



Pengembangan Instrumen Asesmen Berbasis *Higher Order Thinking Skills* pada Materi Komponen Ekosistem dan Interaksinya di SMA Negeri 1 Tondano

Imelda Febrianti Makalalag^{1*}, Arrijani², Anatje Lihiang³

¹⁻³Program Studi Pendidikan Biologi, Universitas Negeri Manado, Indonesia

*Penulis Korespondensi: imeldaamakalalag@gmail.com

Abstract. *The rapid demands of 21st-century education require assessment instruments capable of measuring higher order thinking skills (HOTS), while classroom assessments in many schools are still dominated by lower order thinking skills (LOTS). This study aims to develop a HOTS-based assessment instrument on ecosystem components and their interactions at SMA Negeri 1 Tondano and evaluate its feasibility based on validity, reliability, item difficulty, discriminating power, and student responses. The study used a Research and Development (R&D) approach with the 3-D model consisting of define, design, and develop stages. The subjects involved 32 tenth-grade students and one biology teacher. The instrument consisted of 20 multiple-choice questions and 5 essay questions designed to measure cognitive levels of analysis (C4), evaluation (C5), and creation (C6). Data were collected through expert validation, validity and reliability testing, item analysis, and student response questionnaires. The results showed an expert validation score of 97.2 (very valid), a trial score of 77.25 (good), and a student response score of 89.87 (very valid). Item analysis indicated that most items were valid, reliable, and had balanced difficulty levels. The developed HOTS-based assessment instrument is valid, practical, and effective for biology learning evaluation.*

Keywords: *Assessment Instrument; Biology Learning; Ecosystem; Higher Order Thinking Skills; Research And Development.*

Abstrak. Tuntutan pendidikan abad ke-21 yang semakin berkembang membutuhkan instrumen penilaian yang mampu mengukur kemampuan berpikir tingkat tinggi (*Higher Order Thinking Skills/HOTS*), sementara pelaksanaan penilaian di sekolah masih banyak didominasi oleh kemampuan berpikir tingkat rendah (*Lower Order Thinking Skills/LOTS*). Penelitian ini bertujuan untuk mengembangkan instrumen penilaian berbasis HOTS pada materi komponen ekosistem dan interaksinya di SMA Negeri 1 Tondano serta mengetahui kelayakannya berdasarkan validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan respons siswa. Penelitian ini menggunakan metode *Research and Development* (R&D) dengan model 3-D yang meliputi tahap define, design, dan develop. Subjek penelitian terdiri atas 32 siswa kelas X dan satu guru biologi. Instrumen yang dikembangkan terdiri dari 20 soal pilihan ganda dan 5 soal uraian yang dirancang berdasarkan tingkat kognitif analisis (C4), evaluasi (C5), dan kreasi (C6). Data diperoleh melalui validasi ahli, uji validitas, reliabilitas, analisis butir soal, serta angket respons siswa. Hasil penelitian menunjukkan nilai validasi ahli sebesar 97,2 (sangat valid), nilai uji coba sebesar 77,25 (baik), dan respons siswa sebesar 89,87 (sangat valid). Analisis butir menunjukkan sebagian besar soal valid, reliabel, dan memiliki tingkat kesukaran yang seimbang. Instrumen penilaian berbasis HOTS yang dikembangkan dinyatakan valid, praktis, dan efektif digunakan sebagai alat evaluasi pembelajaran biologi.

Kata kunci: Ekosistem; *Higher Order Thinking Skills*; Instrumen Asesmen; Pembelajaran Biologi; Penelitian Pengembangan.

1. LATAR BELAKANG

Pendidikan pada era abad ke-21 menuntut perubahan yang signifikan dalam pendekatan pengajaran dan pembelajaran. Salah satu hal yang paling mendasar adalah perlunya pengembangan keterampilan berpikir kritis dan analitis pada peserta didik. Keterampilan tersebut dikenal dengan istilah *Higher Order Thinking Skills* (HOTS), yang mencakup kemampuan untuk menganalisis, mengevaluasi, dan menciptakan solusi terhadap berbagai permasalahan yang kompleks (Amalia et al., 2021). Dalam kerangka kurikulum pendidikan nasional yang terus berkembang, keterampilan berpikir tingkat tinggi menjadi bagian yang

sangat penting dalam pembentukan kualitas pendidikan karena mendorong peserta didik untuk berpikir secara lebih mendalam, kreatif, dan reflektif, serta mampu menyelesaikan masalah yang tidak selalu memiliki jawaban yang jelas atau pasti (Mangelep et al., 2026; Winarti & Istiyono, 2025).

Seiring dengan tuntutan pendidikan yang lebih tinggi, instrumen asesmen atau penilaian yang digunakan di kelas juga harus mengalami perubahan. Penilaian tradisional yang hanya mengukur kemampuan kognitif tingkat rendah, seperti mengingat fakta atau prosedur, tidak lagi memadai untuk mengukur perkembangan kemampuan berpikir tingkat tinggi. Oleh karena itu, diperlukan pengembangan instrumen asesmen yang dapat mengukur kemampuan HOTS peserta didik secara lebih akurat. Instrumen asesmen berbasis HOTS dirancang untuk mengevaluasi kemampuan peserta didik dalam berpikir kritis, memecahkan masalah, membuat keputusan, serta menghasilkan gagasan baru yang inovatif, sehingga menuntut peserta didik melakukan analisis, sintesis, evaluasi, dan kreasi, bukan sekadar menghafalan atau pemahaman dasar (Lewis & Smith, 2023).

Mata pelajaran Biologi memegang peran penting dalam mendukung perkembangan keterampilan berpikir kritis dan analitis peserta didik. Namun, dalam proses pembelajarannya kerap dijumpai berbagai permasalahan, khususnya berkaitan dengan konsep-konsep Biologi yang kompleks dan abstrak sehingga menyulitkan peserta didik dalam memahaminya. Kondisi ini diperberat oleh perbedaan tingkat kemampuan berpikir dan gaya belajar peserta didik, yang jika tidak diakomodasi dengan asesmen yang tepat akan menimbulkan kesenjangan pemahaman di dalam kelas (Purnasari et al., 2024; Runturambi & Somba 2026). Pengembangan soal yang berkualitas dan mampu memicu penalaran tingkat tinggi telah menjadi perhatian dalam berbagai penelitian pengembangan, baik pada ranah matematika maupun sains (Kalengkongan et al., 2021; Mangelep, 2013; Mangelep & Kaunang, 2018).

SMA Negeri 1 Tondano menjadi fokus studi ini setelah dilakukan serangkaian wawancara dengan guru dan tenaga kependidikan. Hasil wawancara mengindikasikan bahwa praktik pembelajaran dan penilaian di sekolah tersebut masih dominan menggunakan pendekatan *Lower Order Thinking Skills* (LOTS). Berdasarkan hasil analisis dokumen, soal ujian Biologi yang digunakan sebagian besar masih berada pada level mengingat (C1) dan memahami (C2), serta belum banyak mengaitkan konsep Biologi dengan permasalahan nyata di lingkungan sekitar. Kondisi ini penting untuk ditelaah karena dominasi LOTS berpotensi membatasi pengembangan kompetensi berpikir tingkat tinggi peserta didik.

Meskipun urgensi asesmen berbasis HOTS semakin diakui, masih banyak tantangan dalam implementasinya di kelas, terutama menyangkut bagaimana merancang instrumen yang mampu mengukur keterampilan berpikir tingkat tinggi secara objektif dan valid. Beberapa penelitian terdahulu telah mengembangkan instrumen HOTS, namun sebagian besar difokuskan pada mata pelajaran matematika atau pada jenjang serta materi yang berbeda, sehingga instrumen asesmen HOTS yang secara khusus dirancang untuk materi komponen ekosistem dan interaksinya pada jenjang SMA masih terbatas. Kesenjangan inilah yang menjadi celah (*gap*) sekaligus urgensi penelitian ini.

Berdasarkan uraian tersebut, penelitian ini bertujuan untuk mengembangkan instrumen asesmen berbasis HOTS pada materi komponen ekosistem dan interaksinya di SMA Negeri 1 Tondano, serta mendeskripsikan tingkat kelayakannya ditinjau dari aspek validitas, kepraktisan, dan keefektifan. Kebaruan penelitian ini terletak pada pengembangan instrumen HOTS yang mengintegrasikan stimulus kontekstual dan fenomena ekologis autentik pada materi ekosistem, yang diharapkan mampu menjadi alternatif alat evaluasi bagi guru dalam mengukur sekaligus mendorong kemampuan berpikir tingkat tinggi peserta didik.

2. KAJIAN TEORITIS

Higher Order Thinking Skills (HOTS)

Keterampilan berpikir tingkat tinggi atau *Higher Order Thinking Skills (HOTS)* meliputi kemampuan berpikir kritis, logis, reflektif, metakognitif, dan kreatif yang sangat diperlukan dalam proses pemecahan masalah serta pengambilan keputusan. HOTS umumnya berkembang ketika individu dihadapkan pada situasi yang belum familiar, pertanyaan yang menantang, maupun kondisi yang mengandung ketidakpastian atau dilema.

Menurut (Lewis & Smith, 2023), berpikir tingkat tinggi terjadi ketika seseorang memanfaatkan informasi yang telah tersimpan dalam ingatan kemudian mengintegrasikannya dengan informasi baru melalui proses menghubungkan, menyusun, serta mengembangkan informasi tersebut untuk menemukan kemungkinan jawaban atau solusi atas suatu situasi yang kompleks.

(Tomei, 2025) menambahkan bahwa HOTS mencakup kemampuan mentransformasikan informasi dan gagasan melalui proses analisis, sintesis, generalisasi, hingga penarikan kesimpulan, sedangkan Underbakke et al., (2023) menegaskan bahwa HOTS merupakan kemampuan berpikir strategis dalam memecahkan masalah, menganalisis argumen, dan membuat prediksi.

Taksonomi Bloom Revisi

Taksonomi Bloom merupakan kerangka yang diciptakan oleh Benjamin Bloom dan koleganya untuk mengklasifikasikan tujuan pendidikan berdasarkan tingkat kesulitannya, mulai dari kemampuan berpikir tingkat rendah hingga kemampuan berpikir tingkat tinggi (Bloom et al., 1956). Dalam perkembangannya, taksonomi ini diadaptasi dan direvisi oleh Anderson & Krathwohl sehingga ranah kognitif tersusun atas mengingat (C1), memahami (C2), mengaplikasikan (C3), menganalisis (C4), mengevaluasi (C5), dan mencipta (C6). Kategori C4, C5, dan C6 inilah yang menjadi ranah HOTS dan digunakan sebagai acuan dalam menyusun indikator soal pada penelitian ini (Harefa et al., 2024).

Konstruktivisme dan Pemecahan Masalah

Teori konstruktivisme memandang pembelajaran sebagai proses aktif di mana individu membangun pemahaman sendiri melalui pencarian kompetensi, pengetahuan, dan pengalaman, bukan sekadar menerima informasi secara pasif (Thobroni & Mustofa, 2025; Yaumi & Hum, 2022). Dalam konteks penilaian berbasis HOTS, teori ini menyoroti perlunya peserta didik terlibat dalam tugas-tugas yang menuntut berpikir kritis dan penyelesaian masalah yang kontekstual. Hal ini sejalan dengan prinsip pembelajaran berbasis masalah, yang menuntut peserta didik menghadapi persoalan autentik dan tidak terstruktur (ill-structured problem) sehingga memerlukan strategi penyelesaian dan penalaran produktif. Pengembangan soal pemecahan masalah dengan strategi tertentu telah terbukti mampu mendorong proses berpikir yang lebih kompleks pada peserta didik (Lestari & Yudhanegara, 2024; Mangelep, 2013, 2015; Mangelep & Kaunang, 2018).

Instrumen Asesmen

Instrumen asesmen merupakan alat yang memenuhi persyaratan akademis sehingga dapat digunakan untuk mengukur objek tertentu maupun mengumpulkan data terkait suatu variabel. Suatu instrumen dikatakan baik apabila memiliki tingkat validitas dan reliabilitas yang memadai, baik dari segi validitas isi, konstruk, maupun empiris (Azwar, 2023; Sudjana, 2025). Menurut (Nurkancana, 2023), sebuah alat ukur dianggap valid apabila mampu mengukur dengan tepat apa yang seharusnya diukur, dan dianggap reliabel apabila menghasilkan hasil yang relatif serupa ketika digunakan pada kelompok yang sama pada kesempatan berbeda. Oleh karena itu, pengembangan instrumen asesmen harus dilaksanakan secara sistematis melalui perumusan indikator, penyusunan butir, validasi ahli, serta uji coba lapangan. Penelitian pengembangan perangkat dan instrumen penilaian yang sistematis telah banyak dilakukan pada berbagai konteks pembelajaran dan terbukti menghasilkan produk yang

valid serta layak digunakan (Domu & Mangelep, 2019, 2025; Mangelep, 2017; Runturambi et al., 2026; Somba et al., 2026).

3. METODE PENELITIAN

Penelitian ini menggunakan metode penelitian dan pengembangan (*Research and Development/R&D*) dengan model 3-D yang meliputi tiga tahap, yaitu tahap pendefinisian (*define*), tahap perancangan (*design*), dan tahap pengembangan (*develop*). Model ini dipilih karena sesuai dengan tujuan penelitian, yakni menghasilkan produk berupa instrumen asesmen berbasis HOTS yang valid, reliabel, dan layak digunakan (Maydiantoro, 2025; Somba et al., 2026).

Penelitian dilaksanakan di SMA Negeri 1 Tondano, Kecamatan Tondano Barat, Kabupaten Minahasa. Subjek penelitian adalah peserta didik kelas X yang berjumlah 32 orang serta satu orang guru mata pelajaran Biologi yang dilibatkan sebagai informan dalam proses pengumpulan data. Instrumen yang dikembangkan berupa soal pilihan ganda dan soal uraian berbasis HOTS pada materi komponen ekosistem dan interaksinya, yang disusun berdasarkan indikator kemampuan berpikir tingkat tinggi pada level kognitif C4, C5, dan C6.

Prosedur Pengembangan

Tahap *define* mencakup analisis kurikulum, analisis karakteristik peserta didik, dan analisis instrumen asesmen yang telah digunakan guru untuk mengidentifikasi kebutuhan pengembangan. Tahap *design* meliputi penyusunan kisi-kisi soal, pengembangan butir soal berbasis HOTS melalui penambahan stimulus kontekstual dan fenomena biologis, serta penyusunan pedoman penskoran. Tahap *develop* meliputi validasi ahli, revisi, uji coba terbatas kepada peserta didik, dan analisis butir untuk memperoleh instrumen final yang valid dan reliabel.

Teknik Analisis Data

Analisis validitas isi dilakukan berdasarkan penilaian validator menggunakan skala Likert, dengan skor validasi dihitung menggunakan rumus persentase, yaitu jumlah skor yang diperoleh dibagi jumlah skor tertinggi dikalikan 100. Skor tersebut kemudian dikonversi menjadi kriteria kualitatif kevalidan. Adapun kriteria kevalidan penilaian perangkat pembelajaran disajikan pada Tabel 1.

Tabel 1. Kriteria Kevalidan Penilaian Perangkat Pembelajaran.

Skor	Keterangan
$81 \leq x \leq 100$	Sangat Valid
$61 \leq x < 81$	Valid

Skor	Keterangan
$41 \leq x < 61$	Cukup Valid
$21 \leq x < 41$	Kurang Valid
$0 \leq x < 21$	Tidak Valid

Analisis empiris dilakukan terhadap data hasil uji coba, meliputi uji validitas butir menggunakan korelasi product moment pada taraf signifikansi 5%, uji reliabilitas menggunakan rumus Alpha Cronbach dengan kriteria reliabel apabila koefisien $\geq 0,60$, analisis tingkat kesukaran soal untuk mengelompokkan butir ke dalam kategori mudah, sedang, dan sukar, serta analisis daya pembeda untuk mengetahui kemampuan soal dalam membedakan peserta didik berkemampuan tinggi dan rendah (Azizah & Chaimatusadiah, 2025; Darma, 2023). Selain itu, data kualitatif berupa komentar dan saran validator dianalisis secara deskriptif kualitatif sebagai dasar revisi instrumen (Sugiyono, 2022).

4. HASIL DAN PEMBAHASAN

Tahap Pendefinisian (*Define*)

Tahap pendefinisian bertujuan menetapkan kebutuhan pengembangan instrumen. Hasil analisis kurikulum menunjukkan bahwa materi komponen ekosistem dan interaksinya termasuk dalam kompetensi dasar yang menuntut kemampuan berpikir tingkat tinggi, yaitu menganalisis hubungan komponen biotik dan abiotik, mengevaluasi perubahan ekosistem, serta merancang solusi terhadap permasalahan lingkungan. Analisis karakteristik peserta didik menunjukkan bahwa peserta didik kelas X telah berada pada tahap operasional formal, namun kemampuan berpikir tingkat tingginya belum berkembang optimal karena lebih terbiasa mengerjakan soal menghafal dan memahami. Sementara itu, analisis instrumen asesmen menunjukkan bahwa soal ujian yang digunakan guru sebagian besar masih didominasi soal LOTS pada level C1 dan C2 serta belum mengaitkan konsep dengan permasalahan nyata. Berdasarkan ketiga analisis tersebut, disimpulkan bahwa diperlukan pengembangan instrumen asesmen berbasis HOTS pada materi komponen ekosistem dan interaksinya.

Tahap Perancangan (*Design*)

Pada tahap perancangan disusun kisi-kisi soal yang mengintegrasikan kompetensi dasar, indikator HOTS, serta level kognitif berdasarkan Taksonomi Bloom revisi (C4, C5, dan C6). Kisi-kisi memuat komponen kompetensi dasar, indikator soal, materi, level kognitif, dan bentuk soal, sebagaimana disajikan pada Tabel 2 (disajikan sebagian sebagai perwakilan).

Tabel 2. Kisi-kisi Instrumen Asesmen Berbasis HOTS (Sebagian).

No	Indikator Soal	Materi	Level	Bentuk
1	Menganalisis dampak penebangan terhadap ekosistem	Komponen ekosistem	C4	PG
2	Menganalisis dampak hilangnya predator	Interaksi	C4	PG
3	Menghitung aliran energi trofik	Energi	C4	PG
4	Menganalisis eutrofikasi	Pencemaran	C5	PG
5	Menganalisis dampak deforestasi pada daur air	Daur air	C5	PG
6	Merancang mitigasi lingkungan	Ekosistem	C6	Essay
7	Menghubungkan siklus hidrologi	Daur air	C4	Essay

Pengembangan soal dilakukan dengan memodifikasi soal yang sebelumnya berada pada level LOTS menjadi soal berbasis HOTS melalui penambahan stimulus kontekstual, pengintegrasian data atau kasus biologis nyata, serta perubahan tuntutan soal menjadi analisis, evaluasi, dan kreasi. Sebagai contoh, soal “Apa yang dimaksud dengan rantai makanan?” (C1) dikembangkan menjadi kasus penurunan populasi katak di sawah yang menuntut peserta didik mengevaluasi dampaknya terhadap rantai makanan (C5). Transformasi semacam ini sejalan dengan prinsip pengembangan soal bermutu yang menuntut penalaran, bukan sekadar pengenalan konsep (Mangelep, 2013). Instrumen final terdiri atas 20 butir soal pilihan ganda dan 5 butir soal uraian, dilengkapi kisi-kisi, kunci jawaban, dan pedoman penskoran, dengan skor maksimum keseluruhan 100.

Tahap Pengembangan (Develop)

Validasi Ahli

Instrumen yang telah disusun divalidasi oleh dua validator, yaitu satu orang dosen ahli dan satu orang guru Biologi SMA Negeri 1 Tondano. Hasil penilaian validator disajikan pada Tabel 3.

Tabel 3. Hasil Penilaian Validator.

Validator	Skor Penilaian	Kriteria
Validator 1 (Dosen)	97,7	Sangat Valid
Validator 2 (Guru)	96,6	Sangat Valid
Rata-rata	97,2	Sangat Valid

Berdasarkan hasil validasi diperoleh skor rata-rata keseluruhan sebesar 97,2 dengan kriteria sangat valid. Hal ini menunjukkan bahwa instrumen telah memenuhi aspek kesesuaian isi dengan HOTS, konstruksi soal, aspek kontekstual dan saintifik, serta aspek kebahasaan. Tingginya skor validasi mengindikasikan bahwa stimulus, tuntutan kognitif, dan indikator kompetensi pada setiap butir telah selaras dengan karakteristik HOTS.

Hasil Uji Coba dan Respon Peserta Didik

Uji coba dilakukan terhadap 32 peserta didik dan memperoleh nilai rata-rata sebesar 77,25 yang termasuk dalam kategori baik. Nilai ini mengindikasikan bahwa instrumen dapat digunakan dengan baik dan mampu mengukur aspek yang diteliti secara optimal. Sebagai perwakilan, sebagian hasil uji coba disajikan pada Tabel 4, sedangkan data lengkap 32 responden disajikan pada lampiran.

Tabel 4. Hasil Uji Coba Instrumen (Perwakilan).

No	Kode Siswa	Skor	Kriteria
1	KK	95,5	Sangat Baik
2	KM	88,0	Sangat Baik
3	AA	91,5	Sangat Baik
4	HH	73,5	Baik
5	GMT	74,0	Baik

Selain instrumen tes, angket respon peserta didik juga digunakan untuk menilai kepraktisan dan keberterimaan produk. Berdasarkan hasil pengisian oleh 32 peserta didik diperoleh rata-rata skor 89,87 dengan kategori sangat valid. Hasil uji validitas angket terhadap butir 1 sampai 10 menunjukkan seluruh nilai r hitung $\geq r$ tabel (0,349), sehingga seluruh butir dinyatakan valid, dan instrumen angket termasuk kategori sangat reliabel. Penilaian positif ini menunjukkan bahwa soal dinilai jelas, sesuai materi, dapat dipahami, dan relevan dengan pengalaman belajar peserta didik.

Analisis Butir Soal

Analisis butir dilakukan untuk mengetahui kualitas soal berdasarkan uji validitas, reliabilitas, tingkat kesukaran, dan daya pembeda. Hasil uji validitas terhadap 20 butir soal pilihan ganda disajikan pada Tabel 5.

Tabel 5. Hasil Analisis Validitas Butir Soal Pilihan Ganda.

Butir Soal Valid	Butir Soal Tidak Valid
PG1, PG2, PG3, PG4, PG5, PG6, PG7, PG8, PG9, PG10, PG11, PG12, PG13, dan PG18	PG14, PG15, PG16, PG17, PG19, dan PG20

Berdasarkan hasil analisis, 14 butir soal pilihan ganda dinyatakan valid dan 6 butir dinyatakan tidak valid, sedangkan seluruh 5 butir soal uraian dinyatakan valid. Uji reliabilitas menunjukkan bahwa instrumen memiliki tingkat konsistensi internal yang tinggi sehingga tergolong reliabel. Analisis tingkat kesukaran menunjukkan bahwa 11 butir berada pada kategori sedang dan 9 butir pada kategori mudah, tanpa adanya butir berkategori sukar. Dominasi soal berkategori sedang merupakan kondisi yang ideal karena memungkinkan instrumen mengukur kemampuan peserta didik secara proporsional. Sementara itu, analisis daya pembeda menunjukkan sebagian besar butir berada pada kategori baik hingga sangat baik,

meskipun butir PG14, PG15, PG16, PG17, PG19, dan PG20 memiliki daya pembeda rendah sehingga perlu direvisi.

Revisi Produk Akhir

Butir soal yang belum valid dan berdaya pembeda rendah direvisi dengan mengacu pada hasil analisis butir dan masukan validator. Revisi difokuskan pada memperjelas konteks, memperkuat hubungan sebab-akibat, meningkatkan kualitas distraktor, dan menyesuaikan level kognitif agar lebih sesuai dengan karakteristik HOTS. Sebagai contoh, butir nomor 14 direvisi dengan menambahkan dimensi waktu dan memperjelas dampak ekologis jangka panjang, sehingga peserta didik tidak hanya mengenali hubungan antar komponen ekosistem, tetapi juga menganalisis konsekuensi perubahan populasi terhadap keseimbangan ekosistem. Demikian pula, butir nomor 16 direvisi untuk meningkatkan level berpikir dari sekadar mengenali ciri organisme menjadi menganalisis bentuk adaptasi terhadap faktor abiotik ekstrem.

Pembahasan

Hasil pengembangan menunjukkan bahwa kualitas instrumen tidak hanya ditentukan oleh kesesuaian isi soal dengan indikator pembelajaran, tetapi juga oleh hasil pengujian empiris terhadap validitas, reliabilitas, tingkat kesukaran, daya pembeda, serta respon pengguna. Butir yang valid menunjukkan adanya kesesuaian antara stimulus, tuntutan kognitif, dan indikator kompetensi yang hendak diukur, sedangkan butir yang belum valid mengindikasikan kelemahan pada aspek konstruk, redaksi, atau keterwakilan materi. Temuan ini sejalan dengan konsep HOTS yang menekankan pengukuran kemampuan berpikir pada ranah menganalisis, mengevaluasi, dan mencipta (Rengkuan et al., 2025; Sumampouw, 2021), sehingga revisi butir menjadi penting agar setiap item benar-benar mengukur proses berpikir yang kompleks, bukan sekadar ketepatan bentuk soal.

Reliabilitas yang tinggi menegaskan bahwa instrumen memiliki konsistensi internal yang baik dan mampu memberikan hasil yang relatif ajeg apabila digunakan pada kondisi serupa. Distribusi tingkat kesukaran yang terpusat pada kategori sedang menunjukkan bahwa instrumen memberikan tantangan kognitif yang proporsional sesuai karakteristik HOTS, karena soal yang terlalu mudah cenderung tidak mampu membedakan kemampuan peserta didik, sedangkan soal yang terlalu sukar berpotensi menurunkan fungsi ukur instrumen. Temuan ini konsisten dengan penelitian pengembangan instrumen HOTS sebelumnya, yang menekankan bahwa validitas dan reliabilitas instrumen sangat ditentukan oleh ketepatan konstruk, kejelasan redaksi, serta keterwakilan materi (Latifah et al., 2023; Purnasari et al., 2024).

Keefektifan instrumen tercermin dari nilai rata-rata uji coba sebesar 77,25 (kategori baik) yang menunjukkan bahwa peserta didik mampu merespons tuntutan kognitif pada setiap butir, tidak hanya pada level mengingat atau memahami, tetapi juga pada level analisis, evaluasi, dan pemecahan masalah. Adanya variasi skor menunjukkan bahwa instrumen memiliki daya pembeda yang memadai dalam mengidentifikasi perbedaan tingkat kemampuan peserta didik. Dukungan hasil angket respons yang sangat valid (89,87) memperkuat temuan tersebut, sekaligus menunjukkan bahwa produk asesmen tidak hanya layak secara teoritis, tetapi juga fungsional secara praktis. Dengan demikian, secara keseluruhan instrumen asesmen berbasis HOTS yang dikembangkan telah memenuhi kriteria valid, praktis, dan efektif, meskipun beberapa butir tetap perlu penyempurnaan agar kualitas pengukuran menjadi lebih optimal.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian dapat disimpulkan bahwa instrumen asesmen berbasis HOTS pada materi komponen ekosistem dan interaksinya yang dikembangkan dengan model 3-D telah memenuhi kriteria kelayakan ditinjau dari aspek validitas, kepraktisan, dan keefektifan. Ditinjau dari validitas, instrumen memperoleh skor validasi ahli rata-rata 97,2 dengan kriteria sangat valid, dan sebagian besar butir soal dinyatakan valid serta reliabel setelah dilakukan analisis empiris. Ditinjau dari kepraktisan, instrumen tergolong praktis karena mudah dipahami dan digunakan, yang ditunjukkan oleh skor respons peserta didik sebesar 89,87 dengan kriteria sangat valid. Ditinjau dari keefektifan, instrumen terbukti efektif dengan nilai rata-rata uji coba sebesar 77,25 pada kategori baik serta mampu membedakan tingkat kemampuan peserta didik secara proporsional. Dengan demikian, instrumen ini dinyatakan layak digunakan sebagai alat evaluasi pembelajaran Biologi, khususnya dalam mengukur kemampuan berpikir tingkat tinggi peserta didik.

Berdasarkan simpulan tersebut, disarankan agar pengembangan instrumen selanjutnya memperhatikan distribusi tingkat kesukaran secara lebih seimbang dengan menambahkan butir berkategori sukar, serta merevisi secara berkelanjutan butir yang belum valid dan berdaya pembeda rendah pada aspek konstruksi soal, kejelasan redaksi, dan kualitas distraktor. Bagi pendidik, penggunaan instrumen asesmen berbasis HOTS perlu diintegrasikan secara sistematis dalam pembelajaran agar peserta didik terbiasa menghadapi soal yang menuntut berpikir kritis, analitis, dan evaluatif. Penelitian selanjutnya diharapkan menguji instrumen pada sampel yang lebih luas dan beragam guna meningkatkan generalisasi dan validitas eksternal, serta memposisikan instrumen tidak hanya sebagai alat evaluasi tetapi juga sebagai

bagian integral dari strategi pembelajaran yang bermakna. Keterbatasan penelitian ini terletak pada cakupan subjek yang terbatas pada satu kelas di satu sekolah, sehingga hasilnya belum dapat digeneralisasi secara luas.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Universitas Negeri Manado, khususnya Program Studi Pendidikan Biologi, serta kepada pihak SMA Negeri 1 Tondano, para validator, guru, dan peserta didik yang telah berpartisipasi dalam penelitian ini. Artikel ini merupakan bagian dari skripsi penulis pada Program Studi S1 Pendidikan Biologi, Fakultas Matematika, Ilmu Pengetahuan Alam, dan Kebumian, Universitas Negeri Manado.

DAFTAR REFERENSI

- Amalia, A. R., Rusdi, R., & Kamid, K. (2021). Pengembangan soal matematika bermuatan HOTS setara PISA berkonteks Pancasila. *Jurnal Cendekia: Jurnal Pendidikan Matematika*, 5(1), 1–19. <https://doi.org/10.31004/cendekia.v5i1.386>
- Azwar, S. (2023). *Reliabilitas dan validitas*. Pustaka Pelajar.
- Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain*. David McKay Company.
- Domu, I., & Mangelep, N. O. (2019). Developing of mathematical learning devices based on the local wisdom of the Bolaang Mongondow for elementary school. *Journal of Physics: Conference Series*, 1387(1), Article 012135. <https://doi.org/10.1088/1742-6596/1387/1/012135>
- Domu, I., & Mangelep, N. O. (2025). Development of mathematics learning model based on Mapalus values: Efforts to improve mathematical literacy and students' social skills. *Journal of Cultural Analysis and Social Change*, 1509–1519. <https://doi.org/10.64753/jcasc.v10i4.3038>
- Harefa, E., Afendi, A. R., Karuru, P., Sulaeman, & Wote, A. Y. V. (2024). *Buku ajar: Teori belajar dan pembelajaran*. Deepublish.
- Kalengkongan, L. N., Regar, V. E., & Mangelep, N. O. (2021). Analisis kesalahan siswa dalam menyelesaikan soal cerita pokok bahasan program linear berdasarkan prosedur Newman. *Marisekolah: Jurnal Matematika Riset Edukasi dan Kolaborasi*, 2(2), 31–38. <https://doi.org/10.53682/marisekola.v2i2.1102>
- Latifah, A., Dewi, I. P., & Nugraha, R. (2023). Pengembangan instrumen penilaian berbasis HOTS untuk mengukur kemampuan berpikir kritis peserta didik SMA. *Jurnal Didaktik*, 12(2), 56–65. <https://doi.org/10.36989/didaktik.v9i2.1057>
- Lestari, K. E., & Yudhanegara, M. R. (2024). *Penelitian pendidikan matematika*. Refika Aditama.
- Lewis, A., & Smith, D. (2023). Defining higher order thinking. *Theory Into Practice*, 32(3), 131–137. <https://doi.org/10.1080/00405849309543588>

- Mangelep, N. O. (2013). Pengembangan soal matematika pada kompetensi proses koneksi dan refleksi PISA. *Jurnal Edukasi Matematika*, 4(7), 451–466.
- Mangelep, N. O. (2015). Pengembangan soal pemecahan masalah dengan strategi finding a pattern. In *Konferensi Nasional Pendidikan Matematika-VI (KNPM6): Prosiding* (pp. 104–112).
- Mangelep, N. O. (2017). Pengembangan perangkat pembelajaran matematika pada pokok bahasan lingkaran menggunakan pendekatan PMRI dan aplikasi GeoGebra. *Mosharafa: Jurnal Pendidikan Matematika*, 6(2), 193–200. <https://doi.org/10.31980/mosharafa.v6i2.440>
- Mangelep, N. O., & Kaunang, D. F. (2018). Pengembangan soal matematika realistik berdasarkan kerangka teori Program for International Student Assessment. *Mosharafa: Jurnal Pendidikan Matematika*, 7(3), 455–466. https://journal.institutpendidikan.ac.id/index.php/mosharafa/article/viewFile/mv7n3_14/376
- Mangelep, N. O., Pongoh, F. M., & Bander, S. E. (2026). Digital media and K–12 mathematical literacy: A systematic review of comparative studies. [Nama jurnal tidak tersedia], 9(1), 174–198.
- Maydiantoro, A. (2025). Model-model penelitian pengembangan (research and development). *Jurnal Penelitian Pendidikan Indonesia*, 5(2), 45–52.
- Nurkencana, W. (2023). *Evaluasi pendidikan*. Usaha Nasional.
- Purnasari, P. D., Silvester, & Lumbantobing, W. L. (2024). Pengembangan instrumen asesmen Higher Order Thinking Skills (HOTS) ditinjau dari gaya belajar siswa. *Sebatik*, 25(2), 571–578. <https://doi.org/10.46984/sebatik.v25i2.1607>
- Rengkuan, M. (2025). Enhancing students' critical thinking skills: Developing a biology practicum module using the PjBL model by integrating local Banana Goroho. *Edubiotik: Jurnal Pendidikan, Biologi dan Terapan*, 10(2), 465–482. <https://doi.org/10.33503/ebio.v10i02.2154>
- Runturambi, I., Pulukadang, R. J., Mangelep, N. O., Tuerah, P. E., & Sumarauw, S. J. (2026). Development of a realistic mathematics teaching module on arithmetic sequences and series to support the mathematical literacy of vocational high school students. *International Journal of Social Sciences and Humanities*, 4(2), 556–568.
- Somba, C. M., Domu, I., Mangelep, N. O., Runtu, P. V., & Tumulun, N. K. (2026). Development of a social arithmetic teaching module based on the Indonesian realistic mathematics education approach to support the numeracy skills of junior high school students. *International Journal of Social Sciences and Humanities*, 4(2), 580–594.
- Sudjana, N. (2025). *Penilaian hasil proses belajar mengajar*. Remaja Rosdakarya.
- Sumampouw, H. J. (2021). Media pembelajaran konsep genetika menggunakan *Drosophila melanogaster* isolat lokal dengan aplikasi WhatsApp berbasis kearifan lokal budaya Mapalus. *Jurnal Pendidikan Biologi Indonesia*, 7(3), 178–192.
- Thobroni, M., & Mustofa, A. (2025). *Belajar dan pembelajaran: Teori dan praktik*. Ar-Ruzz Media.
- Tomei, L. A. (2025). *Taxonomy for the technology domain*. IGI Global.

- Underbakke, M., Borg, J. M., & Peterson, P. L. (2023). Predicting students' critical thinking performance in higher education. *Journal of Educational Psychology*, 85(4), 531–540.
- Winarti, & Istiyono, E. (2025). *Taksonomi Higher Order Thinking Skill (HOTS)*. Widya Sari Press.
- Yaumi, M., & Hum, M. (2022). *Prinsip-prinsip desain pembelajaran*. Kencana.