

PENGEMBANGAN INSTRUMEN SOAL *HIGHER ORDER THINKING SKILLS* (HOTS) MATERI JARINGAN DAN ORGAN PADA TUMBUHAN KELAS XI SMA***Development of Higher Order Thinking Skills (HOTS) Assessment in Material of Tissue and Organ Plants for Grade 11th in Senior High School*****Farah Fidia**

Pendidikan Biologi, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Surabaya
e-mail: farah.18035@mhs.unesa.ac.id

Rinie Pratiwi Puspitawati

Pendidikan Biologi, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Surabaya
e-mail: riniepratiwi@unesa.ac.id

Pramita Yakub

Pendidikan Biologi, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Surabaya
e-mail: pramitayakub@unesa.ac.id

Abstrak

Penelitian ini bertujuan untuk menghasilkan instrumen penilaian HOTS yang valid dan reliabel ditinjau dari validitas teoritis, validitas empiris, reliabilitas, tingkat kesukaran, daya pembeda, dan respon peserta didik. Metode penelitian pengembangan ini adalah model pengembangan ADDIE (*analyze, design, development, implementation, dan evaluation*). Instrumen soal HOTS divalidasi teoritis oleh ahli pendidikan, ahli materi, dan guru biologi SMA mengacu pada instrumen berskala Guttman. Uji coba terbatas dilakukan pada 20 peserta didik kelas XI IPA 3 SMAN Puri Mojokerto untuk mengetahui hasil validitas empiris, tingkat kesukaran, daya pembeda, dan respon peserta didik. Hasil validitas teoritis dinyatakan sangat valid dengan nilai sebesar 90,67% pada aspek materi, 90% pada aspek konstruksi, dan 100% pada aspek bahasa. Hasil uji validitas empiris termasuk kategori valid dengan nilai sebesar 80%. Nilai reliabilitas termasuk kategori sangat reliabel dengan nilai sebesar 0,818. Indeks tingkat kesukaran butir soal memperoleh hasil 30% soal sukar, 70% soal sedang. Indeks daya pembeda butir soal menunjukkan 30% kurang, 30% cukup, dan 40% baik. Respon peserta didik diperoleh 92% kategori sangat baik. Dari hasil yang diperoleh instrumen soal HOTS materi jaringan dan organ pada tumbuhan termasuk instrumen yang valid dan reliabel.

Kata Kunci: instrumen penilaian, HOTS, soal uraian, jaringan dan organ tumbuhan.

Abstract

This research aims to produce of HOTS assessment instrument that valid and reliable considered from validity theory, empiric validity, reliability, the difficulty of question, the differences effort of question, and respond of student. The method in this research development is used ADDIE (analyze, design, development, implementation, and evaluation). Validity theory instrument in HOTS question already exams by expert instructor, material expert, and biology teacher in senior high school used instruments in Guttman scale. Limited observation done by 20 student in grade XI IPA 3 SMAN 1 Puri Mojokerto to know the result of empiric validity, reliability, difficulty, the difference, and respond of student. The result of theory validity is obviously very valid if the score 90,67% in material aspect, 90% in construction aspect, 100% in language aspect. The result of empiric validity belongs to valid if the score 80%. The reliability score belongs to very reliable if the score is 0,818. Difficulty level index of a question got 30% difficult question and 70% medium question. The different effort index of question showed 30% is low, 30% enough, and 40% is good. The respond of student got 92% belongs to very good. From the result of instrument in tissue and organ plants belongs to valid and reliable instrument.

Keywords: assessment instrument, HOTS, essay, tissue and organ plants.

PENDAHULUAN

Keterampilan berfikir tingkat tinggi atau *Higher Order Thinking Skills* (HOTS) menjadi salah satu kompetensi yang harus dimiliki oleh semua peserta didik

setelah mengalami proses belajar. Hal tersebut sesuai dengan tuntutan keterampilan-keterampilan pada dunia pendidikan pada abad 21. Menurut Wardani & Ibrahim (2020) kemampuan berfikir tingkat tinggi menuntut peserta didik untuk memperoleh informasi lewat adanya

proses menganalisis, mensintesis, dan mengevaluasi. Proses pemerolehan informasi dilakukan dengan mengaitkan antara pengalaman yang dimiliki dengan materi yang dipelajari. Berfikir tingkat tinggi memiliki keterkaitan dengan kualitas pembelajaran.

Kualitas pembelajaran salah satunya dapat diketahui dari hasil penilaiannya. Penilaian bertujuan untuk mengetahui tingkat keberhasilan peserta didik mencapai kompetensi pembelajaran. Penilaian termasuk komponen penting dalam proses belajar. Upaya peningkatan kualitas pembelajaran dapat dilakukan dengan meningkatkan kualitas penilaian. Penilaian tidak hanya digunakan dalam mengetahui hasil belajar namun, juga dapat digunakan sebagai faktor penting keberhasilan proses belajar peserta didik. Kemampuan berfikir tingkat tinggi dapat diukur dengan menggunakan instrumen penilaian HOTS. Instrumen penilaian HOTS berisi soal-soal dengan ranah C-4 (analisis), C-5 (evaluasi), dan C-6 (kreasi). Penilaian HOTS memiliki beberapa karakteristik yaitu, soal-soal yang disajikan berkaitan dengan permasalahan yang terjadi disekitar peserta didik, serta bentuk soal yang disajikan beragam (Fanani, 2018).

Menurut penelitian yang dilakukan Avina & Winarsih (2020) di SMAN 1 Bangil, dari keseluruhan soal yang diberikan pada peserta didik hanya 30% ranah soal HOTS yang diberikan. Tidak jauh berbeda penelitian yang dilakukan Wardani & Ibrahim (2020) di SMAN 1 Menganti menunjukkan bahwa guru biologi kelas XI yang belum mengembangkan instrumen penilaian HOTS sebesar 66,7% dan 33,3% guru mengembangkan instrumen penilaian HOTS yang belum valid dan belum reliabel. Sebagian besar guru mengalami kesulitan untuk membedakan materi yang dapat digunakan dalam mengembangkan instrumen penilaian HOTS. Ketika kondisi ini tidak ditindak lanjuti maka, akan berdampak pada kemampuan berfikir tingkat tinggi peserta didik yang tidak terasah. Peserta didik akan kesulitan ketika diberikan soal-soal dengan ranah berfikir tingkat tinggi karena hanya terbiasa mengerjakan soal-soal dengan ranah LOTS (*Lower Order Thinking Skills*).

Seperti yang ditunjukkan oleh hasil survei PISA pada 2012 Indonesia ada pada urutan 64 dari 65 negara. Pada tahun 2015 dari 70 negara Indonesia ada di urutan 63. (OECD, 2015). Hasil PISA tahun 2018 menunjukkan bahwa Indonesia berada pada urutan 72 dari 77 negara (OECD, 2018). Rendahnya hasil PISA yang diperoleh dapat disebabkan karena sistem *assessment* yang digunakan dalam PISA berbasis HOTS. Tujuannya agar peserta didik tidak hanya menguasai materi namun, dapat memecahkan masalah yang dihadirkan (Pratiwi, 2019).

Berdasarkan beberapa kejadian tentang kemampuan berfikir tingkat tinggi peserta didik yang rendah maka, saat ini diperlukannya pengembangan instrumen penilaian HOTS. Manfaat instrumen penilaian HOTS bagi peserta didik antara lain yaitu, dapat meningkatkan motivasi belajar, meningkatkan kreativitas, dan meningkatkan rasa percaya diri (Iskandar, 2015). Pemetaan kompetensi dasar mata pelajaran biologi Kurikulum 2013 mengidentifikasi bahwa 23 dari 35 kompetensi 3 menargetkan untuk berfikir tingkat tinggi (Kemendikbud, 2016). Hasil analisis kompetensi dasar 3 Kurikulum 2013 menunjukkan bahwa penyebaran dimensi kognitif didominasi oleh C-3 dan tingkat C-4, diikuti pencapaian tingkat C-2 sebesar 15% (Herlanti, 2015).

Salah satu materi yang menargetkan pencapaian berfikir tingkat tinggi adalah materi jaringan dan organ tumbuhan. Materi biologi kelas XI ini terdapat pada KD. 3.3 Menganalisis keterkaitan antara struktur sel pada jaringan tumbuhan dengan fungsi organ pada tumbuhan. Materi jaringan dan organ tumbuhan adalah materi yang didalamnya mempelajari tentang struktur sel yang ada pada jaringan tumbuhan serta keterkaitan antara struktur tersebut dengan fungsinya pada organ tumbuhan selain itu, berisi tentang organ-organ yang menyusun tumbuhan. Menurut penelitian yang telah dilakukan Baidlowi (2019) di SMAN 1 Tumpang materi jaringan dan organ pada tumbuhan termasuk materi yang sulit untuk dipahami. Guru mengalami kendala untuk membuat indikator dan instrumen tes. Bentuk soal yang diberikan guru kepada peserta didiknya adalah soal-soal menghafal dan mengingat. Kondisi tersebut tidak sesuai dengan tuntutan KD yakni peserta didik diarahkan untuk melakukan analisis.

Berdasarkan uraian masalah tersebut, penelitian ini memiliki tujuan menghasilkan instrumen soal HOTS yang valid dan reliabel pada materi jaringan dan organ tumbuhan. Instrumen tersebut ditinjau dari validitas teoritis, validitas empiris, reliabilitas, tingkat kesukaran butir soal, daya pembeda dan respon peserta didik yang dapat dikerjakan secara *online* dan *offline* agar dapat memfasilitasi ketersediaan soal HOTS.

METODE

Penelitian ini mengacu pada model pengembangan ADDIE (*analysis-design-development-implementation-evaluation*). Analisis hasil penelitian yang digunakan dalam penelitian ini dilakukan secara kualitatif dan kuantitatif. Tahap *analysis* terdiri dari analisis kurikulum, analisis kompetensi dasar, analisis konsep, dan peserta didik. Tahap kedua yaitu *design*. Instrumen

soal HOTS yang dikembangkan terdiri dari 10 soal uraian materi jaringan dan organ tumbuhan. Tahap yang dilakukan sebelum membuat butir soal adalah memperinci kompetensi dasar menjadi beberapa indikator.

Tahap ketiga yang dilakukan adalah *development*. Dalam tahap ini dilakukan pengembangan butir soal sesuai dengan indikator yang telah dikembangkan. Pengembangan instrumen soal HOTS materi jaringan dan organ pada tumbuhan kelas XI SMA dilakukan di Jurusan Biologi FMIPA Unesa pada bulan November – Januari 2022. Sebelum instrumen soal HOTS diujicobakan terbatas ke peserta didik instrumen soal HOTS tersebut diujikan kepada dosen, dan guru biologi SMA. Instrumen yang digunakan meliputi lembar validasi pada aspek materi, konstruksi, dan bahasa.

Uji validitas teoritis yang digunakan yaitu pengukuran berskala Guttman dengan “Ya” ditandai tanda ceklis (✓) dan “Tidak” ditandai tanda (-). Perhitungan validitas teoritis dihitung menggunakan rumus berikut:

$$P = \frac{\sum \text{per kriteria yang diceklis } (\checkmark)}{\sum \text{validator}} \times 100$$

Tabel 1. Kriteria Hasil Validasi Teoritis

Persentase	Kategori
0%-25%	Tidak valid
26%-50%	Kurang valid
51%-75%	Valid
76%-100%	Sangat valid

Tahap keempat adalah *implementation* dan tahap yang terakhir adalah *evaluation*. Setelah soal HOTS diujicobakan terbatas kepada peserta didik hasil yang diperoleh digunakan dalam tahap evaluasi. Instrumen HOTS tersebut diujicobakan pada 20 peserta didik kelas XI IPA 3 SMAN 1 Puri Mojokerto pada bulan Februari 2022. Tahap *evaluation* dilakukan disetiap akhir tahapan model ADDIE yang dilakukan. Hasil tersebut dianalisis untuk mengetahui hasil validitas empiris, nilai reliabilitas, tingkat kesukaran, daya pembeda, dan respon peserta didik. Uji validitas empiris menggunakan bantuan *software* SPSS 23 dengan rumus kolerasi *product-momnt*.

Perhitungan nilai reliabilitas menggunakan rumus *cronbach's alpha* berikut:

$$\alpha = \frac{R}{R-1} \left(1 - \frac{\sum \sigma_i^2}{\sigma^2} \right)$$

Keterangan:

R = jumlah soal

σ_i^2 = varian soal

σ^2 = varian soal

Hasil yang diperoleh diinterpretasikan sesuai kriteria reliabilitas dari Sugiyono (2015) pada **Tabel 2**.

Tabel 2. Kriteria Reliabilitas

Interval reliabilitas	Kriteria
0,00< α ≤0,20	Tidak reliabel
0,20< α ≤0,40	Kurang reliabel
0,40< α ≤0,60	Cukup reliabel
0,60< α ≤0,80	Reliabel
0,80< α ≤1,00	Sangat reliabel

Analisis tingkat kesukaran soal uraian dihitung menggunakan perhitungan berikut:

$$TK = \frac{\bar{X} \text{ skor peserta didik per butir soal}}{\text{Skor maksimal}}$$

Kriteria tingkat kesukaran soal diinterpretasikan sesuai dengan indeks kesukaran butir soal dari Arikunto (2015).

Tabel 3. Indeks Kesukaran Butir Soal

Indeks kesukaran	Kategori
0,00-0,30	Sukar
0,31-0,70	Sedang
0,71-1,00	Mudah

Analisis daya pembeda diketahui dengan dengan melihat indeks daya pembeda soal. Nilai daya pembeda soal dihitung menggunakan rumus berikut:

$$DP = \frac{\bar{X} \text{ Kelompok atas} - \bar{X} \text{ Kelompok bawah}}{\text{Skor maksimal}}$$

Hasil perhitungan yang diperoleh diinterpretasikan pada **Tabel 4**.

Tabel 4. Indeks Daya Pembeda Soal

Indeks daya pembeda	Kategori
0,00-0,19	Kurang
0,20-0,29	Cukup
0,30-0,39	Baik
0,40-1,00	Sangat baik

Pengumpulan data terhadap respon peserta didik terhadap soal HOTS yang diberikan dilakukan dengan menggunakan angket respon. Respon peserta didik dihitung menggunakan rumus berikut:

$$R = \frac{\sum \text{peserta didik yang menjawab "Ya"}}{\sum \text{peserta didik yang mengisi angket}}$$

Hasil yang didapat dianalisis sesuai dengan kriteria berikut:

Tabel 5. Kriteria Respon Peserta Didik

Persentase	Kategori.
0%-20%	Sangat kurang
21%-40%	Kurang
41%-60%	Cukup
61%-80%	Baik
81%-100%	Sangat baik

HASIL DAN PEMBAHASAN

Penelitian ini menghasilkan produk berupa instrumen soal HOTS materi jaringan dan organ tumbuhan yang valid dan reliabel ditinjau dari hasil validitas teoritis, validitas teoritis, validitas empiris, tingkat kesukaran, daya pembeda, dan respon peserta didik. Pengembangan instrumen penilaian dimulai dengan melakukan analisis kurikulum dan kompetensi dasar.

Tahapan yang dilakukan selanjutnya adalah menyusun indikator soal. Indikator yang disusun disesuaikan dengan indikator HOTS, kriteria penyusunan soal HOTS, dimensi kognitif, dan dimensi proses (Rahmawati & Trimulyono, 2022). Indikator soal yang telah ditentukan akan digunakan untuk menyusun butir soal. Butir soal HOTS yang dikembangkan terdiri dari 10 butir soal tipe uraian dengan dimensi kognitif pada ranah C4 dan C5, serta dengan dimensi proses yang digunakan adalah konseptual dan metakognitif. HOTS Butir soal HOTS yang disusun berisikan grafik, gambar, dan tabel terkait kehidupan sehari-hari. Instrumen soal HOTS yang dikembangkan juga dilengkapi dengan pedoman penskoran. Menurut Yudiandani & Asri (2022) pedoman penskoran berfungsi untuk membantu guru dalam memberikan penilaian agar terarah dan objektif. Kisi-kisi dan contoh salah satu butir soal HOTS yang dikembangkan disajikan pada **Tabel 6** dan **gambar 2** berikut:

Tabel 6. Kisi-kisi Butir Soal HOTS

Kompetensi Dasar	Materi	Indikator Pencapaian Kompetensi	Level Kognitif	Dimensi Pengetahuan	No. Soal	Bentuk Soal
3.3 Menganalisis keterkaitan antara struktur sel pada jaringan tumbuhan dengan fungsi organ pada tumbuhan	Struktur jaringan tumbuhan	3.3.1 Menganalisis struktur jaringan akar tumbuhan pada kondisi lingkungan yang berbeda	C4 (Menganalisis)	Konseptual	1	Uraian
		3.3.2 Menentukan respon jaringan tumbuhan akibat adanya perbedaan kondisi lingkungan tumbuh	C4 (Menganalisis)	Konseptual	2, 5, 10	Uraian
		3.3.3 Mengaitkan struktur jaringan tumbuhan terhadap fungsi organ tumbuhan	C5 (Mengevaluasi)	Konseptual	3	Uraian
		3.3.4 Menganalisis keterkaitan lingkaran tahun yang terbentuk dengan kondisi lingkungan tumbuhan tersebut	C5 (Mengevaluasi)	Metakognitif	4	Uraian
		3.3.5 Menganalisis tipe berkas pengangkut suatu tumbuhan	C4 (Menganalisis)	Konseptual	9	Uraian
	Organ tumbuhan	3.3.6 Memprediksi respon organ tumbuhan terhadap pemberaian larutan tertentu	C5 (Mengevaluasi)	Metakognitif	6, 7, 8	Uraian

Rino sedang membantu ayahnya membuat kerajinan dari bahan kayu. Saat memotong kayu diketahui bahwa pada batang pohon tersebut terdapat lingkaran-lingkaran yang berbeda warna seperti pada gambar. Pohon yang dijadikan kerajinan tersebut tumbuh ditempat dengan kadar air yang rendah karena berada pada daerah dengan kondisi kemarau. Lingkaran-lingkaran tahun pada batang pohon yang dilihat Rino berbentuk tipis dan berwarna gelap karena susunan sel yang terbentuk rapat.



Gambar 1. Lingkaran tahun pada batang pohon

Sumber: dunia tumbuhan, tira pustaka

Prediksilah apakah yang akan terjadi pada lingkaran tahun yang terbentuk jika pohon tersebut tumbuh didaerah yang memiliki intensitas curah hujan yang tinggi?

Gambar 2. Contoh salah satu butir soal HOTS yang dikembangkan

Pada soal tersebut peserta didik diarahkan untuk membuat prediksi kemungkinan yang akan terjadi pada lingkaran tahun yang terbentuk sesuai kondisi pada soal. Proses memprediksi yang dilakukan peserta didik termasuk berfikir pada level kognitif C5 (mengevaluasi) (Ramadhan & Wasis, 2013). Proses memprediksi sesuai juga dengan model pemrosesan informasi yaitu, pada tahap penerapan prinsip dimana peserta didik melakukan prediksi terhadap hal yang mungkin terjadi dari suatu permasalahan yang dihadirkan (Siharli, 2015). Ketika melakukan proses prediksi peserta didik sebelumnya harus sudah memiliki konsep terkait dengan proses pembentukan lingkaran tahun sesuai dengan materi yang telah dipelajari. Pernyataan tersebut sesuai dengan Fitri, dkk (2020) bahwa kemampuan memprediksi didasarkan pada konsep sebelumnya yang telah diperoleh peserta didik atau data yang disajikan.

Soal yang disajikan apabila dikaitkan dengan materi dalam proses pembelajaran peserta didik dapat berisikan informasi bahwa proses pembentukan lingkaran tahun terjadi akibat aktivitas pembelahan kambium yang dipengaruhi oleh kondisi lingkungan. Saat intensitas curah hujan rendah pembelahan akan berlangsung lebih lama karena air yang diserap tumbuhan hanya sedikit. Hal tersebut berpengaruh pada lingkaran tahun yang terbentuk susunannya tipis dan berwarna gelap. Dari

stimulus pada soal tentang kondisi lingkaran tahun di kondisi intensitas curah hujan rendah. Peserta didik dapat membuat prediksi ilmiah tentang kondisi lingkaran tahun pada kondisi intensitas curah hujan tinggi dengan mengaitkan informasi yang didapatkan pada soal dengan pengetahuan yang dimiliki. Menurut Widana (2020) soal dengan stimulus sesuai dengan kehidupan sehari-hari berkaitan erat dengan model soal-soal berbasis HOTS. Stimulus yang disajikan memiliki tujuan untuk mendorong peserta didik untuk menganalisis dan literasi.

Pengembangan butir soal yang dilakukan melalui beberapa langkah. Instrumen soal yang dibuat awal merupakan instrumen soal draft 1. Instrumen pada draft 1 yang belum sesuai akan dievaluasi sehingga menghasilkan instrumen soal draft 2. Instrumen soal draft 2 ini yang diujikan pada ahli pendidikan, ahli materi, dan guru SMA. Tahapan setelah proses validasi yaitu instrumen soal direvisi sesuai dengan hasil evaluasi validator sebelum soal tersebut diuji coba pada peserta didik. Uji coba terbatas dilakukan selama 100 menit melalui *googleform*.

Validitas teoritis

Validitas teoritis pada soal HOTS yang dikembangkan diperoleh dari lembar hasil validasi yang diujikan pada ahli pendidikan, ahli materi, dan guru biologi SMA. Hasil validitas teoritis disajikan pada **Tabel 7**.

Tabel 7. Hasil Validitas Teoritis

No. Soal	Persentase rerata aspek yang ditelaah dari 3 validator (%)		
	Materi	Konstruksi	Bahasa
1.	86,67	88,89	100
2.	93,33	83,34	100
3.	86,67	88,89	100
4.	93,33	94,45	100
5.	86,67	88,89	100
6.	93,33	94,45	100
7.	86,67	88,89	100
8.	93,33	94,45	100
9.	86,67	88,89	100
10.	100	88,89	100
Σ	90,67	90,00	100

Terdapat 5 kriteria pada aspek materi yang ditelaah oleh validator. Kriteria tersebut terdiri dari 1) soal sesuai indikator, 2) soal disajikan sesuai kebenaran konsep, 3) kejelasan antara batasan pertanyaan dan jawaban yang diharapkan, 4) isi materi yang ditanyakan sesuai dengan jenjang, 5) butir soal yang dikembangkan relevan pada dimensi kognitif C4-C6 sesuai dengan yang terdapat pada kisi-kisi soal.

Pada aspek konstruksi terdapat 6 kriteria yang ditelaah adalah 1) terdapat petunjuk soal yang mudah dimengerti, 2) stimulus yang digunakan menarik, 3) stimulus (grafik/gambar, kasus, hasil penelitian, dll sesuai dunia nyata) disajikan jelas dan mudah terbaca, 4) stimulus yang digunakan bermakna, 5) kata tanya yang digunakan pada soal menuntut jawaban terurai, 6) terdapat pedoman penskoran.

Pada aspek bahasa terdapat 5 kriteria yang ditelaah yaitu, 1) butir soal menggunakan bahasa Indonesia sesuai kaidah, 2) bahasa disoal komunikatif, 3) kalimat soal tidak menimbulkan penafsiran ganda, 4) kalimat soal tidak menyinggung perasaan, 5) bahasa yang digunakan sesuai untuk tingkat peserta didik.

Berdasarkan data pada butir soal nomor 10 dari 3 aspek yang ditelaah terdapat 2 aspek yang mendapatkan rerata hasil validitas sebesar 100% pada aspek materi dan aspek bahasa. Pada aspek konstruksi dengan kriteria nomor 1 dan 4 mendapatkan hasil rerata validitas sebesar 66,67%. Dari data tersebut butir soal nomor 10 dapat diterima. Butir soal nomor 2 mendapatkan rerata hasil validitas aspek materi sebesar 66,67% pada kriteria nomor 1, sedangkan pada aspek konstruksi terdapat 3 kriteria yang yang mendapat rerata sebesar 66,67% yaitu pada kriteria nomor 1, 3, dan 4. Dari data tersebut butir soal nomor 2 dinyatakan dapat diterima setelah direvisi. Butir soal nomor 4, 6, dan 8 mendapatkan rerata hasil validitas pada aspek materi sebesar 66,67% pada kriteria nomor 3. Aspek konstruksi mendapatkan rerata sebesar 66,67% pada kriteria nomor 1. Dari data tersebut butir soal nomor 4, 6, 8 dinyatakan dapat diterima dengan dilakukan revisi.

Butir soal nomor 3, 5, dan 7 mendapatkan rerata nilai validitas aspek materi sebesar 66,67% pada kriteria nomor 1 dan 3, sedangkan aspek konstruksi mendapatkan rerata sebesar 66,67% pada kriteria nomor 1 dan 4. Dari data tersebut butir soal nomor 4, 6, 8 dinyatakan dapat diterima dengan dilakukan revisi. Butir soal nomor 1 mendapatkan rerata hasil validitas aspek materi sebesar 66,67% pada kriteria 2 dan 3. Rerata hasil validitas aspek konstruksi sebesar 66,67% pada kriteria 1 dan 5. Butir soal nomor 9 mendapatkan rerata sebesar 66,67% aspek materi pada kriteria nomor 3 dan 5, sedangkan aspek konstruksi pada kriteria nomor 1 dan 4. Butir soal nomor 1 dan 9 diterima dengan dilakukan revisi. Hasil keseluruhan dari 10 butir soal yang dikembangkan memperoleh kategori sangat valid dengan aspek materi mendapatkan persentase rerata hasil validitas sebesar 90,67%, aspek konstruksi 90,00%, dan aspek bahasa 100%.

Validitas Empiris

Hasil dari validitas empiris butir soal HOTS ditunjukkan pada **Tabel 8**.

Tabel 8. Hasil Validitas Empiris Butir Soal HOTS

No. Soal	r Hitung	Kategori
1.	0,817	Valid
2.	0,704	Valid
3.	0,481	Valid
4.	0,722	Valid
5.	0,688	Valid
6.	-0,152	Tidak valid
7.	0,348	Tidak valid
8.	0,719	Valid
9.	0,728	Valid
10.	0,769	Valid

Berdasarkan data yang diperoleh diketahui bahwa dari 10 soal yang dikembangkan terdapat 2 soal tidak valid dan 8 soal valid. Soal dinyatakan valid ketika nilai r tabel lebih rendah dibanding r hitung dan sebaliknya ketika r tabel lebih tinggi dibanding r hitung soal akan dianggap tidak valid (Arikunto, 2015). Nilai r tabel pada perhitungan ini adalah 0,444. Soal nomor 6 dan 7 tidak valid karena r hitung yang didapatkan pada soal tersebut berturut-turut sebesar -0,152 dan 0,348 nilai tersebut <0,444. Berdasarkan hasil yang diperoleh nilai r hitungnya lebih rendah dibanding r tabel. Instrumen soal yang dinyatakan valid oleh validator ketika diujicoba pada peserta didik belum pasti hasilnya valid juga (Ulfa & Kuswanti, 2021). Terdapat beberapa faktor yang dapat mempengaruhi soal dinyatakan tidak valid antara lain yaitu, kondisi peserta didik saat mengerjakan soal. Ketika peserta didik yang biasanya hanya mengerjakan soal dengan ranah berfikir C1-C3 dihadapkan pada soal kemampuan berfikir tingkat tinggi maka, peserta didik akan mengalami kesulitan dalam memahami soal. Saat peserta didik dihadapkan pada soal yang berkaitan dengan pemecahan masalah jawaban yang muncul akan beragam tergantung pemahaman masing-masing peserta didik terhadap materi yang telah disampaikan (Nurlia dkk, 2017). Validitas suatu soal juga dipengaruhi oleh tingkat kesukaran dari soal tersebut. Saat tingkat kesukaran tidak sesuai dengan materi yang dikuasai peserta didik maka, peserta didik kesulitan dalam mengerjakan soal dan berpengaruh pada nilai validitas soal (Mahmudah dkk, 2016).

Reliabilitas

Dalam mengembangkan instrumen yang digunakan untuk mengumpulkan data, nilai reliabilitas menjadi salah satu aspek yang harus diketahui. Reliabilitas reliabel saat digunakan akan menghasilkan informasi yang tidak jauh berbeda walaupun penguji atau subjek uji cobanya berbeda-beda (Retnawati, 2016). Nilai reliabilitas disajikan pada **Tabel 9**.

Tabel 9. Hasil reliabilitas Instrumen Soal HOTS

Nilai Reliabilitas	Kriteria
0,818	Sangat reliabel

Dari data tersebut soal HOTS yang dikembangkan termasuk soal dalam kriteria sangat reliabel dengan nilai sebesar 0,818. Ketika nilai reliabilitas suatu instrumen $\leq 0,60$, maka instrumen tersebut termasuk kriteria yang tidak reliabel dan sebaliknya saat suatu instrumen memiliki nilai reliabilitas $> 0,60$ termasuk instrumen yang reliabel (Sugiyono, 2015). Instrumen Soal HOTS yang dinyatakan reliabel menunjukkan bahwa soal HOTS yang dikembangkan akan memiliki hasil pengukuran yang sama walaupun dikerjakan diwaktu yang berbeda, sehingga instrumen soal HOTS tersebut dapat digunakan dalam mengukur kemampuan berfikir tingkat tinggi peserta didik berulang kali.

Tingkat Kesukaran

Besaran indeks tingkat kesukaran berada pada rentang 0,00-1,0. Jika nilai tingkat kesukaran $\leq 0,30$ soal tersebut tingkat kesukarannya tinggi sedangkan jika memiliki nilai $\geq 0,70$ maka soal tersebut tingkat kesukarannya rendah (Arikunto, 2015). Penyusunan butir soal pada instrumen penilaian harus memperhatikan tingkat kesukaran. Hal tersebut dilakukan agar dapat mengetahui hasil belajar peserta didik dan tingkat prestasi yang sesungguhnya (Azizah & Budijastuti, 2022).

Tabel 10. Hasil Tingkat Kesukaran Soal HOTS

No. Soal	Tingkat Kesukaran	Kriteria
1.	0,55	Sedang
2.	0,54	Sedang
3.	0,56	Sedang
4.	0,69	Sedang
5.	0,63	Sedang
6.	0,29	Sukar
7.	0,28	Sukar
8.	0,54	Sedang
9.	0,53	Sedang
10.	0,60	Sedang

Berdasarkan **Tabel 10** dari 10 soal yang dikembangkan tidak terdapat soal dengan kriteria mudah.

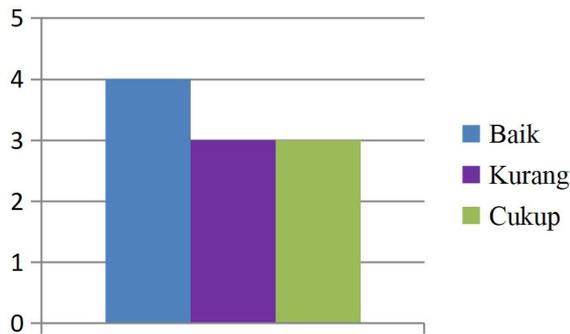
digunakan dalam menentukan konsistensi suatu instrumen penilaian agar instrumen tersebut dapat dipercaya dan sesuai dengan ketentuan. Alat ukur yang Terdapat 8 soal dengan hasil tingkat kesukaran berada pada rentang 0,31-0,70. Soal-soal tersebut termasuk soal dengan tingkat kesukaran sedang. Menurut Wardany (2018) soal yang berada pada kriteria sedang memiliki proporsi yang tepat karena soal tersebut dapat dikerjakan oleh peserta didik yang berada pada kelompok atas maupun kelompok bawah. Terdapat 2 soal dengan hasil tingkat kesukaran $< 0,30$ yaitu soal nomor 6 dan 7. Hasil tingkat kesukaran yang diperoleh berturut-turut adalah 0,288 dan 0,275. Berdasarkan uji coba terbatas faktor yang dapat mempengaruhi soal nomor 6 dan 7 dikatakan memiliki tingkat kesukaran yang tinggi adalah dari hasil validitas teoritis perlu dilakukan perbaikan soal karena soal tersebut menggunakan petunjuk soal yang sulit dipahami peserta didik, stimulus yang digunakan menimbulkan penafsiran yang berbeda, dan butir soal tidak sesuai dengan indikator. Oleh sebab itu, sebagian besar peserta didik kesulitan menjawab soal tersebut. Butir soal dikatakan dapat diterima ketika memiliki tingkat kesukaran dengan kriteria sedang. Butir soal dengan tingkat kesukaran pada kriteria mudah dan sukar dianggap sebagai soal yang kurang baik dan perlu dilakukan revisi terlebih dahulu (Hadi dkk, 2019).

Daya Pembeda

Daya pembeda merupakan kemampuan butir soal untuk mengukur perbedaan antara peserta didik pada kelompok berkemampuan tinggi (kelompok atas) dan peserta didik yang memiliki kemampuan rendah (kelompok bawah) (Masito & Aedi, 2020). Jika soal memiliki nilai daya pembeda mendekati 0,00 maka, soal tersebut memiliki daya pembeda yang kurang dan sebaliknya jika nilai daya pembeda mendekati 1,00 termasuk soal dengan daya pembeda sangat baik.

Tabel 11. Hasil daya Pembeda Soal HOTS

No. Soal	Daya Pembeda	Kriteria
1.	0,30	Baik
2.	0,33	Baik
3.	0,12	Kurang
4.	0,23	Cukup
5.	0,20	Cukup
6.	-0,03	Kurang
7.	0,05	Kurang
8.	0,23	Cukup
9.	0,30	Baik
10.	0,30	Baik



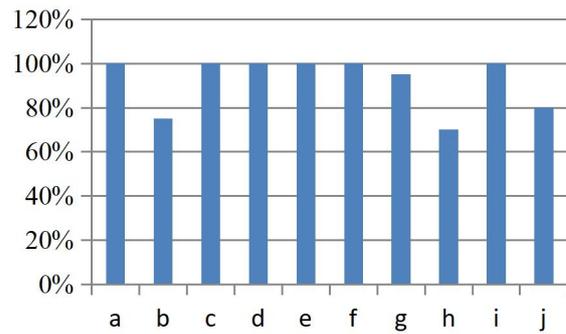
Gambar 3. Proporsi Daya Pembeda Soal HOTS

Berdasarkan data yang diperoleh dari pengembangan instrumen soal HOTS tersebut didapatkan hasil perhitungan indeks daya pembeda soal yaitu, terdapat 30% kurang, 30% baik, dan 40% baik. Soal pada kategori kurang artinya soal tidak dapat membedakan peserta didik yang berada pada kelompok atas dan kelompok bawah (Masito & Aedi, 2020). Soal nomor 3, 6, dan 7 ditolak atau diterima dengan dilakukan revisi terlebih dahulu. Faktor yang dapat mempengaruhi daya pembeda soal kurang adalah soal tersebut memiliki tingkat kesukaran yang tinggi seperti pada soal nomor 6 dan 7.

Soal kriteria cukup artinya soal tersebut diterima namun harus direvisi terlebih dahulu agar mendapatkan hasil yang lebih valid. Terdapat 3 soal dengan kriteria cukup yaitu soal nomor 4, 5, dan 8. Soal kriteria baik terdapat pada soal nomor 1, 2, 9, dan 10. Soal yang baik atau layak digunakan atau diterima jika daya pembeda soal tersebut berada pada kriteria baik atau sangat baik artinya soal tersebut dapat membedakan peserta didik pada kelompok atas dan bawah (Wardani & Ibrahim, 2020).

Respon Peserta Didik

Setelah mengerjakan soal-soal HOTS peserta didik mengisi angket respon dengan memberikan ceklis “ya” atau “tidak” pada setiap pertanyaan pada angket. Hasil respon peserta didik terhadap soal HOTS disajikan pada **Gambar 4**.



Gambar 4. Hasil Respon Peserta Didik Terhadap Soal HOTS

Keterangan:

- a : kesesuaian soal dengan materi
- b : soal menuntut berfikir analisis
- c : soal menuntut berfikir evaluasi
- d : pertanyaan mudah dipahami
- e : petunjuk soal mudah dipahami
- f : tulisan, grafik, gambar, tabel terbaca jelas
- g : pertanyaan menuntut jawaban terurai
- h : alokasi waktu cukup
- i : bahasa komunikatif
- j : istilah-istilah biologi mudah dipahami

Berdasarkan **gambar 4** dari 10 aspek terdapat 6 aspek mendapatkan respon peserta didik sebesar 100%. Aspek tersebut antara lain, materi yang dikembangkan sesuai dengan yang telah dipelajari, saat mengerjakan soal peserta didik dituntut untuk berfikir analisis, saat mengerjakan peserta didik dituntut untuk berfikir evaluasi, pertanyaan mudah dipahami, petunjuk soal mudah dipahami, stimulus (grafik, gambar, tabel, tulisan) yang digunakan terlihat jelas, pertanyaan pada soal menuntut jawaban terurai, alokasi waktu cukup, menggunakan bahasa yang komunikatif, dan soal menggunakan istilah-istilah biologi yang mudah dipahami. Terdapat 75% peserta didik berpendapat bahwa soal yang disajikan mudah untuk dipahami. 80% peserta didik menyatakan bahwa istilah-istilah biologi yang digunakan mudah untuk dipahami. Sebanyak 70% peserta didik menyatakan bahwa waktu yang diberikan cukup untuk menjawab soal uraian tersebut. Secara keseluruhan respon peserta didik memperoleh rerata 92% dengan kategori sangat baik.

Faktor penyebab peserta didik kesusahan memahami istilah-istilah biologi dalam soal adalah alokasi waktu yang diberikan tidak cukup sehingga peserta didik tergesah-gesah serta memahami soal dengan tidak teliti (Putri & Raharjo, 2017). Peserta didik yang tidak terbiasa diberikan soal HOTS dengan tipe esai akan

kesusahan untuk mengerjakan. Faktor yang dapat mempengaruhinya adalah peserta didik malas membaca stimulus soal tipe esai yang panjang. Tidak hanya itu dengan mengerjakan soal tipe pilihan ganda peserta didik hanya tinggal memilih opsi pilihan jawaban yang telah disediakan berbeda dengan tipe soal esai yang hasil bahwa adanya perbedaan hasil sangat signifikan antara peserta didik yang terbiasa mengerjakan soal uraian dengan peserta didik yang tidak terbiasa mengerjakan soal uraian pada UN 2018.

PENUTUP

Simpulan

Berdasarkan hasil penelitian pengembangan instrumen soal HOTS materi jaringan dan organ kelas XI SMA dinyatakan valid dan reliabel. Uji validitas teoritis memperoleh kategori sangat valid. Hasil uji validitas empiris diperoleh 80% soal dinyatakan valid. Hasil uji reliabilitas termasuk sangat reliabel. Tingkat kesukaran butir soal HOTS diperoleh hasil sebesar 30% soal sukar, 70% soal sedang. Daya pembeda butir soal HOTS memperoleh hasil 30% kurang, 30% cukup, dan 40% baik. Respon peserta didik pada kategori sangat baik.

Saran

Penelitian ini dapat ditindak lanjuti dengan mengembangkan instrumen penilaian HOTS pada materi biologi yang lain dan dengan menggunakan tipe soal yang beragam agar peserta didik semakin terbiasa dalam menyelesaikan soal-soal berfikir tingkat tinggi dengan tipe soal yang berbeda-beda.

Ucapan Terima Kasih

Terima kasih kepada Bapak Ahmad Bashri, M.Si. dan Ibu Muji Sri Prastiwi, M.Pd. sebagai dosen validator dan penguji. Guru Biologi SMAN 1 Puri Mojokerto Bapak drh. Linto Purwo dan peserta didik kelas XI IPA 3 SMAN 1 Puri Mojokerto.

DAFTAR PUSTAKA

Arikunto, S. 2015. *Dasar-dasar Evaluasi Pendidikan edisi 2*. Jakarta: Bumi Aksara.

Avina, Y. P & Winarsih. 2020. Pengembangan Instrumen Penilaian Sebagai Contoh Paket Soal *Higher Order Thinking Skills (HOTS) Materi Pencemaran Lingkungan Kelas X SMA. Jurnal BioEdu*. Vol. 9 (1): 217-223.

mengharuskan peserta didik untuk merangkai jawaban sendiri (Rini & Budijastuti, 2022). Hasil tersebut tidak jauh berbeda dengan penelitian Safari (2019) tentang pengaruh kebiasaan peserta didik dalam mengerjakan soal uraian terhadap nilai UN 2018 yang mendapatkan

Azizah, Z. N. & Budijastuti, W. 2022. Pengembangan Instrumen Penilaian Untuk Mengukur Keterampilan Literasi Sains pada Submateri Sistem Peredaran Darah Manusia. *Jurnal BioEdu*. Vol. 11(1): 89-97.

Baidlowi, M. H., Sunarmi, & Sulisetijono. 2019. Pengembangan Instrumen Soal Essay Tipe *Higher Order Thinking Skills* Materi Struktur Jaringan dan Fungsi Organ pada Tumbuhan Kelas XI SMAN 1 Tumpang. *Jurnal pendidikan biologi*. Vol. 10 (2): 57-65.

Fanani, M. Z. 2018. Strategi Pengembangan Soal HOTS dalam Kurikulum 2013. *Journal of Islamic Religious Education*. Vol. 2 (1): 57-65.

Fitri, Herman, Haris, A. 2020. Analisis kemampuan Memprediksi dalam Pembelajaran Fisika Peserta Didik Kelas XII SMAN 9 Makassar. *Jurnal Sains dan Pendidikan Fisika*. Vol. 16 (2): 100-107.

Hadi, K., Dazrullisa, Manurung, B., & Hasruddin. 2019. Tingkat Kesukaran Soal Tes Berfikir Tingkat Tinggi Berbasis Kearifan Lokal pada Materi Keanekaragaman Hayati di SMA. *Bio-Lectura*. Vol. 6 (2): 99-108

Herlanti, Y. 2015. Kesadaran Metakognitif dan Pengetahuan Metakognitif Peserta Didik SMA dalam Mempersiapkan Ketercapaian Standar Kelulusan pada Kurikulum 2013. *Jurnal Cakrawala Pendidikan*. No. 3: 357-367.

Iskandar, H. 2015. *Penyusunan Soal Higher Order Thinking Skills (HOTS) Sekolah Menengah Atas*. Jakarta: Direktorat Pembinaan Sekolah Menengah Atas, Direktorat Jendral Pendidikan Menengah, Kementerian Pendidikan dan Kebudayaan.

Kemendikbud. 2016. *Kerangka Dasar dan Struktur Kurikulum Sekolah Menengah Atas/ Madrasah Aliyah*. Jakarta: Kementerian Pendidikan dan Kebudayaan. (Online), [https://jdih.kemdikbud.go.id/arsip/Salinan Permendikbud Nomor 69 Tahun 2013.pdf](https://jdih.kemdikbud.go.id/arsip/Salinan_Permendikbud_Nomor_69_Tahun_2013.pdf) (Diakses 11 Maret 2021)

Mahmudah, R., dkk. 2016. Analisis Validitas Butir Soal *Certainty of Respon Index (CRI)* untuk Identifikasi Miskonsepsi Materi Tata Surya dan Fenomena Astronomi. *Seminar Nasional*

- Pendidikan Pendidikan Sains Pascasarja Unesa 2016: 579-587.*
- Masitoh, L. & Aedi, W. 2020. Pengemabangan Instrumen Asesmen *Higher Order Thinking Skills* (HOTS) Matematika di SMP. *Jurnal Cendikia*. Vol. 4(2): 886-897.
- Nurlia., dkk. 2017. Hubungan Antara Gaya Belajar dengan Hasil Belajar Biologi Siswa. *Jurnal Pendidikan Biologi*. Vol. 6(2): 321-328.
https://www.oecd.org/pisa/PISA_2018_Insights_and_Interpretations_FINAL_PDF.pdf (Diakses 12 Maret 2021).
- Pratiwi, Indah. 2019. Efek Program PISA Terhadap Kurikulum di Indonesia. *Jurnal Pendidikan dan Kebudayaan*. Vol. 4(1):51-71.
- Purwanto. 2011. *Evaluasi Hasil Belajar*. Yogyakarta: Pustaka Pelajar.
- Putri, B. A. & Raharjo. 2017. Validitas Empiris Soal *High Order Thinking* (HOT) Berbasis *Computer Based Test* (CBT) pada Sub Materi Sistem Indera Siswa Kelas XI SMA. *Jurnal BioEdu*. Vol 6(3): 353-359.
- Rahmadhan, D., Wasis. 2013. Analisis Perbandingan Level Kognitif dan Keterampilan Proses Sains dalam Standar Isi (SI), Soal Ujian Nasional (UN), Soal *Trends In Internasional Mathematic and Science Study* (TIMSS), Soal *Programme for International Student Assessment* (PISA). *Jurnal Inovasi Pendidikan Fisika*. Vol. 2(1): 20-25.
- Retnawati, H. 2016. *Analisis Kuantitatif Instrumen Penilaian*. Yoyakarta: Parama Publishing.
- Rini, A. F., & Budijastuti, W. 2022. Pengembangan Instrumen Soal HOTS Untuk Mengukur Keterampilan Pemecahan Masalah pada Materi Sistem Gerak Manusia. *Jurnal BioEdu*. Vol. 11(1): 127-137.
- Safari. 2019. Pengaruh Kebiasaan Siswa Menjawab Soal Uraian Terhadap Hasil UN 2018. *Indonesian Journal of Education Assessment*. Vol. 2(2): 20-31.
- Sugiyono. 2015. *Metode Penelitian Pendidikan (Pendekatan Kuantitatif, Kualitatif, dan R&D)*. Bandung: CV Alfabeta.
- Suharli. 2015. Teori Belajar dan Model Penerapannya dalam Pembelajaran. *Makalah Pengembangan Model Pembelajaran IPS UPI 2015: 1-13.*
- Ulfa, M., & Kuswati, N. 2021. Development Assessment Instrument Based on Higher Order Thinking Skills of Respiratory System of Grade XI of Senior High School. *Jurnal BioEdu*. Vol. 10(1): 1-11.
- OECD. 2015. PISA 2015 Assessment and Analytical and Financial Literacy. Paris: OECD Publishing. (Online), <https://www.oecd.org/pisa/pisa-2015-results-in-focus.pdf> (Diakses 12 Maret 2021).
- OECD. 2018. PISA 2018: Insthights and Interpretations (Online),
- Wardani, A. N. & Ibrahim, M. 2020. Karakteristik *Higher Order Thinking Skills* (HOTS) Materi Dampak Penyalahgunaan Psikotropika Untuk SMA. *Jurnal BioEdu*. Vol. 9(1): 60-67.
- Wardany, Kusuma. 2018. Kelayakan Instrumen Pengembangan Penilaian *Higher Order Thinking Skills* pada Materi Ekosistem. *Jurnal Pendidikan Sains*. Vol 6(2): 21-31.
- Widana, I. W. 2020. Pengaruh Pemahaman Konsep Asesmen HOTS terhadap Kemampuan Guru Matematika SMA/ SMK Menyusun Soal HOTS. *Jurnal Edukasi Matematika dan Sains*. Vol. 9(1): 66-75.
- Yudiandani, R. E. & Asri, M. T. 2022. Profil dan Validitas Instrumen Penilaian Autentik pada Materi Sel. *Jurnal BioEdu*. Vol 11(1): 1-9.