

**PENGEMBANGAN INSTRUMEN DIAGNOSTIK BENTUK THREE TIER
MULTIPLE CHOICE PADA MATA PELAJARAN MATEMATIKA TINGKAT SD**

Desi Aprianti¹, Tri Suminar²

Penelitian dan Evaluasi Pendidikan Pascasarjana Universitas Negeri Semarang

Alamat e-mail : desiaprianti@students.unnes.ac.id

ABSTRACT

This study aims to develop a Three-Tier Multiple Choice diagnostic instrument on whole number material for fourth-grade elementary school students as an effort to identify and map student misconceptions more accurately. The study uses the Research and Development (R&D) method with the Borg & Gall development model modified into six stages, namely preliminary study, product planning and design, expert validation, small-scale trials, field trials, and product refinement. The research subjects involved fourth-grade elementary school students in small-scale and large-scale trials and teachers as respondents for the practicality of the instrument. The instrument developed consisted of 20 Three-Tier Multiple Choice questions covering three levels, namely answers, reasons, and student confidence levels. Validation results by material, language, and evaluation experts showed that the instrument was in the very valid category with an average validity value of ≥ 0.93 . Reliability tests in large-scale trials produced a reliability coefficient of > 0.895 , which is included in the high category. The results of the practicality test showed that the instrument was easy to use and understand by both teachers and students. The results of the diagnostic test analysis showed that the instrument was able to identify categories of understanding, not understanding, and misconceptions of students, with the discovery of low, medium, and high misconceptions in several indicators of whole numbers. Thus, the developed Three-Tier Multiple Choice diagnostic instrument was declared valid, reliable, and practical, and effective for use as a teacher tool in identifying misconceptions and designing follow-up mathematics learning in elementary schools.

Keywords: Diagnostic Assessment, Three-Tier Multiple Choice, Misconceptions, Whole Numbers

ABSTRAK

Penelitian ini bertujuan untuk mengembangkan instrumen diagnostik bentuk *Three-Tier Multiple Choice* pada materi bilangan cacah untuk peserta didik kelas IV sekolah dasar sebagai upaya mengidentifikasi dan memetakan miskONSEP siswa secara lebih akurat. Penelitian menggunakan metode *Research and Development* (R&D) dengan model pengembangan Borg & Gall yang dimodifikasi menjadi enam tahap, yaitu studi pendahuluan, perencanaan dan desain produk, validasi ahli, uji coba skala kecil, uji coba lapangan, dan penyempurnaan produk. Subjek penelitian

melibatkan siswa kelas IV sekolah dasar pada uji coba skala kecil dan skala besar serta guru sebagai responden kepraktisan instrumen. Instrumen yang dikembangkan terdiri atas 20 butir soal *Three-Tier Multiple Choice* yang mencakup tiga tingkat, yaitu jawaban, alasan, dan tingkat keyakinan siswa. Hasil validasi oleh ahli materi, bahasa, dan evaluasi menunjukkan bahwa instrumen berada pada kategori sangat valid dengan rata-rata nilai validitas $\geq 0,93$. Uji reliabilitas pada uji coba skala besar menghasilkan koefisien reliabilitas $> 0,895$ yang termasuk kategori tinggi. Hasil uji kepraktisan menunjukkan bahwa instrumen mudah digunakan dan dipahami oleh guru maupun siswa. Hasil analisis tes diagnostik menunjukkan bahwa instrumen mampu mengidentifikasi kategori paham, tidak paham, dan miskONSEPSI siswa, dengan ditemukannya miskONSEPSI rendah, sedang, dan tinggi pada beberapa indikator bilangan cacah. Dengan demikian, instrumen diagnostik *Three-Tier Multiple Choice* yang dikembangkan dinyatakan valid, reliabel, dan praktis, serta efektif digunakan sebagai alat bantu guru dalam mengidentifikasi miskONSEPSI dan merancang tindak lanjut pembelajaran matematika di sekolah dasar.

Kata kunci: Asesmen Diagnostik, *Three-Tier Multiple Choice*, MiskONSEPSI, Bilangan Cacah

A. Pendahuluan

Pembelajaran matematika di sekolah dasar memiliki peran penting dalam membangun dasar berpikir logis, sistematis, dan konseptual peserta didik (Prisnanda et al., 2024).. Namun, pada praktiknya pembelajaran matematika masih sering menekankan pada pencapaian hasil akhir berupa jawaban benar atau salah, tanpa menggali secara mendalam proses berpikir siswa. Kondisi tersebut menyebabkan guru kesulitan mengidentifikasi pemahaman konseptual siswa secara utuh, terutama dalam mengungkap miskONSEPSI yang sering muncul pada

konsep-konsep dasar matematika seperti bilangan cacah(Luqman Hakim Abbas, 2016). MiskONSEPSI yang tidak terdeteksi sejak dini berpotensi terbawa ke jenjang pendidikan berikutnya dan berdampak pada rendahnya pemahaman matematika secara berkelanjutan. (Rahman et al., 2017).

Salah satu penyebab utama sulitnya mengidentifikasi miskONSEPSI siswa adalah penggunaan instrumen penilaian yang masih bersifat konvensional, seperti tes pilihan ganda biasa. Instrumen tersebut hanya menilai jawaban akhir siswa tanpa mempertimbangkan alasan dan

tingkat keyakinan terhadap jawaban yang dipilih (Sriyanti et al., 2019).. Akibatnya, siswa yang menjawab benar belum tentu benar-benar memahami konsep, sementara siswa yang menjawab salah belum tentu tidak memahami, tetapi mungkin mengalami kesalahan prosedural. Oleh karena itu, diperlukan instrumen asesmen diagnostik yang mampu mengungkap lebih dalam struktur kognitif siswa, termasuk kesalahan konsep dan tingkat keyakinannya (Juliani et al., 2021).

Instrumen diagnostik bentuk *Three-Tier Multiple Choice* merupakan salah satu alternatif asesmen yang dapat digunakan untuk mengidentifikasi pemahaman konsep dan miskonsepsi siswa secara lebih komprehensif(Wahyudi et al., 2021). Instrumen ini tidak hanya memuat pilihan jawaban, tetapi juga alasan atas jawaban tersebut serta tingkat keyakinan siswa. Dengan demikian, guru dapat membedakan antara siswa yang benar-benar paham, siswa yang tidak paham, dan siswa yang mengalami miskonsepsi (Ritonga & Yastophi, 2019). Penerapan bentuk *Three-Tier Multiple Choice* dalam pembelajaran matematika, khususnya

pada materi bilangan cacah kelas IV sekolah dasar, dinilai relevan karena materi ini menjadi fondasi bagi pemahaman konsep matematika selanjutnya.

Berdasarkan hasil analisis kebutuhan yang dilakukan di sekolah dasar, diketahui bahwa guru masih sangat membutuhkan instrumen diagnostik yang praktis, valid, dan reliabel untuk membantu mengidentifikasi miskonsepsi siswa. Guru juga menyatakan kesulitan dalam menentukan tindak lanjut pembelajaran karena keterbatasan informasi mengenai letak kesalahan konsep siswa. Oleh sebab itu, pengembangan instrumen diagnostik bentuk *Three-Tier Multiple Choice* pada mata pelajaran matematika tingkat sekolah dasar menjadi penting dan mendesak dilakukan sebagai upaya membantu guru dalam merancang pembelajaran yang lebih tepat sasaran.

Berdasarkan uraian tersebut, peneliti tertarik untuk melakukan penelitian dengan judul “Pengembangan Instrumen Diagnostik Bentuk *Three-Tier Multiple Choice* pada Mata Pelajaran Matematika Tingkat Sekolah Dasar”.

Pengembangan instrumen ini diharapkan mampu menghasilkan alat asesmen yang valid, reliabel, dan praktis, serta memberikan gambaran yang akurat mengenai profil pemahaman dan miskonsepsi siswa sehingga dapat digunakan sebagai dasar perbaikan proses pembelajaran matematika di sekolah dasar.

B. Metode Penelitian

Jenis Penelitian yang digunakan adalah penelitian R & D (*Research and Development*) menggunakan model pengembangan Borg & Gall. Menurut Borg & Gall (1983) model pengembangan ini terdiri dari 10 langkah, namun agar lebih praktis dan sesuai dengan kebutuhan peneliti, tahap model ini kemudian disederhanakan menjadi 6 langkah, yaitu: (1) studi pendahuluan, (2) perencanaan dan desain produk, (3) Validasi produk oleh ahli, (4) uji coba skala kecil, (5) uji coba lapangan, dan (6) produk akhir. Penyederhanaan ini dilakukan untuk mempermudah proses pengembangan tanpa menghilangkan esensi dari model aslinya.

Subjek penelitian terdiri atas peserta didik kelas IV sekolah dasar. Subjek pada uji coba skala kecil

berjumlah 25 siswa, sedangkan subjek pada uji coba skala besar melibatkan 109 siswa kelas IV yang berasal dari dua sekolah dasar negeri di Kecamatan Gajah Mungkur, Kota Semarang. Selain siswa, subjek penelitian juga melibatkan guru kelas IV sebagai responden pada tahap analisis kebutuhan dan uji kepraktisan instrumen.

Studi pendahuluan dilakukan melalui wawancara guru dan penyebaran angket untuk mengidentifikasi kebutuhan terhadap instrumen diagnostik yang mampu mengungkap miskonsepsi siswa secara mendalam. Tahap perencanaan dan desain produk mencakup penyusunan kisi-kisi soal, perumusan indikator, serta pengembangan butir soal *Three-Tier Multiple Choice* yang terdiri atas jawaban, alasan, dan tingkat keyakinan.

Validasi produk dilakukan oleh ahli materi, ahli bahasa, dan ahli evaluasi untuk menilai kelayakan instrumen dari aspek isi, kebahasaan, dan konstruksi. Hasil validasi digunakan sebagai dasar revisi instrumen sebelum dilakukan uji coba. Uji coba skala kecil bertujuan untuk

mengetahui keterbacaan dan kejelasan instrumen, sedangkan uji coba lapangan dilakukan untuk menguji validitas empiris, reliabilitas, dan kemampuan instrumen dalam mengidentifikasi miskonsepsi siswa. Tahap terakhir adalah penyempurnaan produk berdasarkan hasil analisis uji coba sehingga dihasilkan instrumen diagnostik yang valid, reliabel, dan praktis untuk digunakan dalam pembelajaran matematika di sekolah dasar.

C. Hasil Penelitian dan Pembahasan

Penelitian ini dilakukan di dua sekolah Dasar Negeri yang menjadi objek penelitian, yaitu SD Negeri 01 Sampangan dan SD Negeri 02 Sampangan yang keduanya berada di Kecamatan Gajah Mungkur, Kota Semarang. Pemilihan kedua sekolah didasarkan pada kesesuaian karakteristiknya dengan tujuan penelitian, yaitu sebagai pelaksana kurikulum merdeka di jenjang SD, serta telah menerapkan pembelajaran matematika pada kelas IV dengan capaian pembelajaran memahami konsep dasar, mengoprasikan pecahan, dan membandingakan serta mengurutkan pecahan.

1. Analisis kebutuhan
Analisis kebutuhan merupakan langkah awal dalam penelitian dan pengembangan instrumen karena menentukan arah pengembangan produk agar sesuai dengan konteks dan kebutuhan pengguna. Pada penelitian ini, analisis kebutuhan dilakukan melalui wawancara dengan 3 guru kelas IV SD dari SD Negeri 1 Sampangan dan SD Negeri 2 Sampangan, dengan tujuan memperoleh gambaran nyata mengenai pelaksanaan asesmen diagnostik, hambatan guru dalam mengidentifikasi miskonsepsi, serta kebutuhan akan instrumen diagnostik inovatif seperti *three tier multiple choice*.

Hasil wawancara menunjukkan bahwa asesmen diagnostik berbentuk *three tier multiple choice* belum dilaksanakan di kedua sekolah tersebut. Kendala utama yang dihadapi adalah belum tersedianya instrument tes diagnostik yang valid serta reliabel untuk mengetahui miskonsepsi siswa, serta kesulitan guru dalam membuat instrumen tes diagnostik berbentuk *three tier multiple choice*.

Temuan di lapangan menunjukkan bahwa semua guru telah

melaksanakan asesmen diagnostik pada awal pembelajaran, namun pelaksanaannya masih bersifat umum, dan belum dirancang secara khusus untuk menelusuri miskonsepsi siswa. Guru cenderung menggunakan Pertanyaan pilihan ganda biasa, Soal uraian singkat, atau dengan pertanyaan terbuka. Asesmen tersebut hanya memberikan informasi apakah siswa benar atau salah, tetapi tidak memberikan petunjuk mengapa siswa menjawab demikian, serta seberapa yakin siswa terhadap jawaban yang dipilih.

Berdasarkan wawancara, seluruh narasumber menyatakan belum pernah mengembangkan instrumen tes diagnostik untuk mengidentifikasi miskonsepsi siswa dan belum mengenal bentuk soal *Three-Tier Multiple Choice*, namun menunjukkan ketertarikan yang sangat tinggi untuk mempelajari dan menggunakan instrumen tersebut. Para guru berpendapat bahwa instrumen soal *three tier multiple choice* sangat diperlukan karena dapat membantu mendeteksi miskonsepsi siswa secara lebih akurat, mempermudah guru mengelompokkan siswa berdasarkan tingkat pemahaman, serta memberikan dasar yang kuat untuk

merancang tindak lanjut pembelajaran yang tepat.

Dengan demikian, hasil analisis kebutuhan menunjukkan bahwa guru membutuhkan instrumen asesmen diagnostik yang valid, reliabel, dan terstruktur, khususnya instrumen *Three-Tier Multiple Choice* pada materi matematika bilangan pecahan untuk mengidentifikasi miskonsepsi siswa dengan lebih efektif. Temuan ini memperkuat urgensi pengembangan instrumen diagnostik dalam penelitian yang sedang dilakukan.

2. Pengembangan Instrumen *Three Tier Multiple Choice*

Instrumen diagnostik yang dikembangkan dalam penelitian ini berbentuk *Three-Tier Multiple Choice* yang terdiri atas tiga tingkatan, yaitu jawaban, alasan, dan tingkat keyakinan. Penyusunan instrumen diawali dengan perumusan indikator soal yang disesuaikan dengan tujuan pembelajaran dan karakteristik materi bilangan cacah kelas IV sekolah dasar. Setiap butir soal dirancang untuk menggali kemungkinan miskonsepsi yang sering dialami siswa, seperti kesalahan dalam operasi hitung, pemahaman nilai tempat, dan konsep perbandingan bilangan.

Penggunaan tiga tingkatan dalam satu soal memberikan keunggulan dibandingkan instrumen konvensional. Jawaban siswa pada tier pertama menunjukkan hasil akhir, tier kedua menggambarkan alasan atau konsep yang digunakan, dan tier ketiga menunjukkan tingkat keyakinan siswa terhadap jawabannya. Kombinasi ketiga tier ini memungkinkan peneliti mengklasifikasikan pemahaman siswa ke dalam kategori paham, tidak paham, dan miskonsepsi secara lebih akurat(Black & London, 2009)

Hal ini sejalan dengan pendapat (Deviana et al., 2024) yang menyatakan bahwa instrumen *three-tier* dapat mengurangi kemungkinan siswa menebak jawaban dan meningkatkan validitas hasil diagnosis miskonsepsi.

3. Validitas

Uji validitas oleh ahli instrument dilakukan terhadap sepuluh indikator yang tergolong ke dalam 3 aspek, diantaranya: aspek materi, aspek konstruksi serta aspek Bahasa. Pada aspek materi indikator yang dinilai sebanyak 4 indikator diantaranya: 1) kesesuaian soal dan capaian pembelajaran dan tujuan pembelajaran, 2) ketepatan soal

dengan tujuan pembelajaran, 3) kebenaran konsep matematika dalam soal dan 4) soal sesuai dengan kisi-kisi. Penilaian pada aspek ini penting untuk memastikan bahwa instrumen benar-benar mencerminkan kompetensi matematika yang ingin diukur serta tidak menyimpang dari standar pembelajaran.

Pengujian Validitas dilakukan oleh lima ahli menggunakan skala 4 tingkat, yaitu dengan skor tertinggi 4 dan skor terendah 1. Indikator penilaian pengembangan instrument oleh ahli terdiri dari 3 aspek dengan 10 indikator, dan dinilai oleh lima ahli/pakar yang terlibat dalam penelitian terdiri dari dua guru kelas IV yang mengajar lebih dari 10 tahun, dan tiga dosen dengan latar belakang pendidikan S2 dan S3 sesuai keahlian masing-masing.

Skor hasil validasi untuk instrument tes *diagnostik three tier multiple choice* yang telah divalidasi pada tabel 1 berikut:

Tabel 1. Rekapitulasi Validasi Instrumen

Aspek	ahli					S1	S2	S3	S4	S5	Σ_E	n (C-1)	%	ket
	1	2	3	4	5									
Materi 1	4	4	4	4	3	3	3	3	3	3	15	15	1	Tinggi
Materi 2	4	3	4	4	3	2	3	3	3	3	14	15	0,93	Tinggi
Materi 3	4	4	3	4	3	3	2	3	3	3	14	15	0,93	Tinggi
Materi 4	4	4	3	4	3	3	2	3	3	3	14	15	0,93	Tinggi
Konstruksi 1	3	4	4	4	2	3	3	3	3	3	14	15	0,93	Tinggi
Konstruksi 2	4	4	4	4	3	3	3	3	3	3	15	15	1	Tinggi
Konstruksi 3	4	3	4	4	3	2	3	3	3	3	14	15	0,93	Tinggi
Bahasa 1	4	4	4	3	2	3	3	2	1	17	15	0,80	Tinggi	
Bahasa 2	3	4	4	3	2	3	3	2	2	12	15	0,80	Tinggi	
Bahasa 3	4	4	4	4	3	3	3	3	3	15	15	1	Tinggi	

Tabel 1 menunjukkan bahwa instrumen diagnostik *Three Tier Multiple Choice* yang dikembangkan berada pada kategori valid dengan nilai Aiken's V sebesar 0,93. Pada tahap validasi, tiga ahli memberikan masukan terkait kesesuaian indikator, kualitas butir soal, ketepatan pengecoh (distractor) pada tier 2, serta kejelasan petunjuk pengerjaan. Revisi dilakukan dengan mengganti alasan pilihan pada tier 2 agar tidak mudah ditebak, memperbaiki redaksi soal, serta menyempurnakan petunjuk pengerjaan agar lebih mudah dipahami oleh siswa. Hasil validasi menunjukkan bahwa instrumen dinilai layak dan dapat digunakan setelah perbaikan. Uji coba skala besar kemudian dilakukan pada 20 butir soal dengan jumlah 84 responden. Setiap respons dianalisis berdasarkan gabungan tiga komponen, yaitu jawaban pilihan ganda (tier 1), alasan (tier 2), dan tingkat keyakinan (tier 3). Analisis dengan sistem skoring 1–0 menghasilkan tiga kategori utama: paham, tidak paham, dan miskONSEPsi.

4. Reliabilitas

Uji reliabilitas dilakukan untuk mengetahui sejauh mana instrumen

tes diagnostik *three-tier multiple choice* memberikan hasil yang konsisten ketika digunakan untuk mengukur objek yang sama pada kondisi yang serupa. Reliabilitas dihitung menggunakan teknik Cronbach's Alpha karena bentuk instrumen berupa tes objektif pilihan ganda dengan skor dikotomis (1 untuk jawaban benar dan 0 untuk jawaban salah). Menurut Azwar (2019), teknik Cronbach's Alpha digunakan untuk mengukur konsistensi internal antarbutir soal dalam satu instrumen pengukuran.

Pengujian reliabilitas pada penelitian ini dilakukan dengan bantuan software SPSS. Berdasarkan output perhitungan diperoleh nilai reliabilitas sebagai berikut:

Reliability Statistics	
Cronbach's Alpha	
,895	20

Gambar 1. Reliabilitas Instrumen

Hasil uji reliabilitas menunjukkan nilai Cronbach's Alpha sebesar 0,895 yang tergolong reliabel, instrumen memiliki tingkat konsistensi yang baik. Hal ini menandakan bahwa instrumen mampu memberikan hasil yang stabil dan dapat dipercaya ketika digunakan pada kelompok siswa yang memiliki

karakteristik serupa (Hilman et al., 2023). Instrumen yang valid dan reliabel merupakan syarat utama agar hasil diagnosis miskonsepsi dapat dijadikan dasar dalam perencanaan pembelajaran remedial (Laksono, 2020).

Temuan ini sesuai dengan pendapat (Wulandari et al., 2023) yang menyatakan bahwa instrumen penelitian yang baik harus memenuhi kriteria valid dan reliabel agar data yang diperoleh akurat dan dapat dipertanggungjawabkan.

5. Kepraktisan Instrumen

Kepraktisan instrumen dinilai berdasarkan respon guru dan siswa setelah penggunaan instrumen pada uji coba skala besar. Hasil angket menunjukkan bahwa instrumen mudah digunakan, petunjuk penggerjaan jelas, dan tampilan soal menarik bagi siswa. Guru menilai bahwa instrumen ini membantu dalam mengidentifikasi miskonsepsi siswa secara lebih cepat dan sistematis dibandingkan asesmen biasa.

Siswa juga menyatakan bahwa meskipun soal terdiri dari tiga bagian, instrumen tetap mudah dipahami dan tidak membingungkan. Hal ini menunjukkan bahwa instrumen

diagnostik *Three Tier Multiple Choice* memiliki tingkat kepraktisan yang baik dan layak digunakan dalam kegiatan asesmen pembelajaran matematika di sekolah dasar. Kepraktisan instrumen merupakan aspek penting dalam penelitian pengembangan, karena instrumen yang baik tidak hanya valid dan reliabel, tetapi juga mudah diterapkan di lapangan (Rosyidah et al., 2020).

6. Interpretasi Hasil

Hasil analisis tes diagnostik menggunakan instrumen *Three-Tier Multiple Choice* pada materi bilangan cacah menunjukkan bahwa tingkat pemahaman konseptual peserta didik kelas IV sekolah dasar masih beragam. Instrumen ini mampu mengelompokkan pemahaman siswa ke dalam tiga kategori utama, yaitu paham, tidak paham, dan miskonsepsi, sehingga memberikan gambaran yang lebih komprehensif dibandingkan tes pilihan ganda konvensional.

Berdasarkan data yang diperoleh, persentase siswa yang berada pada kategori paham menunjukkan variasi pada setiap butir soal. Persentase yang paham dapat dilihat pada gambar 1:



Gambar 1. Persentase siswa yang paham pada setiap butir soal

Dilhat pada gambar 1 terdapat beberapa soal dengan persentase paham yang tinggi, seperti soal nomor 2, 5, dan 6, yang mengindikasikan bahwa konsep bilangan cacah pada indikator tersebut telah dipahami dengan baik oleh sebagian besar siswa. Hal ini diduga karena konsep yang diujikan bersifat dasar dan sering muncul dalam pembelajaran maupun latihan soal. Namun demikian, masih ditemukan sejumlah soal dengan kategori paham rendah, khususnya pada soal nomor 3, 7, 9, 10, 13, 15, 16, 17, dan 18. Kondisi ini menunjukkan bahwa siswa belum sepenuhnya memahami konsep yang lebih kompleks, terutama yang menuntut penalaran dan penerapan konsep dalam situasi yang berbeda.

Pada kategori miskonsepsi, hasil analisis menunjukkan bahwa miskonsepsi masih mendominasi pada beberapa butir soal. Persentase

siswa yang mengalami miskonsepsi dapat dilihat pada gambar 2:

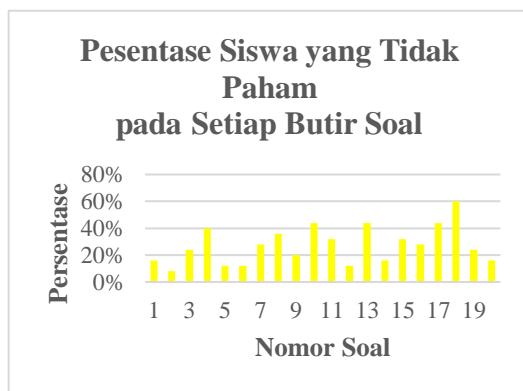


Gambar 2. Persentase siswa yang mengalami miskonsepsi pada setiap butir soal

Dari gambar 2 terdapat 10 soal yang termasuk dalam kategori miskonsepsi rendah, 9 soal dalam kategori miskonsepsi sedang, dan 1 soal dalam kategori miskonsepsi tinggi, yaitu soal nomor 16. Miskonsepsi ini ditandai oleh jawaban siswa yang salah namun disertai alasan yang diyakini benar dengan tingkat keyakinan tinggi. Temuan ini mengindikasikan bahwa siswa memiliki pemahaman konseptual yang keliru tetapi sudah tertanam kuat dalam struktur kognitifnya.

Miskonsepsi paling banyak ditemukan pada soal-soal yang berkaitan dengan nilai tempat, operasi hitung berjenjang, serta pemecahan masalah kontekstual, yang memang dikenal sebagai konsep abstrak dan berpotensi menimbulkan kesalahan pemahaman.

Sementara itu, kategori tidak paham juga masih cukup signifikan.



Gambar 3. Persentase siswa yang tidak paham

Sebanyak 12 butir soal termasuk dalam kriteria tidak paham tingkat rendah dan 8 soal berada pada tingkat sedang. Kondisi ini menunjukkan bahwa pada beberapa konsep tertentu, siswa belum memiliki pemahaman awal yang memadai. Ketidaktahuan ini ditandai dengan jawaban dan alasan yang salah serta tingkat keyakinan yang rendah, sehingga siswa cenderung menebak jawaban. Temuan ini mengindikasikan perlunya penguatan konsep dasar sebelum siswa diarahkan pada pembelajaran yang lebih kompleks.

Secara keseluruhan, hasil tes diagnostik menunjukkan bahwa instrumen *Three-Tier Multiple Choice* yang dikembangkan mampu mengungkap tidak hanya tingkat kebenaran jawaban siswa, tetapi juga alasan dan keyakinan mereka

terhadap jawaban tersebut. Dengan demikian, guru dapat membedakan antara siswa yang benar-benar paham, siswa yang tidak paham, dan siswa yang mengalami miskonsepsi. Informasi ini sangat penting sebagai dasar dalam merancang tindak lanjut pembelajaran yang tepat sasaran, seperti pemberian remedial, pengayaan, maupun perbaikan strategi pembelajaran agar miskonsepsi tidak terus berlanjut pada materi berikutnya.

E. Kesimpulan

Berdasarkan temuan penelitian, dapat disimpulkan bahwa instrumen diagnostik berbentuk *Three-Tier Multiple Choice* yang dikembangkan pada materi bilangan cacah untuk peserta didik kelas IV sekolah dasar dinyatakan valid, reliabel, dan praktis untuk digunakan dalam asesmen diagnostik matematika. Instrumen ini terbukti mampu mengidentifikasi secara lebih mendalam tingkat pemahaman siswa, tidak hanya berdasarkan jawaban benar atau salah, tetapi juga melalui alasan yang diberikan serta tingkat keyakinan siswa terhadap jawabannya. Hasil uji coba menunjukkan bahwa instrumen *Three-Tier Multiple Choice* dapat

mengungkap tiga kategori pemahaman siswa, yaitu paham, tidak paham, dan miskonsepsi, termasuk variasi tingkat miskonsepsi rendah, sedang, hingga tinggi pada beberapa indikator bilangan cacah. Dengan demikian, instrumen ini memberikan gambaran yang komprehensif mengenai profil konseptual siswa dan dapat membantu guru dalam mendiagnosis kesalahan konsep yang tidak teridentifikasi melalui tes pilihan ganda konvensional.

Berdasarkan hasil penelitian tersebut, disarankan agar guru sekolah dasar mulai memanfaatkan instrumen diagnostik berbasis *Three-Tier Multiple Choice* secara berkala, khususnya pada materi matematika yang berpotensi menimbulkan miskonsepsi seperti bilangan cacah. Penggunaan instrumen ini dapat menjadi dasar bagi guru dalam merancang tindak lanjut pembelajaran yang lebih tepat sasaran, baik melalui remedial, pengayaan, maupun perbaikan strategi pembelajaran. Selain itu, guru disarankan untuk mengintegrasikan hasil assesmen diagnostik dengan pendekatan pembelajaran yang berpusat pada siswa agar sesuai dengan proses berpikir dan tingkat pemahaman

mereka. Bagi peneliti selanjutnya, disarankan untuk mengembangkan instrumen *Three-Tier Multiple Choice* pada materi matematika lain atau jenjang kelas yang berbeda serta melibatkan subjek penelitian yang lebih luas guna memperoleh hasil yang lebih representatif dan memperkuat generalisasi temuan penelitian.

DAFTAR PUSTAKA

- Andrian, D., & Effendi, L. A. (2023). Instrument and evaluation of students' involvement in learning Mathematics in 21-century skills at senior high schools of Riau province. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 27(1), 14–25.
<https://doi.org/10.21831/pep.v27i1.58176>
- Black, P., & London, C. (2009). *Developing the theory of formative assessment*. 1–40.
- Deviana, T., Mn, F. R., Sulistyani, N., & Maulyda, M. A. (2024). Pengembangan Instrumen Assesmen Diagnostik Matematika Untuk Mendesain Pembelajaran Berdiferensiasi Di Sekolah Dasar. 13(4), 1269–1280.
- Hilman, I., Akmal, R., & Nugraha, F. (2023). Analisis Gaya Belajar Peserta Didik Melalui Assessment Diagnostik Non Kognitif Pada Pembelajaran Diferensiasi di Sekolah Dasar. *Naturalistic: Jurnal Kajian Dan*

- Penelitian Pendidikan Dan Pembelajaran, 8(1), 161–167. <https://doi.org/10.35568/naturalistic.v8i1.3911>
- Laksono, P. J. (2020). Pengembangan Three Tier Multiple Choice Test Pada Materi Kesetimbangan Kimia Mata Kuliah Kimia Dasar Lanjut. *Orbital: Jurnal Pendidikan Kimia*, 4(1), 44–63. <https://doi.org/10.19109/ojk.v4i1.5649>
- Mutmainna, D., Mania, S., & Sriyanti, A. (2018). Pengembangan Instrumen Tes Diagnostik Pilihan Ganda Dua Tingkat Untuk Mengidentifikasi Pemahaman Konsep Matematika. *MaPan*, 6(1), 56–69. <https://doi.org/10.24252/mapan.2018v6n1a6>
- Rosyidah, A. N. K., Maulyda, M. A., & Oktaviyanti, I. (2020). Miskonsepsi Matematika Mahasiswa PGSD Pada Penyelesaian Operasi Hitung Bilangan Bulat. *Jurnal Ilmiah KONTEKSTUAL*, 2(01), 15–21. <https://doi.org/10.46772/kontekstual.v2i01.244>
- Wahyudi, F., Didik, L. A., & Bahtiar, B. (2021). Pengembangan Instrumen Three Tier Test Diagnostik Untuk Menganalisis Tingkat Pemahaman Dan Miskonsepsi Siswa Materi Elastisitas. *Relativitas: Jurnal Riset Inovasi Pembelajaran Fisika*, 4(2), 48. <https://doi.org/10.29103/relativitas.v4i2.5184>
- Wulandari, G. A. P. T. W., Putrayasa, I. B., & Martha, I. N. (2023). Efektivitas Asesmen Diagnostik dalam Pembelajaran Berdiferensiasi pada Pelajaran Bahasa Indonesia. *Nusantara: Jurnal Pendidikan Indonesia*, 3(3), 433–448. <https://doi.org/10.14421/njpi.2023.v3i3-5>