masih mengalami kesulitan dalam menyelesaikan permasalahan yang membutuhkan kemampuan penalaran, analisis, evaluasi, serta pengembangan gagasan secara kreatif.

Permasalahan tersebut tidak hanya berkaitan dengan kemampuan peserta didik, tetapi juga dipengaruhi oleh kualitas instrumen evaluasi yang digunakan dalam proses pembelajaran. Gradini (2021) menjelaskan bahwa instrumen evaluasi

Biocaster : Jurnal Kajian Biologi
E-ISSN 2808-277X; P-ISSN 2808-3598
Volume 6, Issue 4, October 2026; Page, 2461-2480
Email: biocasterjournal@gmail.com

yang disusun dan digunakan oleh guru masih didominasi oleh soal-soal yang mengukur kemampuan berpikir tingkat rendah (*Lower Order Thinking Skills* atau LOTS). Kondisi serupa juga ditemukan berdasarkan hasil studi pendahuluan yang dilakukan di salah satu SMP di Kabupaten Serang. Hasil observasi menunjukkan bahwa sebagian besar instrumen evaluasi yang digunakan di sekolah masih berorientasi pada pengukuran kemampuan mengingat, memahami, dan menerapkan. Meskipun soal-soal yang mengukur kemampuan berpikir tingkat tinggi (*Higher Order Thinking Skills* atau HOTS) telah mulai diterapkan, proporsinya masih relatif terbatas sehingga belum dapat mengukur keterampilan berpikir tingkat tinggi peserta didik secara optimal.

Hasil wawancara dengan guru IPA menunjukkan bahwa penyusunan instrumen penilaian berbasis HOTS masih menjadi tantangan tersendiri. Kesulitan tersebut terutama berkaitan dengan kemampuan guru dalam mengembangkan stimulus yang kontekstual, merumuskan indikator soal yang sesuai dengan tingkat kognitif C4-C6, serta menyusun rubrik penilaian yang tepat. Kondisi ini menunjukkan bahwa keterbatasan kemampuan guru dalam mengembangkan instrumen evaluasi berbasis HOTS yang valid dan sesuai dengan karakteristik peserta didik menjadi salah satu faktor yang perlu mendapat perhatian dalam upaya meningkatkan kualitas evaluasi pembelajaran.

Mata pelajaran Ilmu Pengetahuan Alam (IPA) pada jenjang Sekolah Menengah Pertama (SMP) memiliki peran strategis dalam mengembangkan kemampuan berpikir tingkat tinggi (*Higher Order Thinking Skills* atau HOTS) peserta didik. Pembelajaran IPA tidak hanya berorientasi pada penguasaan fakta, konsep, dan teori ilmiah, tetapi juga menekankan pengembangan pola pikir yang ilmiah, sistematis, dan logis dalam memahami serta menganalisis berbagai fenomena alam (Manassero-Mas & Vázquez-Alonso, 2022). Karakteristik pembelajaran IPA yang berkaitan erat dengan fenomena kehidupan sehari-hari memberikan peluang bagi peserta didik untuk menghubungkan konsep-konsep ilmiah dengan berbagai permasalahan nyata di lingkungan sekitarnya.

Salah satu materi IPA yang memiliki keterkaitan kontekstual tersebut adalah interaksi makhluk hidup dengan lingkungannya. Materi ini mencakup berbagai fenomena ekologis, seperti pencemaran sungai, kebakaran hutan, perubahan keseimbangan ekosistem, dan terganggunya rantai makanan. Beragam fenomena tersebut dapat dimanfaatkan sebagai stimulus kontekstual dalam penyusunan instrumen tes berbasis HOTS karena memungkinkan peserta didik untuk mengidentifikasi permasalahan, menganalisis hubungan sebab-akibat, mengevaluasi dampak perubahan lingkungan, serta merumuskan solusi berdasarkan konsep-konsep ilmiah yang telah dipelajari.

Seiring dengan perkembangan teknologi, proses penilaian pembelajaran pun mengalami perubahan. Penilaian tidak lagi harus dilakukan secara manual menggunakan kertas, melainkan dapat dilakukan secara digital menggunakan berbagai *platform* berbasis teknologi. Salah satu *platform* yang banyak digunakan oleh guru dan siswa saat ini adalah *quizizz*. *Quizizz* merupakan sebuah aplikasi penilaian berbasis permainan (*game-based assessment*) yang memungkinkan guru untuk membuat soal kuis secara daring dan siswa dapat menjawabnya menggunakan perangkat masing-masing, seperti *smartphone* atau komputer.

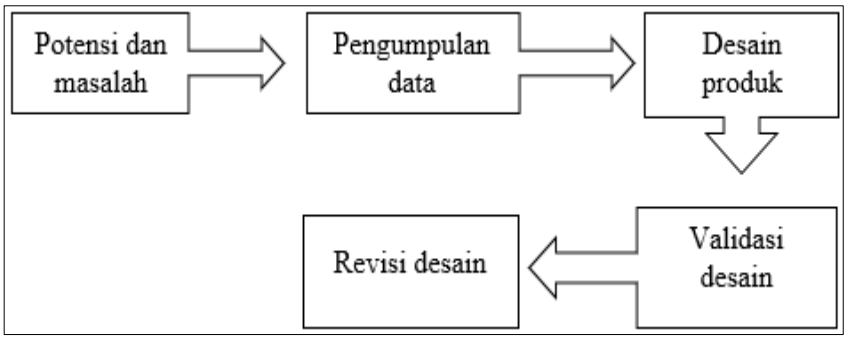
Biocaster : Jurnal Kajian Biologi
E-ISSN 2808-277X; P-ISSN 2808-3598
Volume 6, Issue 4, October 2026; Page, 2461-2480
Email: biocasterjournal@gmail.com

Quizizz juga menyediakan laporan hasil yang langsung dapat dilihat oleh guru, sehingga guru dapat dengan cepat mengetahui perkembangan pemahaman siswa. Dengan berbagai keunggulan tersebut, *quizizz* menjadi *platform* yang sangat potensial untuk dijadikan media dalam mengimplementasikan instrumen tes berbasis HOTS, terutama di era pembelajaran digital seperti saat ini (Hamidah & Wulandari, 2021; Nasution *et al.*, 2022; Pamelia & Hariani, 2022).

Berdasarkan uraian tersebut, pengembangan instrumen tes kemampuan berpikir tingkat tinggi berbasis *quizizz* menjadi salah satu alternatif yang dapat mendukung pelaksanaan evaluasi pembelajaran IPA, khususnya pada materi interaksi makhluk hidup dengan lingkungannya. Oleh karena itu, penelitian ini bertujuan untuk menghasilkan instrumen tes kemampuan berpikir tingkat tinggi (*Higher Order Thinking Skills* atau HOTS) berbasis *quizizz* pada materi interaksi makhluk hidup dengan lingkungannya untuk siswa kelas VII SMP serta menguji tingkat validitas produk yang dikembangkan melalui penilaian para ahli.

METODE

Penelitian ini menggunakan metode penelitian dan pengembangan (*Research and Development* atau R&D), yaitu pendekatan penelitian yang bertujuan untuk menghasilkan suatu produk tertentu serta menguji kelayakan produk tersebut sebelum digunakan. Metode R&D diterapkan untuk mengembangkan produk berupa instrumen tes maupun menyempurnakan produk yang telah tersedia agar sesuai dengan kebutuhan dan tujuan penggunaannya. Model pengembangan yang digunakan dalam penelitian ini mengacu pada model Borg & Gall (1983) yang telah dimodifikasi oleh Sugiyono. Pelaksanaan penelitian dibatasi hingga tahap kelima, yaitu revisi desain, dengan tahapan pengembangan yang meliputi: 1) identifikasi potensi dan masalah; 2) pengumpulan data; 3) desain produk; 4) validasi desain; dan 5) revisi desain. Adapun alur pelaksanaan penelitian berdasarkan tahapan pengembangan tersebut disajikan pada Gambar 1.



Gambar 1. Alur Penelitian Model Pengembangan.

Penelitian ini dilaksanakan di Provinsi Banten. Tahap analisis kebutuhan dilakukan di salah satu Sekolah Menengah Pertama (SMP) di Kabupaten Serang, sedangkan tahap validasi produk dilaksanakan di dua SMP yang berada di kabupaten berbeda di wilayah Provinsi Banten. Subjek pada tahap validasi melibatkan delapan validator yang terdiri atas dua ahli materi, dua ahli evaluasi, dua ahli media, dan dua guru IPA. Pemilihan validator didasarkan pada kompetensi

Biocaster : Jurnal Kajian Biologi
E-ISSN 2808-277X; P-ISSN 2808-3598
Volume 6, Issue 4, October 2026; Page, 2461-2480
Email: biocasterjournal@gmail.com

dan keahlian masing-masing sesuai dengan bidang yang divalidasi. Seluruh validator berperan dalam memberikan penilaian terhadap kelayakan produk berdasarkan aspek keahlian masing-masing.

Produk yang dikembangkan dalam penelitian ini berupa instrumen tes *Higher Order Thinking Skills* (HOTS) berbasis *quizizz* pada materi interaksi makhluk hidup dengan lingkungannya. Instrumen tersebut dirancang sebagai alat evaluasi untuk mengukur kemampuan berpikir tingkat tinggi peserta didik pada jenjang SMP. Data penelitian terdiri atas data kuantitatif dan kualitatif. Data kuantitatif diperoleh dari skor hasil penilaian validator melalui lembar validasi, sedangkan data kualitatif diperoleh melalui wawancara serta masukan dan saran yang diberikan oleh validator sebagai dasar untuk melakukan perbaikan dan penyempurnaan produk.

Penilaian kelayakan produk dilakukan menggunakan lembar angket validasi dengan menerapkan skala penilaian (*rating scale*). Skor yang diperoleh dari hasil penilaian validator selanjutnya diolah dan dikonversi ke dalam bentuk persentase serta diinterpretasikan berdasarkan kriteria penilaian yang telah ditetapkan. Hasil analisis tersebut digunakan untuk menentukan tingkat kelayakan instrumen tes HOTS berbasis *quizizz* yang dikembangkan.

Tabel 1. Skala Penilaian Menggunakan *Rating Scale*.

No.	Keterangan	Skor
1	Sangat Tidak Baik	1
2	Kurang Baik	2
3	Cukup Baik	3
4	Sangat Baik	4

Sumber: Riduwan & Sunarto (2013).

Perhitungan persentase dari angket validasi diolah menggunakan rumus berikut ini.

$$\mathbf{NP} = \frac{R}{SM} \mathbf{\ x\ 100\%}$$

Keterangan:
NP = Persentase Penilaian (%);
R = Skor yang diperoleh;
SM = Skor maksimum; dan
100% = Bilangan tetap.

Tabel 2. Tingkat Pencapaian.

No.	Tingkat Pencapaian	Keterangan
1	80,1% - 100%	Sangat Valid
2	60,1% - 80%	Valid
3	40,1% - 60%	Kurang Valid
4	0% - 40%	Tidak Valid

Sumber: Sugiyono (2015).

Produk tersebut mendapat nilai yang sangat baik dari para ahli jika hasil persentase lebih dari 80,1%. Produk yang dikembangkan cocok untuk digunakan sebagai bahan evaluasi setelah proses pembelajaran. Produk tersebut dapat mendukung proses penilaian hasil belajar peserta didik secara lebih optimal.

Biocaster : Jurnal Kajian Biologi
E-ISSN 2808-277X; P-ISSN 2808-3598
Volume 6, Issue 4, October 2026; Page, 2461-2480
Email: biocasterjournal@gmail.com

HASIL DAN PEMBAHASAN

Penelitian ini merupakan penelitian pengembangan (*Research and Development* atau R&D) yang mengacu pada prosedur pengembangan Sugiyono, dengan pelaksanaan penelitian dibatasi sampai tahap kelima, yaitu revisi desain. Penelitian ini menghasilkan produk berupa instrumen tes *Higher Order Thinking Skills* (HOTS) berbasis *quizizz* pada materi interaksi makhluk hidup dengan lingkungannya untuk siswa kelas VII. Produk yang dikembangkan terdiri atas 10 butir soal yang disusun berdasarkan indikator kemampuan berpikir tingkat tinggi pada ranah kognitif C4 (menganalisis), C5 (mengevaluasi), dan C6 (mencipta). Penyusunan instrumen tes juga disesuaikan dengan capaian dan tujuan pembelajaran materi interaksi makhluk hidup dengan lingkungannya serta dilengkapi dengan stimulus yang relevan dengan konteks kehidupan sehari-hari.

Selanjutnya, produk melalui tahap validasi oleh empat validator, yaitu ahli materi, ahli evaluasi, ahli media, dan guru IPA. Validasi tersebut bertujuan untuk menilai kelayakan isi materi, konstruksi dan kualitas soal, aspek tampilan serta penggunaan media *quizizz*, sehingga diperoleh instrumen tes yang valid dan layak digunakan sebagai alat evaluasi untuk mengukur kemampuan berpikir tingkat tinggi siswa kelas VII.

Tabel 3. Hasil Validasi oleh Validator Instrumen Tes HOTS Berbasis *Quizizz*.

No.	Validator	Rata-rata	Kriteria
1	Ahli Materi	85,34%	Sangat Valid
2	Ahli Evaluasi	90,84%	Sangat Valid
3	Ahli Media	81,46%	Sangat Valid
4	Guru IPA	85,24%	Sangat Valid

Hasil validasi menunjukkan bahwa instrumen tes HOTS berbasis *quizizz* pada materi interaksi makhluk hidup dengan lingkungannya memperoleh penilaian sangat valid dari seluruh validator. Meskipun demikian, terdapat perbedaan persentase penilaian antara ahli materi, ahli evaluasi, ahli media, dan guru IPA. Perbedaan tersebut mengindikasikan bahwa setiap validator memiliki fokus penilaian yang berbeda sesuai dengan bidang keahliannya. Ahli evaluasi memberikan persentase tertinggi sebesar 90,84%, yang menunjukkan bahwa instrumen telah memenuhi kriteria penilaian dari aspek evaluasi pembelajaran. Sedangkan penilaian ahli media memperoleh persentase terendah sebesar 81,46%. Hasil tersebut mengindikasikan bahwa aspek penyajian dan penggunaan media *quizizz* masih memiliki ruang untuk penyempurnaan dibandingkan dengan aspek lainnya. Kondisi ini menunjukkan bahwa kualitas instrumen tidak hanya ditentukan oleh kesesuaian substansi soal, tetapi juga dipengaruhi oleh kualitas penyajian dan kemudahan penggunaan media dalam pelaksanaan evaluasi pembelajaran.

Perbedaan hasil penilaian tersebut menjadi dasar dalam melakukan perbaikan terhadap produk yang dikembangkan. Revisi instrumen perlu diarahkan terutama pada aspek media yang memperoleh persentase paling rendah, tanpa mengabaikan masukan dari ahli materi, ahli evaluasi, dan guru IPA. Penyempurnaan dapat mencakup keterbacaan tampilan, kejelasan petunjuk pengerjaan, konsistensi penyajian soal, serta kemudahan navigasi dalam penggunaan *quizizz*. Perbaikan tersebut bertujuan agar instrumen tidak hanya

Biocaster : Jurnal Kajian Biologi
E-ISSN 2808-277X; P-ISSN 2808-3598
Volume 6, Issue 4, October 2026; Page, 2461-2480
Email: biocasterjournal@gmail.com

memenuhi kelayakan dari segi isi dan konstruksi soal, tetapi juga dapat digunakan secara efektif oleh peserta didik. Dengan demikian, hasil validasi memberikan informasi bahwa instrumen yang dikembangkan telah memenuhi kriteria kelayakan, sekaligus menunjukkan aspek yang perlu mendapatkan perhatian untuk meningkatkan kualitas produk.

Temuan ini memperlihatkan bahwa pengembangan instrumen tes HOTS memerlukan penilaian secara menyeluruh, baik dari aspek materi, evaluasi, media, maupun keterpakaian dalam pembelajaran. Penilaian dari berbagai validator menjadi penting karena instrumen yang baik tidak hanya harus sesuai dengan kompetensi dan karakteristik materi, tetapi juga perlu mampu mengukur kemampuan berpikir tingkat tinggi peserta didik secara tepat. Oleh karena itu, hasil validasi yang berada pada kategori sangat valid menunjukkan bahwa instrumen tes HOTS berbasis *quizizz* pada materi interaksi makhluk hidup dengan lingkungannya layak digunakan sebagai perangkat evaluasi pembelajaran IPA setelah mempertimbangkan dan menindaklanjuti saran perbaikan dari para validator.

Hasil Validasi Ahli Materi

Hasil validasi dari ahli materi memperoleh persentase sebesar 85,34% termasuk kategori "sangat valid", yang diperoleh dari hasil rata-rata lima belas sub aspek dari 3 aspek yaitu kelayakan isi, kesesuaian materi dengan HOTS, dan kebahasaan.

Tabel 4. Hasil Validasi Ahli Materi.

No.	Validator	Persentase	Kriteria
1	Kelayakan Isi	85,83%	Sangat Valid
2	Kesesuaian Materi dengan HOTS	82,71%	Sangat Valid
3	Kebahasaan	87,50%	Sangat Valid
Keseluruhan		85,34%	Sangat Valid

Validasi materi dilakukan untuk mengetahui tingkat kelayakan isi materi yang disajikan dalam instrumen tes HOTS berbasis *quizizz* berdasarkan beberapa subaspek penilaian. Hasil validasi oleh ahli materi menunjukkan bahwa instrumen memperoleh persentase kelayakan keseluruhan sebesar 85,34%, sehingga termasuk dalam kategori "sangat valid". Hasil tersebut menunjukkan bahwa materi yang disajikan dalam instrumen tes telah memenuhi kriteria kelayakan berdasarkan aspek-aspek yang dinilai oleh ahli materi.

Hasil penilaian terhadap aspek kelayakan isi memperoleh persentase rata-rata sebesar 85,83% dengan kategori "sangat valid". Penilaian tersebut didasarkan pada enam subaspek yang menjadi komponen penilaian kelayakan isi instrumen. Subaspek kesesuaian Capaian Pembelajaran (CP) dan Tujuan Pembelajaran (TP) memperoleh persentase sebesar 87,50% dengan kategori "sangat valid". Subaspek keluasan dan kedalaman materi memperoleh persentase sebesar 81,50% dengan kategori "sangat valid". Selanjutnya, subaspek ketepatan konsep dan kebenaran konsep IPA masing-masing memperoleh persentase sebesar 87,50% dengan kategori "sangat valid", sedangkan subaspek kesesuaian tingkat kesulitan dan keabstrakan soal memperoleh persentase sebesar 84,00%. Hasil penilaian tersebut menunjukkan bahwa aspek kelayakan isi instrumen tes HOTS telah memenuhi kriteria validitas dengan kategori "sangat valid".

Biocaster : Jurnal Kajian Biologi
E-ISSN 2808-277X; P-ISSN 2808-3598
Volume 6, Issue 4, October 2026; Page, 2461-2480
Email: biocasterjournal@gmail.com

Ditinjau dari keterkaitan antara CP dan indikator HOTS, hasil validasi tersebut menunjukkan bahwa rumusan capaian pembelajaran memiliki peran penting sebagai landasan dalam menentukan tingkat kognitif yang diukur melalui instrumen tes. Tingkatan kognitif tersebut mencakup kemampuan menganalisis (C4), mengevaluasi (C5), dan mencipta (C6), sehingga butir soal yang dikembangkan tidak hanya mengukur penguasaan konsep, tetapi juga kemampuan berpikir tingkat tinggi peserta didik. Hal ini sejalan dengan pendapat Rahmawati & Trimulyono (2022) yang menjelaskan bahwa pengembangan instrumen evaluasi perlu memperhatikan kesesuaian antara capaian pembelajaran, tujuan pembelajaran, dan indikator kemampuan yang hendak diukur.

Hasil penilaian pada aspek kesesuaian materi dengan *Higher Order Thinking Skills* (HOTS) yang mencakup enam subaspek memperoleh persentase rata-rata sebesar 82,71% dengan kategori "sangat valid". Persentase tersebut diperoleh berdasarkan hasil penilaian terhadap masing-masing subaspek, yaitu kemampuan menganalisis sebesar 87,50%, kemampuan menelaah sebesar 86,50%, kemampuan menilai sebesar 87,50%, kemampuan mengevaluasi sebesar 77,50%, kemampuan mencipta sebesar 76,50%, serta kesesuaian stimulus HOTS sebesar 81,50%. Hasil penilaian menunjukkan bahwa subaspek menganalisis, menelaah, dan menilai termasuk dalam kategori "sangat valid", sedangkan subaspek mengevaluasi dan mencipta berada pada kategori "valid". Sedangkan subaspek stimulus HOTS memperoleh kategori "sangat valid". Temuan tersebut menunjukkan bahwa materi dan butir soal yang dikembangkan telah memenuhi kriteria kesesuaian dengan karakteristik kemampuan berpikir tingkat tinggi, meskipun masih diperlukan penyempurnaan pada beberapa indikator, khususnya kemampuan mengevaluasi dan mencipta.

Berdasarkan hasil validasi, terdapat sejumlah saran dan masukan dari validator ahli sebagai bahan penyempurnaan instrumen tes yang dikembangkan. Salah satu validator menyarankan agar instrumen tes disusun dengan memperhatikan tuntutan kemampuan berpikir tingkat tinggi yang harus dimiliki peserta didik dalam menyelesaikan soal HOTS. Sedangkan validator lainnya memberikan penilaian positif terhadap penggunaan stimulus yang dinilai telah bersifat kontekstual dan relevan dengan fenomena yang terjadi di lingkungan sekitar. Penggunaan stimulus kontekstual tersebut dapat mendorong peserta didik untuk menghubungkan pengetahuan yang dimiliki dengan permasalahan nyata serta mengembangkan kemampuan berpikir kritis dan analitis. Hasil ini sejalan dengan penelitian Aryadi & Margunayasa (2022) yang menyatakan bahwa instrumen HOTS yang telah memenuhi aspek kelayakan dapat digunakan sebagai alat evaluasi dalam pembelajaran IPA.

Hasil penilaian pada subaspek kebahasaan memperoleh persentase rata-rata sebesar 87,50% dengan kategori "sangat valid". Penilaian tersebut mencakup tiga indikator, yaitu kejelasan materi, tidak adanya makna ganda, dan penggunaan bahasa yang mudah dipahami. Ketiga indikator tersebut memperoleh persentase sebesar 87,50% dan termasuk dalam kategori "sangat valid". Hasil ini menunjukkan bahwa penggunaan bahasa dalam instrumen tes HOTS telah memenuhi kriteria kebahasaan yang baik, ditinjau dari aspek kejelasan penyampaian informasi, ketepatan penggunaan istilah, serta kemudahan pemahaman oleh peserta didik.

Biocaster : Jurnal Kajian Biologi
E-ISSN 2808-277X; P-ISSN 2808-3598
Volume 6, Issue 4, October 2026; Page, 2461-2480
Email: biocasterjournal@gmail.com

Kejelasan bahasa dan tidak adanya makna ganda menjadi aspek penting untuk meminimalkan kemungkinan terjadinya kesalahan penafsiran ketika peserta didik memahami dan menjawab setiap butir soal. Temuan ini sejalan dengan pendapat Ismail *et al.* (2021) yang menyatakan bahwa penggunaan bahasa dalam instrumen pembelajaran perlu disesuaikan dengan tingkat perkembangan peserta didik, antara lain melalui pemilihan istilah yang umum digunakan dan mudah dipahami. Dengan demikian, persentase yang diperoleh menunjukkan bahwa aspek kebahasaan pada instrumen tes HOTS yang dikembangkan telah memenuhi kriteria kelayakan dan dinyatakan sangat valid untuk digunakan sebagai instrumen evaluasi pembelajaran.

Hasil Validasi Ahli Evaluasi

Hasil validasi yang dilakukan oleh ahli evaluasi memperoleh persentase rata-rata sebesar 90,84% dan termasuk dalam kategori "sangat valid". Persentase tersebut diperoleh berdasarkan hasil penilaian terhadap lima belas subaspek yang mencakup tiga aspek utama, yaitu aspek substansi, konstruksi, dan kebahasaan. Hasil tersebut menunjukkan bahwa instrumen tes yang dikembangkan telah memenuhi kriteria kelayakan dari segi evaluasi.

Tabel 5. Hasil Validasi Ahli Evaluasi.

No.	Validator	Persentase	Kriteria
1	Substansi	93,25%	Sangat Valid
2	Konstruksi	88,96%	Sangat Valid
3	Kebahasaan	90,32%	Sangat Valid
Keseluruhan		90,84%	Sangat Valid

Validasi ahli evaluasi dilakukan untuk menilai kelayakan instrumen tes *Higher Order Thinking Skills* (HOTS) berbasis *quizizz* berdasarkan sejumlah subaspek penilaian yang mencakup kualitas substansi, konstruksi, dan kebahasaan. Hasil validasi oleh ahli evaluasi menunjukkan bahwa instrumen tes memperoleh persentase keseluruhan sebesar 90,84%, sehingga termasuk dalam kategori "sangat valid". Persentase tersebut diperoleh berdasarkan hasil perhitungan rata-rata seluruh subaspek yang dinilai oleh validator ahli evaluasi. Hasil ini menunjukkan bahwa instrumen tes yang dikembangkan telah memenuhi kriteria kelayakan untuk digunakan sebagai alat evaluasi pembelajaran.

Hasil penilaian pada aspek substansi memperoleh persentase rata-rata sebesar 93,25% dan termasuk dalam kategori "sangat valid". Penilaian tersebut diperoleh berdasarkan lima subaspek yang dinilai. Subaspek kesesuaian soal dengan tujuan pembelajaran memperoleh persentase sebesar 87,50% dengan kategori "sangat valid". Subaspek kesesuaian soal dengan materi interaksi makhluk hidup dengan lingkungannya memperoleh persentase sebesar 100% dengan kategori "sangat valid". Selanjutnya, subaspek kemampuan soal dalam mengukur keterampilan *Higher Order Thinking Skills* (HOTS) memperoleh persentase sebesar 87,50% dengan kategori "sangat valid". Penggunaan stimulus yang kontekstual memperoleh persentase sebesar 96,50% dengan kategori "sangat valid", sedangkan subaspek soal berbasis pemecahan masalah memperoleh persentase sebesar 95,00% dengan kategori yang sama. Berdasarkan hasil penilaian tersebut, instrumen tes HOTS berbasis *quizizz* dinyatakan sangat layak untuk digunakan sebagai alat evaluasi pembelajaran. Instrumen ini dapat membantu guru

Biocaster : Jurnal Kajian Biologi
E-ISSN 2808-277X; P-ISSN 2808-3598
Volume 6, Issue 4, October 2026; Page, 2461-2480
Email: biocasterjournal@gmail.com

dalam mengukur dan mengevaluasi keterampilan berpikir tingkat tinggi peserta didik, khususnya pada materi interaksi makhluk hidup dengan lingkungannya.

Hasil penilaian pada aspek substansi memperoleh persentase rata-rata sebesar 88,96% dengan kategori "sangat valid". Persentase tersebut diperoleh berdasarkan penilaian terhadap enam subaspek. Subaspek kejelasan pokok soal memperoleh persentase sebesar 87,50% dengan kategori "sangat valid". Hasil tersebut menunjukkan bahwa pokok soal dalam instrumen tes HOTS telah disajikan secara jelas, ringkas, dan tegas sehingga dapat meminimalkan potensi kesalahpahaman peserta didik dalam memahami soal. Subaspek penggunaan kalimat efektif memperoleh persentase sebesar 81,50% dengan kategori "sangat valid". Hal ini menunjukkan bahwa kalimat yang digunakan dalam instrumen tes HOTS telah memenuhi kaidah keefektifan bahasa dan relatif mudah dipahami oleh peserta didik. Selanjutnya, subaspek kejelasan perintah soal memperoleh persentase sebesar 86,50% dengan kategori "sangat valid", yang mengindikasikan bahwa petunjuk pengerjaan soal telah dirumuskan secara jelas sehingga peserta didik dapat memahami tuntutan jawaban yang diharapkan.

Subaspek kesesuaian soal dengan bentuk uraian memperoleh persentase sebesar 100,00% dengan kategori "sangat valid". Hasil tersebut menunjukkan bahwa seluruh butir soal yang dikembangkan telah sesuai dengan karakteristik soal berbentuk uraian (*essay*), sehingga memberikan kesempatan kepada peserta didik untuk mengemukakan jawaban secara sistematis berdasarkan pemahaman dan kemampuan berpikirnya. Subaspek keterstrukturan jawaban memperoleh persentase sebesar 91,50% dengan kategori "sangat valid". Capaian tersebut menunjukkan bahwa jawaban yang diharapkan dalam instrumen tes telah dirumuskan secara terstruktur dan selaras dengan tuntutan setiap butir soal. Sedangkan subaspek tingkat kesulitan soal memperoleh persentase sebesar 97,50% dengan kategori "sangat valid". Hasil ini mengindikasikan bahwa tingkat kesulitan soal telah disesuaikan dengan karakteristik instrumen berbasis *Higher Order Thinking Skills* (HOTS), yang menuntut peserta didik untuk mengembangkan kemampuan menganalisis (C4), mengevaluasi (C5), dan mencipta (C6).

Berdasarkan keseluruhan hasil penilaian tersebut, aspek substansi dinyatakan memenuhi kriteria "sangat valid". Dengan demikian, instrumen tes HOTS berbasis *quizizz* yang dikembangkan telah memenuhi kelayakan dari aspek substansi dan dapat digunakan sebagai salah satu alat evaluasi dalam pembelajaran IPA. Temuan ini sejalan dengan pendapat Zainuddin *et al.* (2020) yang menyatakan bahwa validitas instrumen evaluasi dapat dibuktikan melalui hasil penilaian ahli terhadap aspek-aspek penyusunnya, termasuk kesesuaian substansi dan konstruksi instrumen. Hasil penilaian tersebut menjadi dasar untuk menentukan kelayakan instrumen sebelum digunakan dalam proses pembelajaran.

Hasil penilaian pada aspek kebahasaan memperoleh persentase rata-rata sebesar 90,32% dan termasuk dalam kategori "sangat valid". Penilaian tersebut diperoleh berdasarkan hasil rata-rata dari dua validator ahli evaluasi terhadap empat subaspek kebahasaan. Subaspek pernyataan tidak menimbulkan makna ganda memperoleh persentase sebesar 90,00% dengan kategori "sangat valid". Hasil tersebut menunjukkan bahwa penggunaan bahasa dalam instrumen tes telah disusun secara jelas dan tidak menimbulkan penafsiran ganda sehingga peserta didik dapat

Biocaster : Jurnal Kajian Biologi
E-ISSN 2808-277X; P-ISSN 2808-3598
Volume 6, Issue 4, October 2026; Page, 2461-2480
Email: biocasterjournal@gmail.com

memahami maksud setiap butir soal dengan tepat. Subaspek ketepatan penggunaan kalimat memperoleh persentase sebesar 85,00% dengan kategori "sangat valid". Capaian tersebut mengindikasikan bahwa struktur kalimat yang digunakan dalam instrumen tes telah memenuhi kaidah kebahasaan serta disusun secara efektif sehingga mudah dipahami oleh peserta didik. Selanjutnya, subaspek kesesuaian bahasa dengan tingkat perkembangan peserta didik SMP memperoleh persentase sebesar 87,50% dan termasuk dalam kategori "sangat valid". Hasil ini menunjukkan bahwa penggunaan bahasa dalam instrumen tes HOTS telah disesuaikan dengan tingkat perkembangan kognitif dan kemampuan pemahaman peserta didik jenjang SMP. Sedangkan subaspek tidak terdapat kesalahan penulisan memperoleh persentase tertinggi, yaitu sebesar 95,00% dengan kategori "sangat valid".

Hasil tersebut menunjukkan bahwa penulisan butir soal telah memenuhi aspek ketepatan teknis, baik dari segi pemilihan ukuran huruf, penggunaan spasi, maupun ketepatan penulisan. Aspek tipografi, khususnya ukuran huruf, perlu diperhatikan karena dapat memengaruhi kenyamanan dan kemampuan membaca peserta didik dalam memahami informasi tertulis (Nikmah *et al.*, 2018). Dengan demikian, hasil penilaian pada aspek kebahasaan menunjukkan bahwa instrumen tes HOTS yang dikembangkan telah memenuhi kelayakan bahasa dari segi kejelasan, ketepatan kalimat, kesesuaian dengan tingkat perkembangan peserta didik, serta ketepatan penulisan.

Hasil Validasi Ahli Media

Validasi oleh ahli media dilakukan untuk menilai kelayakan media berdasarkan tiga aspek utama, yaitu aspek penyajian, kegrafikan, dan kemudahan penggunaan. Penilaian tersebut mencakup 13 subaspek yang digunakan sebagai indikator untuk mengukur kualitas dan kelayakan media yang dikembangkan.

Tabel 6. Hasil Validasi Ahli Media.

No.	Validator	Persentase	Kriteria
1	Penyajian	76,88%	Valid
2	Kegrafikan	80,00%	Valid
3	Kemudahan Penggunaan	87,50%	Sangat Valid
Keseluruhan		81,46%	Sangat Valid

Validasi ahli media dilakukan untuk mengetahui tingkat kelayakan aspek tampilan dan penyajian instrumen tes HOTS yang dikembangkan melalui *platform quizizz*. Hasil validasi ahli media memperoleh persentase keseluruhan sebesar 81,46%, sehingga termasuk dalam kategori "sangat valid". Penilaian tersebut diperoleh berdasarkan hasil perhitungan terhadap seluruh aspek yang dinilai oleh validator ahli media. Adapun uraian hasil validasi ahli media disajikan secara terperinci berdasarkan masing-masing aspek penilaian sebagai berikut:

Hasil penilaian pada aspek penyajian memperoleh persentase rata-rata sebesar 76,88% dengan kategori "valid". Penilaian tersebut diperoleh berdasarkan hasil rata-rata dari dua validator ahli media terhadap empat subaspek penyajian. Subaspek kejelasan tulisan pada butir soal memperoleh persentase sebesar 85,00% dengan kategori "sangat valid". Hasil tersebut menunjukkan bahwa tulisan yang digunakan dalam instrumen tes memiliki tingkat keterbacaan yang baik sehingga memudahkan peserta didik dalam membaca dan memahami informasi yang

Biocaster : Jurnal Kajian Biologi
E-ISSN 2808-277X; P-ISSN 2808-3598
Volume 6, Issue 4, October 2026; Page, 2461-2480
Email: biocasterjournal@gmail.com

disajikan. Kejelasan tulisan menjadi salah satu aspek penting karena informasi yang disampaikan secara jelas dapat meminimalkan kesalahpahaman pembaca dalam memahami maksud soal.

Subaspek penempatan gambar memperoleh persentase sebesar 72,50% dengan kategori "valid". Meskipun demikian, hasil penilaian menunjukkan bahwa masih terdapat beberapa gambar yang memiliki resolusi rendah sehingga tampak kurang jelas atau buram. Oleh karena itu, kualitas dan resolusi gambar perlu ditingkatkan agar informasi visual dapat tersampaikan secara optimal. Hal tersebut sejalan dengan pernyataan Kemendikbud (2019) bahwa penyajian stimulus, tata letak, gambar, dan elemen visual perlu dirancang secara jelas serta fungsional untuk mendukung proses berpikir tingkat tinggi peserta didik.

Selanjutnya, subaspek kejelasan penyajian stimulus memperoleh persentase sebesar 71,50% dengan kategori "valid". Hasil tersebut menunjukkan bahwa stimulus yang digunakan telah mengarah pada penyajian informasi yang kontekstual dan relevan dengan kehidupan sehari-hari peserta didik. Stimulus kontekstual dapat dikembangkan melalui berbagai permasalahan yang dekat dengan lingkungan peserta didik, termasuk permasalahan yang terdapat di satuan pendidikan, adat istiadat, budaya, kasus lokal, maupun potensi dan keunggulan daerah. Hal ini sejalan dengan pendapat Fanani (2018) yang menyatakan bahwa stimulus dalam penyusunan soal dapat bersumber dari berbagai permasalahan kontekstual yang berkaitan dengan kehidupan dan lingkungan peserta didik. Adapun subaspek kemudahan tampilan *quizizz* sebagai media penyaji soal memperoleh persentase sebesar 79,00% dengan kategori "valid". Hasil tersebut menunjukkan bahwa tampilan media *quizizz* dinilai cukup mudah digunakan dalam menyajikan instrumen tes, sehingga dapat mendukung keterlaksanaan evaluasi pembelajaran secara lebih praktis.

Hasil penilaian pada aspek kegrafikan memperoleh persentase rata-rata sebesar 80,00% dan termasuk dalam kategori "sangat valid". Penilaian tersebut diperoleh berdasarkan hasil evaluasi terhadap lima subaspek kegrafikan. Subaspek kejelasan gambar memperoleh persentase sebesar 66,50% dengan kategori "valid". Hasil tersebut menunjukkan bahwa gambar yang disajikan dalam instrumen tes telah memenuhi kriteria kelayakan, meskipun masih memerlukan penyempurnaan dari segi kejelasan visual. Hal ini sejalan dengan pendapat Kustandi *et al.* (2021) yang menyatakan bahwa gambar harus memiliki kualitas yang baik, menarik, dan mudah dipahami oleh peserta didik karena keberadaannya berperan penting dalam membantu memahami materi yang dipelajari.

Selanjutnya, subaspek kesesuaian ikon dan ilustrasi memperoleh persentase sebesar 79,00% dengan kategori "valid", sedangkan subaspek tampilan media interaktif memperoleh persentase sebesar 75,00% dengan kategori "valid". Subaspek kesesuaian ukuran gambar, tabel, atau grafik memperoleh persentase sebesar 80,00% dan termasuk dalam kategori "valid". Meskipun ketiga subaspek tersebut telah memenuhi kriteria validitas, beberapa perbaikan tetap dilakukan berdasarkan saran dan masukan dari validator. Perbaikan tersebut meliputi peningkatan kejelasan gambar, penyesuaian ukuran gambar agar proporsional dengan teks, serta penyempurnaan keterangan gambar sehingga lebih representatif dalam menggambarkan fenomena yang disajikan. Upaya perbaikan ini bertujuan

Biocaster : Jurnal Kajian Biologi
E-ISSN 2808-277X; P-ISSN 2808-3598
Volume 6, Issue 4, October 2026; Page, 2461-2480
Email: biocasterjournal@gmail.com

untuk meningkatkan kualitas tampilan visual instrumen tes agar informasi yang disampaikan dapat dipahami secara lebih jelas oleh peserta didik. Hal tersebut sejalan dengan Kustandi *et al.* (2021) yang menekankan pentingnya kualitas visual dalam mendukung penyampaian informasi pembelajaran.

Sedangkan subaspek penggunaan variasi huruf memperoleh persentase sebesar 87,50% dengan kategori "sangat valid". Hasil tersebut menunjukkan bahwa penggunaan variasi huruf pada instrumen tes telah memenuhi kriteria kelayakan, terutama dalam hal pemilihan dan konsistensi jenis huruf yang digunakan. Variasi huruf yang tidak berlebihan dinilai sesuai dengan karakteristik penyajian soal karena dapat membantu menjaga keterbacaan dan konsistensi tampilan instrumen. Hal ini sejalan dengan Rudolph *et al.* (2019) yang menyatakan bahwa penggunaan variasi huruf secara berlebihan dalam penyusunan soal perlu dihindari karena dapat mengurangi efektivitas penyajian informasi. Dengan demikian, aspek kegrafikan instrumen tes HOTS berbasis *quizizz* telah memenuhi kriteria sangat valid, meskipun beberapa penyempurnaan visual tetap dilakukan untuk meningkatkan kualitas dan keterbacaan instrumen tes.

Hasil penilaian pada aspek kemudahan penggunaan memperoleh persentase rata-rata sebesar 87,50% dengan kategori "sangat valid". Penilaian tersebut diperoleh berdasarkan hasil rata-rata dari dua validator ahli media terhadap empat subaspek kemudahan penggunaan. Subaspek kemudahan akses memperoleh persentase sebesar 100,00% dengan kategori "sangat valid". Hasil tersebut menunjukkan bahwa media *quizizz* dapat diakses dengan mudah oleh peserta didik. Temuan ini sejalan dengan pendapat Arsyad (2020) yang menyatakan bahwa media pembelajaran yang baik harus memiliki tingkat kemudahan penggunaan yang tinggi agar peserta didik dapat lebih memusatkan perhatian pada materi pembelajaran daripada prosedur pengoperasian media.

Subaspek kejelasan petunjuk penggunaan memperoleh persentase sebesar 100,00% dengan kategori "sangat valid". Hasil tersebut mengindikasikan bahwa petunjuk penggunaan yang disajikan dalam media telah disusun secara jelas sehingga mudah dipahami oleh peserta didik. Selanjutnya, subaspek media membantu proses evaluasi memperoleh persentase sebesar 75,00% dengan kategori "valid". Meskipun memperoleh persentase terendah dibandingkan dengan subaspek lainnya, hasil tersebut menunjukkan bahwa media *quizizz* telah memenuhi kriteria kelayakan untuk mendukung pelaksanaan evaluasi pembelajaran. Namun demikian, diperlukan penyempurnaan berdasarkan saran dan masukan dari validator agar fungsi media dalam menunjang proses evaluasi dapat dioptimalkan. Sedangkan subaspek kemudahan penggunaan tombol atau menu memperoleh persentase sebesar 100,00% dengan kategori "sangat valid". Hasil ini menunjukkan bahwa seluruh tombol dan menu yang tersedia dalam media dapat dioperasikan dengan mudah oleh peserta didik.

Hasil penilaian pada aspek kemudahan penggunaan menunjukkan bahwa media *quizizz* memiliki tingkat aksesibilitas dan kemudahan pengoperasian yang sangat baik. Meskipun demikian, subaspek media membantu proses evaluasi masih memerlukan perbaikan agar pemanfaatan media dapat lebih optimal dalam mendukung kegiatan evaluasi pembelajaran. Hasil ini sejalan dengan pendapat Anidhea & Sulaiman (2024) yang menyatakan bahwa *quizizz* merupakan *platform*

Biocaster : Jurnal Kajian Biologi
E-ISSN 2808-277X; P-ISSN 2808-3598
Volume 6, Issue 4, October 2026; Page, 2461-2480
Email: biocasterjournal@gmail.com

teknologi pembelajaran yang menyediakan kuis interaktif secara daring serta memungkinkan peserta didik untuk belajar secara mandiri maupun kolaboratif melalui lingkungan pembelajaran yang kompetitif dan menyenangkan.

Hasil Validasi Guru IPA

Hasil validasi yang dilakukan oleh guru IPA memperoleh persentase rata-rata sebesar 85,24% dan termasuk dalam kategori sangat valid. Validasi tersebut bertujuan untuk mengetahui tingkat kepraktisan instrumen tes *Higher Order Thinking Skills* (HOTS) sebelum diujicobakan kepada peserta didik. Penilaian dilakukan oleh dua orang guru IPA yang memiliki pengalaman dalam melaksanakan pembelajaran IPA di tingkat SMP. Proses penilaian mengacu pada tiga aspek penilaian yang digunakan sebagai dasar untuk mengevaluasi kepraktisan instrumen tes, baik dari segi kesesuaian maupun kemudahan penggunaannya dalam kegiatan evaluasi pembelajaran.

Tabel 7. Hasil Validasi Guru IPA.

No.	Validator	Persentase	Kriteria
1	Kelayakan Isi	85,42%	Sangat Valid
2	Kebahasaan	85,32%	Sangat Valid
3	Tampilan	85,00%	Sangat Valid
Keseluruhan		85,24%	Sangat Valid

Validasi oleh guru IPA dilakukan untuk mengetahui tingkat kepraktisan instrumen tes HOTS yang dikembangkan sebelum diujicobakan kepada peserta didik. Hasil penilaian dari dua guru IPA memperoleh persentase rata-rata sebesar 85,24%, sehingga termasuk dalam kategori "sangat valid". Penilaian tersebut didasarkan pada tiga aspek, yaitu kelayakan isi, kebahasaan, dan tampilan. Ketiga aspek tersebut digunakan sebagai indikator untuk menilai kesesuaian instrumen dari segi materi, penggunaan bahasa, serta penyajian instrumen. Uraian mengenai hasil penilaian validasi guru IPA selanjutnya disajikan berdasarkan masing-masing aspek sebagai berikut:

Hasil penilaian pada aspek kelayakan isi memperoleh persentase rata-rata sebesar 85,24% dan termasuk dalam kategori "sangat valid". Penilaian tersebut didasarkan pada enam subaspek. Subaspek kesesuaian materi dengan capaian pembelajaran memperoleh persentase sebesar 86,25% dengan kategori "sangat valid". Hasil ini menunjukkan bahwa materi yang disajikan dalam instrumen tes telah sesuai dengan capaian pembelajaran yang ditetapkan. Selanjutnya, subaspek kesesuaian materi dengan tujuan pembelajaran memperoleh persentase sebesar 85,00% dengan kategori "sangat valid", yang mengindikasikan bahwa materi dalam instrumen telah selaras dengan tujuan pembelajaran yang hendak dicapai.

Subaspek karakteristik peserta didik memperoleh persentase sebesar 86,25% dengan kategori "sangat valid". Hasil tersebut menunjukkan bahwa pengembangan instrumen tes telah mempertimbangkan karakteristik dan kebutuhan peserta didik. Pada subaspek indikator HOTS, diperoleh persentase sebesar 87,50% dengan kategori "sangat valid". Capaian ini menunjukkan bahwa instrumen yang dikembangkan telah mengacu pada indikator kemampuan berpikir tingkat tinggi (*higher order thinking skills*), yang mencakup kemampuan menganalisis, mengevaluasi, dan mencipta. Sedangkan subaspek stimulus yang mendukung

Biocaster : Jurnal Kajian Biologi
E-ISSN 2808-277X; P-ISSN 2808-3598
Volume 6, Issue 4, October 2026; Page, 2461-2480
Email: biocasterjournal@gmail.com

pemahaman peserta didik memperoleh persentase sebesar 86,25% dengan kategori "sangat valid". Hasil tersebut mengindikasikan bahwa stimulus yang disajikan telah mendukung pemahaman peserta didik terhadap permasalahan yang diujikan. Penggunaan stimulus kontekstual dapat membantu peserta didik menghubungkan konsep pembelajaran dengan permasalahan nyata sehingga mendorong keterlibatan dalam proses berpikir tingkat tinggi (Kemendikbud, 2019).

Adapun subaspek produk yang membantu proses pembelajaran IPA memperoleh persentase sebesar 81,25% dengan kategori "sangat valid". Hasil ini menunjukkan bahwa instrumen tes HOTS berbasis *quizizz* dinilai mampu mendukung pelaksanaan evaluasi dalam pembelajaran IPA. Keenam subaspek yang dinilai oleh guru IPA memperoleh kategori "sangat valid". Temuan tersebut menunjukkan bahwa instrumen yang dikembangkan telah memenuhi kriteria kelayakan isi serta relevan dengan capaian pembelajaran, tujuan pembelajaran, karakteristik peserta didik, dan kebutuhan pembelajaran IPA di sekolah. Dengan demikian, instrumen tes HOTS berbasis *quizizz* dinyatakan layak digunakan sebagai instrumen evaluasi untuk mengukur kemampuan berpikir tingkat tinggi peserta didik.

Hasil penilaian pada aspek kebahasaan memperoleh persentase rata-rata sebesar 85,32% dan termasuk dalam kategori "sangat valid". Penilaian tersebut didasarkan pada hasil evaluasi terhadap empat subaspek kebahasaan. Subaspek bahasa mudah dipahami oleh peserta didik memperoleh persentase sebesar 85,00% dengan kategori "sangat valid". Hasil ini menunjukkan bahwa bahasa yang digunakan dalam instrumen tes telah disesuaikan dengan tingkat pemahaman peserta didik sehingga pesan dan informasi yang disampaikan dapat dipahami dengan baik. Hal tersebut sejalan dengan pendapat Jailani (2018) yang menyatakan bahwa bahasa memiliki peran penting dalam mengekspresikan pikiran anak, membangun hubungan sosial, serta membantu anak menyampaikan dan memahami informasi.

Subaspek penggunaan kaidah bahasa Indonesia memperoleh persentase sebesar 86,25% dengan kategori "sangat valid". Hasil tersebut menunjukkan bahwa penggunaan bahasa dalam instrumen tes telah mengacu pada kaidah bahasa Indonesia yang baik dan benar. Selanjutnya, subaspek bahasa tidak menimbulkan makna ganda memperoleh persentase sebesar 82,50% dengan kategori "sangat valid". Temuan ini sejalan dengan pendapat Bantani *et al.* (2024) yang menyatakan bahwa keefektifan kalimat diperlukan untuk memastikan kejelasan informasi serta menghindari munculnya penafsiran yang berbeda atau makna ganda. Sedangkan subaspek penggunaan istilah yang sesuai dengan materi IPA memperoleh persentase sebesar 87,50%, yang menunjukkan bahwa istilah-istilah dalam instrumen tes telah sesuai dengan konteks dan materi pembelajaran IPA.

Berdasarkan hasil penilaian keempat subaspek tersebut, aspek kebahasaan dinyatakan sangat valid. Meskipun demikian, subaspek bahasa tidak menimbulkan makna ganda memperoleh persentase terendah dibandingkan dengan ketiga subaspek lainnya. Hal ini mengindikasikan bahwa masih terdapat beberapa istilah atau ungkapan yang kurang familiar bagi peserta didik sehingga berpotensi menimbulkan perbedaan penafsiran terhadap informasi dalam butir soal. Oleh karena itu, penyederhanaan penggunaan istilah dan pemilihan kalimat yang lebih

Biocaster : Jurnal Kajian Biologi
E-ISSN 2808-277X; P-ISSN 2808-3598
Volume 6, Issue 4, October 2026; Page, 2461-2480
Email: biocasterjournal@gmail.com

komunikatif perlu diperhatikan agar setiap butir soal memiliki makna yang jelas, tidak ambigu, serta sesuai dengan tingkat perkembangan kognitif dan kemampuan berbahasa peserta didik.

Hasil penilaian pada subaspek tampilan produk memperoleh persentase sebesar 82,50%, yang menunjukkan bahwa tampilan produk yang dikembangkan telah memenuhi kriteria kelayakan dengan baik. Pada subaspek tata letak komponen yang tersusun secara sistematis, diperoleh persentase sebesar 84,00%. Hasil tersebut mengindikasikan bahwa komponen-komponen dalam media telah disusun secara terstruktur sehingga memudahkan peserta didik dalam mengoperasikan media. Selanjutnya, subaspek gambar, warna, dan tampilan yang mendukung pembelajaran memperoleh persentase sebesar 86,50%. Capaian ini menunjukkan bahwa penggunaan elemen visual pada media mampu mendukung penyampaian materi serta membantu memperjelas informasi yang disajikan dalam soal. Sedangkan subaspek kemudahan penggunaan produk memperoleh persentase sebesar 87,50%, yang mengindikasikan bahwa media yang dikembangkan mudah dioperasikan oleh peserta didik.

Berdasarkan hasil penilaian validator guru IPA, media *quizizz* yang dikembangkan dinilai memiliki tampilan yang menarik dan berpotensi meningkatkan motivasi serta antusiasme peserta didik dalam mengikuti kegiatan evaluasi pembelajaran. Penggunaan *quizizz* sebagai *platform* penyajian soal memberikan pengalaman evaluasi yang lebih interaktif dibandingkan dengan metode evaluasi konvensional. Penyajian soal melalui media digital yang dilengkapi dengan unsur visual dan ilustrasi dapat menciptakan suasana evaluasi yang lebih menarik sehingga mendorong keterlibatan peserta didik dalam mengerjakan soal. Ilustrasi yang disajikan turut membantu peserta didik dalam memahami konteks permasalahan yang terdapat pada setiap butir soal. Dengan demikian, aspek tampilan dan kemudahan penggunaan media *quizizz* tidak hanya mendukung keterbacaan dan pengoperasian produk, tetapi juga berkontribusi dalam menciptakan pengalaman evaluasi pembelajaran yang lebih interaktif dan menarik bagi peserta didik.

Revisi Desain

Revisi desain dilakukan setelah instrumen tes *Higher Order Thinking Skills* (HOTS) berbasis *quizizz* melalui tahapan validasi oleh para ahli, yang meliputi ahli materi, ahli evaluasi, ahli media, serta guru IPA sebagai validator. Tahap revisi ini bertujuan untuk menyempurnakan instrumen yang telah dikembangkan berdasarkan saran dan masukan dari para validator agar produk yang dihasilkan memenuhi prinsip pengembangan instrumen HOTS serta sesuai dengan karakteristik peserta didik kelas VII SMP.

Perbaikan instrumen difokuskan pada beberapa aspek, yaitu kelayakan isi, konstruksi soal, kebahasaan, tampilan media, dan rubrik penilaian. Masukan yang diberikan oleh para validator mencakup penyempurnaan redaksi soal, penggunaan stimulus yang lebih kontekstual, perbaikan tampilan media *quizizz*, serta pengembangan rubrik penilaian yang lebih terperinci dan sistematis. Seluruh proses revisi dilakukan dengan tetap mempertahankan tujuan pengukuran dan substansi instrumen yang telah dirancang sebelumnya. Dengan demikian, revisi desain dapat meningkatkan kualitas instrumen, baik dari segi isi maupun penyajiannya.

Biocaster : Jurnal Kajian Biologi
E-ISSN 2808-277X; P-ISSN 2808-3598
Volume 6, Issue 4, October 2026; Page, 2461-2480
Email: biocasterjournal@gmail.com

Berdasarkan hasil perbaikan sesuai dengan saran dan masukan validator, instrumen tes HOTS berbasis *quizizz* yang dikembangkan dinyatakan layak digunakan sebagai instrumen evaluasi pembelajaran IPA untuk mengukur kemampuan berpikir tingkat tinggi peserta didik.

SIMPULAN

Berdasarkan hasil penelitian pengembangan instrumen tes *Higher Order Thinking Skills* (HOTS) berbasis *quizizz* pada materi interaksi makhluk hidup dengan lingkungannya untuk siswa kelas VII, dapat disimpulkan bahwa instrumen yang dikembangkan memiliki tingkat kelayakan dengan kategori sangat valid. Hal tersebut dibuktikan melalui hasil penilaian dari masing-masing validator, yaitu ahli materi sebesar 84,92%, ahli evaluasi sebesar 90,75%, ahli media sebesar 81,35%, dan guru IPA sebesar 85,27%. Hasil penilaian tersebut menunjukkan bahwa instrumen telah memenuhi kriteria kelayakan berdasarkan berbagai aspek yang dievaluasi, meliputi kelayakan isi, kesesuaian materi dengan indikator HOTS, substansi konstruksi, penyajian, kegrafikan, kemudahan penggunaan, dan tampilan produk.

Proses pengembangan instrumen juga melibatkan tahap revisi berdasarkan saran dan masukan yang diberikan oleh para validator. Revisi tersebut dilakukan sebagai upaya penyempurnaan kualitas produk, baik dari aspek substansi materi, konstruksi instrumen, maupun tampilan media, sehingga instrumen yang dihasilkan sesuai dengan tujuan pengembangan dan karakteristik peserta didik. Hasil validasi dan tahapan pengembangan menunjukkan bahwa instrumen tes HOTS berbasis *quizizz* pada materi interaksi makhluk hidup dengan lingkungannya dinyatakan sangat valid dan layak digunakan sebagai alternatif instrumen evaluasi pembelajaran untuk mengukur kemampuan berpikir tingkat tinggi siswa kelas VII SMP.

SARAN

Berdasarkan hasil penelitian pengembangan instrumen tes *Higher Order Thinking Skills* (HOTS) berbasis *quizizz* pada materi interaksi makhluk hidup dengan lingkungannya, terdapat beberapa saran yang dapat dijadikan sebagai bahan pertimbangan untuk pemanfaatan dan pengembangan penelitian selanjutnya. Bagi guru, instrumen tes HOTS yang telah dikembangkan diharapkan dapat menjadi salah satu alternatif dalam melaksanakan evaluasi pembelajaran. Guru juga disarankan untuk mengembangkan instrumen tes HOTS pada materi pembelajaran yang lebih beragam dengan cakupan materi yang lebih luas, sehingga peserta didik terbiasa mengerjakan soal-soal yang mengukur kemampuan berpikir tingkat tinggi. Penerapan instrumen tersebut diharapkan dapat mendorong peningkatan kemampuan peserta didik dalam menganalisis, mengevaluasi, dan menciptakan solusi terhadap permasalahan yang berkaitan dengan materi pembelajaran.

Bagi pembaca dan peneliti selanjutnya, hasil penelitian pengembangan instrumen tes HOTS berbasis *quizizz* ini diharapkan dapat menjadi referensi dan sumber informasi pendukung dalam melaksanakan penelitian sejenis. Penelitian selanjutnya disarankan untuk mengembangkan instrumen tes HOTS pada materi, jenjang pendidikan, dan karakteristik peserta didik yang berbeda agar diperoleh

produk pengembangan dengan cakupan penerapan yang lebih luas. Peneliti juga perlu memastikan bahwa pemahaman mengenai kemampuan berpikir tingkat tinggi dan indikator HOTS yang digunakan telah sesuai dengan landasan teori serta konteks penelitian. Apabila penelitian melibatkan guru atau narasumber melalui wawancara, peneliti juga perlu memastikan kesamaan pemahaman terhadap konsep dan indikator HOTS agar informasi yang diperoleh relevan dengan tujuan penelitian.

UCAPAN TERIMA KASIH

Penulis menyampaikan terima kasih dan penghargaan yang sebesar-besarnya kepada Bapak dan Ibu Dosen serta seluruh Staf Jurusan Pendidikan IPA, Fakultas Keguruan dan Ilmu Pendidikan, Universitas Sultan Ageng Tirtayasa, atas ilmu, bimbingan, dukungan, serta arahan yang telah diberikan selama proses pelaksanaan penelitian ini. Apresiasi juga disampaikan kepada seluruh pihak yang telah memberikan bantuan, motivasi, dan dukungan, baik secara langsung maupun tidak langsung, sehingga penelitian ini dapat diselesaikan dengan baik.

DAFTAR RUJUKAN

Anidhea, N. O., & Sulaiman, P. K. A. (2024). *Systematic Literature Review* (SLR): Efektivitas Penggunaan Media *Quizizz* terhadap Hasil Belajar Siswa SMP pada Pembelajaran IPA. In *Seminar Nasional Hasil Riset dan Pengabdian* (pp. 1-12). Surabaya, Indonesia: Universitas PGRI Adi Buana Surabaya.

Arsyad, A. (2020). *Media Pembelajaran (Edisi Revisi, Cetakan ke-22)*. Depok: Rajawali Pers.

Aryadi, K. S., & Margunayasa, I. G. (2022). Instrumen Penilaian *High Order Thinking Skills* (HOTS) pada Pembelajaran IPA. *Indonesian Journal of Instruction, 3*(1), 34-41. https://doi.org/10.23887/iji.v3i1.44761

Bantani, R. A., Berlian, L., & Kurniasih, S. (2024). Pengembangan E-Modul Tema Kemagnetan dan Pemanfaatannya dalam Produk Teknologi untuk Melatih Kemampuan Berpikir Kreatif Siswa. *Jurnal Ilmiah Pendidikan IPA, 6*(3), 1053-1066. https://doi.org/10.29100/.v6i3.4933

Borg, W. R., & Gall, M. D. (1983). *Educational Research: An Introduction (4th Ed.)*. Edinburgh: Longman.

Fanani, M. Z. (2018). Strategi Pengembangan Soal HOTS pada Kurikulum 2013. *Edudeena : Journal of Islamic Religious Education*, *2*(1), 57-76. https://doi.org/10.30762/ed.v2i1.582

Fitria, Y., Safnowandi, S., & Fajri, S. R. (2022). Pengaruh Model Pembelajaran Inkuiri Terbimbing (*Guided Inquiry*) Berbasis Saintifik terhadap Kemampuan Berpikir Kritis Siswa. *Biocaster : Jurnal Kajian Biologi, 2*(3), 128-141. https://doi.org/10.36312/bjkb.v2i3.97

Gradini, E. (2021). Penyusunan Soal Tes Bermuatan *Lower* dan *Higher-Order Thinking Skills* pada Guru SMP. *Jurnal Inovasi Hasil Pengabdian Masyarakat (JIPEMAS), 4*(3), 300-312. https://doi.org/10.33474/jipemas.v4i3.10121

Hamidah, M. H., & Wulandari, S. 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

Ichsan, I. Z., Hasanah, R., Aini, S., Ristanto, R. H., & Miarsyah, M. (2019). *Higher Order Thinking Skills Assessment Based on Environmental Problem (HOTS-AEP)*: Mendesain Evaluasi Pembelajaran Abad 21. *Jurnal Biotek, 7*(1), 14-26. https://doi.org/10.24252/jb.v7i1.7939

Ismail, R., Rifma, R., & Fitria, Y. (2021). Pengembangan Bahan Ajar Tematik Berbasis Model PjBL di Sekolah Dasar. *Jurnal Basicedu, 5*(2), 958-965. https://doi.org/10.31004/basicedu.v5i2.808

Jailani, M. S. (2018). Perkembangan Bahasa Anak dan Implikasinya dalam Pembelajaran. *Innovatio: Journal for Religious-Innovation Studies, XVIII*(1), 15-26. https://doi.org/10.30631/innovatio.v18i1.36

Kementerian Pendidikan dan Kebudayaan. (2019). *Indonesia Educational Statistics in Brief 2018/2019*. Jakarta: Pusat Data dan Statistik Pendidikan dan Kebudayaan, Sekretariat Jenderal.

Kustandi, C., Zianadezdha, A., Fitri, A. K., Farhan, M., & Agustia L, N. (2021). Pemanfaatan Media Visual dalam Tercapainya Tujuan Pembelajaran. *Akademika, 10*(02), 291-299. https://doi.org/10.34005/akademika.v10i02.1402

Manassero-Mas, M. A., & Vázquez-Alonso, A. (2022). An Empirical Analysis of the Relationship between Nature of Science and Critical Thinking through Science Definitions and Thinking Skills. *SN Social Sciences, 2*, 1-27. https://doi.org/10.1007/s43545-022-00546-x

Nasution, D. K., Manurung, Y. H., Arda, M., & Handayani, S. (2022). Pelatihan *Digital Assesment* Berbasis *Kahoot* dan *Quizizz* dalam Pembelajaran Daring (*Online*) untuk Guru Muhamadiyah 61 Medan. *Ihsan: Jurnal Pengabdian Masyarakat, 4*(2), 113-118. https://doi.org/10.30596/ihsan.v4i2.10486

Nikmah, N., Akil, M., & Iskandar, I. (2018). *The Effects Font Size on Student's Reading Comprehension*. Makassar: University of Makassar.

Pamelia, A. A., & Hariani, D. (2022). The Development of HOTS Assessment Instrument Using Quizizz During Covid-19 Pandemic in 11th Grade Biology Second Semester. *Berkala Ilmiah Pendidikan Biologi (BioEdu), 11*(1), 107-115. https://doi.org/10.26740/bioedu.v11n1.p107-115

Rahmawati, D. E., & Trimulyono, G. (2022). Validitas Instrumen Penilaian *Higher Order Thinking Skills* (HOTS) pada Materi Keanekaragaman Hayati. *Berkala Ilmiah Pendidikan Biologi (BioEdu), 11*(1), 138-147. https://doi.org/10.26740/bioedu.v11n1.p138-147

Ratnasari, B., Haris, A., & Aziz, A. (2021). Studi Kemampuan Berfikir Tingkat Tinggi Fisika di SMA. *Jurnal Sains dan Pendidikan (JSPF), 4*(1), 57-68. https://doi.org/10.35580/jspf.v17i1.19063

Riduwan, R., & Sunarto, S. (2013). *Pengantar Statistika untuk Penelitian Pendidikan, Sosial, Ekonomi, Komunikasi, dan Bisnis*. Bandung: CV. Alfabeta.

Rudolph, M. J., Daugherty, K. K., Ray, M. E., Shuford, V. P., Lebovitz, L., & DiVall, M. V. (2019). Best Practices Related to Examination Item Construction and Post-Hoc Review. *American Journal of Pharmaceutical*

Education, 83(7), 1492-1503. https://doi.org/10.5688/ajpe7204

Sugiyono, S. (2015). *Metode Penelitian Pendidikan: Pendekatan Kuantitatif, Kualitatif, dan R&D*. Bandung: CV. Alfabeta.

Wero, L., Laksana, D. N. L., & Lawe, Y. U. (2021). Integrasi Konten dan Konteks Budaya Lokal Etnis Ngada dalam Bahan Ajar Multilingual untuk Pembelajaran Siswa Sekolah Dasar. *Mimbar PGSD UNDIKSHA, 9*(3), 515-522. https://doi.org/10.23887/jjpgsd.v9i3.40867

Zainuddin, Z., Chu, S. K. W., Shujahat, M., & Perera, C. J. (2020). The Impact of Gamification on Learning and Instruction: A Systematic Review of Empirical Evidence. *Educational Research Review, 30*, 100326. https://doi.org/10.1016/j.edurev.2020.100326