



Pengembangan Instrumen Three-Tier Test Diagnostik pada Materi Sistem Koordinasi untuk Mengidentifikasi Miskonsepsi Peserta Didik di SMAN 2 Sinjai

¹A. Aida Ariya, ²Abd. Muis, ³*Ismail

^{1,2,3}Program Studi Pendidikan Biologi, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Makassar, Makasar, Indonesia.

*Corresponding Author e-mail: ismail6131@unm.ac.id

Received: May 2025; Revised: May 2025; Accepted: June 2025; Published: June 2025

Abstrak: Tujuan dari penelitian ini adalah untuk mengembangkan instrumen tes diagnostik dan mengidentifikasi tingkat miskonsepsi yang dimiliki oleh siswa terhadap materi sistem koordinasi. Desain pengembangan penelitian ini menggunakan uji coba pada siswa kelas XI SMAN 12 Sinjai dan uji lapangan pada siswa kelas XI SMAN 2 Sinjai. Metode yang digunakan meliputi pengembangan soal tes, wawancara dan dokumentasi. Instrumen yang dihasilkan meliputi kisi-kisi, tes, dan kunci jawaban. Pengembangan soal tes telah melalui validasi oleh validator ahli, dan dilakukan uji reliabilitas, tingkat kesukaran dan daya beda terhadap semua uji coba lapangan (one to one, small group dan field tes). Dari total 30 butir soal yang telah dinyatakan valid, hanya 21 butir yang lolos hingga tahap akhir, yaitu uji coba lapangan (field test). Temuan penelitian ini mengungkapkan bahwa rata-rata sebesar 40,47% dari 50 siswa mengalami miskonsepsi, yang terbagi ke dalam dua indikator utama. Indikator dengan tingkat miskonsepsi tertinggi terkait dengan materi sistem koordinasi adalah pembahasan mengenai struktur sistem saraf pusat dan saraf tepi, serta struktur dan fungsi alat indera, dengan rata-rata miskonsepsi sebesar 56% pada butir soal yang menguji indikator tersebut. Berdasarkan hasil pengembangan instrumen tes diagnostik three-tier pada materi sistem koordinasi, dapat disimpulkan bahwa instrumen ini efektif dan layak digunakan untuk mendeteksi miskonsepsi yang dialami siswa pada topik tersebut.

Kata Kunci: Tes diagnostik; three-tier; materi sistem koordinasi; miskonsepsi

Abstract: The aim of this study was to develop a diagnostic test instrument and to identify the level of misconceptions held by students regarding the topic of the coordination system. The development design of this research involved trials with Grade XI students at SMAN 12 Sinjai and field testing with Grade XI students at SMAN 2 Sinjai. The methods employed included test item development, interviews, and documentation. The resulting instruments consisted of test blueprints, test items, and answer keys. The test items underwent expert validation and were subjected to reliability testing, difficulty level analysis, and item discrimination index evaluation through all trial stages (one-to-one, small group, and field testing). Out of 30 test items that were validated, only 21 advanced to the final stage, namely the field test. The findings revealed that, on average, 40.47% of the 50 students experienced misconceptions, which were categorized into two main indicators. The highest misconception rates in the coordination system topic were found in the indicators related to the structure of the central and peripheral nervous systems, as well as the structure and function of sensory organs, with an average misconception rate of 56% for the relevant test items. Based on the results of developing the three-tier diagnostic test on the coordination system material, it can be concluded that the instrument is effective and appropriate for identifying students' misconceptions on the topic.

Keywords: Diagnostic test; three-tier; coordination system material; misconception

How to Cite: Ariya, A., Muis, A., & Ismail, I. (2025). Pengembangan Instrumen Three-Tier Test Diagnostik pada Materi Sistem Koordinasi untuk Mengidentifikasi Miskonsepsi Peserta Didik di SMAN 2 Sinjai. *Bioscientist: Jurnal Ilmiah Biologi*, 13(2), 1281-1297. doi:<https://doi.org/10.33394/bioscientist.v13i2.15044>



<https://doi.org/10.33394/bioscientist.v13i2.15044>

Copyright©2025, Ariya et al

This is an open-access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) License.



PENDAHULUAN

Pembelajaran biologi memiliki peran penting dalam membantu peserta didik memahami struktur dan fungsi tubuh makhluk hidup serta fenomena kehidupan lainnya. Namun, banyak konsep dalam biologi yang bersifat abstrak dan kompleks sehingga seringkali menimbulkan kesulitan pemahaman, bahkan mengarah pada terbentuknya miskonsepsi. Miskonsepsi, atau pemahaman yang tidak sesuai dengan konsep ilmiah yang telah disepakati, dapat menghambat pemahaman lanjutan karena

konsep-konsep dalam biologi saling terkait satu sama lain (Ojose, 2015). Salah satu topik yang rawan menimbulkan miskonsepsi adalah sistem koordinasi. Pemahaman awal yang keliru pada materi ini akan berdampak pada kesalahan pemahaman konsep-konsep lanjutan, sesuai dengan pandangan konstruktivisme yang menekankan bahwa pembentukan pengetahuan baru sangat bergantung pada pengertian sebelumnya.

Penelitian terkait miskonsepsi menjadi penting untuk mendeteksi dan mengatasi kesalahan pemahaman sejak dini. Hal ini bertujuan agar pembelajaran berikutnya dapat berlangsung lebih efektif dan tidak terbebani oleh pemahaman yang salah (Hasyim *et al.*, 2018). Dalam konteks pembelajaran biologi, siswa umumnya dikategorikan ke dalam tiga kelompok berdasarkan pemahaman mereka, yaitu siswa yang benar-benar memahami konsep, siswa yang mengalami miskonsepsi, dan siswa yang tidak memahami konsep sama sekali. Miskonsepsi dapat muncul dalam berbagai bentuk, seperti kesalahan dalam menghubungkan gagasan, penerapan konsep yang keliru, atau interpretasi intuitif yang tidak tepat. Sarlina (2017) menjelaskan bahwa miskonsepsi dicirikan oleh tiga hal utama: (1) adanya perbedaan antara pemahaman individu dengan konsep ilmiah yang sebenarnya; (2) keyakinan yang tinggi terhadap pemahaman yang salah tersebut; dan (3) kecenderungan mempertahankan pemahaman tersebut meskipun diberikan informasi yang benar (Ibrahim, 2018).

Untuk mengidentifikasi miskonsepsi siswa secara akurat, diperlukan instrumen diagnostik yang tidak hanya menilai jawaban akhir, tetapi juga mengeksplorasi alasan di balik pilihan siswa serta tingkat keyakinannya. Tes diagnostik tiga tingkat hadir sebagai solusi yang lebih komprehensif dibandingkan dengan tes dua tingkat, karena mencakup tiga komponen utama: (1) jawaban pilihan ganda, (2) alasan yang mendasari pilihan tersebut, dan (3) tingkat keyakinan siswa terhadap jawabannya (Gurel *et al.*, 2015). Pendekatan ini memungkinkan guru untuk membedakan secara lebih akurat antara siswa yang mengalami miskonsepsi dan siswa yang sekadar belum memahami konsep. Dengan demikian, intervensi pembelajaran dapat dirancang secara lebih tepat sasaran (Mubarak *et al.*, 2016).

Tes diagnostik tiga tingkat juga berfungsi sebagai alat untuk: (1) mendiagnosis miskonsepsi siswa secara lebih mendalam; (2) mengidentifikasi bagian-bagian materi yang memerlukan penekanan lebih dalam proses pembelajaran; dan (3) merancang strategi pembelajaran yang efektif guna mengatasi miskonsepsi yang terjadi. Salah satu topik biologi yang dinilai cukup kompleks dan berisiko tinggi menyebabkan miskonsepsi adalah sistem koordinasi, khususnya sistem saraf. Materi ini sering dianggap abstrak dan dipenuhi dengan istilah-istilah ilmiah yang kompleks, sehingga menjadi tantangan tersendiri bagi peserta didik. Ketidakhahaman terhadap konsep-konsep sistem saraf dapat berdampak negatif pada keberlanjutan pembelajaran dan hasil belajar secara keseluruhan (Saputri *et al.*, 2016).

Beberapa studi terdahulu menunjukkan tingginya angka miskonsepsi pada berbagai topik biologi. Simatupang (2021) mencatat bahwa sebanyak 65,6% siswa kelas XII di SMA Negeri 1 Masjid Raya mengalami miskonsepsi pada materi asam dan basa. Sementara itu, penelitian Nurhalisah menunjukkan bahwa dalam topik klasifikasi makhluk hidup, hanya 23,04% siswa yang memahami konsep dengan baik, sedangkan 38,43% mengalami miskonsepsi, dan sisanya tidak memahami konsep sama sekali. Meskipun efektivitas tes diagnostik tiga tingkat telah dibuktikan dalam berbagai konteks, penerapannya secara spesifik pada materi sistem koordinasi masih sangat terbatas.

Berdasarkan latar belakang tersebut, penelitian ini difokuskan pada pengembangan instrumen tes diagnostik tiga tingkat untuk materi sistem koordinasi,

yang mencakup integrasi antara sistem saraf, sistem endokrin, dan sistem indera. Kompleksitas hubungan antar submateri ini menjadikan sistem koordinasi sebagai salah satu materi yang sangat potensial menimbulkan miskonsepsi. Penelitian ini bertujuan untuk mengembangkan dan mengevaluasi kelayakan instrumen tes diagnostik tiga tingkat pada materi sistem koordinasi bagi siswa kelas XI di SMAN 2 Sinjai. Evaluasi dilakukan melalui analisis validitas, reliabilitas, tingkat kesukaran, dan daya beda instrumen, sebagai dasar untuk mendeteksi miskonsepsi siswa secara lebih akurat dan mendalam.

METODE

Penelitian ini merupakan penelitian dan pengembangan (Research and Development) yang bertujuan utama untuk merancang, mengembangkan, serta menguji validitas dan reliabilitas instrumen tes diagnostik tiga tingkat (three-tier test) pada materi sistem koordinasi. Pengembangan instrumen ini ditujukan untuk mendeteksi secara lebih akurat pemahaman konseptual dan miskonsepsi siswa dalam pembelajaran Biologi, khususnya pada konsep-konsep kompleks seperti sistem saraf, sistem endokrin, dan sistem indera. Model pengembangan yang digunakan dalam penelitian ini mengacu pada model pengembangan pendidikan menurut Tessmer, yang terdiri dari dua tahapan utama, yaitu tahap evaluasi pendahuluan (preliminary evaluation) dan tahap evaluasi formatif (formative evaluation).

Tahap evaluasi formatif meliputi beberapa langkah sistematis, antara lain evaluasi diri (*self-evaluation*), pengembangan prototipe (*prototyping*), uji coba kelompok kecil (*small group evaluation*), dan uji coba lapangan (*field test*). Penelitian ini dilaksanakan pada semester genap tahun ajaran 2022/2023, tepatnya pada tanggal 5 hingga 15 Juni 2023. Subjek penelitian adalah peserta didik kelas XI di dua sekolah menengah atas, yaitu SMAN 12 Sinjai dan SMA Negeri 2 Sinjai. Lokasi pelaksanaan penelitian bertempat di Jalan Andi Massalinri dan Jalan Persatuan Raya No. 50 B, Kecamatan Sinjai Selatan, Kabupaten Sinjai. Dengan menggunakan pendekatan ini, penelitian diharapkan dapat menghasilkan instrumen asesmen diagnostik yang valid, reliabel, dan kontekstual sesuai dengan karakteristik siswa serta kebutuhan pembelajaran Biologi di tingkat SMA.

Tahap Preliminary (Tahap Persiapan).

Pada tahap awal, penelitian ini akan melibatkan pengkajian berbagai sumber referensi yang relevan. Setelah teori dan informasi terkumpul, langkah selanjutnya adalah menentukan lokasi dan subjek uji coba. Hal ini dilakukan dengan menghubungi kepala sekolah dan guru biologi di sekolah yang dipilih sebagai lokasi uji coba. Selain itu, akan dilakukan wawancara dengan guru Biologi untuk memperoleh gambaran mengenai pelaksanaan pembelajaran berdasarkan Kurikulum 2013, penggunaan instrumen tes diagnostik, serta tingkat pemahaman konsep siswa di sekolah tersebut. Wawancara ini akan dilakukan menggunakan pendekatan semi-terstruktur, sehingga memungkinkan fleksibilitas dalam menggali informasi yang lebih mendalam. Data hasil wawancara akan dianalisis secara deskriptif kualitatif untuk mengidentifikasi permasalahan pembelajaran yang berkaitan dengan miskonsepsi, serta digunakan sebagai dasar dalam merancang pengembangan instrumen tes diagnostik tiga tingkat.

Tahap Formative Evaluation (Tahap Penilaian Formatif).

Tahapan ini dilaksanakan setelah tahap preliminary dinyatakan sepenuhnya siap. Beberapa langkah yang dilakukan dalam tahap ini antara lain sebagai berikut:

- a. Self Evaluation (Evaluasi Diri) merupakan tahap di mana peneliti melakukan penilaian awal terhadap desain instrumen tes diagnostik yang akan dikembangkan.

Tahapan ini mencakup:

- 1) Analisis Kurikulum: Peneliti menelaah dokumen Kurikulum 2013 dan buku teks Biologi menggunakan teknik studi dokumen, guna memastikan kesesuaian indikator soal dengan kompetensi dasar serta mendapatkan gambaran konsep yang rawan miskonsepsi.
 - 2) Analisis Peserta Didik: Dilakukan melalui pengumpulan data nilai ulangan peserta didik, dan memahami kelas-kelas apakah tersebar atau apakah ad kelas unggulan.
 - 3) Analisis Materi: Peneliti menyusun pokok bahasan sistem koordinasi secara sistematis berdasarkan dokumen kurikulum dan referensi ilmiah, untuk mempermudah pengembangan butir soal diagnostik. Teknik yang digunakan adalah studi dokumen dan analisis kebutuhan materi.
- b. Pada tahap Desain, kegiatan yang dilakukan meliputi penentuan jumlah soal yang akan dibuat, penyusunan kisi-kisi soal tes diagnostik, serta pembuatan butir soal pilihan ganda tiga tingkat lengkap dengan kunci jawaban. Proses pembuatan soal mengikuti prinsip pengembangan instrumen, termasuk validitas isi dan validitas konstruk, untuk memastikan butir soal sesuai dengan tujuan pengukuran dan materi yang diajarkan. Selanjutnya, desain tes yang telah dibuat akan dikomunikasikan dengan dosen pembimbing, yang berperan memberikan masukan akademis dan teknis guna meningkatkan kualitas instrumen. Prototype instrumen hasil pengembangan ini kemudian dievaluasi berdasarkan aspek konten, konstruk, dan bahasa agar memenuhi standar kelayakan dan keakuratan.
- c. Pada tahap prototyping, produk yang telah dirancang akan diuji coba dalam tiga kelompok untuk memastikan kualitasnya:
- 1) *Expert Review* (Ulasan Para Ahli). Pada tahap ini, prototipe awal instrumen tes diagnostik tiga tingkat diserahkan kepada para ahli untuk dilakukan evaluasi dan validasi. Proses validasi menggunakan angket penilaian yang mencakup aspek isi, konstruk, dan bahasa, dengan skala Likert (1–5) untuk menilai kelayakan setiap butir soal. Para validator memberikan umpan balik berupa saran dan masukan tertulis melalui lembar validasi yang telah disediakan. Seluruh tanggapan tersebut dijadikan dasar untuk merevisi instrumen agar lebih valid dan sesuai dengan standar yang ditetapkan. Setelah revisi dilakukan dan seluruh aspek memperoleh penilaian yang memadai, instrumen dinyatakan valid dan layak untuk digunakan dalam pengujian.
 - 2) *One-to-one* (Perorangan). Hasil revisi dari tahap penilaian ahli kemudian diujicobakan kepada tiga peserta didik yang tidak termasuk dalam subjek penelitian utama. Ketiga peserta didik tersebut mewakili kategori kemampuan tinggi, sedang, dan rendah. Mereka diminta untuk menjawab instrumen tes diagnostik tiga tingkat yang telah dimodifikasi. Setelah itu, komentar dan tanggapan dari peserta didik dianalisis dan digunakan oleh peneliti untuk menyempurnakan desain tes diagnostik pilihan ganda tiga tingkat. Revisi yang dihasilkan pada tahap ini menghasilkan *Prototype II* sebagai bentuk pengembangan instrumen berikutnya.
 - 3) *Small Group* (kelompok kecil). Uji coba prototype II akan dilakukan pada enam siswa non-subjek penelitian, yang terdiri dari dua siswa dengan kemampuan tinggi, dua siswa dengan kemampuan sedang, dan dua siswa dengan kemampuan rendah. Pemilihan peserta dilakukan oleh guru mata pelajaran berdasarkan pengamatan terhadap keseharian siswa dan nilai rapor. Hasil dari uji coba tahap one-to-one digunakan sebagai dasar revisi untuk menyusun prototype II yang kemudian diuji dalam kelompok kecil ini. Selama uji coba, data dikumpulkan melalui observasi langsung dan pencatatan komentar atau

tanggapan lisan dari peserta didik. Peserta diminta memberikan masukan terkait kejelasan soal, tingkat kesulitan, serta keterpahaman terhadap tes diagnostik tiga tingkat yang diuji. Komentar dan tanggapan peserta didik dianalisis secara kualitatif oleh peneliti, kemudian digunakan untuk merevisi desain instrumen menuju tahap berikutnya. Perbaikan dilakukan pada aspek redaksi soal, pilihan jawaban, tingkat kompleksitas, maupun format penyajian, agar instrumen lebih akurat dalam mengidentifikasi miskonsepsi. Pada tahap ini, peneliti berharap dapat menghasilkan tes diagnostik tiga tingkat yang lebih valid dan efektif dalam menemukan miskonsepsi siswa. Prototype hasil revisi ini selanjutnya disebut sebagai prototype III, yang merupakan penyempurnaan dari prototype II hasil evaluasi tahap one-to-one.

- 4) *Field Tes* (Uji Coba Lapangan). Pada tahap ini, *Prototype III* yang telah direvisi berdasarkan hasil uji coba kelompok kecil diuji cobakan kepada subjek penelitian, yaitu peserta didik kelas XI SMAN 2 Sinjai. Pengumpulan data dilakukan dengan meminta peserta didik mengerjakan tes diagnostik tiga tingkat dalam kondisi terkontrol. Data yang diperoleh dari uji lapangan dianalisis melalui beberapa tahapan, yaitu analisis reliabilitas untuk mengukur konsistensi instrumen, analisis validitas untuk memastikan bahwa instrumen benar-benar mengukur konsep yang dimaksud, serta analisis butir soal untuk menilai kualitas setiap butir soal. Hasil dari analisis ini menjadi dasar penentuan kelayakan instrumen sebelum digunakan secara lebih luas dalam konteks pembelajaran biologi di sekolah menengah atas.

Metode pengumpulan data dalam penelitian ini digunakan untuk menguji validitas instrumen, yang mencakup validitas isi dan validitas empiris. Validitas isi diperoleh melalui penilaian kualitatif oleh validator ahli, yang mengevaluasi kesesuaian butir soal dengan materi dan tujuan pembelajaran. Validitas empiris, di sisi lain, dianalisis secara statistik berdasarkan data hasil tes untuk menilai sejauh mana butir soal berkorelasi dengan kemampuan peserta didik. Validitas isi berfokus pada kualitas konten soal, sedangkan validitas empiris mencerminkan efektivitas soal berdasarkan bukti empiris.

1. Uji Kevalidan

Validitas empirik ditentukan melalui uji validitas butir soal menggunakan program statistik SPSS versi 25.0. Uji ini dilakukan dengan menganalisis korelasi antara skor setiap item dengan skor total (corrected item-total correlation), yang menghasilkan nilai r hitung. Sampel yang digunakan dalam validasi ini pada tahap *field tes* terdiri dari 30 peserta didik kelas XI yang memiliki karakteristik heterogen dari segi kemampuan akademik. Nilai r hitung kemudian dibandingkan dengan nilai r tabel pada taraf signifikansi 5% ($\alpha = 0,05$) dengan derajat kebebasan ($df = n - 2$). Jika r hitung > r tabel, maka butir soal dinyatakan valid. Sebaliknya, jika r hitung < r tabel, maka butir soal dianggap tidak valid dan akan direvisi atau dihapus sesuai kebutuhan pengembangan instrumen. Selain uji empirik, validitas isi juga dianalisis berdasarkan penilaian para validator, yang terdiri atas dua bentuk data: data kuantitatif berupa skor penilaian yang dihitung rata-ratanya dalam bentuk persentase, dan data kualitatif berupa kritik serta saran sebagai acuan dalam melakukan perbaikan instrumen. Pada lembar penilaian, setiap item dinilai menggunakan skala Likert 1–5, dengan kategori: 5 (Sangat Baik), 4 (Baik), 3 (Cukup), 2 (Kurang), dan 1 (Sangat Kurang).

Tabel 1. Kriteria kevalidan.

Nilai	Interpretasi
0.80-1.00	Sangat Baik
0.60-0.79	Baik
0.40-0.59	Cukup
0.20-0.39	Kurang
0.00-1.99	Sangat Kurang

(Sumber: Riyani *et al.*, 2017)

2. Uji Reliabilitas.

Pengujian reliabilitas dilakukan dengan memasukkan data ke dalam Microsoft Excel dan dianalisis menggunakan rumus KR-20. Rumus KR-20 digunakan untuk menguji reliabilitas tes pilihan ganda yang berskala dikotomis (benar-salah). KR-20 dipilih karena lebih sesuai untuk jenis item tes diagnostik tiga tingkat yang menggunakan jawaban tertutup, dibandingkan dengan Cronbach's Alpha yang lebih umum untuk skala non-dikotomis. Uji reliabilitas dilakukan terhadap data hasil jawaban 30 peserta didik kelas XI yang memiliki tingkat kemampuan beragam. Perhitungan dilakukan menggunakan Microsoft Excel, dengan mempertimbangkan proporsi jawaban benar dan varians total skor. Nilai reliabilitas yang tinggi menunjukkan bahwa instrumen memiliki konsistensi internal yang baik.

Tabel 2. Uji reliabilitas

Nilai	Kriteria
0.80-1.00	Sangat Kuat
0.60-0.79	Kuat
0.40-0.59	Sedang
0.20-0.39	Rendah
0.00-0.19	Sangat Rendah

(Sumber: Riyani *et al.*, 2017)

3. Analisis Tingkat Kesukaran

Uji tingkat kesukaran dilakukan untuk mengetahui sejauh mana suatu butir soal dapat dijawab dengan benar oleh peserta tes. Tingkat kesukaran dihitung sebagai proporsi antara jumlah peserta yang menjawab benar dengan total peserta tes. Semakin besar proporsi ini, semakin mudah soal tersebut, dan sebaliknya. Perhitungan dilakukan menggunakan SPSS versi 25.0. Tingkat kesukaran yang ideal berada pada rentang sedang, karena mampu membedakan antara peserta didik yang menguasai materi dan yang belum.

Tabel 3. Kriteria kesukaran soal

Nilai	Kriteria
0,00-0,20	Sangat Sukar
0,21-0,40	Sukar
0,41-0,60	Sedang
0,61-0,80	Mudah
0,81-1,00	Sangat Mudah

(Sumber: Riyani *et al.*, 2017)

4. Analisis Daya Beda

Indeks daya pembeda soal mengukur bagaimana sebaiknya sebuah soal membedakan tingkat kemampuan siswa. Indeks daya pembeda soal bernilai – 1,00

sampai + 1,00. Hasil analisis daya beda dilakukan perbandingan pada hasil uji validitas di aplikasi SPSS *statistic* 25.0.

Tabel 4. Penafsiran Angka Indeks

Nilai	Kriteria
0.00-0.20	Lemah
0.21-0.40	Sedang
0.41-0.70	Kuat
0.71-1.00	Sangat Kuat

(Sumber: Riyani *et al.*, 2017)

Analisis Tes Diagnostik *Three-Tier*

Tes diagnostik *three-tier* merupakan alat penilaian yang efektif untuk mengidentifikasi tingkat pemahaman konseptual dan miskonsepsi siswa secara lebih akurat dibandingkan tes konvensional. Model ini dikembangkan pertama kali oleh Treagust (1988) melalui kerangka "diagnostic multiple choice instrument" dan kemudian dimodifikasi menjadi bentuk tiga tingkat oleh Caleon dan Subramaniam (2010).

a. Tahap pertama penilaian diberikan pada tingkat pertama soal. Setiap jawaban benar diberi skor 1 dan jika salah diberi nilai 0.

Tabel 5. Kriteria penskoran tahap 1

Kriteria	Skor
Benar	1
Salah	0

b. Tahap kedua dilakukan penilaian berdasarkan hasil dari tingkat pertama dan kedua.

Tabel 6. Kriteria penskoran tahap 2

Tingkat Pertama	Tingkat Kedua	Skor
Benar (1)	Benar (1)	1
Benar (1)	Salah (0)	0
Salah (0)	Benar (1)	0
Salah (0)	Salah (0)	0

c. Tahap ketiga dilakukan dengan melihat jawaban secara keseluruhan untuk setiap tingkatan.

Tabel 7. Kriteria penskoran tahap 3

Tingkat Pertama	Tingkat Kedua	Skor	Tingkat Ketiga	Skor
Benar (1)	Benar (1)	1	Yakin (1)	1
Benar (1)	Benar (1)	1	Tidak Yakin (0)	0
Benar (1)	Salah (0)	0	Yakin (1)	0
Benar (1)	Salah (0)	0	Tidak Yakin (0)	0
Salah (0)	Benar (1)	0	Yakin (1)	0
Salah (0)	Benar (1)	0	Tidak Yakin (0)	0
Salah (0)	Salah (0)	0	Yakin (1)	0
Salah (0)	Salah (0)	0	Tidak Yakin (0)	0

2. Analisis Kombinasi Jawaban Tes Diagnostik *Three-Tier*

Analisis kombinasi jawaban pada tes diagnostik tiga tingkat digunakan sebagai alat untuk mengidentifikasi tingkat pemahaman konsep peserta didik, apakah termasuk dalam kategori memahami konsep, tidak memahami konsep, atau mengalami miskonsepsi.

Tabel 9. Kombinasi Jawaban Tes Diagnostic .

Analisis Tingkat Sosial	Tipe Jawaban			Kategori
	Tingkat Pertama	Tingkat Kedua	Tingkat Ketiga	
Tes diagnostik <i>three-tier</i>	Jawaban benar (1)	Alasan benar (1)	Yakin (1)	Paham konsep
	Jawaban benar (1)	Alasan benar (1)	Tidak yakin (0)	
	Jawaban salah (0)	Alasan benar (1)	Tidak yakin (0)	
	Jawaban benar (1)	Alasan salah (0)	Tidak yakin (0)	Tidak paham konsep
	Jawaban salah (0)	Alasan salah (0)	Tidak yakin (0)	
	Jawaban salah (0)	Alasan benar (1)	Yakin (1)	
	Jawaban salah (0)	Alasan salah (0)	Yakin (1)	Miskonsepsi
	Jawaban benar (1)	Alasan salah (0)	Yakin (1)	

HASIL DAN PEMBAHASAN

Instrumen tes diagnostik *three-tier* yang telah mengalami proses revisi dan penilaian pada tahap *expert review*, selanjutnya diuji cobakan melalui tahapan *one to one* dan kelompok kecil (*small group*) di SMAN 12 Sinjai, kemudian dilanjutkan dengan uji lapangan (*field test*) di SMAN 2 Sinjai.

Tabel 10. Persentase uji validitas tahap *expert review*

Validator	Nilai Validitas	Rata-Rata Nilai Validitas	Kategori
1	86,15	85,38	Sangat Valid
2	84,61		

Hasil penilaian dari tahap *expert review* menunjukkan bahwa instrumen memperoleh nilai validitas yang tinggi berdasarkan analisis kuantitatif dan kualitatif. Secara kuantitatif, skor rata-rata yang diberikan oleh para ahli menunjukkan bahwa instrumen termasuk dalam kategori valid dan layak untuk digunakan. Sementara itu, secara kualitatif, para ahli memberikan masukan berupa kritik dan saran perbaikan yang digunakan sebagai acuan untuk merevisi butir-butir soal sebelum diujicobakan ke peserta didik. Revisi dilakukan untuk memastikan bahwa setiap item tes memiliki kejelasan isi, kesesuaian dengan indikator konsep, dan kemampuan untuk mendeteksi miskonsepsi secara akurat.

Tahap *One to One* (Perorangan)

a. Uji Kevalidan

Tahap *one to one* dilakukan pada siswa kelas XI di SMA 2 Sinjai. Tahapan *One to one* dengan melibatkan 3 siswa dengan kategori kemampuan berbeda yang dipilih oleh guru biologi yaitu dengan kemampuan rendah, kemampuan sedang, dan kemampuan tinggi. 3 siswa tersebut diberikan instrumen tes diagnostik *three-tier* yang telah melalui tahap *expert review* (ulasan para ahli) dengan soal sebanyak 30 butir soal pada materi sistem koordinasi yang telah dikembangkan.

Tabel 11. Hasil analisis validitas tahap *one to one*

Kategori	Butir Soal	Jumlah Soal	Persentase
Valid	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 16, 17, 18, 20, 21, 22, 24, 25, 28, 29	24	80%

Kategori	Butir Soal	Jumlah Soal	Persentase
Tidak Valid	15, 19, 23, 26, 27, 30	6	20%

Dari total 30 butir soal yang dikembangkan, sebanyak 24 butir soal dinyatakan valid berdasarkan hasil perhitungan nilai korelasi item-total dengan kriteria $r_{hitung} > r_{tabel}$, pada taraf signifikansi 0,05, yaitu sebesar 0,996. Dengan demikian, 6 butir soal (20%) tidak memenuhi kriteria validitas karena memiliki $r_{hitung} \leq r_{tabel}$, sehingga tidak digunakan pada tahap selanjutnya. Ketidakvalidan 6 butir soal pada tahap one-to-one disebabkan oleh rendahnya keterkaitan antara skor item dengan skor total, yang dapat mengindikasikan ketidakjelasan konstruksi soal atau ketidaksesuaian dengan indikator konsep. Butir-butir ini kemudian direvisi atau dieliminasi untuk meningkatkan kualitas instrumen. Sementara itu, reliabilitas yang dihitung menggunakan rumus KR-20 pada tahap one-to-one mencapai nilai maksimum 1,00. Hal ini dapat terjadi karena jumlah peserta yang sangat sedikit ($n=3$), sehingga menghasilkan konsistensi internal yang tampak sempurna. Namun, pada saat uji lapangan dengan jumlah peserta yang lebih besar dan karakteristik yang lebih bervariasi, reliabilitas KR-20 menurun menjadi 0,53. Penurunan ini mencerminkan bahwa instrumen perlu penyempurnaan lebih lanjut agar memiliki kestabilan dan konsistensi yang baik dalam populasi yang lebih luas.

b. Uji Reliabilitas

Hasil analisis reliabilitas pada tahap *one-to-one* sebagaimana disajikan pada Tabel 12 berikut ini.

Tabel 12. Hasil analisis reliabilitas tahap *one to one*

Jumlah Soal	Nilai Reliabilitas KR 20	Interpretasi reliabilitas	Keterangan
30	1,00	0.80-1.00	Sangat Kuat

Hasil analisis reliabilitas pada tahap *one-to-one* menunjukkan bahwa instrumen memperoleh nilai reliabilitas sebesar 1,00, yang termasuk dalam kategori sangat tinggi. Nilai ini mengindikasikan bahwa butir-butir soal memiliki konsistensi internal yang sangat kuat dan dapat diandalkan untuk digunakan dalam tahap uji awal.

c. Uji Tingkat Kesukaran

Hasil analisis tingkat kesukaran butir soal sebagaimana disajikan pada Tabel 13 berikut ini.

Tabel 13. Hasil analisis taraf kesukaran tahap *one to one*

Kriteria	Nomor Soal	Jumlah Soal	Persentase
Sukar	2, 6, 19, 26, 27, 30	6	20%
Mudah	1, 3, 4, 5, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 20, 21, 22, 24, 25, 28, 29	23	76,66%
Sangat Mudah	23	1	3.34%

Hasil analisis tingkat kesukaran pada tahap *one-to-one* menunjukkan bahwa dari 30 butir soal yang diuji, sebanyak 6 butir soal (20%) tergolong dalam kategori sukar. Sementara itu, 23 butir soal (76,66%) berada dalam kategori mudah, dan 1 butir soal (3,34%) dikategorikan sebagai sangat mudah. Temuan ini memberikan gambaran awal mengenai sebaran tingkat kesukaran butir soal, yang selanjutnya digunakan sebagai dasar untuk revisi dan penyempurnaan instrumen sebelum tahap uji berikutnya.

d. Uji Daya Beda

Hasil analisis tingkat kesukaran butir soal sebagaimana disajikan pada Tabel 14 berikut ini.

Tabel 14. Hasil analisis daya beda tahap *one to one*

Kriteria	Nomor Soal	Jumlah Soal	Persentase
Negatif	15, 23	2	6,66%
Lemah	19, 26	2	6,66%
Kuat	30	1	3,33%
Sangat Kuat	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 16, 17, 18, 20, 21, 22.	20	66,66%

Hasil analisis daya pembeda pada tahap *one-to-one* menunjukkan bahwa dari 30 butir soal yang diuji, terdapat 2 butir soal (6,66%) yang masuk dalam kategori negatif. Artinya, butir soal tersebut memiliki indeks daya pembeda negatif, yang menunjukkan bahwa peserta dengan skor rendah justru menjawab benar soal tersebut. Selain itu, terdapat 2 butir soal (6,66%) dengan daya pembeda berkategori lemah, 1 butir soal (3,33%) berkategori kuat, dan 20 butir soal (66,66%) berkategori sangat kuat. Hasil ini menjadi dasar untuk mempertimbangkan revisi atau penggantian terhadap butir soal yang tidak memberikan kontribusi optimal dalam membedakan kemampuan peserta didik.

Tahap *Small Group* (Kelompok Kecil)

a. Uji kevalidan

Uji kevalidan pada tahap *small group* dilaksanakan pada peserta didik kelas XI di SMAN 12 Sinjai, dengan melibatkan siswa yang berbeda dari tahap *one-to-one*. Pada tahap ini, sebanyak enam peserta didik dipilih oleh guru mata pelajaran Biologi berdasarkan kemampuan akademik mereka, yang mencakup dua siswa dengan kemampuan tinggi, dua siswa dengan kemampuan sedang, dan dua siswa dengan kemampuan rendah. Keenam siswa tersebut diberikan instrumen tes diagnostik *three-tier* yang telah direvisi berdasarkan hasil uji coba sebelumnya, dengan jumlah 24 butir soal yang berfokus pada materi sistem koordinasi.

Tabel 15. Hasil analisis validitas tahap *small group*

Kategori	Butir Soal	Jumlah Soal	Persentase
Valid	1, 2, 3, 4, 5, 6, 7, 8, 9, 11, 13, 14, 15, 16, 18, 19, 20, 21, 22, 23, 24	21	87,5%
Tidak Valid	10,12,17	3	12,5%

Hasil analisis validitas pada tahapan *small group* menunjukkan bahwa dari 24 butir soal yang diuji pada enam siswa, sebanyak 21 butir soal dinyatakan valid dengan nilai $r_{hitung} > r_{tabel}$ pada taraf signifikansi 0,05 ($r_{tabel} = 0,811$ atau sebesar 87,5%). Tiga butir soal dinyatakan tidak valid karena memiliki $r_{hitung} \leq r_{tabel}$ dengan persentase sebesar 12,5%. Soal-soal yang tidak valid tersebut tidak dilanjutkan ke tahap berikutnya, karena dianggap tidak mampu mengukur kemampuan siswa secara tepat. Ketidakvalidan kemungkinan disebabkan oleh ambiguitas dalam redaksi soal atau ketidaksesuaian antara indikator dan item. Oleh karena itu, soal-soal tersebut perlu direvisi sebelum dipertimbangkan untuk digunakan kembali.

b. Uji Reliabilitas

Hasil analisis tingkat reliabilitas tahap *small group* sebagaimana disajikan pada Tabel 14 berikut ini.

Tabel 14. Hasil analisis reliabilitas tahap *small group*

Jumlah Soal	Nilai Reliabilitas KR 20	Interpretasi reliabilitas	Keterangan
24	0,94	0.80-1.00	Sangat Kuat

Hasil analisis reliabilitas pada tahap *small group* menunjukkan nilai reliabilitas sebesar 0,94 yang tergolong dalam kategori sangat kuat. Nilai ini menunjukkan bahwa instrumen tes diagnostik three-tier memiliki konsistensi internal yang tinggi dan layak digunakan untuk mengukur pemahaman siswa secara andal.

c. Uji Tingkat Kesukaran

Analisis taraf kesukaran pada uji *small group* dengan banyaknya soal yaitu 24 butir soal.

Tabel 15. Hasil analisis taraf kesukaran tahap *small group*

Kriteria	Nomor soal	Jumlah Soal	Persentase
Sukar	16, 18, 19, 20, 21	5	20,83%
Sedang	4, 6, 15, 17, 23	5	20,83%
Mudah	1, 2, 3, 5, 7, 8, 9, 10, 11, 12, 13, 14, 22, 24	14	58,33%

Hasil analisis pada tahap *small group* menunjukkan bahwa dari 24 butir soal, sebanyak 5 soal (20,83%) tergolong sukar, 5 soal (20,83%) tergolong sedang, dan 14 soal (58,33%) tergolong mudah. Klasifikasi ini mencerminkan variasi tingkat kesukaran yang cukup seimbang untuk mengukur kemampuan siswa secara menyeluruh.

d. Uji Daya Beda

Hasil analisis daya beda dilakukan perbandingan hasil uji validitas di aplikasi SPSS *statistic* 25.0. Setelah melalui penyeleksian soal di tahap *one to one*.

Tabel 16. Hasil analisis daya beda tahap *small group*

Kriteria	Nomor Soal	Jumlah Soal	Persentase
Negatif	10, 12, 17	3	12,5%
Sangat Kuat	1, ,2, 3, 4, 5, 6, 7, 8, 9, 11, 13, 14, 15, 16, 18, 19, 20, 21, 22, 23, 24	21	87,5%

Hasil analisis tahap *small group* menunjukkan bahwa dari 24 butir soal, 3 soal (12,5%) memiliki daya beda negatif, mengindikasikan bahwa siswa dengan skor rendah justru menjawab benar. Sementara itu, 21 soal (87,5%) termasuk kategori sangat kuat karena mampu membedakan dengan baik antara siswa berkemampuan tinggi dan rendah.

Tahap *Field Tes* (Uji Lapangan)

a. Uji Kevalidan

Tahap *field test* dilaksanakan pada siswa kelas XI Mipa 1 dan 2 SMA 2 Sinjai dengan melibatkan 50 siswa tanpa klasifikasi kemampuan oleh guru. Instrumen tes diagnostik *three-tier* yang digunakan terdiri dari 21 butir soal valid hasil seleksi pada

tahap *one-to-one* dan *small group*, yang mengukur pemahaman siswa pada materi sistem koordinasi.

Tabel 17. Hasil analisis validitas tahap *field tes*

Kategori	Butir Soal	Jumlah Soal	Persentase
Valid	1, 2, 3, 4, 5, 6, 7, 8, 10, 11,12, 13, 14, 16, 17, 18,19	19	90,48%
Tidak Valid	9,15	2	9,52%

Hasil analisis validitas pada tahap *field test* menunjukkan bahwa dari 21 butir soal yang sebelumnya dinyatakan valid pada tahap *small group*, sebanyak 19 soal tetap valid (90,48%) dengan nilai rhitung > rtabel (0,285) pada taraf signifikansi 0,05. Sementara itu, 2 soal (9,52%) dinyatakan tidak valid karena nilai rhitung \leq rtabel , sehingga tidak digunakan pada tahap selanjutnya.

b. Uji Reliabilitas

Hasil analisis reliabilitas tahap Field Tes sebagaimana disajikan pada Tabel 18 berikut ini.

Tabel 18. Hasil analisis reliabilitas tahap *field tes*

Jumlah Soal	Nilai Reliabilitas KR 20	Interpretasi reliabilitas	Keterangan
21	0,53	0.40-0,59	Sedang

Hasil analisis reliabilitas pada tahap field test menunjukkan nilai sebesar 0,53, yang termasuk dalam kategori sedang, namun tetap menunjukkan bahwa instrumen cukup reliabel untuk digunakan dalam pengukuran diagnostik pada konteks ini.

c. Uji Tingkat Kesukaran

Hasil analisis taraf kesukaran pada *field tes* dengan banyaknya soal yaitu 21 butir soal sebagaimana disajikan pada Tabel 19.

Tabel 19. Hasil analisis taraf kesukaran tahap *field tes*

Kriteria	Nomor Soal	Jumlah Soal	Persentase
Sukar	2,11,17,18	4	19,05%
Sedang	1, 3, 4, 5, 6, 7, 8, 9, 10, 12, 13, 14, 15, 16, 19, 20, 21	17	80,95%

Hasil analisis pada tahap *field test* menunjukkan bahwa dari 21 butir soal, sebanyak 4 soal (19,05%) tergolong dalam kriteria sukar, sedangkan 17 soal (80,95%) tergolong sedang. Tidak terdapat soal dengan kategori sangat mudah atau sangat sukar pada tahap ini.

d. Uji Daya Beda

Hasil analisis daya beda dilakukan dengan membandingkan hasil uji validitas menggunakan aplikasi SPSS Statistics 25.0, berdasarkan butir soal yang telah diseleksi pada tahap *small group*. Analisis ini bertujuan untuk mengukur sejauh mana setiap butir soal mampu membedakan antara siswa berkemampuan tinggi dan rendah.

Tabel 20. Hasil analisis daya beda tahap *field tes*

Kriteria	Nomor Soal	Jumlah Soal	Persentase
Lemah	15	1	4,76%
Sedang	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 16, 17, 18, 19, 20, 21	20	95,23%

Hasil tes lapangan (*field test*) terhadap instrumen diagnostik tiga tingkat pada materi sistem koordinasi menunjukkan bahwa dari 21 butir soal, sebanyak 95,23% respon siswa termasuk dalam kategori paham konsep, miskonsepsi, dan tidak paham konsep, sedangkan 4,76% sisanya berasal dari soal dengan kriteria daya beda lemah.

Tabel 21. Persentase kategori kombinasi jawaban peserta didik

Kategori	Jumlah	Persentase
Paham Konsep	271	25,80 %
Tidak Paham Konsep	353	33,61 %
Miskonsepsi	425	40,47 %

Jumlah siswa yang termasuk dalam kategori paham konsep adalah 25,80 persen, persentase siswa yang tidak paham konsep adalah 33,61 persen, dan persentase siswa yang mengalami miskonsepsi adalah 40,47 persen. Menurut hasil analisis data tes diagnostik tiga tingkat untuk mengidentifikasi pemahaman konsep dan miskonsepsi siswa, yaitu: (1) Paham konsep, persentase hasil analisis kategori pemahaman konsep peserta didik menunjukkan bahwa butir soal yang paling banyak dalam kategori paham konsep adalah butir soal nomor 1 dimana persentase paham konsep sebesar 42 %. (2) Tidak Paham Konsep, persentase hasil analisis kategori pemahaman konsep peserta didik menunjukkan bahwa butir soal yang paling banyak dalam kategori tidak paham konsep adalah butir soal nomor 10 dimana persentase tidak paham konsep sebesar 50 %. (3) Miskonsepsi, dapat dilihat persentase pemahaman konsep butir soal nomor 12 dan 20 pada Persentase hasil analisis kategori miskonsepsi pada peserta didik menunjukkan bahwa butir soal yang paling banyak dalam kategori miskonsepsi adalah butir soal ke ke-12 dan 20 dengan persentase miskonsepsi masing-masing sebesar 56%.

Salah satu soal yaitu nomor 12 mengenai jalur busur refleksi, diketahui bahwa persentase siswa yang mengalami miskonsepsi tergolong tinggi. Hal ini menunjukkan bahwa soal tersebut memunculkan pemahaman yang keliru pada sebagian besar peserta didik. Soal ini menanyakan jalur yang dilalui oleh impuls pada refleksi lutut, yang secara biologis seharusnya mengikuti urutan: reseptor di lutut – saraf sensorik – sumsum tulang belakang – saraf motorik – efektor (kaki). Jawaban yang benar adalah pilihan B, dengan alasan bahwa jalur refleksi tidak melalui otak melainkan hanya melibatkan sumsum tulang belakang. Namun, banyak siswa justru memilih jawaban yang menyebutkan keterlibatan otak seperti pada opsi C dan D, atau terganggu oleh istilah yang membingungkan seperti “saraf konektor menyilang” pada opsi E.

Tingginya tingkat miskonsepsi ini menunjukkan bahwa konsep dasar mengenai jalur refleksi belum dipahami secara menyeluruh oleh siswa. Beberapa kemungkinan penyebabnya antara lain kompleksitas konsep sistem saraf yang melibatkan beberapa jenis sel saraf dengan fungsi berbeda (sensorik, intermediat, dan motorik), serta penggunaan bahasa soal dan pilihan jawaban yang mengandung istilah teknis yang kurang familiar bagi siswa. Selain itu, alasan yang dipilih siswa dalam tier ketiga, seperti “selalu diakhiri oleh sensorik” atau “jalurnya tidak melalui kaki”, semakin

menegaskan adanya miskonsepsi mengenai fungsi dan arah kerja saraf dalam sistem refleksi. Kesalahan ini bukan sekadar tidak paham, tetapi merupakan bentuk pemahaman keliru yang telah terbentuk dan perlu diluruskan melalui pembelajaran yang lebih menekankan pada visualisasi jalur refleksi serta penekanan pada perbedaan antara refleksi dan respon sadar.

Selama tahap penilaian ahli, isi butir soal tes diagnostik three-tier telah melalui proses validasi. Beberapa butir soal direvisi agar sesuai dengan indikator dan konsep yang telah disarankan oleh para validator. Perbaikan dilakukan dengan menyesuaikan materi pada sistem koordinasi serta penggunaan bahasa yang baik dan benar. Setelah proses revisi, butir-butir soal dinyatakan layak untuk diujicobakan kepada peserta didik kelas XI SMA.

1) Tahap *One to One*

Di SMA 2 Sinjai, tahap uji coba one-to-one dilakukan terhadap tiga peserta didik yang dipilih oleh guru biologi berdasarkan kategori kemampuan akademik, yaitu tinggi, sedang, dan rendah. Berdasarkan hasil analisis terhadap 30 butir soal yang diujikan, sebanyak 24 butir dinyatakan valid, sementara 6 butir soal tidak valid dan tidak digunakan pada tahap selanjutnya. Instrumen tes diagnostik three-tier yang telah diuji menggunakan perangkat lunak Microsoft Excel menunjukkan reliabilitas sangat tinggi pada tahap ini, yang mengindikasikan bahwa instrumen cukup stabil dan konsisten untuk mengidentifikasi miskonsepsi siswa pada materi sistem koordinasi.

Selain validitas dan reliabilitas, analisis tingkat kesukaran soal juga dilakukan. Hasil menunjukkan bahwa dari 30 butir soal, sebanyak 6 butir termasuk kategori sukar, 23 butir termasuk kategori mudah, dan 1 butir termasuk kategori sangat mudah. Uji daya pembeda dilakukan untuk menilai kemampuan soal dalam membedakan antara siswa dengan kemampuan tinggi dan rendah. Dari analisis tersebut diperoleh bahwa 20 butir soal memiliki daya beda sangat kuat, 1 butir tergolong kuat, 2 butir lemah, dan 2 butir memiliki daya beda negatif. Soal dengan indeks daya pembeda yang tinggi menunjukkan kualitas yang baik karena mampu mengidentifikasi secara efektif perbedaan kemampuan akademik antar peserta didik.

2) Tahap *Small Group*

Setelah pelaksanaan uji coba tahap one-to-one, dilakukan analisis dan revisi terhadap butir soal. Sebanyak enam soal yang tidak memenuhi kriteria validitas dihilangkan, sehingga tersisa 24 soal dari total 30 butir soal awal untuk diuji pada tahap kelompok kecil. Uji coba tahap ini melibatkan enam peserta didik dengan kategori kemampuan yang berbeda, yaitu dua siswa berkemampuan tinggi, dua siswa berkemampuan sedang, dan dua siswa berkemampuan rendah. Berdasarkan hasil analisis, dari 24 butir soal yang diujikan, sebanyak 21 soal dinyatakan valid, sedangkan 3 soal tidak valid dan karenanya dieliminasi dari instrumen.

Instrumen tes diagnostik tiga tingkat yang telah direvisi menunjukkan tingkat reliabilitas yang sangat tinggi pada tahap small group. Analisis reliabilitas dilakukan menggunakan Microsoft Excel, dan hasilnya menunjukkan bahwa instrumen cukup konsisten dalam mengukur pemahaman siswa. Selain itu, analisis tingkat kesukaran menunjukkan bahwa dari 24 soal, sebanyak 5 soal tergolong sukar, 5 soal berada pada kategori sedang, dan 14 soal tergolong mudah. Sementara itu, berdasarkan analisis daya pembeda, ditemukan bahwa 21 soal memiliki daya pembeda sangat kuat, sedangkan 3 soal menunjukkan daya pembeda negatif.

Selanjutnya, pada tahap field test, sebanyak 50 siswa dari dua kelas di SMA 2 Sinjai dengan kemampuan akademik yang beragam dilibatkan sebagai responden. Instrumen yang digunakan merupakan hasil revisi dari tahap sebelumnya, terdiri atas 21 soal yang telah lolos validasi. Hasil uji validitas menunjukkan bahwa 19 soal

dinyatakan valid, sedangkan 2 soal tidak memenuhi kriteria validitas. Dari segi tingkat kesukaran, 4 soal tergolong sukar dan 17 soal berada pada kategori sedang. Hasil analisis reliabilitas pada tahap ini menunjukkan bahwa instrumen memiliki reliabilitas pada kategori sedang, yang menandakan bahwa instrumen tersebut cukup konsisten untuk mendeteksi miskonsepsi siswa pada materi sistem koordinasi. Validitas suatu butir soal menjadi indikator utama dalam menentukan kelayakan penggunaannya. Soal yang tidak valid tidak dapat dijadikan alat ukur yang representatif. Sebaliknya, jika suatu soal dinyatakan valid, maka penilaiannya dapat dilanjutkan dengan mempertimbangkan aspek tingkat kesukaran dan daya pembeda. Suatu soal dianggap berkualitas baik apabila memenuhi minimal dua dari tiga indikator tersebut, sehingga layak digunakan dalam evaluasi pembelajaran.

Hasil tes menunjukkan bahwa kemampuan siswa dalam memahami konsep masih rendah, karena jumlah siswa yang mengalami miskonsepsi atau tidak memahami konsep lebih besar dibandingkan dengan siswa yang benar-benar memahami materi. Analisis miskonsepsi mengidentifikasi bahwa butir soal ke-12 dan ke-20 merupakan yang paling sering memunculkan miskonsepsi, terutama pada konsep susunan saraf pusat dan saraf tepi serta struktur dan fungsi sistem indera. Hal ini menunjukkan bahwa siswa keliru dalam memahami hubungan antar komponen sistem saraf dan fungsinya. Miskonsepsi semacam ini dapat berdampak serius dalam pembelajaran biologi, karena sistem koordinasi merupakan materi dasar yang berkaitan erat dengan konsep lain seperti sistem indera, gerak, dan regulasi tubuh. Jika miskonsepsi tidak segera diperbaiki, pemahaman siswa terhadap materi lanjutan akan semakin terhambat dan berdampak pada rendahnya integrasi konsep dalam sistem biologi secara menyeluruh.

Hasil wawancara dengan guru mata pelajaran biologi menunjukkan bahwa salah satu penyebab utama munculnya miskonsepsi siswa adalah kurangnya perhatian saat proses pembelajaran berlangsung serta kecenderungan siswa melupakan materi yang telah dipelajari. Temuan ini sejalan dengan hasil tes diagnostik tiga tingkat yang menunjukkan bahwa sebanyak 40,47% siswa mengalami miskonsepsi pada materi sistem koordinasi. Secara teoretis, Rohmah et al. (2023) menyatakan bahwa proses pembelajaran yang mengabaikan pengetahuan awal atau prakonsepsi siswa dapat memicu terbentuknya miskonsepsi. Prakonsepsi ini sering bersumber dari pemahaman yang dibentuk sendiri oleh siswa berdasarkan pengamatan atau sumber tidak valid yang mereka anggap benar. Selain itu, menurut Ritonga et al. (2021), miskonsepsi juga dapat muncul akibat kurangnya penguasaan konsep oleh guru saat mengajar, pengalaman belajar sebelumnya yang keliru, serta penyajian informasi dalam buku teks yang tidak tepat. Oleh karena itu, penting untuk mengidentifikasi dan mereduksi miskonsepsi siswa sebelum mereka mempelajari konsep baru agar tidak mengganggu pemahaman terhadap materi selanjutnya.

KESIMPULAN

Berdasarkan hasil penelitian, dapat disimpulkan bahwa pengembangan tes diagnostik tiga tingkat pada materi sistem koordinasi berhasil menghasilkan instrumen berjumlah 21 butir soal yang telah melalui tahapan validasi oleh ahli, uji coba perorangan, uji kelompok kecil, dan uji lapangan. Instrumen ini memenuhi kriteria soal yang baik, ditinjau dari aspek validitas, reliabilitas, tingkat kesukaran, dan daya pembeda. Hasil analisis data menunjukkan distribusi pemahaman peserta didik di SMAN 2 Sinjai terdiri atas 25,80% paham konsep, 33,61% tidak paham konsep, dan 40,47% mengalami miskonsepsi. Proporsi miskonsepsi yang cukup tinggi menunjukkan adanya kebutuhan untuk memperbaiki strategi pembelajaran, terutama

dalam menyampaikan konsep-konsep sistem koordinasi yang bersifat kompleks. Dengan demikian, instrumen tes diagnostik tiga tingkat ini efektif dalam mengidentifikasi tingkat pemahaman dan miskonsepsi peserta didik secara lebih mendalam.

REKOMENDASI

Instrumen diagnostik tiga tingkat ini dapat dimanfaatkan untuk mengidentifikasi miskonsepsi siswa pada materi sistem koordinasi dalam pembelajaran biologi di tingkat sekolah menengah atas. Hasil dari instrumen ini dapat digunakan oleh guru untuk mengevaluasi pemahaman siswa, sehingga kesulitan dalam memahami materi selanjutnya dapat diminimalisasi. Selain itu, guru diharapkan mampu mengembangkan instrumen sejenis pada materi dasar lainnya guna meningkatkan pemahaman konseptual siswa secara menyeluruh. Meskipun demikian, agar instrumen ini menghasilkan data yang lebih akurat dan representatif, beberapa aspek teknis dan substansi perlu disempurnakan lebih lanjut.

UCAPAN TERIMA KASIH

Ucapan terima kasih disampaikan kepada Tuhan Yang Maha Esa atas rahmat dan karunia-Nya sehingga penelitian dan penyusunan artikel ini dapat diselesaikan dengan baik. Penghargaan yang tulus juga ditujukan kepada orang tua atas doa dan dukungannya, serta kepada pihak sekolah, guru, dan peserta didik yang telah memberikan bantuan dan kerja sama selama proses pelaksanaan penelitian ini.

DAFTAR PUSTAKA

- Arifin, M., Kadarohman, A., & Rusdiana, D. (2015). Pengembangan Tes Diagnostik Three-Tier Pada Materi Larutan Elektrolit Dan Nonelektrolit. *Jurnal Pendidikan IPA Indonesia*, 4(2), 179–185.
- Fauziah, F., & Sahidu, H. (2018). Identifikasi Miskonsepsi Siswa Pada Materi Listrik Dinamis Menggunakan Certainty of Response Index (CRI). *Jurnal Pendidikan Fisika Dan Teknologi*, 4(1), 1–6.
- Gurel, D. E. R. Y. A., Eryilmaz, A., & Mcdermott, L. (2015). A Review And Comparison Of Diagnostic Instruments To Identify Students' Misconceptions In Science. *Eurasia Journal Of Mathematics Science And Technology Education*, 11(5).
- Hasyim, W., Suwono, H., & Susilo, H. (2018). Three-Tier Test To Identify Students' Misconception Of Human Reproduction System. *Jurnal Pendidikan Sains*, 6(2), 48–54.
- Hidayati, R., & Halim, A. (2016). Penerapan Tes Diagnostik Untuk Mengidentifikasi Miskonsepsi Siswa Kelas XI Pada Materi Fluida Statis. *Jurnal Pendidikan Sains Indonesia*, 4(1), 67–74.
- Ibrahim, M. (2018). *Perubahan Konsepsi IPA*. Sidoarjo: Zifatama Jawa.
- Mubarak, S., Susilaningih, E., & Cahyono, E. (2016). Pengembangan Tes Diagnostik Three Tier Multiple Choice Untuk Mengidentifikasi Miskonsepsi Peserta Didik Kelas Xi. *Journal Of Innovative Science Education*, 5(2), 101–110.
- Novita, D., & Kartowagiran, B. (2019). Pengembangan Instrumen Tes Diagnostik Berbasis Three-Tier Untuk Mengidentifikasi Miskonsepsi Pada Materi Momentum Dan Impuls. *Jurnal Inovasi Pendidikan IPA*, 5(1), 1–9.
- Nurhalisah, N. (2019). *Pengembangan Instrumen Tes Diagnostik Pilihan Ganda Beralasan Pada Materi Biologi Kelas Vii Mts Madani Alauddin Kabupaten Gowa*

- (Skripsi, Universitas Islam Negeri Alauddin Makassar). Uin Alauddin Makassar Repository.
- Ojose, B. (2015). *Common Misconceptions In Mathematics: Strategies To Correct Them*. University Press Of America.
- Ritonga, N., Sari, N. F., & Gultom, H. S. B. (2021). Miskonsepsi Guru Biologi Pada Materi Sistem Respirasi Di Sma Negeri Se-Kabupaten Labuhanbatu. *Jurnal Pendidikan Sains Dan Aplikasinya*, 4(1), 1-5.
- Riyani, R., Maizora, S., & Hanifah, H. (2017). Uji Validitas Pengembangan Tes Untuk Mengukur Kemampuan Pemahaman Relasional Pada Materi Persamaan Kuadrat Siswa Kelas Viii Smp. *Jurnal Penelitian Pembelajaran Matematika Sekolah (Jp2ms)*, 1(1), 60–65.
- Rohmah, M., Priyono, S., & Sari, R. S. (2023). Analisis Faktor-Faktor Penyebab Miskonsepsi Peserta Didik Sma. *Utility: Jurnal Ilmiah Pendidikan Dan Ekonomi*, 7(01), 39-47.
- Rusilowati, A., & Wahyudi, W. (2017). Identifikasi Miskonsepsi Siswa Menggunakan Diagnostik Tes Two-Tier Pada Materi Gaya Dan Gerak. *Unnes Physics Education Journal*, 6(1), 33–40.
- Saputri, L. A., Muldayanti, N. D., & Setiadi, A. E. (2016). Analisis Miskonsepsi Siswa Dengan Certainty of Response Index (Cri) Pada Submateri Sistem Saraf Di Kelas Xi Ipa Sma Negeri 1 Selimbau. *Jurnal Bioeducation*, 3(2), 53–62.
- Sarlina. (2015). Miskonsepsi Siswa Terhadap Pemahaman Konsep Matematika Pada Pokok Bahasan Persamaan Kuadrat Siswa Kelas X5 SMA Negeri 11 Makassar. *Jurnal Matematika dan Pembelajaran*, 3 (3). (DOI: <https://doi.org/10.24252/mapan.2015v3n2a5>).
- Simatupang, L. F. (2021). Identifikasi Miskonsepsi Siswa Menggunakan Tes Diagnostik Three Tier Pada Materi Asam Basa Di SMA Negeri I Mesjid Raya (*Doctoral Dissertation, Uin Ar-Raniry*).
- Suryawati, E., & Osman, K. (2017). Contextual Learning: Innovative Approach Toward the Development of Students' Scientific Attitude and Natural Science Performance. *Eurasia Journal of Mathematics, Science and Technology Education*, 13(8), 3659–3676.
- Trianto. (2017). *Pengembangan Model Pembelajaran Terpadu*. Jakarta: Prestasi Pustaka.
- Wulandari, T., & Suryani, N. (2021). Analisis Miskonsepsi Siswa Menggunakan Model Certainty of Response Index (CRI) Pada Materi Sistem Ekskresi. *Jurnal Pendidikan Biologi*, 10(1), 11–18.
- Yulianti, E., & Widodo, A. (2019). Analisis Miskonsepsi Siswa SMA Pada Materi Bioteknologi Dengan Menggunakan Tes Diagnostik Two-Tier. *Jurnal Pendidikan Biologi Indonesia*, 5(1), 75–84.