

PENGEMBANGAN INSTRUMEN PENILAIAN UNTUK MENGUKUR KEMAMPUAN LITERASI SISWA SMA KELAS XI PADA MATERI SISTEM EKSRESI*Development of Assessment Instruments to Measure Science Literacy Skills 11th of Grade High School Students on Excretory System Material***Ulfi Habibatul Husna**

Pendidikan Biologi, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Surabaya

E-mail: ulfi.20067@mhs.unesa.ac.id**Nur Ducha**

Biologi, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Surabaya

E-mail: nurducha@unesa.ac.id**Abstrak**

Melatih dan mengukur kemampuan literasi sains, diperlukan pengembangan instrumen penilaian yang selaras dengan kompetensi abad ke-21. Pada pengukuran kemampuan literasi sains, diperlukan pengembangan instrumen penilaian yang selaras. Penelitian ini bertujuan menghasilkan instrumen penilaian yang layak digunakan berdasarkan validitas teoritis dan empiris instrumen penilaian literasi sains pada materi Sistem Ekskresi, serta mendeskripsikan profil kemampuan literasi sains siswa. Model pengembangan yang digunakan adalah ADDIE (*Analysis, Design, Development, Implementation, Evaluate*). Validitas teoritis dilakukan oleh dosen ahli, berdasarkan aspek materi, konstruksi, serta bahasa. Validitas empiris dianalisis dari hasil uji coba di kelas XI-2 SMA Labschool Unesa 1 secara terbatas yang mencakup validitas empiris butir soal, reliabilitas, tingkat kesukaran, serta daya pembeda soal. Hasil penelitian yang diperoleh yaitu validitas teoritis sebesar 96,5% dengan interpretasi sangat valid, validitas empiris butir soal diperoleh 18 soal valid serta 2 soal tidak valid. Instrumen penilaian terbukti reliabel dengan rata-rata skor 0,71. Tingkat kesukaran soal objektif mencakup 1 soal mudah, 7 soal sedang, dan 1 soal sulit, sedangkan soal subjektif, 10 soal sedang dan 1 soal sulit. Daya pembeda soal objektif menunjukkan 3 soal kurang dan 6 soal cukup, sedangkan soal subjektif, 4 soal kurang, 6 soal cukup, dan 1 soal baik. Profil literasi sains siswa menunjukkan sebanyak 5 siswa berada di kategori tinggi, 18 siswa kategori sedang, dan 7 siswa kategori rendah. Kesimpulan penelitian ini yakni instrumen penilaian untuk mengukur kemampuan literasi sains siswa SMA kelas XI pada materi Sistem Ekskresi valid secara teoritis dan empiris dalam mengukur kemampuan literasi sains siswa.

Kata Kunci: Instrumen Penilaian, Kemampuan Literasi Sains, Sistem Ekskresi**Abstract**

Science education in 21st century aims to prepare students to use scientific understanding to solve everyday problems through science literacy. To train and measure science literacy skills, it is necessary to develop assessment instruments that are in accordance with 21st century competencies. This study aims to describe the theoretical and empirical validity of science literacy assessment instruments on excretory system material, and describe the profile of students' science literacy skills. This research method used the ADDIE development. Theoretical validity was conducted by expert lecturers, based on material, construction, and language aspects. Empirical validity was analyzed based on the results of a limited trial in class XI-2 SMA Labschool Unesa 1, which included empirical validity of the items, reliability, difficulty level, and question differentiation. The results showed a theoretical validity of 96.5%, with 18 valid and 2 invalid empirical questions. The instrument was reliable with a score of 0.71. Objective questions included 1 easy, 7 medium, and 1 difficult question, while subjective questions included 10 medium and 1 difficult question. The distinguishing power of objective questions included 3 lacking and 6 sufficient questions; subjective questions included 4 lacking, 6 sufficient, and 1 good question. Students science literacy profile showed 5 students were in the high category, 18 students were in the medium category, and 7 students were in the low category. The study concluded that the assessment instrument was valid for measuring the science literacy skills of grade XI students on excretory system material.

Keywords: Assessment Instrument, Science Literacy Skills, Excretory System Material

PENDAHULUAN

Pendidikan sains abad-21 bertujuan untuk mempersiapkan siswa agar menggunakan pemahaman ilmiah dalam memecahkan masalah sehari-hari (Fuadi dkk., 2020). Kemampuan sains juga sebagai modal penting untuk siswa dalam menyelesaikan berbagai permasalahan di zaman global (Barus dkk., 2024). Literasi sains mencakup keterampilan untuk menganalisis pertanyaan, mendapatkan pengetahuan baru, mengidentifikasi fenomena ilmiah, menyimpulkan dari bukti ilmiah dengan menerapkan konsep ilmiah (Fuadi dkk., 2020). Kurikulum merdeka fokus pada materi esensial seperti literasi dan numerasi, menjadikan literasi sains sebagai komponen penting dalam penilaian (Putri, 2020). Dengan demikian, literasi sains dibutuhkan dalam implementasi kurikulum merdeka.

Programme for International Student Assessment (PISA) adalah sebuah studi global yang diadakan oleh *Organisation for Economic Cooperation and Development* (OECD) guna mengukur kemampuan siswa berusia 15 tahun pada literasi sains dan matematika. Tes PISA bukan hanya mengukur kemampuan siswa untuk mengingat informasi, tetapi kemampuan siswa untuk menganalisis juga diukur, menyelesaikan masalah, dan menerapkan pengetahuan mereka dalam situasi yang baru. PISA dilakukan setiap tiga tahun sekali dan bertujuan untuk mengevaluasi sejauh mana siswa di berbagai negara telah menguasai kemampuan tersebut. Hasil Indonesia pada tes PISA tahun 2009, 2012, 2015, 2018, dan 2022 mencapai skor sebesar 383, 382, 403, 396, dan 383 (OECD, 2023). Pada tahun 2022, terlihat Indonesia mengalami penurunan skor tes PISA. Selain itu, peringkat Indonesia jika dibandingkan negara lain masih rendah. Indonesia ada di urutan 68 dari 77 negara yang berpartisipasi. Hal tersebut dapat diartikan bahwa kemampuan literasi sains siswa Indonesia masih rendah berdasarkan PISA. Literasi sains siswa yang masih rendah dapat disebabkan karena proses belajar mengajar belum mengarah pada keterampilan tersebut (Daniah, 2020). Oleh karena itu, instrumen penilaian yang sudah mengacu pada literasi sains dapat dimanfaatkan dalam melakukan evaluasi proses pembelajaran dengan menganalisis sejauh mana siswa telah mencapai setiap indikator literasi sains yang digunakan. Dengan demikian, guru dapat mengidentifikasi indikator yang belum diintegrasikan dalam proses pembelajaran (Nurfadillah dkk., 2023).

Kemampuan literasi sains adalah salah satu kompetensi yang dibutuhkan di abad-21. Untuk itu, diperlukan instrumen penilaian yang dapat mengukur

kemampuan literasi sains siswa guna mengetahui sejauh mana kemampuan tersebut telah dikuasai oleh siswa. Tingkat kemampuan literasi sains siswa bisa diketahui dari hasil yang diperoleh dari instrumen penilaian yang digunakan. Hasil penilaian ini dapat digunakan sebagai dasar evaluasi dalam proses pembelajaran (Khayati dan Raharjo, 2020). Dengan menganalisis ketercapaian literasi sains siswa, maka dapat mengevaluasi sejauh mana proses pembelajaran sudah sesuai dengan tuntutan abad-21 khususnya dalam hal literasi sains. Oleh karena itu, instrumen penilaian merupakan komponen penting dalam pembelajaran. Fuadi dkk. (2020) berpendapat bahwa penilaian literasi sains harus dilakukan untuk mendapatkan informasi bagaimana ketercapaian konsep sains yang telah dipelajari siswa dan menerapkannya dalam pemecahan masalah sehari-hari.

Hasil penelitian dari Chasanah dkk. (2022) menunjukkan bahwa metode evaluasi biologi di sekolah masih sederhana dan masih konseptual, sehingga belum dapat digunakan untuk mengevaluasi proses pembelajaran yang berbasis literasi sains. Hal tersebut sejalan dengan wawancara dan observasi pada salah satu guru biologi di SMA Labschool Unesa 1, yang mendapatkan hasil bahwa untuk instrumen penilaian sumatif untuk mata pelajaran biologi, soal-soal yang diberikan pada siswa belum berbasis literasi sains. Sekitar 80% soal yang digunakan pada penilaian sumatif adalah soal dengan tingkatan kognitif rendah (C1-C3) dan tidak terdapat literatur maupun data dari penelitian atau fenomena sehari-hari yang dapat dianalisis oleh siswa, yang hal tersebut adalah karakteristik dari soal literasi sains. Oleh sebab itu, instrumen tersebut tidak memenuhi standar kompetensi literasi sains yang ditentukan oleh PISA.

Instrumen penilaian yang difungsikan dalam mengukur literasi sains hendaknya berisi soal yang dapat secara akurat mendeteksi apakah siswa dapat memahami dan menerapkan konsep dalam menemukan jawabannya (Nufus dkk., 2017). Karakteristik instrumen penilaian literasi sains menurut PISA, memuat soal yang disusun menggunakan 3 indikator yaitu: menjelaskan fenomena ilmiah, merancang serta mengevaluasi penyelidikan ilmiah, dan menafsirkan data serta bukti secara ilmiah (OECD, 2023).

Contoh mata pelajaran yang membutuhkan literasi sains adalah biologi. Materi biologi mengandung konsep ilmiah yang abstrak dan memerlukan proses ilmiah untuk memahaminya (Fadilah, 2020). Materi pada mata pelajaran biologi yang bersifat kontekstual adalah Sistem Ekskresi, yang mencakup struktur organ, proses pembentukan urin dan keringat, serta kelainan pada

Sistem Ekskresi. Materi ini relevan dengan kehidupan sehari-hari karena berhubungan dengan organ dan bioproses dalam tubuh manusia (Komalasari dkk., 2019). Oleh karena itu, materi Sistem Ekskresi cocok digunakan dalam proses pembelajaran yang berbasis literasi sains, sehingga diperlukan juga instrumen penilaian yang dapat mengevaluasi apakah proses pembelajaran sudah sesuai dengan indikator literasi sains yang ditentukan. Selain itu, belum ada penelitian sebelumnya yang mengembangkan instrumen penilaian literasi sains dengan sub indikator kompetensi sains dari PISA yang menggunakan materi Sistem Ekskresi, sehingga penggunaan materi Sistem Ekskresi ini termasuk dalam kebaruan pada penelitian yang dilakukan.

Instrumen penilaian untuk mengukur literasi sains dirancang dengan soal-soal berjenjang kognitif tinggi (C4, C5, C6) sesuai taksonomi Bloom dan menggunakan aspek kompetensi sains PISA. Menurut Setiawan (2019), *framework* PISA adalah dokumen komprehensif mengenai literasi sains yang digunakan secara internasional untuk evaluasi literasi sains. Instrumen ini berisi 20 soal yang dilengkapi bacaan, tabel, dan grafik terkait fenomena sehari-hari yang berhubungan dengan Sistem Ekskresi, sehingga mendorong pemahaman proses ilmiah pada siswa. Instrumen berbasis *paper based test* ini terdiri dari soal objektif dan subjektif.

Keberhasilan pembelajaran Biologi dapat diukur dari penilaian hasil belajar siswa. Menurut Nuraisyah (2023), hasil belajar adalah aspek penting guna menilai keberhasilan siswa, proses, dan apakah tujuan pembelajaran telah tercapai. Oleh karena itu, mengembangkan instrumen penilaian yang layak digunakan berdasarkan validitas secara teoritis dan empiris guna mengukur kemampuan literasi sains siswa kelas XI SMA pada materi Sistem Ekskresi merupakan tujuan dari penelitian ini. Diharapkan dengan adanya penelitian pengembangan instrumen penilaian ini, proses pembelajaran di dalam kelas dapat dievaluasi agar lebih mengarah pada pembelajaran yang berbasis literasi sains, sehingga siswa lebih terampil untuk mengerjakan soal yang mengharuskan penerapan literasi sains.

METODE

Jenis penelitian ini yakni *Research and Development* (R&D), merupakan penelitian yang dapat menghasilkan produk sekaligus mengevaluasi seberapa efektif produk tersebut (Saadah, 2021). Pengembangan instrumen penilaian ini memiliki tujuan dalam menghasilkan instrumen penilaian yang layak serta berfungsi dalam mengukur tingkat literasi sains siswa pada materi Sistem Ekskresi. Proses pengembangan serta validasi

dilaksanakan di Prodi Pendidikan Biologi UNESA, sedangkan uji coba secara terbatas dilakukan di SMA Labschool Unesa 1 pada bulan Juni 2024.

Model pengembangan ADDIE (*Analysis, Design, Development, Implementation, and Evaluation*) digunakan untuk penelitian ini. Tahap "*Analysis*", dilakukan dengan menganalisis kebutuhan pengembangan instrumen penilaian yang meliputi analisis kurikulum, siswa, dan materi. Tahap "*Design*" melibatkan perancangan produk instrumen penilaian yang disesuaikan pada tujuan penyusunan instrumen yang sudah ditetapkan. Tahap "*Development*", produk dilakukan dengan menyusun instrumen penilaian yang sudah dirancang sebelumnya. Butir soal disusun sesuai dengan indikator literasi sains oleh PISA yang meliputi: menjelaskan fenomena ilmiah; mengevaluasi serta merancang penyelidikan ilmiah; dan menafsirkan data ke bentuk yang lain. Setelah itu, draf hasil pengembangan divalidasi oleh dosen ahli materi serta ahli pendidikan, yang mana hasil dari validasi ini menjadi dasar untuk melakukan revisi produk. Tahap "*Implementation*" melibatkan uji coba pada siswa kelas XI SMA Labschool Unesa 1 sebanyak 30 siswa secara terbatas pada 2 Juni 2024 guna menguji validitas empiris instrumen penilaian dan profil literasi sains siswa. Tahap "*Evaluation*" dilakukan di setiap tahap untuk perbaikan produk dan menguji kelayakan instrumen penilaian.

Instrumen penelitian yang digunakan berupa produk instrumen penilaian literasi sains serta lembar validasi. Lembar instrumen penilaian literasi sains memiliki 20 soal dengan 2 tipe soal yang berbeda yakni objektif serta subjektif. Tipe soal objektif terdiri dari bentuk soal pilihan ganda, jawaban centang, dan Benar/Salah atau Ya/Tidak, sedangkan tipe soal subjektif terdiri dari bentuk soal uraian. Penilaian pada lembar validasi mencakup tiga aspek yakni materi, konstruk, serta bahasa. Setiap aspek penilaian dinilai menggunakan skala Guttman (Tabel 1):

Tabel 1. Kategori skala Guttman

Kategori	Skor
Ya (√)	1
Tidak (-)	0

Setelah itu, skor yang didapatkan dari dua validator dihitung memakai rumus di bawah ini untuk menentukan persentase kevalidan:

$$N(\%) = \frac{\sum \text{jumlah skor yang menjawab ya}}{\sum \text{skor maksimal}} \times 100\% \dots \dots \dots (1)$$

Hasil persentase skor validasi kelayakan materi, konstruk, serta bahasa lalu ditafsirkan berdasarkan kriteria persentase validitas dalam Tabel 2 berikut:

Tabel 2. Kriteria Validitas

Validitas (%)	Kriteria
81,50 ≤ N ≤ 100,0	Sangat valid
62,75 ≤ N ≤ 81,49	Valid
44,00 ≤ N ≤ 62,74	Cukup valid
25,00 ≤ N ≤ 43,99	Kurang
00,00 ≤ N ≤ 24,99	Tidak valid

Setelah proses validasi dengan para validator dan dinyatakan valid, maka instrumen penilaian akan diimplementasikan untuk siswa kelas XI di SMA Labschool UNESA 1, tahun ajaran 2023/2024 kelas XI-1. Hasil yang didapat lalu dianalisis secara kuantitatif yang meliputi validitas empiris pada butir soal, reliabilitas, tingkat kesukaran soal, serta daya pembeda soal.

Validitas empiris pada butir soal dianalisis menggunakan metode *Pearson Product Moment Correlation* dengan bantuan *software SPSS 25*. Analisis ini dilakukan pada dua tipe soal (soal objektif dan soal subjektif), dilanjutkan dengan membandingkan hasil uji dengan batas signifikansi. Tingkat signifikansi yang digunakan adalah 5%, yang menunjukkan tingkat kesalahan 5% dan tingkat kepercayaan 95%. Pada taraf signifikansi 5%, batas signifikansi koefisien korelasi untuk butir soal pada r tabel yang diujicobakan pada 30 siswa (N = 30) adalah 0,361. Menurut Arikunto (2018) soal dapat dinyatakan valid bila r hitung bernilai lebih tinggi dari r tabel. Dengan demikian, butir soal berarti valid secara empiris jika koefisien korelasi r hitung > 0,361.

Perhitungan nilai reliabilitas soal dapat digunakan teknik Cronbach's Alfa pada program SPSS 25. Hasil perhitungan tersebut kemudian dikategorikan dalam beberapa kategori tingkat reabilitas instrumen berdasarkan Tabel 3 berikut:

Tabel 3. Kategori Reliabilitas Soal

Interval	Kategori
0,80 < N ≤ 1,00	Sangat reliabel
0,60 < N ≤ 0,80	Reliabel
0,40 < N ≤ 0,60	Cukup reliabel
0,20 < N ≤ 0,40	Kurang reliabel
0,00 < N ≤ 0,20	Tidak reliabel

Tingkat kesukaran soal diketahui melalui indeks kesukaran soal. Soal dikategorikan sebagai soal mudah jika indeksnya mendekati 0 dan dikategorikan soal sukar jika indeksnya mendekati 1. Tipe soal objektif dihitung kesukaran soalnya memakai rumus yakni:

$$P = \frac{B}{JS} \dots\dots\dots(2)$$

Keterangan :

P = indeks kesukaran

B = total siswa menjawab benar

JS = total seluruh siswa yang mengikuti tes

Rumus untuk butir soal dengan tipe soal subjektif

menggunakan rumus sebagai berikut :

$$Indeks\ kesukaran\ soal = \frac{rata-rata\ skor}{skor\ maksimum} \dots\dots\dots(3)$$

Hasil skor yang diperoleh dikategorikan menurut Tabel 4 indeks kesukaran soal berikut ini :

Tabel 4. Kategori Kesukaran Soal

Indeks Kesukaran	Kategori
K < 0,30	Soal sukar
0,30 ≤ K ≤ 0,70	Soal sedang
K > 0,70	Soal mudah

Setelah uji tingkat kesukaran soal, dilakukan analisis daya pembeda soal berfungsi mengelompokkan siswa kemampuan tinggi serta siswa kemampuan rendah. Koefisien daya beda butir soal mempunyai rentang nilai yakni 0 hingga 1. Indeks daya beda soal untuk tipe soal objektif dapat dihitung menggunakan rumus berikut :

$$DB = \frac{BA}{JA} - \frac{BB}{JB} = PA - PB \dots\dots\dots(4)$$

Keterangan :

DB = Indeks daya pembeda

BA = Total siswa golongan atas dengan jawaban benar

BB = Total siswa golongan bawah dengan jawaban benar

JA = Total golongan atas

JB = Total golongan bawah

PA = Proporsi golongan dengan jawaban benar

PB = Proporsi golongan dengan jawaban salah

Indeks daya pembeda soal pada tipe subjektif dihitung menggunakan rumus sebagai berikut :

$$DB = \frac{(mean\ kel.\ atas) - (mean\ kel.\ bawah)}{skor\ maksimum\ total} \dots\dots\dots(5)$$

Hasil skor yang diperoleh dapat diinterpretasikan berdasarkan Tabel 5 berikut:

Tabel 5. Kategori Daya Pembeda Soal

Indeks Daya pembeda	Kategori
0,71 – 1,00	Sangat baik
0,41 – 0,70	Baik
0,21 – 0,40	Cukup
0,00 – 0,20	Kurang

Kategori literasi sains siswa didapatkan dari menghitung rata-rata nilai siswa dengan standar deviasi. Hasil pengerjaan tes oleh siswa dinilai dengan memberikan skor tiap butir soal yang telah dijawab, kemudian menghitung total skor yang didapatkan pada keseluruhan tes. Setelah itu, nilai siswa didapatkan dengan mengubah skor keseluruhan tes pada skala 0-100.

$$NS = \frac{Skor\ yang\ diperoleh\ siswa}{skor\ total} \times 100 \dots\dots\dots(6)$$

Keterangan :

NS = Nilai siswa

Rata-rata nilai kelas, dihitung dengan rumus :

$$x = \frac{\sum xi}{n} \dots\dots\dots(7)$$

Keterangan :

x = rata-rata nilai kelas

∑xi = total nilai seluruh siswa

n = jumlah siswa

Rumus mencari nilai Standar Deviasi yaitu:

$$S = \sqrt{\frac{\sum(xi-x)^2}{n-1}} \dots\dots\dots(8)$$

Keterangan:

S = standar deviasi

xi = nilai tiap siswa

x = rata-rata nilai kelas

n = jumlah siswa

Kemudian, skor yang diperoleh dapat ditentukan kriterianya berdasarkan Tabel 6 berikut:

Tabel 6. Kategori literasi sains siswa

Ketentuan	Kategori
Nilai siswa $> x + SD$	Tinggi
$x - SD \leq$ Nilai siswa $\leq x + SD$	Sedang
Nilai siswa $< x - SD$	Rendah

HASIL DAN PEMBAHASAN

Penelitian pengembangan ini menghasilkan instrumen penilaian yang dimaksudkan guna mengukur kemampuan siswa dalam literasi sains pada konteks Sistem Ekskresi. Hasil penelitian meliputi: 1) pengembangan instrumen penilaian, 2) validitas teoritis pada instrumen penilaian, 3) validitas empiris pada instrumen penilaian, dan 4) analisis ketercapaian literasi sains siswa. Penelitian ini menggunakan model ADDIE yaitu, *Analysis, Design, Development, Implementation, dan Evaluation*. Pada tahap *Analysis*, melibatkan identifikasi terhadap capaian pembelajaran untuk menyusun spesifikasi butir soal dan indikator literasi sains yang sesuai untuk materi Sistem Ekskresi. Kisi-kisi soal termuat pada Tabel 7.

Tabel 7. Spesifikasi Butir Soal Instrumen Penilaian untuk Mengukur Kemampuan Literasi Sains Siswa pada Materi Sistem Ekskresi

Capaian Pembelajaran	Indikator Literasi Sains	Sub Indikator Literasi Sains	No. Soal
Menganalisis hubungan terkait struktur dan fungsi organ dalam sistem organ, serta mempelajari gangguan yang terjadi pada sistem organ tersebut.	Menjelaskan fenomena secara ilmiah	Mengingat serta menjelaskan fenomena ilmiah dengan kemampuan mengimplementasikan pengetahuan ilmiah yang sesuai.	1,2,3
		Mengidentifikasi, menggunakan, serta menghasilkan model dan merepresentasi kannya	4,5
		Membuat serta mengoreksi prediksi yang tepat.	6,7, dan 8
	Mengevaluasi serta merancang penyelidikan	Mengidentifikasi pertanyaan yang dieksplorasi dalam studi ilmiah	10, 11

Capaian Pembelajaran	Indikator Literasi Sains	Sub Indikator Literasi Sains	No. Soal
	ilmiah	Menjelaskan dan menilai berbagai metode yang digunakan oleh ilmuwan guna memverifikasi ketepatan dan objektivitas data	12, 13
		Mengubah data dari satu representasi ke representasi lainnya	14, 15
		Menganalisis, menginterpretasikan, serta menarik kesimpulan data	16, 17, 18
		Mengidentifikasi asumsi, bukti, serta penalaran pada teks sains	19, 20

Berdasarkan Tabel 7 yaitu spesifikasi butir soal instrumen penilaian menggunakan indikator literasi sains dari PISA tahun 2022. Dari ketiga indikator yang telah disebutkan, diambil 8 sub indikator yang digunakan untuk menyusun butir soal. Sari (2021) menyatakan bahwa dengan menggunakan indikator literasi sains PISA, siswa bukan hanya ditekankan di pemahaman materi sains, tetapi juga pada kemampuannya berpikir untuk memahami fenomena ilmiah mendapatkan informasi dan pengetahuan baru, menganalisis pertanyaan, pengimplementasian metode ilmiah dalam konteks nyata, dan menganalisis kesimpulan pada data.

Tahapan selanjutnya yakni *Design*, yang dilakukan adalah perancangan butir soal instrumen penilaian yang disesuaikan dengan capaian pembelajaran fase F khususnya pada materi Sistem Ekskresi dan indikator literasi sains menurut PISA. Instrumen penilaian ini menerapkan literasi sains di seluruh butir soal dengan indikator literasi sains dari PISA. Setiap soal mengharuskan siswa untuk membaca dan menganalisis artikel, tabel, dan grafik untuk menjawab soal dengan benar. Selain itu, setiap soal juga dilengkapi dengan stimulus berupa fenomena ilmiah yang sering ditemui di kehidupan nyata, yang sesuai dengan konsep materi Sistem Ekskresi. Jumlah butir soal sebanyak 20, meliputi 4 bentuk soal yang berbeda yaitu pilihan ganda, benar salah/ya tidak, jawaban centang, dan uraian.

Tahapan selanjutnya adalah *Development*, yang dilakukan yakni mengembangkan instrumen penilaian. Dimulai dari mengembangkan draf awal, yang kemudian ditelaah oleh dosen pembimbing, sehingga menghasilkan draf 1. Draft 1 kemudian ditelaah oleh validator pada saat seminar proposal, kemudian direvisi sehingga menghasilkan draf 2. Draft 2 ini siap dilakukan validasi secara teoritis oleh validator mencakup aspek materi,

konstruksi, dan bahasa. Menurut Arifin (2019) validitas teoritis sangat penting untuk memastikan apakah instrumen penilaian yang dikembangkan layak untuk diimplementasikan kepada siswa. Hasil validitas teoritis termuat pada Tabel 8.

Tabel 8. Hasil Validitas Teoritis pada Instrumen Penilaian

Aspek	Rata-Rata (%)	Interpretasi
Materi	89,7	Sangat valid
Konstruksi	99,3	Sangat valid
Bahasa	97,7	Sangat valid
Rata-rata	95,6	Sangat valid

Berdasarkan Tabel 8, penilaian berdasarkan telaah dari validator menunjukkan bahwa butir tes pada instrumen penilaian untuk mengukur literasi sains siswa SMA kelas XI pada materi Sistem Ekskresi sangat valid, dengan persentase rata-rata mencapai 95,6%, yang berarti layak digunakan untuk mengukur kemampuan literasi sains siswa. Sugiyono (2022) berpendapat yakni instrumen penilaian dengan validitas tinggi mampu mengukur sesuai dengan tujuan instrumen dan memberikan hasil pengukuran yang tepat.

Meskipun demikian, beberapa revisi diperlukan sebelum instrumen penilaian diujicobakan secara terbatas kepada siswa. Berikut merupakan hasil revisi produk instrumen penilaian (Tabel 9).

Tabel 9. Saran dan Masukan beserta Hasil Revisi pada Produk Instrumen Penilaian

No	Saran dan Masukan	Hasil Revisi																								
1.	<p>Pilihan jawaban perlu diperbaiki karena terdapat kesalahan konsep</p> <table border="1"> <thead> <tr> <th>Nomor</th> <th>Alasan</th> </tr> </thead> <tbody> <tr> <td>A. 2</td> <td>Karena tugasnya adalah menambahkan kembali glukosa pada urin</td> </tr> <tr> <td>B. 2</td> <td>Karena berfungsi dalam memfiltrasi darah agar glukosa tidak lolos bersama urin</td> </tr> <tr> <td>C. 1</td> <td>Karena berfungsi dalam penyerapan kembali glukosa pada darah.</td> </tr> <tr> <td>D. 5</td> <td>Karena berfungsi dalam memfiltrasi darah agar glukosa tidak lolos bersama urin</td> </tr> <tr> <td>E. 6</td> <td>Karena berfungsi dalam penyerapan kembali glukosa pada darah</td> </tr> </tbody> </table>	Nomor	Alasan	A. 2	Karena tugasnya adalah menambahkan kembali glukosa pada urin	B. 2	Karena berfungsi dalam memfiltrasi darah agar glukosa tidak lolos bersama urin	C. 1	Karena berfungsi dalam penyerapan kembali glukosa pada darah.	D. 5	Karena berfungsi dalam memfiltrasi darah agar glukosa tidak lolos bersama urin	E. 6	Karena berfungsi dalam penyerapan kembali glukosa pada darah	<p>Pilihan jawaban sudah diperbaiki sesuai dengan konsep yang sesuai</p> <table border="1"> <thead> <tr> <th>Nomor</th> <th>Alasan</th> </tr> </thead> <tbody> <tr> <td>A. 2</td> <td>Karena tugasnya adalah menambahkan kembali glukosa pada urin</td> </tr> <tr> <td>B. 2</td> <td>Karena berfungsi dalam memfiltrasi darah yang terlalu banyak mengandung glukosa</td> </tr> <tr> <td>C. 1</td> <td>Karena berfungsi dalam penyerapan kembali glukosa pada darah</td> </tr> <tr> <td>D. 5</td> <td>Karena berfungsi dalam memfiltrasi darah agar glukosa tidak lolos bersama urin</td> </tr> <tr> <td>E. 6</td> <td>Karena berfungsi dalam penyerapan kembali glukosa pada darah</td> </tr> </tbody> </table>	Nomor	Alasan	A. 2	Karena tugasnya adalah menambahkan kembali glukosa pada urin	B. 2	Karena berfungsi dalam memfiltrasi darah yang terlalu banyak mengandung glukosa	C. 1	Karena berfungsi dalam penyerapan kembali glukosa pada darah	D. 5	Karena berfungsi dalam memfiltrasi darah agar glukosa tidak lolos bersama urin	E. 6	Karena berfungsi dalam penyerapan kembali glukosa pada darah
Nomor	Alasan																									
A. 2	Karena tugasnya adalah menambahkan kembali glukosa pada urin																									
B. 2	Karena berfungsi dalam memfiltrasi darah agar glukosa tidak lolos bersama urin																									
C. 1	Karena berfungsi dalam penyerapan kembali glukosa pada darah.																									
D. 5	Karena berfungsi dalam memfiltrasi darah agar glukosa tidak lolos bersama urin																									
E. 6	Karena berfungsi dalam penyerapan kembali glukosa pada darah																									
Nomor	Alasan																									
A. 2	Karena tugasnya adalah menambahkan kembali glukosa pada urin																									
B. 2	Karena berfungsi dalam memfiltrasi darah yang terlalu banyak mengandung glukosa																									
C. 1	Karena berfungsi dalam penyerapan kembali glukosa pada darah																									
D. 5	Karena berfungsi dalam memfiltrasi darah agar glukosa tidak lolos bersama urin																									
E. 6	Karena berfungsi dalam penyerapan kembali glukosa pada darah																									
2.	<p>Pilihan jawaban pada soal perlu direvisi karena terdapat kesalahan konsep materi dan kalimat terlalu berbelit.</p> <table border="1"> <thead> <tr> <th>Pernyataan</th> <th>Centang (✓)</th> </tr> </thead> <tbody> <tr> <td>Ketika tubuh terhidrasi dengan baik, tubuh memiliki kemampuan untuk mengencerkan konsentrasi urea dalam keringat karena ada lebih banyak cairan yang tersedia untuk dieliminasi</td> <td></td> </tr> <tr> <td>Ketika tubuh tidak terhidrasi dengan baik, ginjal cenderung mempertahankan urea dalam darah untuk menjaga keseimbangan cairan dan elektrolit. Akibatnya, kadar urea dalam keringat bisa meningkat.</td> <td></td> </tr> <tr> <td>Ketika tubuh terhidrasi dengan baik, ginjal dapat membuang urea dengan lebih efisien, yang menghasilkan konsentrasi urea dalam keringat yang lebih rendah.</td> <td></td> </tr> <tr> <td>Ketika tubuh tidak terhidrasi dengan baik, tubuh cenderung mempertahankan konsentrasi urea yang lebih rendah dalam keringat untuk membantu menghilangkan kelebihan nitrogen dari tubuh</td> <td></td> </tr> </tbody> </table>	Pernyataan	Centang (✓)	Ketika tubuh terhidrasi dengan baik, tubuh memiliki kemampuan untuk mengencerkan konsentrasi urea dalam keringat karena ada lebih banyak cairan yang tersedia untuk dieliminasi		Ketika tubuh tidak terhidrasi dengan baik, ginjal cenderung mempertahankan urea dalam darah untuk menjaga keseimbangan cairan dan elektrolit. Akibatnya, kadar urea dalam keringat bisa meningkat.		Ketika tubuh terhidrasi dengan baik, ginjal dapat membuang urea dengan lebih efisien, yang menghasilkan konsentrasi urea dalam keringat yang lebih rendah.		Ketika tubuh tidak terhidrasi dengan baik, tubuh cenderung mempertahankan konsentrasi urea yang lebih rendah dalam keringat untuk membantu menghilangkan kelebihan nitrogen dari tubuh		<p>Pilihan jawaban sudah direvisi sesuai dengan konsep materi dan menggunakan kalimat yang efektif.</p>														
Pernyataan	Centang (✓)																									
Ketika tubuh terhidrasi dengan baik, tubuh memiliki kemampuan untuk mengencerkan konsentrasi urea dalam keringat karena ada lebih banyak cairan yang tersedia untuk dieliminasi																										
Ketika tubuh tidak terhidrasi dengan baik, ginjal cenderung mempertahankan urea dalam darah untuk menjaga keseimbangan cairan dan elektrolit. Akibatnya, kadar urea dalam keringat bisa meningkat.																										
Ketika tubuh terhidrasi dengan baik, ginjal dapat membuang urea dengan lebih efisien, yang menghasilkan konsentrasi urea dalam keringat yang lebih rendah.																										
Ketika tubuh tidak terhidrasi dengan baik, tubuh cenderung mempertahankan konsentrasi urea yang lebih rendah dalam keringat untuk membantu menghilangkan kelebihan nitrogen dari tubuh																										

No	Saran dan Masukan	Hasil Revisi										
		<table border="1"> <thead> <tr> <th>Pernyataan</th> <th>Centang (✓)</th> </tr> </thead> <tbody> <tr> <td>Konsumsi banyak air dapat menurunkan konsentrasi urea dalam plasma darah, sehingga penurunan konsentrasi urea dalam plasma darah juga menyebabkan konsentrasi urea dalam keringat menjadi lebih rendah.</td> <td></td> </tr> <tr> <td>Konsumsi sedikit air menyebabkan ginjal cenderung mempertahankan cairan dalam tubuh untuk menjaga keseimbangan cairan dan elektrolit. Akibatnya urea yang dikeluarkan ginjal juga lebih sedikit, sehingga urea lebih banyak dikeluarkan melalui keringat</td> <td></td> </tr> <tr> <td>Konsumsi banyak air menyebabkan produksi keringat meningkat dan keringat yang dihasilkan lebih encer, dengan konsentrasi zat terlarut yang lebih rendah, termasuk urea</td> <td></td> </tr> <tr> <td>Konsumsi sedikit air menyebabkan kelenjar keringat cenderung mempertahankan cairan dengan mengurangi produksi keringat, sehingga urea yang dikeluarkan bersama keringat juga ikut berkurang.</td> <td></td> </tr> </tbody> </table>	Pernyataan	Centang (✓)	Konsumsi banyak air dapat menurunkan konsentrasi urea dalam plasma darah, sehingga penurunan konsentrasi urea dalam plasma darah juga menyebabkan konsentrasi urea dalam keringat menjadi lebih rendah.		Konsumsi sedikit air menyebabkan ginjal cenderung mempertahankan cairan dalam tubuh untuk menjaga keseimbangan cairan dan elektrolit. Akibatnya urea yang dikeluarkan ginjal juga lebih sedikit, sehingga urea lebih banyak dikeluarkan melalui keringat		Konsumsi banyak air menyebabkan produksi keringat meningkat dan keringat yang dihasilkan lebih encer, dengan konsentrasi zat terlarut yang lebih rendah, termasuk urea		Konsumsi sedikit air menyebabkan kelenjar keringat cenderung mempertahankan cairan dengan mengurangi produksi keringat, sehingga urea yang dikeluarkan bersama keringat juga ikut berkurang.	
Pernyataan	Centang (✓)											
Konsumsi banyak air dapat menurunkan konsentrasi urea dalam plasma darah, sehingga penurunan konsentrasi urea dalam plasma darah juga menyebabkan konsentrasi urea dalam keringat menjadi lebih rendah.												
Konsumsi sedikit air menyebabkan ginjal cenderung mempertahankan cairan dalam tubuh untuk menjaga keseimbangan cairan dan elektrolit. Akibatnya urea yang dikeluarkan ginjal juga lebih sedikit, sehingga urea lebih banyak dikeluarkan melalui keringat												
Konsumsi banyak air menyebabkan produksi keringat meningkat dan keringat yang dihasilkan lebih encer, dengan konsentrasi zat terlarut yang lebih rendah, termasuk urea												
Konsumsi sedikit air menyebabkan kelenjar keringat cenderung mempertahankan cairan dengan mengurangi produksi keringat, sehingga urea yang dikeluarkan bersama keringat juga ikut berkurang.												
3.	Pilihan jawaban tidak homogen	Pilihan jawaban sudah homogen										

<p>Berdasarkan bacaan di atas, apa yang menyebabkan volume urin lebih banyak ketika berada pada suhu dingin?</p> <p>a. Suhu tubuh meningkat, sehingga ginjal mengalami tekanan dan menyebabkan produksi urin meningkat</p> <p>b. Ketika suhu tubuh menurun, pembuluh darah menyempit, aliran darah di glomerulus dan tubulus ginjal meningkat sehingga memicu otak dan tubuh mengirimkan sinyal bagi ginjal untuk melakukan filtrasi lebih cepat dan menyebabkan urin keluar lebih banyak.</p> <p>c. Ketika suhu tubuh menurun, maka terjadi vasodilatasi (pelebaran pembuluh darah), menyebabkan tekanan darah meningkat sehingga ginjal juga lebih cepat dalam melakukan filtrasi darah</p> <p>d. Saat suhu tubuh menurun, aquaporin akan merangsang air agar masuk ke dalam sel, mekanisme ini menyebabkan terdapat banyak air pada aliran darah sehingga kadar air pada sel tubuh dan aliran darah tidak seimbang dan ginjal membuang kelebihan cairan melalui urin.</p>	<p>Berdasarkan bacaan di atas, apa yang menyebabkan volume urin lebih banyak ketika berada pada suhu dingin?</p> <p>a. Suhu tubuh menurun, sehingga ginjal mengalami tekanan dan menyebabkan produksi urin meningkat</p> <p>b. Ketika suhu tubuh menurun, pembuluh darah menyempit, aliran darah di glomerulus dan tubulus ginjal meningkat sehingga memicu otak dan tubuh mengirimkan sinyal bagi ginjal untuk melakukan filtrasi lebih cepat dan menyebabkan urin keluar lebih banyak.</p> <p>c. Ketika suhu tubuh menurun, maka terjadi vasodilatasi (pelebaran pembuluh darah), menyebabkan tekanan darah meningkat sehingga ginjal juga lebih cepat dalam melakukan filtrasi darah</p> <p>d. Saat suhu tubuh menurun, aquaporin akan merangsang air agar masuk ke dalam sel, mekanisme ini menyebabkan terdapat banyak air pada aliran darah sehingga kadar air pada sel tubuh dan aliran darah tidak seimbang dan ginjal membuang kelebihan cairan melalui urin.</p>																	
<p>4. Kalimat pada artikel yang disajikan pada soal menimbulkan penafsiran ganda</p> <p>Hasil penelitian menunjukan bahwa tingkat produktivitas keringat pada subyek laki-laki maupun perempuan pada lingkungan outdoor lebih banyak dari pada lingkungan indoor karena suhu di outdoor lebih tinggi dari suhu di indoor. Hal tersebut karena suhu berpengaruh terhadap produksi keringat. Oleh sebab itu, saat kita berada di outdoor atau dalam kondisi lingkungan yang sangat panas, suhu tubuh manusia akan menjadi tinggi, sehingga merangsang kelenjar kulit untuk menyeimbangkan suhu tubuh.</p> <p>Sumber : diadaptasi dari Aji & Ashadi, 2019.</p>	<p>Data yang disajikan pada soal sudah tidak menimbulkan penafsiran ganda.</p> <p>Tabel 1. Data kondisi lingkungan</p> <table border="1"> <thead> <tr> <th rowspan="2">Keadaan lingkungan</th> <th colspan="2">Rata-rata</th> </tr> <tr> <th>Kelompok 1</th> <th>Kelompok 2</th> </tr> </thead> <tbody> <tr> <td>Suhu (C)</td> <td>28,1</td> <td>29,1</td> </tr> <tr> <td>Kelembaban (%)</td> <td>80,2</td> <td>90,2</td> </tr> </tbody> </table> <p>Tabel 2. Rata-rata rasio keringat</p> <table border="1"> <thead> <tr> <th>Kelompok</th> <th>Rata-rata Keringat</th> </tr> </thead> <tbody> <tr> <td>Kelompok 1</td> <td>0,49 liter</td> </tr> <tr> <td>Kelompok 2</td> <td>0,62 liter</td> </tr> </tbody> </table> <p>Sumber : diadaptasi dari Andriana, dkk (2019)</p>	Keadaan lingkungan	Rata-rata		Kelompok 1	Kelompok 2	Suhu (C)	28,1	29,1	Kelembaban (%)	80,2	90,2	Kelompok	Rata-rata Keringat	Kelompok 1	0,49 liter	Kelompok 2	0,62 liter
Keadaan lingkungan	Rata-rata																	
	Kelompok 1	Kelompok 2																
Suhu (C)	28,1	29,1																
Kelembaban (%)	80,2	90,2																
Kelompok	Rata-rata Keringat																	
Kelompok 1	0,49 liter																	
Kelompok 2	0,62 liter																	

Menurut Tabel 9, pada aspek materi, terdapat kesalahan konsep pada soal mengenai Sistem Ekskresi khususnya pada proses fisiologis pengeluaran zat sisa metabolisme pada ginjal dan kulit yang harus diperbaiki untuk mencegah miskonsepsi pada siswa. Kesalahan ini dapat mengganggu pengukuran kemampuan siswa secara akurat (Rovita, 2020). Selain itu, terdapat soal yang terlalu rumit untuk tingkat kemampuan siswa SMA kelas XI dan perlu disederhanakan. Menurut Nurhalimah dkk. (2022) soal yang terlalu rumit dapat membuat siswa

kesulitan menjawab dan kurang semangat untuk mencoba. Pada aspek konstruksi, beberapa pilihan jawaban memerlukan revisi agar lebih homogen dan logis. Menurut Husna dkk. (2023) homogenitas pilihan jawaban penting untuk membuat pilihan jawaban berfungsi sebagai pengecoh yang efektif. Terakhir, pada aspek bahasa, beberapa soal menimbulkan penafsiran ganda dan perlu diperbaiki untuk mencegah kesalahpahaman. Menurut Febienti dkk. (2021), dalam mengembangkan instrumen penilaian, butir soal harus memiliki batasan pertanyaan yang jelas untuk mencegah penafsiran ganda.

Instrumen penilaian yang sudah valid secara teoritis, selanjutnya diujicobakan terbatas pada 30 siswa kelas XI-2 di SMA Labschool Unesa 1. Hasil yang diperoleh digunakan untuk menganalisis validitas empiris instrumen penilaian yang mencakup evaluasi terhadap validitas butir soal secara empiris, reliabilitas, tingkat kesukaran, serta daya pembeda soal dalam menentukan kesesuaian soal sebagai produk tes (Alfajri dkk., 2019). Hasil distribusi uji validitas butir soal secara empiris termuat dalam Tabel 10.

Tabel 10. Hasil Validitas Empiris pada Butir Soal

Indeks Validitas	Butir Soal	Jumlah	Persentase (%)
>0,361	1, 2, 3, 4, 5, 6, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20	18	90
<0,361	7, 8	2	10

Berdasarkan Tabel 10 yaitu hasil validitas empiris butir soal dengan menggunakan *software* SPSS 25, darikeseluruhan soal, ada 18 soal valid dan 2 soal tidak valid menurut perhitungan validitas empiris menggunakan SPSS 25. Soal valid pada Tabel 10 mempunyai nilai r hitung lebih besar daripada r tabel ($R_{count} > 0,361$), sedangkan pada soal yang dikatakan tidak valid, memiliki nilai r hitung lebih kecil dari nilai r tabel ($R_{count} < 0,361$). Sudjana (2016) menyatakan bahwa soal valid harus mampu mengukur kemampuan literasi sains siswa. Namun, soal nomor 7 dan 8 dinyatakan tidak valid dan perlu direvisi. Hal ini terjadi karena banyak siswa yang menjawab salah pada kedua soal tersebut, sehingga menyebabkan ketidakvalidan pada soal. Kedua soal ini termasuk dalam tingkat kognitif yang tinggi (C5). Soal nomor 7 meminta siswa untuk mengevaluasi pendapat tentang kandungan keringat, sedangkan soal nomor 8 meminta evaluasi terhadap pendapat mengenai gagal ginjal. Siswa mengalami kesulitan dalam menjawab kedua soal ini karena selama proses pembelajaran, mereka jarang dilatih untuk melakukan evaluasi terhadap suatu pendapat. Mahmudah dan Setyarsih (2024) menyatakan

bahwa validitas soal juga dipengaruhi oleh benar salahnya jawaban yang diberikan siswa, jika soal berada di ranah kognitif tinggi (C4-C6), maka besar kemungkinan banyak siswa yang menjawab salah. Fidias dan Puspitawati (2022) menambahkan bahwa siswa yang terbiasa dengan soal ranah kognitif rendah (C1-C3) akan kesulitan dengan soal ranah kognitif tinggi (C4-C6).

Instrumen penilaian yang dikembangkan perlu diuji reliabilitasnya. Magdalena dkk. (2021) berpendapat bahwa perhitungan reliabilitas bertujuan untuk mengetahui konsistensi suatu instrumen penilaian dalam mengukur yang semestinya diukur. Perhitungan nilai reliabilitas soal menggunakan metode Cronbach's Alpha pada *software* SPSS 25. Hasil perhitungan reliabilitas soal termuat dalam Tabel 11 berikut.

Tabel 11. Hasil Reliabilitas Instrumen Penilaian

Jenis Soal	Nilai Reliabilitas	Keterangan
Soal Objektif	0,72	Reliabel
Soal Subjektif	0,70	Reliabel
Rata-rata	0,71	Reliabel

Menurut Tabel 11 yaitu hasil perhitungan reliabilitas soal menggunakan SPSS menunjukkan bahwa nilai reliabilitas berada dalam interval $0,60 < \alpha \leq 0,80$, yang termasuk dalam kategori soal reliabel (Tabel 3). Hasil ini mengindikasikan bahwa butir soal pada instrumen penilaian berada pada kategori reliabel. Hal tersebut sesuai dengan pernyataan oleh Sugiyono (2022), bahwa sebuah instrumen dianggap reliabel jika koefisien reliabilitasnya mencapai minimal 0,6. Instrumen penilaian yang dinyatakan reliabel menurut Safitri dkk. (2024) adalah jika instrumen penilaian mampu memperoleh hasil yang tetap jika digunakan berulang terhadap siswa yang sama.

Instrumen penilaian yang sudah diuji reliabilitasnya kemudian diuji tingkat kesukaran soalnya untuk mengetahui mengetahui suatu butir soal yang diteskan berada pada soal mudah, sedang, atau sulit, berdasarkan total siswa dengan jawaban benar pada satu soal (Ndiung dan Jediut, 2020). Hasil uji tingkat kesukaran soal termuat dalam Tabel 12.

Tabel 12. Hasil Tingkat Kesukaran Butir Soal

Tipe Soal	Interpretasi Tingkat esukaran	Jumlah Butir Soal
Objektif	Mudah	1
	Sedang	7
	Sukar	1
Subjektif	Mudah	0
	Sedang	10
	Sukar	1

Berdasarkan Tabel 12 tentang hasil analisis tingkat kesukaran butir soal, sebagian besar soal mempunyai kategori sedang, yang dapat diartikan bahwa instrumen penilaian yang dikembangkan baik, karena memiliki soal sedang dengan jumlah lebih banyak daripada soal mudah

dan sukar. Menurut Safitri dkk. (2024), sebuah soal dinyatakan baik jika tingkat kesukarannya tidak tinggi atau tidak rendah. Soal mudah tidak dapat membuat siswa berpikir lebih dalam untuk mengerjakannya, sementara soal sulit bisa menyebabkan siswa kehilangan motivasi untuk mengerjakan soal tersebut karena di luar batas pengetahuan mereka.

Daya pembeda soal juga diuji pada instrumen penilaian yang dikembangkan. Analisis daya pembeda soal bertujuan dalam mengukur suatu butir soal apakah bisa mengelompokkan siswa menguasai kompetensi dan yang belum melalui kriteria tertentu (Arifin, 2019). Hasil perhitungan daya pembeda soal tertera dalam Tabel 13.

Tabel 13. Hasil Daya Pembeda Butir Soal

Tipe Soal	Interpretasi Daya Pembeda	Jumlah Butir Soal
Objektif	Kurang	3
	Cukup	6
	Baik	0
Subjektif	Kurang	4
	Cukup	6
	Baik	1

Berdasarkan Tabel 13 mengenai perhitungan daya pembeda soal, dapat diketahui bahwa sebagian besar soal merupakan soal kategori cukup. Hal ini, dapat diartikan bahwa soal-soal tersebut masih dapat digunakan, tetapi perlu dilakukan revisi. Sejalan dengan pernyataan dari Fidias dan Puspitawati (2022), bahwa soal yang memiliki daya pembeda soal kategori cukup, berarti soal itu dapat diterima, tetapi dilakukan revisi dahulu supaya memperoleh hasil yang lebih valid.

Sebanyak 7 butir soal memiliki kategori daya pembeda kurang. Menurut Afrida dkk. (2022), soal dengan daya pembeda rendah bisa dijawab benar oleh siswa dengan kemampuan tinggi maupun rendah, atau tidak bisa dijawab oleh keduanya. Masito dan Aedi (2020) berpendapat yakni soal dengan daya pembeda rendah tidak bisa membedakan siswa ke dalam kelompok atas dan bawah. Faktor yang mempengaruhi rendahnya daya beda soal antara lain tingginya tingkat kesukaran soal karena berada di tingkat kognitif C4-C6. Hal ini sejalan dengan pendapat Sholihah dkk. (2021), yakni soal dengan materi terlalu sulit dapat membuat banyak siswa hanya menebak jawabannya, sehingga daya pembeda menjadi rendah.

Selain analisis butir soal, profil kemampuan literasi sains siswa juga dianalisis melalui uji coba secara terbatas instrumen penilaian yang telah dilakukan berdasarkan nilai yang diperoleh tiap siswa. Soal yang sudah teruji validitasnya digunakan dalam menganalisis kemampuan literasi sains siswa, sedangkan soal yang tidak valid secara empiris (nomor 7 dan 8) tidak digunakan untuk menganalisis kemampuan literasi sains

siswa. Kemampuan literasi sains dikategorikan dalam 3 kategori yakni tinggi, sedang, dan rendah. Kategori literasi sains siswa termuat pada Tabel 14.

Tabel 14. Kategori Literasi Sains Siswa

Ketentuan	Kategori	Jumlah Siswa
$N > 68,6$	Tinggi	4
$32,6 \leq N \leq 68,6$	Sedang	19
$N < 32,6$	Rendah	7
Jumlah		30

Berdasarkan Tabel 14 yaitu kategori literasi sains siswa, diketahui yakni dari 30 siswa partisipan uji coba terbatas, paling banyak adalah siswa dengan kemampuan literasi sains pada kategori sedang, diikuti oleh siswa dengan kemampuan literasi sains pada kategori rendah, dan paling sedikit adalah siswa pada dengan kategori literasi sains tinggi. Menurut Nurfadhillah dkk. (2023), hasil ini berarti bahwa kemampuan siswa dalam pengerjaan soal literasi sains yang menuntut pemahaman serta analisis sudah cukup baik.

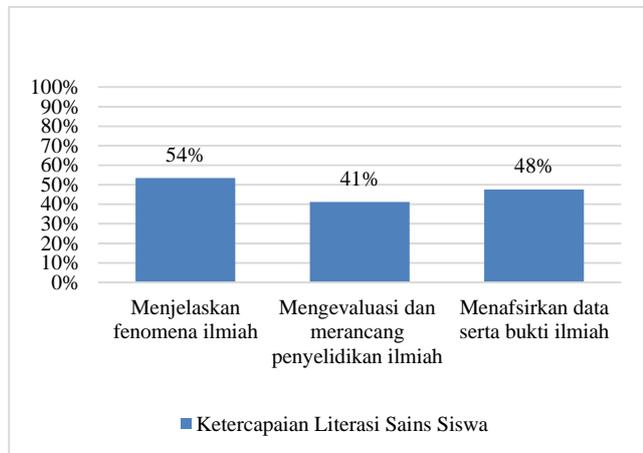
Terdapat beberapa indikator kompetensi sains beserta sub indikatornya yang digunakan dalam pengembangan instrumen penilaian. Hasil ketercapaian siswa pada tiap indikator dan sub indikator literasi sains termuat pada Tabel 15.

Tabel 15. Ketercapaian Literasi Sains Siswa

Indikator Kompetensi Sains	Sub Indikator	Persentase (%)	Kategori
Menjelaskan fenomena secara ilmiah	Mengingat serta menjelaskan fenomena ilmiah dengan kemampuan mengimplementasikan pengetahuan ilmiah yang sesuai.	50	Sedang
	Mengidentifikasi, menggunakan, serta membuat model dan merepresentasi kannya	59,6	Sedang
	Membuat dan mengoreksi prediksi yang tepat.	50,6	Sedang
Rata-rata		53,5	Sedang
Mengevaluasi dan merancang penyelidikan ilmiah	Mengidentifikasi pertanyaan yang dieksplorasi dalam studi ilmiah	48,3	Sedang
	Menjelaskan dan menilai berbagai metode yang digunakan oleh ilmuwan guna memverifikasi kebenaran data dan objektivitas.	34	Sedang
Rata-rata		41,15%	Sedang
Menafsirkan data dan bukti ilmiah	Mengubah data dari ke representasi lain	48,6	Sedang
	Menganalisis, menginterpretasikan, serta menarik kesimpulan data	59,5	Sedang
	Mengidentifikasi	35	Sedang

	asumsi, bukti, serta penalaran pada teks sains		
	Rata-rata	47,6	Sedang

Berdasarkan Tabel 15 maka proporsi ketercapaian literasi sains siswa pada ketiga indikator termuat pada Gambar 1.



Gambar 1. Proporsi Ketercapaian Literasi Sains Siswa

Berdasarkan Gambar 1 dan Tabel 15 dapat diketahui bahwa ketercapaian siswa tertinggi berada pada indikator menjelaskan fenomena ilmiah, diikuti oleh indikator menafsirkan data serta bukti ilmiah yang berada di urutan kedua tertinggi, sementara indikator mengevaluasi serta merancang penyelidikan ilmiah memperoleh ketercapaian siswa terendah.

Pada indikator menjelaskan fenomena ilmiah, peneliti menggunakan tiga sub indikator untuk mengukur ketercapaian literasi sains siswa, yang mana ketercapaian siswa pada masing-masing indikator tersebut berada di kategori sedang. Indikator ini memperoleh persentase tertinggi di antara indikator lainnya, menunjukkan bahwa siswa cukup mampu mengingat dan mengimplementasikan pengetahuan ilmiah dalam kehidupan nyata. Hal dapat terjadi karena guru sudah menimplementasikan indikator tersebut dalam proses pembelajarannya, sehingga siswa terbiasa dengan soal-soal kontekstual. Menurut Sutrisna (2021), ketercapaian literasi sains siswa dipengaruhi oleh pemberian instrumen penilaian yang fokus terhadap kemampuan literasi sains. Selain itu, diantara ketiga indikator lain, soal dengan indikator ini adalah soal yang paling mudah dipahami oleh siswa. Berdasarkan pernyataan dari OECD (2019), soal dengan indikator menjelaskan fenomena ilmiah mewajibkan siswa mengingat pengetahuan terkait permasalahan yang berhubungan dengan kejadian nyata serta mengimplementasikan pengetahuan tersebut dalam mengidentifikasi serta menjelaskan suatu fenomena.

Pada indikator mengevaluasi serta merancang penyelidikan ilmiah, peneliti menggunakan dua sub indikator untuk mengukur ketercapaian literasi sains siswa, yang mana ketercapaian siswa pada masing-masing indikator tersebut berada di kategori sedang, sehingga siswa mempunyai kemampuan yang memadai untuk mengevaluasi sebuah penyelidikan ilmiah serta merumuskan langkah dalam menjawab pertanyaan secara ilmiah. Meskipun demikian, indikator tersebut memperoleh persentase terendah dibandingkan dengan dua indikator lainnya. Faktor yang mempengaruhinya adalah ketidakbiasaan siswa di SMA Labschool Unesa 1 dalam menjawab pertanyaan yang mengeksplorasi dan mengevaluasi metode ilmiah. Rosidi (2021) menyatakan bahwa rendahnya ketercapaian siswa dalam mengevaluasi serta merancang penyelidikan ilmiah menunjukkan kurangnya kemampuan siswa dalam pemecahan masalah yang memerlukan pemikiran komprehensif. Tulaiya (2020) menegaskan bahwa kemampuan mengevaluasi serta merancang penyelidikan ilmiah berhubungan erat pada pengetahuan sains dasar mereka. Rachmadani (2022) menambahkan bahwa rendahnya ketercapaian indikator tersebut juga disebabkan oleh rendahnya kemampuan prosedural siswa.

Pada indikator menafsirkan data dan bukti ilmiah, peneliti menggunakan tiga sub indikator guna mengukur ketercapaian literasi sains siswa, yang mana ketercapaian siswa pada masing-masing indikator tersebut berada di kategori sedang, menunjukkan bahwa kemampuan siswa telah mencapai kemampuan yang memadai dalam menganalisis data serta argumen yang disajikan pada tabel dan grafik, serta menyimpulkannya secara ilmiah. Namun, ketercapaian literasi sains siswa pada indikator ini masih lebih rendah daripada pada indikator menjelaskan fenomena secara ilmiah. Irwan dkk. (2019) berpendapat bahwa lebih rendahnya ketercapaian siswa pada indikator menafsirkan data serta bukti secara ilmiah daripada indikator lainnya berarti bahwa siswa masih perlu mengembangkan pemahaman dan kemampuan membuat argumen atau kesimpulan yang lebih mendalam pada suatu data hasil penelitian untuk menjawab soal yang menggunakan indikator ini.

Dari hasil analisis perhitungan ketercapaian literasi sains siswa pada tiap indikator, maka dapat dinyatakan bahwa secara keseluruhan, profil kemampuan literasi sains siswa kelas XI di SMA Labschool Unesa 1 berada di kategori sedang. Berdasarkan pendapat dari Putranta dan Supahar (2019) ketercapaian literasi sains siswa pada kategori sedang bisa diartikan yakni siswa menunjukkan kemampuan literasi sains yang memadai dalam

menggunakan konsep dan keterampilan proses sains untuk menghadapi fenomena nyata sehari-hari.

PENUTUP

Simpulan

Berdasarkan uraian di atas, bisa disimpulkan yakni instrumen penilaian yang dikembangkan dapat dinyatakan layak digunakan dalam mengukur kemampuan literasi sains siswa SMA kelas XI pada materi Sistem Ekskresi. Hasil uji validitas secara teoritis yang ditinjau dari aspek materi, konstruk, serta bahasa memperoleh skor 96,5% yang mempunyai interpretasi sangat valid. Validitas empiris pada butir didapatkan hasil bahwa 18 soal valid serta 2 soal tidak valid. Instrumen penilaian dikatakan reliabel pada rata-rata skor 0,71. Uji tingkat kesukaran soal pada instrumen penilaian pada tipe soal objektif terdiri dari 1 soal kategori mudah, 7 soal kategori sedang, dan 1 soal sukar, sedangkan pada tipe soal subjektif terdiri dari 20 soal sedang dan 1 soal kategori sukar. Uji daya pembeda soal pada instrumen penilaian pada tipe soal objektif terdiri dari 6 soal ada di kategori cukup serta 3 soal ada di kategori kurang, sedangkan pada tipe soal subjektif sebanyak 1 soal ada di kategori baik, 6 soal ada di kategori cukup, serta 4 soal ada di kategori kurang. Profil kemampuan literasi sains siswa SMA kelas XI IPA-2 di SMA Labschool Unesa menunjukkan sebanyak 5 siswa ada di kategori tinggi, 18 siswa ada di kategori sedang, dan 7 ada di kategori rendah.

Saran

Penelitian lanjutan diperlukan untuk mengembangkan instrumen penilaian literasi sains pada topik biologi selain Sistem Ekskresi, supaya siswa terbiasa dengan soal-soal yang memerlukan literasi sains, mencakup penguasaan konsep dan kemampuan mengidentifikasi masalah terkait fenomena sehari-hari menggunakan konsep-konsep ilmiah. Selain itu, perlu penelitian lanjutan untuk mengimplementasikan instrumen penilaian dalam skala lebih luas, guna menilai kelayakannya pada subjek berbeda. Penelitian lanjutan perlu dilakukan guna merevisi soal yang kurang layak berdasarkan validitas empiris, daya pembeda, dan tingkat kesulitan. Kajian lebih mendalam tentang konsep materi Sistem Ekskresi juga diperlukan untuk menghindari kekeliruan konsep pada soal yang dikembangkan.

Ucapan Terima Kasih

Peneliti menyampaikan terimakasih untuk Bapak Raharjo, M.Si., dan Ibu Nur Kuswanti, M.Sc. St., sebagai penguji dan validator yang sudah memberikan

bimbingan, penilaian serta masukan pada produk instrumen penilaian yang dikembangkan. Selain itu, terimakasih kepada guru Biologi SMA Labschool Unesa 1, Ibu Ekris Sulianti S.Pd., M.Si. dan Ibu Shinta Dwi Martika, S.Pd., yang sudah berkenan memberikan izin penelitian, serta siswa-siswi kelas XI-2 SMA Labschool Unesa 1 yang sudah berpartisipasi pada uji coba produk instrumen penilaian.

DAFTAR PUSTAKA

- Afina, D.R., Hayati, M.N. and Fatkhurrohman, M.A., 2021. Profil Capaian Kompetensi Literasi Sains Siswa SMP Negeri Kota Tegal Menggunakan PISA. *PSEJ (Pancasakti Science Education Journal)*, 6(1), pp.10-21.
- Arifin, Z. 2019. *Evaluasi Pembelajaran: Prinsip, Teknik, dan Prosedur*. Bandung: PT Remaja Rosdakarya.
- Arikunto. 2018. *Dasar-Dasar Evaluasi Pendidikan (edisi 3)*. Jakarta: Bumi Aksara.
- Chasanah, N., Widodo, W. and Suprpto, N., 2022. Pengembangan Instrumen Asesmen Literasi Sains untuk Mendeskripsikan Profil Siswa. *PENDIPA Journal of Science Education*, 6(2), pp.474-483.
- Daniah, D., 2020. Pentingnya Inkuiri Ilmiah pada Praktikum dalam Pembelajaran IPA untuk Peningkatan Literasi Sains Mahasiswa. *Pionir: Jurnal Pendidikan*, 9(1), pp.144-153
- Fadillah, E. N. 2017. Pengembangan Instrumen Penilaian untuk Mengukur Keterampilan Proses Sains Siswa SMA. *Didaktika Biologi: Jurnal Penelitian Pendidikan Biologi*, 1(2), 123-134.
- Febianti, D., Mirandah, S., Aulia, A., Pratiwi, S., Putri, R.I.I., Simarmata, R.H. and Nuraeni, Z., 2021. Kemampuan Menyelesaikan Soal *Open-Ended* Materi Segiempat bagi Calon Guru Matematika. *Jurnal Pendidikan Matematika Undiksha*, 12(2), pp.26-33.
- Fidia, F., and Puspitawati, R.P., 2022. Pengembangan Instrumen Soal *Higher Order Thinking Skills* (HOTS) Materi Jaringan dan Organ pada Tumbuhan Kelas XI SMA. *Berkala Ilmiah Pendidikan Biologi (BioEdu)*, 11(3), pp.745-754.
- Fuadi, H., Robbia, A.Z., Jamaluddin, J. and Jufri, A.W., 2020. Analisis Faktor Penyebab Rendahnya Kemampuan Literasi Sains Peserta Didik. *Jurnal Ilmiah Profesi Pendidikan*, 5(2), pp.108-116.
- Hadiprayitno, G., Muhlis., dan Artayasa, I.P. 2020. Pendampingan Guru Biologi dalam

- Penyusunan Instrumen Penilaian Berorientasi HOTS di Kabupaten Lombok Barat. *Jurnal Pengabdian Magister Pendidikan IPA*, 3(2), pp.143-148.
- Hartanto, Rudi. 2014. Pengembangan Instrumen Asesmen untuk Menilai Materi Ekologi Berbasis Kurikulum 2013 Sekolah Menengah Atas. *Skripsi Tidak Dipublikasikan*. Surabaya: FMIPA Unesa.
- Hidayani, S., Jamaluddin, J., Ramdani, A. 2021. Pemanfaatan Hasil Pengembangan Instrumen untuk Penilaian Literasi Sains Siswa pada Mata Pelajaran IPA di SMPN 2 Mataram. *Jurnal Pengabdian Magister Pendidikan IPA*, 4(1), pp.73-77.
- Husna, N., Nurmila, N. and Rochman, C., 2023. Analysis of The Final School Test Question Instrument for Islamic Religious Education and Ethics. *Educan: Jurnal Pendidikan Islam*, 7(2), pp.168-180.
- Indri, I.L., 2023. Pengaruh Metode Praktikum dengan Asesmen Kinerja terhadap Keterampilan Proses Sains dan Sikap Ilmiah pada Materi Mikrobiologi di SMK SMTI Bandar Lampung. *Doctoral dissertation, UIN Raden Intan Lampung*.
- Irwan, A.P., Usman, & Amin, B. D. 2019. Analisis Kemampuan Literasi Sains Peserta Didik Ditinjau Dari Kemampuan Menyelesaikan Soal Fisika di SMAN 2 Bulukumba. *Jurnal Sains dan Pendidikan Fisika (JSPF)*, 15(3), pp.17-24.
- Johari, J., Sahari, J., Wahab, D. A., Abdullah, S., Abdullah, S., Omar, M. Z., & Muhamad, N. (2011). Difficulty Index of Examinations and Their Relation to the Achievement of Programme Outcomes. *Procedia-Social and Behavioral Sciences*, 18, pp.71-80.
- Magdalena, I., Islami, N. F., Rasid, E. A., & Diasty, N. T. 2020. Tiga Ranah Taksonomi Bloom dalam Pendidikan. *EDISI*, 2(1), pp.132-139.
- Ndiung, S. and Jediut, M., 2020. Pengembangan Instrumen Tes Hasil Belajar Matematika Peserta Didik Sekolah Dasar Berorientasi pada Berpikir Tingkat Tinggi". *Premiere Educandum: Jurnal Pendidikan Dasar Dan Pembelajaran*, 10(1), pp.94-107.
- Nufus, S. H., Gani, A., & Suhendrayatna, S. (2017). "Pengembangan Instrumen Penilaian Sikap Berbasis Kurikulum 2013 pada Pembelajaran Kimia SMA". *Jurnal Pendidikan Sains Indonesia*, 5(1), pp.44-51.
- Nuraisyah, A.S., 2023. "Korelasi antara Kesiapan Belajar dan Disiplin Belajar terhadap Hasil Belajar Peserta Didik (Studi Korelasional Di Kelas X MIPA Pada Mata Pelajaran Biologi SMA Negeri 3 Tasikmalaya Tahun Ajaran 2021/2022)". (*Doctoral dissertation, Universitas Siliwangi*).
- Nurfadillah, T. and Elvia, R., 2023. "Pengembangan Instrumen Tes Kimia Berbasis Literasi Sains untuk Mengukur Literasi Sains Siswa". *ALOTROP*, 7(1), pp.44-56.
- Nurhalimah, S., Hidayati, Y., Rosidi, I. and Hadi, W.P., 2022. "Hubungan Antara Validitas Item dengan Daya Pembeda dan Tingkat Kesukaran Soal Pilihan Ganda". *Natural Science Education Research (NSER)*, 4(3), pp.249-257.
- Nurlia., Hala, Y., Muchtar, R., Jumadi, O., & Taiyeb, M.A. 2017. "Hubungan Antara Gaya Belajar dan Minat Belajar dengan Hasil Belajar Biologi Siswa". *Jurnal Pendidikan Biologi*.6(2), pp.321-328.
- OECD. 2019. *PISA 2018 Assessment and Analytical Framework*. Paris: OECD Publishing.
- OECD. 2023. *PISA Assesment 2022 and Analytical Framework*. Paris : OECD Publishing.
- Putranta, H., & Supahar, S. 2019. "Development of Physics-Tier Tests (PysTT) to Measure Students' Conceptual Understanding and Creative Thinking Skills: A Qualitative Synthesis". *Journal for the Education of Gifted Young Scientists*, 7(3), pp.747-775.
- Putri, O. D., Nevrita, N., & Hindrasti, N. E. K. 2019. "Pengembangan Instrumen Penilaian Keterampilan Berpikir Kritis Siswa Sma Pada Materi Sistem Pencernaan. *BIOEDUKASI (Jurnal Pendidikan Biologi)*, 10(1), pp.14-27.
- Rahmadani, F., Setiadi, D., Yamin, M. and Kusmiyati, K., 2022. Analisis Kemampuan Literasi Sains Biologi Peserta Didik SMA Kelas X di SMAN 1 Kuripan. *Jurnal Ilmiah Profesi Pendidikan*, 7(4b), pp.2726-2731.
- Sa'adah, R. N. 2021. *Metode Penelitian R&D (Research and Development) Kajian Teoretis dan Aplikatif*. CV Literasi Nusantara Abadi.
- Safitri, I., Lestarani, D., Imtikanah, R.D.N.W., Akbarini, N.R., Sari, M.W., Fitrah, M. and Hapsan, A., 2024. *Teori Pengukuran dan Evaluasi*. CV. Ruang Tentor.
- Sari, N., & Yogica, R. 2021. Multimedia Interaktif Bermuatan Game Tebak Kata tentang Materi Sistem Reproduksi pada Manusia untuk Siswa

- Kelas XI SMA. *Journal for Lesson and Learning Studies*, 4(3), pp.357-363.
- Setiawan, A. R. 2019. Menyusun Instrumen Penilaian untuk Pembelajaran Topik Lingkungan Berorientasi Literasi Sainifik. In *Seminar Nasional Fisika*, 1(1), pp. 7-14).
- Sudjana, N. 2016. *Penilaian Hasil Proses Belajar Mengajar*. Bandung: PT Remaja Rosdakarya
- Sugiyono. 2022. *Metode Penelitian Pendidikan (Pendekatan Kuantitatif, Kualitatif, dan R&D)*. Bandung: CV Alfabeta.
- Sutrisna, N., 2021. Analisis Kemampuan Literasi Sains Peserta Didik SMA di Kota Sungai Penuh, *Jurnal Inovasi Penelitian*, 1(12), pp.2683-2694.
- Riduwan. 2015. *Skala Pengukuran Variabel-Variabel Penelitian*. Bandung: Alfabeta.
- Rosidi, I., 2021. Profil Literasi Sains Aspek Kompetensi Siswa Pondok Pesantren di Masa Pandemi dengan Menggunakan Penilaian Berbasis Digital. *Natural Science Education Research (NSER)*, 4(1), pp.1-9.
- Rovita, C.A., 2020. Pengembangan Alat Evaluasi Pembelajaran Matematika Berbasis *Two Tier Multiple Choice* Menggunakan *Ispring Suite 9*". *Doctoral dissertation, Universitas Muhammadiyah Gresik*.
- Tulaiya, W. 2020. Analisis Kemampuan Literasi Sains Peserta Didik SMA/MA di Kabupaten Sumenep. *IPF: Inovasi Pendidikan Fisika*, 9(3), pp.417-42.