

PENGEMBANGAN TES DIAGNOSTIK *THREE-TIER*
UNTUK MENGIDENTIFIKASI MISKONSEPSI PADA
MATERI GEOMETRI

SKRIPSI

Oleh:
NURUL FITHROTUZ ZAIDAH
NIM D04216027



UIN SUNAN AMPEL
S U R A B A Y A

UNIVERSITAS ISLAM NEGERI SUNAN AMPEL SURABAYA
FAKULTAS TARBIYAH DAN KEGURUAN
JURUSAN PENDIDIKAN MATEMATIKA DAN IPA
PRODI PENDIDIKAN MATEMATIKA
2020

PERNYATAAN KEASLIHAN TULISAN

Saya yang bertanda tangan dibawah ini:

Nama : NURUL FITHROTUZ ZAIDAH
NIM : D04216027
Jurusan/Program Studi : PMIPA/PENDIDIKAN MATEMATIKA
Fakultas : TARBIYAH DAN KEGURUAN

Menyatakan dengan sesungguhnya bahwa skripsi yang saya tulis ini benar-benar tulisan saya. Dan bukan merupakan plagiasi baik sebagian atau seluruhnya.

Apabila dikemudian hari terbukti atau dapat dibuktikan bahwa skripsi ini hasil plagiasi, baik sebagian atau seluruhnya, maka saya bersedia menerima sanksi atau perbuatan tersebut sesuai dengan ketentuan yang berlaku.

Surabaya, 10 Juni 2020
Yang membuat pernyataan,



Nurul Fithrotuz Zaidah
NIM. D04216027

PERSETUJUAN PEMBIMBING

Skripsi oleh:

Nama : Nurul Fithrotuz Zaidah

NIM : D04216027

Judul : PENGEMBANGAN TES DIAGNOSTIK *THREE-TIER* UNTUK
MENGIDENTIFIKASI MISKONSEPSI PADA MATERI GEOMETRI

Ini telah diperiksa dan disetujui untuk diujikan.

Surabaya, 27 April 2020

Pembimbing I,



Dr. Suparto, M.Pd.I
NIP. 196904021995031002

Pembimbing II,



Lisanul Uswah Sadieda, S.Si., M.Pd
NIP. 198309262006042002

PENGESAHAN TIM PENGUJI


Skripsi oleh Nurul Fithrotuz Zaidah telah dipertahankan di depan Tim Penguji
Skripsi

Surabaya, 10 Juni 2020

Mengesahkan, Fakultas Tarbiyah dan Keguruan
Universitas Islam Negeri Sunan Ampel Surabaya



Dekan

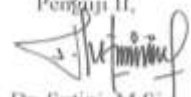

Dr. Ali Mas'ud, M.Ag., M.Pd.I
NIP. 196301231993031002

Tim Penguji


Penguji I,


Dr. Siti Lailiyah, M.Si
NIP. 198409282009122007

Penguji II,


Dr. Sutini, M.Si
NIP. 197701032009122001

Penguji III,


Dr. Suparto, M.Pd.I
NIP. 196904021995031002

Penguji IV,


Lisatul Ujwah Sadigda, S.Si., M.Pd
NIP. 198309262006042002



KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL SURABAYA
PERPUSTAKAAN

Jl. Jend. A. Yani 117 Surabaya 60237 Telp. 031-8431972 Fax.031-8413300
E-Mail: perpus@uin-sby.ac.id

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademika UIN Sunan Ampel Surabaya, yang bertanda tangan di bawah ini, saya:

Nama : NURUL FITHROTUZ ZAIDAH
NIM : 009216027
Fakultas/Jurusan : TARBIYAH DAN KEGURUAN / PMT
E-mail address : nurul.fithrotuz.06.91@gmail.com

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Perpustakaan UIN Sunan Ampel Surabaya, Hak Bebas Royalti Non-Ekklusif atas karya ilmiah :

Skripsi Tesis Desertasi Lain-lain (.....)

yang berjudul :

PENGEMBANGAN TES DIAGNOSTIK THREE-TIER UNTUK
MENGENAL MENGENAL MISKONSEPSI PADA MATERI GEOMETRI

berserta perangkat yang diperlukan (bila ada). Dengan Hak Bebas Royalti Non-Ekklusif ini Perpustakaan UIN Sunan Ampel Surabaya berhak menyimpan, mengalih-media/format-kan, mengelolanya dalam bentuk pangkalan data (database), mendistribusikannya, dan menampilkan/mempublikasikannya di Internet atau media lain secara *fulltext* untuk kepentingan akademik tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan UIN Sunan Ampel Surabaya, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini yang saya buat dengan sebenarnya.

Surabaya, 14 Juni 2020

Penulis

(Nurul Fithrotuz Z.)
nama terang dan tanda tangan

PENGEMBANGAN TES DIAGNOSTIK *THREE-TIER* UNTUK MENDENTIFIKASI MISKONSEPSI PADA MATERI GEOMETRI

Oleh:
NURUL FITHROTUZ ZAIDAH

ABSTRAK

Geometri digunakan oleh setiap orang dalam kehidupan sehari-hari. Namun, masih banyak siswa yang mengalami kesulitan dalam memahami konsep geometri, sehingga dapat mempengaruhi pola pikir siswa ketika menyelesaikan permasalahan geometri. Hal ini sering mengakibatkan miskonsepsi pada siswa. Salah satu upaya untuk mengidentifikasi miskonsepsi yang terjadi yakni dengan melakukan tes diagnostik *three-tier*. Penelitian ini memiliki tujuan untuk mengetahui kevalidan, reliabilitas, dan kualitas butir soal tes diagnostik pilihan ganda *three-tier* serta hasil tes diagnostik pilihan ganda *three-tier* dalam mengidentifikasi miskonsepsi pada materi geometri.

Jenis penelitian ini merupakan penelitian pengembangan. Penelitian pengembangan ini mengadaptasi model pengembangan Tessmer yang terdiri dari dua tahapan yaitu tahap *preliminary* dan tahap *formative evaluation* yang meliputi *self evaluation*, *prototyping*, *small group*, dan *field test*. Subjek penelitian ini adalah 40 siswa dari kelas IX-C SMP Negeri 13 Surabaya. Data penelitian yang dikembangkan diperoleh dari lembar validasi dan tes diagnostik pilihan ganda *three-tier*. Data lembar validasi yang diperoleh kemudian dianalisis menggunakan rumus Rata-Rata Tiap Validator (RTV) dan data tes diagnostik pilihan ganda *three-tier* dianalisis dengan menggunakan program *SPSS Versi 16.0*.

Hasil penelitian ini menunjukkan: 1) Terdapat 8 dari 10 butir soal yang dinyatakan memiliki kevalidan yang baik dan layak untuk diuji coba; 2) Tes diagnostik pilihan ganda *three-tier* yang dikembangkan dinyatakan reliabel dengan rata-rata kategori tinggi sebesar 0,641; 3) Kualitas 10 butir soal tes diagnostik pilihan ganda *three-tier* yang dikembangkan yang meliputi daya pembeda soal, tingkat kesukaran dan keefektifan pengecoh termasuk kategori baik, hanya butir soal nomor 3 dengan daya pembeda, tingkat kesukaran dan keefektifan pengecoh yang mempunyai kategori tidak baik, butir soal nomor 4 dan 5 dengan daya pembeda soal yang tidak baik, serta butir soal nomor 9 dan 10 dengan keefektifan pengecoh alasan yang tidak berfungsi dengan baik; 4) Secara umum siswa mengalami jenis miskonsepsi teoritikal pada indikator mendefinisikan bangun datar segiempat dan jajargenjang.

Kata Kunci: Geometri, Miskonsepsi, Tes Diagnostik Pilihan Ganda *Three-Tier*

DAFTAR ISI

SAMPUL DALAM	i
PERSETUJUAN PEMBIMBING	ii
PENGESAHAN TIM PENGUJI	iii
PERNYATAAN KEASLIHAN TULISAN	iv
MOTTO	v
HALAMAN PERSEMBAHAN	vi
ABSTRAK	vii
KATA PENGANTAR	viii
DAFTAR ISI	x
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiv
DAFTAR LAMPIRAN	xv
BAB I PENDAHULUAN	1
A. Latar Belakang	1
B. Rumusan Masalah	8
C. Tujuan Penelitian dan Pengembangan	9
D. Spesifikasi Produk	9
E. Manfaat Pengembangan	10
F. Batasan Masalah	11
G. Definisi Operasional Variabel	11
BAB II KAJIAN PUSTAKA	13
A. Tes Diagnostik	13
B. Pilihan Ganda Tiga Tingkat (<i>Three-Tier Multiple Choice</i>)	14
1. Pengertian Pilihan Ganda <i>Three-Tier</i>	14
2. Langkah Pengembangan Tes Diagnostik Pilihan Ganda <i>Three-Tier</i>	16
C. Miskonsepsi	20
1. Pengertian Miskonsepsi	20
2. Penyebab Miskonsepsi	21
3. Jenis Miskonsepsi Materi Bangun Datar Segiempat	24
D. Tinjauan Materi Bangun Datar Segiempat	28
1. Definisi Segiempat	28
2. Macam-Macam Segiempat	28

E. CRI (<i>Certainty of Response Index</i>)	34
F. Validitas	38
G. Reliabilitas	43
H. Kualitas Butir Soal	47
1. Daya Pembeda Soal	47
2. Tingkat Kesukaran	49
3. Pengecoh atau Distraktor	51
BAB III METODE PENELITIAN	54
A. Model Penelitian dan Pengembangan	54
B. Prosedur Penelitian dan Pengembangan	55
C. Uji Coba Produk	60
D. Teknik Pengambilan Data	61
E. Instrumen Pengambilan Data	62
F. Teknik Pengambilan Data	63
BAB IV HASIL PENELITIAN DAN PEMBAHASAN	76
A. Deskripsi Data	76
B. Analisis Data	89
C. Revisi Produk	105
D. Kajian Produk Akhir	109
BAB V PENUTUP	144
A. Simpulan	144
B. Saran	145
DAFTAR PUSTAKA	146
LAMPIRAN	150

DAFTAR TABEL

Tabel 2.1	Penyebab Miskonsepsi	24
Tabel 2.2	Jenis dan Indikator Miskonsepsi	26
Tabel 2.3	Skala CRI dan Kriteria	36
Tabel 2.4	Kriteria Jawaban CRI Kategori Rendah dan Tinggi	37
Tabel 2.5	Modifikasi Kategori Tingkat Pemahaman Siswa	38
Tabel 2.6	Kriteria Validitas Instrumen Tes	43
Tabel 2.7	Interpretasi Reliabilitas	48
Tabel 2.8	Klasifikasi Interpretasi Daya Pembeda	50
Tabel 2.9	Klasifikasi Interpretasi Tingkat Kesukaran	51
Tabel 2.10	Klasifikasi Interpretasi Indeks Pengecoh	53
Tabel 3.1	Aspek Penilaian yang Menjadi Fokus <i>Prototype</i>	57
Tabel 3.2	Kriteria Penilaian Kevalidan Perangkat Pembelajaran	64
Tabel 3.3	Interpretasi Koefisien Korelasi	66
Tabel 3.4	Kategori Penafsiran Koefisien Reliabilitas	68
Tabel 3.5	Klasifikasi Interpretasi Daya Pembeda	69
Tabel 3.6	Klasifikasi Interpretasi Tingkat Kesukaran	71
Tabel 3.7	Klasifikasi Interpretasi Indeks Pengecoh	72
Tabel 3.8	Kriteria Penilaian Pilihan Jawaban	73
Tabel 3.9	Kriteria Penilaian Pilihan Alasan	73
Tabel 3.10	Interpretasi Skala CRI	73
Tabel 3.11	Interpretasi Hasil	74
Tabel 3.12	Kriteria Penilaian Tingkat Miskonsepsi	75
Tabel 4.1	Daftar Nama Validator Tes Diagnostik Pilihan Ganda <i>Three-Tier</i> Beserta Kodenya	76
Tabel 4.2	Data Hasil Validasi Para Ahli Tes Diagnostik Pilihan Ganda <i>Three-Tier</i>	77
Tabel 4.3	Data Hasil Validasi Empirik Tes Diagnostik Pilihan Ganda <i>Three-Tier</i>	79
Tabel 4.4	Data Hasil Reliabilitas Tes Diagnostik Pilihan Ganda <i>Three-Tier</i>	80
Tabel 4.5	Data Hasil Daya Pembeda Soal Tes Diagnostik Pilihan Ganda <i>Three-Tier</i>	83
Tabel 4.6	Data Hasil Tingkat Kesukaran Tes Diagnostik Pilihan Ganda <i>Three-Tier</i>	84
Tabel 4.7	Data Hasil Pengecoh Pilihan Jawaban Tes Diagnostik Pilihan Ganda <i>Three-Tier</i>	86

Tabel 4.8	Data Hasil Pengecoh Pilihan Alasan Tes Diagnostik Pilihan Ganda <i>Three-Tier</i>	87
Tabel 4.9	Hasil Analisis Data Validasi Para Ahli	89
Tabel 4.10	Hasil Analisis Validasi Empirik Tes Diagnostik Pilihan Ganda <i>Three-Tier</i>	91
Tabel 4.11	Hasil Analisis Reliabilitas Tes Diagnostik Pilihan Ganda <i>Three-Tier</i>	92
Tabel 4.12	Hasil Analisis Daya Pembeda Soal Tes Diagnostik Pilihan Ganda <i>Three-Tier</i>	94
Tabel 4.13	Hasil Analisis Tingkat Kesukaran Soal Tes Diagnostik Pilihan Ganda <i>Three-Tier</i>	95
Tabel 4.14	Hasil Analisis Pengecoh Jawaban Tes Diagnostik Pilihan Ganda <i>Three-Tier</i>	97
Tabel 4.15	Hasil Analisis Pengecoh Alasan Tes Diagnostik Pilihan Ganda <i>Three-Tier</i>	97
Tabel 4.16	Data Jumlah Jawaban Benar yang Diperoleh Siswa	99
Tabel 4.17	Persentase Paham Konsep, Miskonsepsi, dan Tidak Paham Konsep Berdasarkan Butir Soal	101
Tabel 4.18	Perhitungan Persentase Miskonsepsi Berdasarkan Indikator Pencapaian Kompetensi dan Indikator Soal	102
Tabel 4.19	Hasil Revisi Kisi-Kisi Tes Diagnostik Pilihan Ganda <i>Three-Tier</i>	105
Tabel 4.20	Hasil Revisi Soal Tes Diagnostik Pilihan Ganda <i>Three-Tier</i> ..	106
Tabel 4.21	Hasil Revisi Kunci Jawaban dan Pedoman Penskoran Tes Diagnostik Pilihan Ganda <i>Three-Tier</i>	108

DAFTAR GAMBAR

Gambar 2.1 Persegi	29
Gambar 2.2 Persegi Panjang	30
Gambar 2.3 Trapesium	31
Gambar 2.4 Belah Ketupat	32
Gambar 2.5 Jajargenjang	33
Gambar 2.6 Layang-Layang	34
Gambar 2.7 Segiempat Sembarang	34
Gambar 3.1 Alur Tahapan Tessmer	55
Gambar 3.2 Alur Tahapan Pengembangan Tessmer	60
Gambar 4.1 Bukti Komentar/Saran Validator Ahli	75
Gambar 4.2 Grafik Perbandingan Persentase Paham Konsep, Miskonsepsi, dan Tidak Paham Konsep	98

DAFTAR LAMPIRAN

1. Lembar <i>Prototype</i> I	150
2. Lembar <i>Prototype</i> II	155
3. Lembar <i>Prototype</i> III	160
4. Lembar Validasi Tes Diagnostik Pilihan Ganda <i>Three-Tier</i>	165
5. Hasil Validasi Menurut Penilaian Para Ahli	168
6. Data Hasil Tes Diagnostik Pilihan Ganda <i>Three-Tier</i>	177
7. Hasil Uji Validitas Empiris Berbantuan <i>SPSS Versi 16.0</i>	178
8. Hasil Analisis Uji Validitas Empiris	181
9. Hasil Uji Reliabilitas Berbantuan <i>SPSS Versi 16.0</i>	183
10. Hasil Analisis Uji Reliabilitas	184
11. Hasil Analisis Daya Pembeda Butir Soal	185
12. Hasil Analisis Tingkat Kesukaran Butir Soal.....	187
13. Hasil Analisis Pengecoh Butir Soal.....	189
14. Hasil Analisis Data Hasil Tes Diagnostik Pilihan Ganda <i>Three-Tier</i>	192
15. Kisi-Kisi Tes Diagnostik Pilihan Ganda <i>Three-Tier</i>	193
16. Soal Tes Diagnostik Pilihan Ganda <i>Three-Tier</i>	205
17. Kunci Jawaban Tes Diagnostik Pilihan Ganda <i>Three-Tier</i>	214
18. Pedoman Penskoran Tes Diagnostik Pilihan Ganda <i>Three-Tier</i>	217
19. Surat Keterangan Melakukan Penelitian	218
20. Surat Keterangan Telah Melakukan Penelitian	219
21. Surat Tugas Pembimbing	220
22. Kartu Konsultasi Pembimbing	221
23. Daftar Riwayat Hidup	222
24. Dokumentasi	223

BAB I PENDAHULUAN

A. Latar Belakang

Geometri adalah salah satu cabang materi matematika yang mempelajari hubungan antara titik-titik, garis-garis, sudut-sudut, bidang-bidang serta bidang datar dan bangun ruang.¹ Budiarto menyatakan bahwa tujuan pembelajaran geometri adalah untuk mengembangkan kemampuan berpikir logis, mengembangkan intuisi keruangan, menanamkan pengetahuan untuk menunjang materi yang lain, dan dapat membaca serta menginterpretasikan argumen-argumen matematik.²

Geometri merupakan cabang matematika yang perlu dikaji dan dipelajari secara mendalam, karena geometri digunakan oleh setiap orang dalam kehidupan sehari-hari.³ Menurut Sulaiha terdapat 3 alasan tentang pentingnya geometri: pertama, geometri merupakan salah satu ilmu yang dapat mengaitkan matematika dengan bentuk fisik dunia nyata. Kedua, geometri merupakan salah satu ilmu yang memungkinkan ide-ide dari bidang matematika yang lain untuk digambar. Ketiga, geometri dapat memberi contoh yang tak tunggal tentang sistem matematika.⁴ Namun masih banyak siswa yang berasumsi bahwa matematika adalah pelajaran yang sulit.⁵ Nafiah mengemukakan bahwasanya kesulitan siswa dalam geometri karena kesalahan-kesalahan dalam menyelesaikan masalah geometri.⁶ Hal ini didukung oleh hasil penelitian yang dilakukan oleh Basuki, bagi kebanyakan siswa kesulitan yang dialami dalam menyelesaikan sebuah pertanyaan tentang bidang datar adalah kurang memahami

¹ Mega Teguh Budiarto, "Pembelajaran Geometri dan Berpikir Geometri" (Prosiding Seminar Nasional Matematika FMIPA ITS Surabaya, Surabaya, 2000), hal. 439.

² Ibid.

³ Christine Wulandari, "Menanamkan Konsep Bentuk Geometri", *Jurnal Pengabdian Masyarakat Ipteks*, 3:1, 2017, hal. 2.

⁴ Sulaiha, Tesis: "*Profil Berpikir Geometri Siswa MTs Pada Materi Bangun Segiempat Ditinjau dari Gaya Kognitif*", (Surabaya: Program Pascasarjana (PPs) Universitas Negeri Surabaya, 2016), hal. 1.

⁵ Malikatun Ngilman Nafiah, Endah Budi Rahaju, "Identifikasi Tahap Pemahaman Geometri Siswa Berdasarkan Teori Van Hiele Ditinjau dari Perbedaan Gender Pada Materi Persegi Panjang Kelas VII SMP", *Jurnal Ilmiah Pendidikan Matematika*, 2:6, 2017, hal. 293.

⁶ Ibid.

konsepnya.⁷ Oleh karena itu, pemahaman konsep sangat penting dalam pembelajaran matematika terutama pada pembelajaran geometri.

Para pendidik kerap kali menjumpai bahwa para siswa memiliki konsep awal sebagai pengetahuan sebelum siswa memasuki ruang pembelajaran yang mana konsep tersebut belum sama atau berbeda dengan konsep ilmiah, konsep itulah yang dinamakan prakonsepsi.⁸ Salah satu faktor rendahnya prestasi belajar siswa dikarenakan adanya suatu prakonsepsi yang berbeda dengan konsep ilmiah. Jika siswa diajarkan tentang perkalian bukan hanya membentuk kemampuan berhitung perkalian secara cepat saja melainkan pemahaman tentang sesuatu yang disebut dengan perkalian itu sendiri. Pemberian rumus-rumus menghitung cepat tanpa menanamkan pemahaman konsep kepada siswa akan menjadi bumerang bagi pendidik di kemudian hari.⁹ Ketidaksesuaian atau kekeliruan dalam pemahaman konsep yang diakui oleh para ahli inilah yang sering disebut miskonsepsi.

Miskonsepsi adalah suatu konsep yang tidak sesuai dengan konsep yang diakui para ahli.¹⁰ Konsep yang dibawa siswa dapat sesuai dengan konsep ilmiah tetapi juga dapat tidak sesuai dengan konsep ilmiah. Konsep awal yang dimiliki siswa disebut dengan konsepsi. Konsep awal atau konsepsi yang tidak sesuai dengan konsep ilmiah disebut sebagai miskonsepsi. Salah satu kecenderungan yang menyebabkan siswa gagal dalam mengerjakan soal-soal matematika dengan baik yaitu karena siswa kesulitan dalam memahami konsep dan kurang menggunakan nalar yang baik dalam menyelesaikan soal atau tes yang diberikan.¹¹

⁷ Novila Rahmad Basuki, "Analisis Kesulitan Siswa SMK Pada Materi Pokok Geometri dan Alternatif Pemecahannya" (*Prosiding Seminar Nasional Pendidikan Matematika Universitas Muhammadiyah Surakarta*, 2012).

⁸ Arifatul Ahla Ainus Salamah, Skripsi: "Analisis Miskonsepsi Siswa Menggunakan Pendekatan Kognitif Menurut Teori Piaget pada Materi Optik Kelas VIII MTs NU MU'ALLIMAT Kudus". (Semarang: UIN Walisongo Semarang, 2015), hal. 1.

⁹ Muhammad Hariwijaya-Sutan Surya, *Adventure in Math Tes IQ Matematika*, (Jakarta: Oryza, 2012), hal. 42.

¹⁰ Paul Suparno., *Miskonsepsi dan Perubahan Konsep Pendidikan Fisika*. (Jakarta : PT. Grasindo, 2013), hal. 8.

¹¹ Dian Mutmainna, dkk., "Pengembangan Tes diagnostik Pilihan Ganda Tiga tingkat Untuk Mengidentifikasi Pemahaman Konsep Matematika", *Jurnal Matematika dan Pembelajaran*, 6:1, 2018, hal. 58.

Penelitian Budiarto dan Artiono menunjukkan bahwa: (1) 22% dari 54 siswa menggunakan “yang dibuktikan sebagai yang diketahui”, (2) 19,4% dari 42 guru SMP dan SMU Surabaya mengalami kesulitan menyelesaikan masalah “buktikan bahwa ...”, (3) 24% dari 62 siswa peserta mata kuliah geometri tidak dapat menuliskan apa yang diketahui dan apa yang ditanyakan dari suatu permasalahan yang diberikan, dan adanya miskonsepsi siswa dalam memahami konsep-konsep geometri.¹² Di samping itu, juga menunjukkan koneksi antara konsep-konsep geometri yang lemah, yaitu tidak dapat mengaitkan pengetahuan satu dengan pengetahuan lain dalam geometri apalagi dengan bidang lain dalam matematika di luar geometri.¹³

Materi geometri dipandang sebagai salah satu pokok bahasan yang cukup sulit karena geometri bersifat abstrak.¹⁴ Konsep segiempat merupakan salah satu materi kajian geometri dalam matematika, dimana siswa masih mengalami kesulitan terutama dalam mengungkapkan definisi bangun datar segiempat. Penelitian yang dilakukan oleh Budiarto dan Ningrum menunjukkan bahwa siswa mengalami miskonsepsi pada definisi bangun datar segiempat, sebagian besar mereka menganggap bahwa segiempat selalu dalam bentuk beraturan saja.¹⁵ Hasil penelitian Farida juga mengungkapkan bahwa siswa mengalami miskonsepsi pada simbol dan istilah matematika pada materi bangun datar segiempat.¹⁶ Pemahaman konsep, prinsip, dan operasi dinilai sangat penting, karena konsep ini berkaitan dengan konsep geometri. Jika konsep bangun datar segiempat siswa mengalami miskonsepsi atau bahkan belum menguasai, maka siswa akan mengalami kesulitan dalam memahami konsep selanjutnya.

¹² Mega Teguh Budiarto, Rudianto Artiono, “Geometri dan Permasalahan Dalam Pembelajarannya (Suatu Penelitian Meta Analisis)”, *Jurnal Magister Pendidikan Matematika*, 1:1, 2019, hal. 10.

¹³ Ibid.

¹⁴ Mega Teguh Budiarto, Rachmania Widya Ningrum, “Miskonsepsi Siswa SMP pada Materi Bangun Datar Segiempat dan Alternatif Mengatasinya”, *MATHEdunesa Jurnal Ilmiah Pendidikan Matematika*, 1:5, 2016, hal. 60.

¹⁵ Ibid, hal. 59.

¹⁶ Anisatul Farida, “Analisis Miskonsepsi Siswa Terhadap Simbol dan Istilah Matematika pada Konsep Hubungan Bangun Datar Segiempat Melalui Permainan dengan Alat Peraga (SD Muhammadiyah 1 Surakarta)” (*Prosiding Konferensi Nasional Penelitian Matematika dan Pembelajarannya Universitas Negeri Semarang, 2016*), hal. 286.

Adanya miskonsepsi ini jelas akan sangat menghambat proses penerimaan dan asimilasi pengetahuan-pengetahuan baru dalam diri siswa, sehingga akan menghalangi keberhasilan siswa dalam proses belajar.¹⁷ Hirarki matematika sebagai suatu bidang yang memiliki keterkaitan pembahasan satu dengan lainnya menyebabkan timbulnya miskonsepsi merupakan hal yang sangat fatal.

Atas kenyataan ini, semestinya sekolah harus berperan membantu memecahkan masalah siswa dikarenakan pemahaman konsep merupakan komponen yang penting dalam prinsip pembelajaran matematika. Beberapa metode yang dilakukan oleh guru dalam mengevaluasi siswa guna mengetahui miskonsepsi siswa diantaranya dengan melakukan wawancara, membuat peta konsep oleh siswa, tes essay, dan tes diagnostik.¹⁸

Wawancara merupakan cara efektif untuk menggali pemahaman siswa secara mendalam dan terperinci, namun wawancara akan sangat menyulitkan apabila digunakan pada responden yang banyak.¹⁹ Pemahaman konsep merupakan cara belajar yang mengembangkan proses belajar yang bermakna yang akan meningkatkan pemahaman siswa dan daya ingat belajarnya akan tetapi perlunya waktu yang cukup lama dalam menyusun, sedangkan waktu yang tersedia terbatas.²⁰ Tes essay merupakan tes yang terdiri dari suatu pertanyaan atau suatu suruhan yang menghendaki jawaban berupa uraian-uraian yang relatif panjang, sehingga akan menyulitkan bagi guru untuk mengetahui kelemahan atau kelebihan setiap siswa dan membutuhkan waktu yang lama dalam penilaian.²¹ Sedangkan tes diagnostik adalah tes yang digunakan untuk mengetahui kelemahan-kelemahan siswa sehingga

¹⁷ Klammer Joel, *An Overview of Techniques for Identifying, Acknowledging and Overcoming Alternate Conceptions in Physics Education, 1997/98 Klingenstein Project Report*. (Teachers College-Columbia University, 1998), hal.7.

¹⁸ Dessy Rositasari, Skripsi: *Pengembangan Tes Diagnostik Three-tier Untuk Mendeteksi Miskonsepsi Siswa SMA Pada Topik Asam Basa*, (Jakarta: UIN Syarifuddin Jakarta), 2014, hal. 5.

¹⁹ Ayla Cetin Dinar, Omer Geban, "Developmet of a Three-Tier Test to Assess High School Students Understanding of Acids and Bases", (*Procedia Sosial and Behavioral Sciences 15, 2011*), hal. 600.

²⁰ Novak, Gowin, *Learning How to Learn*, (Cambridge: Cambridge University Press, 1985), hal. 18.

²¹ Purwanto, *Evaluasi Hasil Belajar*, (Yogyakarta: Pustaka Pelajar, 2013), hal. 58.

berdasarkan hal tersebut dapat dilakukan penanganan yang tepat.²² Oleh karena itu, peneliti menggunakan tes diagnostik karena tes tersebut menjadi salah satu alat pengukuran yang baik untuk menilai pemahaman konsep matematika siswa, dimana kebanyakan siswa mengalami kelemahan maupun kesulitan dalam menyelesaikan persoalan.

Brueckner & Melby menjelaskan bahwa tes diagnostik digunakan untuk menentukan elemen-elemen dalam suatu mata pelajaran yang memiliki kelemahan-kelemahan khusus dan menyediakan alat untuk menemukan penyebab kekurangan tersebut.²³ Tes diagnostik merupakan sarana yang ditujukan untuk mengungkap miskonsepsi yang dimiliki siswa berdasarkan informasi kesalahan yang dibuatnya, sehingga dapat diberi tindak lanjut yang sesuai dengan hasil tes tersebut.

Tes diagnostik yang telah banyak digunakan antara lain *one-tier*, *two-tier*, dan *three-tier*.²⁴ *Two-tier multiple choice* merupakan alat tes yang cukup sukses mendiagnosis miskonsepsi siswa dan mudah untuk dinilai, tetapi *two-tier test* tidak dapat membedakan miskonsepsi dengan kurangnya pengetahuan (*lack of knowledge*).²⁵ *Three tier test* menggunakan cara yang sederhana dan mudah untuk mengidentifikasi miskonsepsi dan membedakannya dengan kurangnya pengetahuan (*lack of knowledge*), yaitu dengan menambahkan tingkat keyakinan jawaban yang dipilih siswa.²⁶

Diagnostik *three tier test* diprediksi dapat mengidentifikasi miskonsepsi siswa lebih akurat dibandingkan dengan tes diagnostik *one tier* atau *two tier*.²⁷ Tes diagnostik *three-tier* akan memungkinkan bagi guru untuk mengidentifikasi miskonsepsi

²² Suharsimi Arikunto, *Prosedur Penelitian Suatu Pendekatan Praktek*, (Jakarta: Rineka Cipta, 2012), hal. 48.

²³ Leo John Brueckner, Ernest O. Melby, *Diagnostic and Remedial Teaching*, (Boston: Houghton Mifflin Company, 1981), hal. 73.

²⁴ Harika Ozge Arslan, et.al., "A Three-Tier Diagnostic Test to Assess Pre-Service Teacher's Misconceptions About Global Warming, Greenhouse Effect, Ozone Layer Depletion, and Acid Rain", *International Journal of Science Education*, 34:11, 2012.

²⁵ Saleem Hasan, et.al., "Misconceptions and the Certainty of Response Index (CRI)", *Physics Education Research American Journal of Physics*, 34:5, 1999, hal. 294-299.

²⁶ Aliefman Hakim, et.al., "Student Concept Understanding of Natural Product Chemistry in Primary and Secondary Metabolites Using the Data Collecting Technique of Modified CRI", *International Online Journal of Education Sciences*, 4:3, 2012, hal. 544-553.

²⁷ Harika Ozge Arslan, et.al., Op. Cit, hal. 1667-1686.

dengan cara yang mudah dan akurat sehingga dapat memberikan penguasaan siswa terhadap materi yang telah disampaikan, sedangkan bagi siswa akan mengubah konsep yang salah menuju konsep yang benar atau memperbaiki miskonsepsi yang terjadi pada dirinya dengan konsep ilmiah.

Pengembangan tes diagnostik bentuk pilihan ganda *three-tier* ini sudah dilakukan oleh beberapa peneliti di Indonesia, di antaranya yaitu dilakukan oleh Nabilah tentang instrumen diagnostik *three-tier* yang berfungsi untuk mendeteksi miskonsepsi siswa kelas VII SMPN 24 Makassar pada materi pecahan.²⁸ Tes diagnostik dikembangkan dengan mengkombinasi kemampuan berpikir siswa dan alasan jawaban siswa dengan tingkat keyakinan dalam menjawab. Struktur tes diagnostik tiga tingkat menggunakan Borg dan Gall yang terdiri atas studi pendahuluan, perancangan draft produk, pengembangan produk, dan uji coba produk. Sedangkan pada penelitian ini, peneliti mengembangkan tes diagnostik pilihan ganda *three-tier* dengan menggunakan model pengembangan *tipe formative evaluation Tessmer* untuk mengidentifikasi miskonsepsi pada materi geometri yang terdiri dari dua tahapan yaitu tahap *preliminary* dan tahap *formative evaluation* yang meliputi *self evaluation, prototyping, small group* dan *field test*.

Ada beberapa penelitian pengembangan tes diagnostik tiga tingkat pada pembelajaran IPA, salah satunya yakni penelitian yang dilakukan Aisy mengenai pengembangan instrumen tes diagnostik pilihan ganda tiga tingkat pada materi Kimia yaitu konsep redoks dengan subjek penelitian siswa SMA kelas X yang menghasilkan sebuah produk tes diagnostik tiga tingkat yang terdiri dari 40 soal pilihan ganda.²⁹ Sedangkan pada penelitian ini menghasilkan tes diagnostik *three-tier* pada materi Geometri yang terdiri dari 10 butir soal dengan subjek penelitian siswa SMP kelas IX

²⁸ Lu'lu Tu'yikan Nabilah, Skripsi: *Pengembangan Instrumen Diagnostik Three-Tier Pada Materi Pecahan Kelas VII SMPN 24 Makassar*, (Makassar: Universitas Negeri Makassar, 2019), hal. 32.

²⁹ Ziadatul Aisy, Skripsi: *Pengembangan Instrumen Tes Diagnostik Pilihan Ganda Tiga Tingkat Untuk Mengungkap Miskonsepsi Peserta Didik Kelas X Materi Redoks*, (Semarang: UIN Walisongo, 2018), hal. 79.

Di samping itu, penelitian yang dilakukan oleh Astutik tentang gerak melingkar beraturan.³⁰ Hasil penelitian berupa setiap butir soal tes diagnostik *three-tier* terdiri atas tiga rangkaian soal bertingkat yaitu tingkat pertama soal pilihan ganda dengan 4 pilihan jawaban, kemudian soal tingkat kedua berupa 3 pilihan jawaban alasan tertutup serta 1 pilihan jawaban alasan terbuka untuk soal tingkat pertama, serta tingkat ketiga berupa tingkat keyakinan responden dengan skala enam. Sedangkan pada penelitian ini menghasilkan setiap butir soal tes diagnostik pilihan ganda *three-tier* terdiri atas rangkaian soal bertingkat yaitu tingkat pertama soal pilihan ganda dengan 4 pilihan jawaban, kemudian soal tingkat kedua berisi 4 pilihan alasan tertutup serta tingkat ketiga berupa kolom CRI dengan skala empat.

Berdasarkan latar belakang yang telah dikemukakan dan belum ada tes diagnostik mengenai pembelajaran geometri maka peneliti tertarik untuk melakukan penelitian dengan judul “**Pengembangan Tes Diagnostik *Three-Tier* Untuk Mengidentifikasi Miskonsepsi Pada Materi Geometri**”.

B. Rumusan Masalah

Berdasarkan kondisi yang telah dipaparkan dalam latar belakang, pertanyaan dalam peneliti ini dirumuskan sebagai berikut:

1. Bagaimana kevalidan tes diagnostik pilihan ganda *three-tier* untuk mengidentifikasi miskonsepsi pada materi geometri?
2. Bagaimana reliabilitas tes diagnostik pilihan ganda *three-tier* untuk mengidentifikasi miskonsepsi pada materi geometri?
3. Bagaimana kualitas butir soal tes diagnostik pilihan ganda *three-tier* untuk mengidentifikasi miskonsepsi pada materi geometri?
4. Bagaimana jenis miskonsepsi siswa pada materi geometri dari hasil tes diagnostik pilihan ganda *three-tier*?

C. Tujuan Penelitian dan Pengembangan

Berdasarkan rumusan masalah yang telah disebutkan, maka tujuan penelitian ini adalah:

³⁰ Widi Astutik, Skripsi: *Pengembangan Instrumen Three-Tier Multiple Choice Diagnostic Test Untuk Mengidentifikasi Miskonsepsi Siswa SMA Materi Gerak Melingkar Beraturan*, (Semarang: UIN Walisongo, 2018), hal. 5.

1. Untuk mendeskripsikan kevalidan tes diagnostik pilihan ganda *three-tier* untuk mengidentifikasi miskonsepsi pada materi geometri.
2. Untuk mendeskripsikan reliabilitas tes diagnostik pilihan ganda *three-tier* untuk mengidentifikasi miskonsepsi pada materi geometri.
3. Untuk mendeskripsikan kualitas butir soal tes diagnostik pilihan ganda *three-tier* untuk mengidentifikasi miskonsepsi pada materi geometri.
4. Untuk mendeskripsikan jenis miskonsepsi siswa pada materi geometri dari hasil tes diagnostik pilihan ganda *three-tier*.

D. Spesifikasi Produk

Produk yang dihasilkan dalam pengembangan ini adalah tes diagnostik pilihan ganda *three-tier* untuk mengidentifikasi miskonsepsi pada materi geometri dengan spesifikasi produk yang dikembangkan adalah sebagai berikut:

1. Produk yang dikembangkan adalah soal-soal tes diagnostik pilihan ganda *three-tier* dilengkapi dengan kisi-kisi soal, kunci jawaban, dan pedoman penskoran.
2. Setiap butir soal tes diagnostik pilihan ganda *three-tier* terdiri atas rangkaian soal bertingkat yaitu tingkat pertama soal pilihan ganda dengan 4 pilihan jawaban, kemudian soal tingkat kedua berisi 4 pilihan alasan tertutup serta tingkat ketiga berupa kolom CRI dengan skala empat.
3. Tes diagnostik pilihan ganda *three-tier* disusun berdasarkan materi yang telah ditentukan pada kompetensi dasar kelas VII.

E. Manfaat Pengembangan

Berdasarkan tujuan penelitian, adapun beberapa manfaat dari penelitian ini sebagai berikut:

1. Bagi Sekolah
Dengan adanya penelitian ini dapat membantu pihak sekolah dalam menentukan kebijakan penerapan metode evaluasi yang tepat untuk mengatasi miskonsepsi siswa supaya terwujudnya tujuan pembelajaran yang diharapkan.
2. Bagi Guru
Dapat membantu guru dalam menentukan alat ukur alternatif yang digunakan untuk mengidentifikasi miskonsepsi siswa

sehingga dapat memberikan tindak lanjut yang tepat untuk mengatasi miskonsepsi yang terjadi pada siswa.

3. Bagi Peneliti

Dapat menambah wawasan dan pengalaman peneliti tentang bagaimana mengembangkan tes diagnostik pilihan ganda *three-tier* pada materi geometri.

F. Batasan Masalah

Agar tidak terjadi salah penafsiran terhadap maksud dari penelitian ini, didefinisikan terkait rumusan masalah sebagai berikut:

1. Materi tes diagnostik pilihan ganda *three-tier* pada materi geometri sesuai dengan kurikulum 2013 SMP/MTs edisi revisi 2017 mengenai bangun datar segiempat.
2. Kualitas butir soal tes diagnostik pilihan ganda *three-tier* meliputi daya pembeda butir soal, tingkat kesukaran butir soal dan pengecoh atau distraktor.

G. Definisi Operasional Variabel

Agar tidak terjadi salah penafsiran terhadap maksud dari penelitian ini, didefinisikan istilah-istilah rumusan masalah sebagai berikut:

1. Tes diagnostik pilihan ganda *three-tier* adalah tes yang bertujuan untuk mengetahui kelemahan siswa yang terdiri dari tiga tahap (*tier*) pilihan, yakni *tier* pertama berisi sejumlah pilihan jawaban, *tier* kedua berisi sejumlah pilihan alasan untuk jawaban yang dipilih pada *tier* pertama, dan *tier* ketiga berisi tingkat keyakinan dalam menjawab *tier* pertama dan *tier* kedua.
2. Proses pengembangan tes diagnostik *three-tier* untuk mengidentifikasi miskonsepsi pada materi geometri dalam penelitian ini adalah keseluruhan proses perancangan yang meliputi tahap *preliminary* dan tahap *formative evaluation* yang terdiri dari *self evaluation*, *prototyping*, dan *field test*.
3. Miskonsepsi merupakan suatu kesalahan dalam menentukan klasifikasi antar bangun datar, menjelaskan hubungan antar bangun datar dan mendefinisikan konsep bangun datar.
4. CRI (*Certainty of Response Index*) merupakan ukuran tingkat keyakinan responden dalam menjawab setiap pertanyaan atau soal yang diberikan.

5. Validitas adalah suatu ukuran yang menunjukkan tingkat kevalidan atau kesahihan suatu tes.
6. Reliabilitas adalah sebuah tes yang apabila diujikan berkali-kali mendapatkan hasil yang relatif tetap.
7. Daya pembeda soal merupakan kemampuan suatu butir soal yang dapat membedakan antara siswa yang telah menguasai materi dengan siswa yang belum menguasai materi dan mempunyai indeks deskriminasi antara 0,400 sampai 0,700.
8. Tingkat kesukaran soal adalah angka yang menunjukkan proporsi siswa yang menjawab benar pada setiap butir soal dengan nilai indeks kesukaran berada di atas atau sama dengan 0,400.
9. Pengecoh atau distraktor soal adalah pilihan jawaban yang bukan kunci jawaban dan dipilih minimal 5% dari peserta tes.



BAB II KAJIAN PUSTAKA

A. Tes Diagnostik

Diagnosis adalah proses yang kompleks dalam suatu usaha untuk menarik kesimpulan dari hasil-hasil pemeriksaan gejala-gejala, perkiraan penyebab, pengamatan dan penyesuaian dengan kategori secara baik.³¹ Menurut Depdiknas, diagnostik merupakan suatu kegiatan yang dilakukan untuk mengidentifikasi gejala-gejala yang ditimbulkan.³² Seperti halnya pada bidang kedokteran yakni pekerjaan seorang dokter, sebelum menentukan penyakit dan obat yang tepat bagi pasien, seorang dokter akan melakukan pemeriksaan lebih teliti. Begitu pula dalam dunia pendidikan, pemeriksaan awal juga harus dilakukan oleh seorang guru untuk mengetahui kesulitan belajar yang dihadapi siswa dalam pembelajaran agar bisa memberikan bentuk bantuan yang tepat kepada siswa.

Menurut Ebel dan Fribie, tes diagnostik adalah rancangan untuk mengetahui kekurangan-kekurangan yang khusus atau kegagalan-kegagalan dalam belajar pada beberapa subjek atau pelajaran seperti membaca dan aritmatika.³³ Suwanto menjelaskan tes diagnostik merupakan tes yang digunakan untuk mengetahui kelemahan atau miskonsepsi pada topik tertentu dalam pembelajaran sehingga dari hasil tes didapat masukan tentang respon siswa untuk memperbaiki kelemahannya.³⁴ Pendapat senada dinyatakan oleh Linn dan Gronlud, tes diagnostik adalah tes yang dirancang untuk mengetahui sebab kegagalan siswa dalam belajar.³⁵

Tes diagnostik merupakan tes yang dibentuk untuk mendapatkan informasi khusus dari jawaban siswa, sehingga dapat diidentifikasi kelemahan pada pola pikir siswa tersebut. Berdasarkan penjelasan di atas, tes diagnostik adalah suatu tes yang bertujuan untuk mengetahui kelemahan atau kesulitan siswa dalam kegiatan belajar.

³¹ Suwanto, *Pengembangan Tes Diagnostik dalam Pembelajaran*, (Yogyakarta: Pustaka Pelajaran, 2013), hal. 90.

³² Badan Standar Nasional Pendidikan, *Panduan Penilaian Kelompok Mata Pelajaran Ilmu*, Jakarta, 2007, hal. 3.

³³ Robbert L. Ebel., David A. Fribie, *Essentials of Educational Measurment*, (New Jersey: Prentice-Hall Inc, 1991), hal. 375.

³⁴ Suwanto, Op. Cit., hal. 114.

³⁵ Robert L. Linn, Norman E. Gronlud, *Measurement and Assessment in Teaching*, (New York: Mac Milian Publishing Co, 2008), hal.13.

B. Pilihan Ganda Tiga Tingkat (*Three-Tier Multiple Choice*)

1. Pengertian Pilihan Ganda *Three-Tier*

Tes diagnostik *three tier multiple choice* merupakan pengembangan dari tes diagnostik *two tier multiple choice*.³⁶ Menurut Syahrul dan Setyarsih, *three-tier diagnostic test* merupakan tes diagnostik yang tersusun dari tiga tingkatan soal.³⁷ Tingkat pertama (*one-tier*) berupa pilihan ganda biasa, tingkat kedua (*two-tier*) berupa pilihan alasan, dan tingkat ketiga (*three-tier*) berupa pertanyaan penegasan tentang keyakinan dari jawaban yang telah dipilih pada tingkat satu dan dua.³⁸

Siswa yang menjawab dengan benar dan yakin atas jawabannya pada *two-tier test* menunjukkan bahwa ia memang paham terhadap konsep tertentu, siswa yang yakin dengan jawabannya walaupun jawaban tersebut salah menunjukkan bahwa ia mengalami miskonsepsi, sedangkan siswa yang menjawab salah dan tidak yakin atas jawabannya bukan berarti ia mengalami miskonsepsi, tetapi ia mengalami *lack of knowledge*.³⁹

Diagnostik *three tier test* diprediksi dapat mengidentifikasi miskonsepsi siswa lebih akurat dibandingkan dengan tes diagnostik *one tier* atau *two tier*.⁴⁰ Pesman dan Eryilmaz menyatakan bahwa *three tier test* dapat dianggap sebagai instrumen yang lebih valid dan dapat diandalkan untuk penilaian prestasi atau miskonsepsi.⁴¹ Sedangkan menurut penelitian yang dilakukan Arslan, et.al

³⁶ Syarifatul Mubarak, dkk., “Pengembangan Tes Diagnostik Three Tier Multiple Choice Untuk Mengidentifikasi Miskonsepsi Peserta Didik Kelas XI”, *Journal of Innovative Science Education*, 5:2, 2016, hal. 102.

³⁷ Dimas Ardiansyah Syahrul, Woro Setyarsih, “Identifikasi Miskonsepsi dan Penyebab Miskonsepsi Siswa dengan Three-Tier Diagnostic Test Pada Materi Dinamika Rotasi”, *Jurnal Inovasi Pendidikan Fisika (JIPF)*, 4:3, 2015, hal. 68.

³⁸ Zubeyde Demet Kirbulut, “Using Three-Tier Diagnostic Test to Assess Student’s Misconceptions of Sytates of Matter”, *Eurasia Journal of Mathematics, Science & Technology Education*, 10:5, 2014, hal. 509-521.

³⁹ Harika Ozge Arslan, et.al., “A Three-Tier Diagnostic Test to Assess Pre-Service Teacher’s Misconceptions About Global Warming, Greenhouse Effect, Ozone Layer Depletion, and Acid Rain”, *International Journal of Science Education*, 34:11, 2012.

⁴⁰ Harika Ozge Arslan, et.al., “A Three-Tier Diagnostic Test to Assess Pre-Service Teacher’s Misconceptions About Global Warming, Greenhouse Effect, Ozone Layer Depletion, and Acid Rain”, *International Journal of Science Education*, 34:11, 2012.

⁴¹ Haki Pesman, Ali Eryilmaz, “Development of a Three-Tier Test to Assess Misconceptions About Simple Electric Circuits”, *The Journal of Educational Research*, 103:3, 2010, hal. 209.

menyimpulkan bahwa instrumen diagnostik *three tier test* yang valid dan reliabel tidak hanya bisa mengidentifikasi miskonsepsi guru dalam mengajar tetapi juga miskonsepsi siswa dalam belajar.⁴² Hal ini sejalan dengan penelitian yang dilakukan Taslidere menyimpulkan bahwa *three-tier test* yang dikembangkan adalah alat ukur yang reliabel dan valid untuk menginvestigasi pemahaman konseptual dan miskonsepsi siswa.⁴³

Berdasarkan dari penjelasan di atas, maka dapat disimpulkan bahwasanya tes diagnostik pilihan ganda *three-tier* adalah tes yang bertujuan untuk mengetahui kelemahan siswa yang terdiri dari tiga tahap (*tier*) pilihan, yakni *tier* pertama berisi sejumlah pilihan jawaban, *tier* kedua berisi sejumlah pilihan alasan untuk jawaban yang dipilih pada *tier* pertama, dan *tier* ketiga berisi tingkat keyakinan dalam menjawab *tier* pertama dan *tier* kedua.

2. Langkah Pengembangan Tes Diagnostik Pilihan Ganda *Three-Tier*

Secara luas pengembangan tes diagnostik untuk mengidentifikasi miskonsepsi siswa dalam bidang tertentu terdiri dari 10 tahapan. Adapun penjelasan dari setiap tahap pengembangan antara lain:⁴⁴

a. Tahap pendefinisian (*Defining the content*)

Terdapat empat langkah pertama yang berkaitan dengan penetapan batas konsep dan mengidentifikasi pernyataan pengetahuan proporsional dan pengembangan peta konsep sebagai berikut:

- 1) Mengidentifikasi pernyataan pengetahuan proporsional yang telah dijelaskan oleh Finley dan Stewart dengan menguraikan pengembangan kurikulum dan pengajaran. Pada tahap ini, langkah pertama yang dilakukan adalah menelaah kurikulum pembelajaran dalam mengembangkan tes diagnostik pilihan ganda *three-tier*.

⁴² Harika Ozge Arslan, et.al., Op. Cit.

⁴³ Erdal Taslidere, "Development and Use of a Three-Tier Diagnostic Test to Assess High School Student's Misconceptions About the Photoelectric Effect", *Journal Research in Science & Technological Education*, 34:2, 2016.

⁴⁴ David F. Treagust, "Development and Use of Diagnostic Test to Evaluate Student's Misconceptions in Science", *International Journal of Science Education*, 10:2, 1988, hal. 161-164.

- 2) Setelah menelaah kurikulum, tahap selanjutnya adalah mengembangkan peta konsep yang berhubungan dengan materi yang akan diteliti berdasarkan prosedur yang dijelaskan oleh Novaik yakni peta konsep hanya terdiri atas dua konsep yang dihubungkan oleh satu kata penghubung untuk membentuk proposisi atau pernyataan dalam bentuk kalimat yang dapat dinilai benar dan salahnya, kemudian peta konsep disusun secara urutan tingkatan, semakin ke bawah maka konsep lebih spesifik.
- 3) Menghubungkan pengetahuan proporsional ke peta konsep. Proporsional pernyataan pengetahuan terkait langsung dengan peta konsep untuk memastikan konten yang dianalisis konsisten secara internal dengan pemeriksaan reliabilitas. Dalam artian pada tahap ini memastikan adanya keterkaitan antara pernyataan pengetahuan konsep yang akan dianalisis dengan peta konsep.
- 4) Memvalidasi isi. Pada tahap ini ialah memvalidasi pernyataan mengenai pengetahuan konsep dan peta konsep yang telah dianalisis dan dibuat. Validasi dapat dilakukan oleh pendidik sains, guru, dosen, ilmuwan, atau semacamnya dengan pengetahuan menyeluruh tentang materi yang akan dianalisis. Apabila terdapat perbedaan atau penyimpangan pada konsep maka dapat direvisi atau dimodifikasi, sehingga tidak ada pertanyaan yang dikembangkan untuk tes pilihan ganda yang tidak ada hubungan dengan konsep yang diajarkan.

b. Memperoleh informasi mengenai miskonsepsi siswa (*Obtaining information about student's misconceptions*)

Pengembangan tes diagnostik guna mengevaluasi miskonsepsi siswa melibatkan pemeriksaan menyeluruh dari literatur yang relevan dan berstruktur kognitif, wawancara dengan siswa tentang pemahaman mereka mengenai konsep tersebut serta memperoleh tanggapan dari siswa.

- 1) Memeriksa literatur atau penelitian yang terkait. Sebelum memulai untuk mengidentifikasi masalah kesalahpahaman di area subjek, penting untuk

memeriksa literatur/referensi yang berkaitan dengan penelitian pada miskonsepsi dan mencari informasi sebanyak-banyaknya mengenai penelitian miskonsepsi. Tujuan pada tahapan ini adalah sebagai wadah untuk membangun informasi dasar dalam mengembangkan pertanyaan pilihan ganda untuk mengidentifikasi miskonsepsi pada topik tersebut.

- 2) Melakukan wawancara tidak terstruktur. Mengadakan tanya-jawab tidak terstruktur secara langsung kepada siswa yang sudah mempelajari topik yang akan dianalisis untuk mendapatkan perspektif pemahaman siswa secara luas mengenai topik tersebut, sehingga terjadinya sebuah diskusi yang dapat membantu peneliti mengidentifikasi hal-hal yang mengalami miskonsepsi dan memberikan kontribusi dalam proses pengembangan tes pilihan ganda dengan alasan terbuka.
- 3) Mengembangkan soal pilihan ganda dengan alasan terbuka. Setiap soal yang dikembangkan berdasarkan topik yang telah dipelajari yang mana terdiri dari butir soal dari konsep yang terkait dan menyebutkan miskonsepsi-miskonsepsi yang telah didapat dari literatur/penelitian terdahulu dan hasil dari wawancara tidak terstruktur. Setiap soal pilihan ganda diikuti dengan kolom untuk alasan dengan bebas atas jawaban yang telah dipilihnya. Setelah menyusun butir soal-soal tersebut, diuji cobakan kepada suatu kelas. Alhasil, miskonsepsi yang didapatkan akan lebih jelas terlihat pada alasan yang diberikan oleh siswa.

c. Pengembangan tes diagnostik (*Developing a diagnostic test*)

Mengkonstruksi uji dengan melibatkan pengembangan tiga tingkatan, tingkat pertama memerlukan tanggapan isi dan tingkat kedua membutuhkan alasan untuk tanggapan tersebut.

- 1) Mengembangkan tes diagnostik tiga tingkat. Tes diagnostik tiga tingkat adalah soal pilihan ganda yang dikembangkan dengan desain yang sebanding dengan format '*Test of Logical Thinking*' yang terdiri dari *tier* pertama yaitu pilihan ganda jawaban sedangkan *tier* kedua merupakan pilihan ganda alasan yang kemungkinan alasan dari jawaban *tier* pertama. Alasan terdiri dari jawaban yang

- benar, miskonsepsi yang teridentifikasi, dan alasan yang salah jika dibutuhkan. Bagian dari setiap item dalam tes dikembangkan dari tanggapan siswa pada alasan yang diberikan serta informasi yang terkumpul dari hasil wawancara dan literatur penelitian terdahulu.
- 2) Membuat kisi-kisi tes. Pada tahapan ini bertujuan untuk memastikan bahwa pengujian tes diagnostik dapat mencakup konsep-konsep pengetahuan dan konsep-konsep dalam peta konsep yang mendasari topik.
 - 3) Perbaikan berkelanjutan. Penyempurnaan tiga tingkat berturut-turut pada item pilihan ganda dengan kelas yang berbeda memastikan tes secara keseluruhan dapat digunakan untuk mendiagnosis kesalahan/miskonsepsi siswa dalam topik tertentu. Perkembangan tes tersebut sampai saat ini telah menunjukkan bahwa masing-masing item dapat berhasil disempurnakan untuk meningkatkan sifat diagnostiknya untuk mengidentifikasi miskonsepsi.

C. Miskonsepsi

1. Pengertian Miskonsepsi

Miskonsepsi adalah suatu interpretasi konsep yang tidak bisa diterima secara teori dalam sebuah pernyataan. Suwanto menyatakan miskonsepsi siswa adalah refleksi pemikiran siswa atau kegagalan dalam menerapkan kurikulum.⁴⁵ Sedangkan menurut Modell, Michael dan Wonderoth menyatakan bahwa miskonsepsi merupakan pemahaman suatu konsep atau prinsip yang tidak konsisten dengan penafsiran atau pandangan yang berlaku umum tentang konsep tersebut.⁴⁶

Menurut Suparno, miskonsepsi sebagai pengertian yang tidak akurat akan konsep, penggunaan konsep yang salah, klasifikasi contoh-contoh yang salah, kekacauan konsep-konsep yang berbeda dan hubungan hierarkis konsep-konsep yang tidak benar.⁴⁷ Berdasarkan beberapa definisi di atas, maka dapat disimpulkan bahwa miskonsepsi merupakan suatu kesalahan dalam menentukan

⁴⁵ Suwanto. *Pengembangan Tes Diagnostik Dalam Pembelajaran*, (Yogyakarta: Pustaka Pelajar, 2013), hal. 76-77.

⁴⁶ Ibid.

⁴⁷ Paul Suparno, *Miskonsepsi & Perubahan Konsep dalam Pendidikan Fisika*, (Jakarta: PT. Grasindo, 2013), hal. 34-52.

klasifikasi antar bangun datar, menjelaskan hubungan antar bangun datar dan mendefinisikan konsep bangun datar.

2. Penyebab Miskonsepsi

Suwarto menyatakan bahwa miskonsepsi terjadi karena kesalahan yang dilakukan seseorang dalam membangun konsepsi berdasarkan informasi lingkungan fisik di sekitarnya atau teori yang telah diterima.⁴⁸ Miskonsepsi pada siswa terjadi ketika siswa mengikuti proses kegiatan belajar mengajar di kelas karena adanya kesalahan menerjemahkan konsep-konsep yang merupakan hal baru bagi siswa tersebut.⁴⁹ Dahar mengajukan beberapa hal, antara lain:⁵⁰

- 1) Miskonsepsi bersifat pribadi. Bila dalam suatu kelas anak-anak disuruh menulis tentang percobaan yang sama (mungkin hasil demonstrasi guru), mereka memberikan berbagai interpretasi. Setiap anak “melihat” dan menginterpretasikan eksperimen itu menurut caranya sendiri. Setiap anak mengontruksi kebermaknaannya sendiri.
- 2) Miskonsepsi memiliki sifat yang stabil. Kerap kali terlihat bahwa gagasan anak yang berbeda dengan gagasan ilmiah ini tetap dipertahankan anak, walaupun guru sudah berusaha memberikan suatu kenyataan yang berlawanan.
- 3) Bila menyangkut koherensi, anak tidak merasa butuh pandangan yang koheren sebab interpretasi dan prediksi tentang peristiwa-peristiwa alam praktis kelihatannya cukup memuaskan. Kebutuhan akan koherensi dan kriteria untuk koherensi menurut persepsi anak tidak sama dengan yang dipersepsi ilmuwan.

Bentuk miskonsepsi pada umumnya dapat berupa gagasan yang salah atau konsep awal yang kurang tepat. Berikut ini terdapat beberapa pernyataan yang berhubungan dengan miskonsepsi menurut penelitian yang relevan antara lain:⁵¹

- 1) Miskonsepsi siswa terjadi sebagai akibat perbedaan budaya, agama, dan bahasa.

⁴⁸ Suwarto, Op. Cit., hal. 78.

⁴⁹ Ibid.

⁵⁰ Ratna Wilis Dahar, *Teori-Teori Belajar & Pembelajaran*, (Bandung: Erlangga, 2011), hal.154.

⁵¹ Gusti Ayu Dewi Setiawati, dkk., “Identifikasi Miskonsepsi Dalam Materi Fotosintesis dan Respirasi Tumbuhan Pada Siswa Kelas IX SMP di Kota Denpasar”, *Jurnal Bakti Saraswati (BJS)*, 3:2. 2014, hal. 21.

- 2) Sebelum pembelajaran berlangsung miskonsepsi sudah terdapat dalam pikiran siswa dan sangat sulit untuk mengubahnya.
- 3) Berbagai miskonsepsi dapat terjadi saat menjelaskan suatu fenomena alam.
- 4) Miskonsepsi dapat terjadi setelah pembelajaran berlangsung.

Miskonsepsi bisa terjadi apabila konsepsi yang diterima oleh seseorang terhadap suatu materi berbeda dengan konsepsi yang diterima oleh para ahli dalam bidang tertentu. Menurut Suparno dalam Cahyani menyatakan bahwa faktor penyebab miskonsepsi adalah sebagai berikut:⁵²

Tabel 2.1
Penyebab Miskonsepsi

Sebab Utama	Sebab Khusus
Siswa	<ol style="list-style-type: none"> 1) Prakonsepsi 2) Pemikiran asosiatif (proses asimilasi, akomodasi, dan akulturasi) 3) Pemikiran humanistik (berbagai jalan pikiran yang berbeda) 4) Alasan yang tidak lengkap 5) Kemampuan siswa, minat belajar siswa 6) Pengalaman belajar siswa 7) Bahasa sehari-hari yang berbeda 8) Teman diskusi yang salah 9) Penjelasan orangtua atau orang lain yang salah 10) Konteks hidup siswa (tv, radio, dan film yang memberikan informasi yang salah) 11) Perasaan senang atau tidak (senang, bebas, atau tertekan)
Guru	<ol style="list-style-type: none"> 1) Tidak menguasai bahan ajar 2) Tidak membiarkan siswa mengungkapkan alasan atau ide 3) Komunikasi antara siswa dan guru yang tidak berjalan dengan baik 4) Metode mengajar hanya ceramah dan meminta

⁵² Fatmawati Nur Indah Cahyani, Skripsi: “*Analisis Miskonsepsi Siswa Materi Bangun Datar Segiempat Dibedakan dari Gaya Kognitif Siswa*”, (Surabaya: Universitas Islam Negeri Sunan Ampel Surabaya, 2018), hal. 12.

	<p>anak mencatat</p> <p>5) Memberikan materi langsung berupa rumus tanpa diawali dengan cara mendapatkannya</p> <p>6) Tidak mengungkapkan kemungkinan miskonsepsi yang dapat terjadi pada materi yang akan diajarkan</p> <p>7) Tidak mengkoreksi jawaban siswa yang salah</p>
Buku Teks	<p>1) Penjelasan yang salah</p> <p>2) Salah tulis, terutama dalam rumus dan notasi</p> <p>3) Tingkat penulisan buku yang terlalu tinggi, baik dari segi bahasa maupun materi</p> <p>4) Tidak tahu membaca buku teks</p>

3. Jenis Miskonsepsi pada Materi Bangun Datar Segiempat

Menurut L.S. Cox dalam Cahyani menyebutkan miskonsepsi ditinjau dari sifatnya dikelompokkan menjadi 4 yaitu: (1) miskonsepsi yang sistematis (*systematic error*), yaitu kesalahan yang terjadi jika siswa membuat kesalahan dengan pola yang sama pada sekurang-kurangnya tiga soal dari lima soal yang diberikan; (2) miskonsepsi yang random (*random error*) yaitu kesalahan yang terjadi jika siswa membuat kesalahan dengan pola yang berbeda pada sekurang-kurangnya tiga soal dari lima soal yang diberikan; (3) miskonsepsi yang diakibatkan dari kecerobohan adalah kesalahan yang terjadi jika siswa hanya membuat dua kesalahan dari lima soal yang diberikan; (4) miskonsepsi yang tidak dapat dimasukkan dalam salah satu tipe di atas misalnya lembar data yang tidak lengkap.⁵³

Menurut Salirawati memberikan penjelasan mengenai beberapa jenis miskonsepsi pada siswa antara lain:⁵⁴

- 1) Miskonsepsi klasifikasional, yaitu merupakan bentuk miskonsepsi yang didasarkan atas kesalahan klasifikasi fakta-fakta ke dalam bagan-bagan yang terorganisir. Sebagai contoh, mengelompokkan bangun datar segiempat dan bukan segiempat yang kurang tepat.
- 2) Miskonsepsi korelasional, yaitu merupakan bentuk miskonsepsi yang didasarkan atas kesalahan mengenai kejadian-kejadian khusus yang saling berhubungan, atau observasi-observasi yang terdiri atas dugaan-dugaan terutama berbentuk formulasi prinsip-prinsip umum.

⁵³ Fatmawati Nur Indah Cahyani, Op. Cit., hal. 14.

⁵⁴ Das Salirawati, Disertasi: “*Pengembangan Instrumen Pendeteksi Miskonsepsi Kimia pada Peserta Didik SMA*”, (Yogyakarta: PPs-UNY, 2011), hal. 36.

Sebagai contoh, merepresentasikan setiap soal ke dalam bentuk gambar sesuai perintah soal dengan tidak tepat.

- 3) Miskonsepsi teoritikal, yaitu merupakan bentuk miskonsepsi yang didasarkan atas kesalahan dalam mempelajari fakta-fakta atau kejadian-kejadian dalam sistem yang terorganisir. Sebagai contoh, mendefinisikan jajargenjang yang tidak sesuai dengan pengertiannya.

Adapun jenis dan indikator miskonsepsi yang diadaptasi dari jurnal Ainiyah yang mana disajikan dalam bentuk Tabel 2.2 yakni:⁵⁵

Tabel 2.2
Jenis dan Indikator Miskonsepsi

Jenis Miskonsepsi	Deskripsi	Indikator Miskonsepsi	Indikator Materi Segiempat
Miskonsepsi klasifikasi	Bentuk miskonsepsi yang didasarkan atas kesalahan klasifikasi fakta-fakta ke dalam bagan-bagan yang terorganisir	Siswa melakukan kesalahan dalam menentukan klasifikasi antar bangun datar	1.1 Siswa melakukan kesalahan dalam menentukan persegi termasuk belah ketupat 1.2 Siswa melakukan kesalahan dalam menentukan persegi termasuk persegi panjang 1.3 Siswa melakukan kesalahan dalam menentukan persegi panjang termasuk jajargenjang 1.4 Siswa melakukan kesalahan dalam menentukan belah ketupat termasuk layang-layang
Miskonsepsi korelasional	Bentuk miskonsepsi	Siswa tidak dapat	2.1 Siswa tidak dapat mempresentasikan

⁵⁵ Lutfia Afifatul Ainiyah, "Identifikasi Miskonsepsi Siswa Dalam Materi Geometri Pada Pembelajaran Matematika Siswa Kelas VIII SMP Negeri 1 Punggalen", diakses dari journal.student.uny.ac.id, pada tanggal 14 Desember 2019.

Jenis Miskonsepsi	Deskripsi	Indikator Miskonsepsi	Indikator Materi Segiempat
	yang didasarkan atas kesalahan mengenai kejadian-kejadian khusus yang saling berhubungan, atau observasi-observasi yang terdiri atas dugaan-dugaan terutama berbentuk formulasi prinsip-prinsip umum	menjelaskan hubungan antar bangun datar	soal ke dalam bentuk jawaban 2.2 Siswa tidak dapat menjelaskan hubungan antara belah ketupat dan persegi 2.3 Siswa tidak dapat menjelaskan hubungan antara persegi dan persegi panjang 2.4 Siswa tidak dapat menjelaskan hubungan antara persegi panjang dan jajargenjang 2.5 Siswa tidak dapat menjelaskan hubungan antara belah ketupat dan layang-layang 2.6 Siswa melakukan kesalahan dalam menerapkan hubungan antara rumus yang digunakan dengan permasalahan yang terdapat dalam soal
Miskonsepsi teoritikal	Bentuk miskonsepsi yang didasarkan atas kesalahan dalam	Siswa tidak mampu mendefinisikan konsep bangun datar	3.1 Siswa melakukan kesalahan dalam menyelesaikan soal yang berhubungan dengan segiempat 3.2 Siswa kurang tepat dalam

Jenis Miskonsepsi	Deskripsi	Indikator Miskonsepsi	Indikator Materi Segiempat
	mempelajari fakta-fakta atau kejadian-kejadian dalam sistem yang teroganisir		mendefinisikan persegi 3.3 Siswa kurang tepat dalam mendefinisikan persegi panjang 3.4 Siswa kurang tepat dalam mendefinisikan belah ketupat 3.5 Siswa kurang tepat dalam mendefinisikan layang-layang 3.6 Siswa kurang tepat dalam mendefinisikan jajargenjang 3.7 Siswa melakukan kesalahan dalam menggunakan rumus luas dan keliling persegi dan persegi panjang

D. Tinjauan Materi Bangun Datar Segiempat

1. Definisi Segiempat

Segiempat merupakan suatu bidang datar yang dibentuk oleh empat garis lurus sebagai sisinya dan empat sudut. Segiempat adalah salah satu bentuk dasar dalam geometri yang paling terkenal. Dalam trigonometri, setiap sudut dalam bangun segiempat diberi nama berdasarkan nama titik-titik sudutnya. Misal sebuah segiempat memiliki sudut A, B, C, dan D maka dinamakan segiempat ABCD.

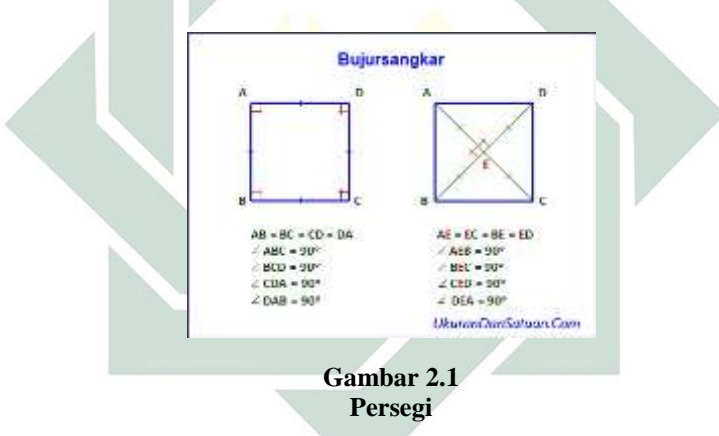
2. Macam-Macam Segiempat

Adapun bangun datar segiempat dibedakan berdasarkan keteraturan sifat sisi, sudut, dan diagonal antara lain:

1) Persegi atau bujur sangkar

Persegi atau yang biasa disebut bujur sangkar adalah suatu segiempat dengan semua sisinya sama panjang dan semua sudutnya sama besar dan siku-siku (sudut 90°). Sifat-sifat persegi yaitu:

- Memiliki empat sisi yang berhadapan sama panjang.
- Mempunyai empat buah sudut siku-siku (sudut 90°).
- Mempunyai dua pasang sisi sejajar dan sama panjang.
- Mempunyai dua diagonal bidang yang sama panjang.
- Setiap sudutnya dibagi dua sama ukuran oleh diagonal-diagonalnya.
- Diagonal-diagonalnya berpotongan saling tegak lurus.
- Rumus keliling persegi adalah $K = 4s$ sedangkan rumus luas persegi adalah $L = s \times s = s^2$ dengan s adalah panjang sisi persegi



Gambar 2.1
Persegi

2) Persegi panjang

Persegi panjang merupakan suatu segiempat dimana sisi-sisi yang berhadapan sejajar dan sama panjang serta semua sudutnya membentuk sudut siku-siku. Persegi panjang memiliki sifat sebagai berikut:

- Memiliki sisi yang berhadapan sejajar dan sama panjang serta setiap sudutnya siku-siku (sudut 90°).
- Memiliki dua diagonal bidang yang sama panjang dan saling berpotongan sehingga membagi diagonal-diagonal tersebut menjadi dua bagian yang sama panjang.

- c) Rumus keliling persegi panjang adalah $K = p + p + l + l = 2p + 2l = 2(p + l)$ sedangkan luas persegi panjang adalah $L = p \times l$.

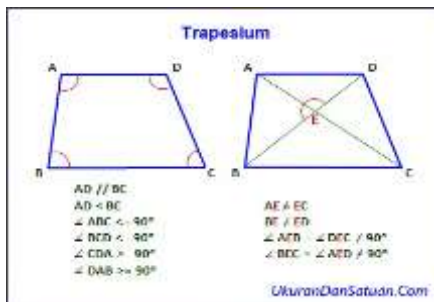


Gambar 2.2
Persegi Panjang

3) Trapesium

Trapesium adalah segiempat yang mempunyai tepat sepasang sisi yang berhadapan sejajar. Adapun sifat-sifat trapesium yaitu:

- Memiliki sepasang sisi sejajar yang tidak sama panjang dan jumlah sudut berdekatan adalah 180° .
- Trapesium siku-siku, salah satu kakinya tegak lurus terhadap sisi sejajarnya.
- Kedua diagonalnya saling berpotongan sehingga membagi dua diagonal-diagonal tersebut menjadi dua bagian tidak sama panjang.
- Rumus keliling trapesium ABCD adalah $K = \overline{AD} + \overline{CD} + \overline{BC} + \overline{AB}$ sedangkan untuk menghitung luas trapesium adalah $L = \frac{1}{2} \times (a + b) \times t$ dimana t adalah tinggi trapesium.



Gambar 2.3
Trapezium

- 4) Belah ketupat
 Belah ketupat adalah segiempat yang semua sisinya sama panjang dan sudut yang berhadapan sama besar. Adapun sifat belah ketupat yaitu:
- Semua sisinya kongruen dan sisi yang berhadapan sejajar.
 - Kedua diagonalnya tidak sama panjang melainkan berpotongan saling tegak lurus sehingga membagi diagonal-diagonal tersebut menjadi dua bagian yang sama panjang.
 - Jumlah ukuran dua sudut yang berdekatan adalah 180° .
 - Keliling belah ketupat adalah jumlah keempat sisinya dan luas belah ketupat adalah $L = \frac{1}{2} \times \overline{AC} \times \overline{BD} = \frac{1}{2} \times d_1 \times d_2$ dimana d_1 adalah diagonal satu dan d_2 adalah diagonal dua.

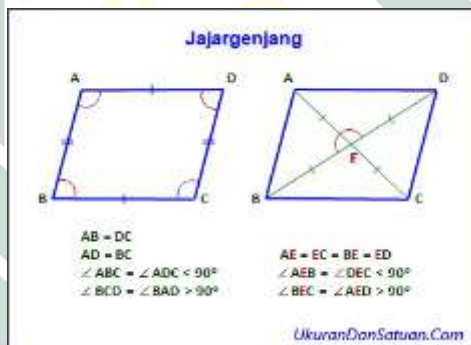


Gambar 2.4
Belah Ketupat

5) Jajargenjang

Jajargenjang adalah segiempat dengan 2 pasang sisi yang berhadapan sama panjang dan sejajar. Sifat-sifat jajargenjang antara lain:

- Mempunyai sisi-sisi yang berhadapan sejajar dan sama panjang serta mempunyai sudut-sudut yang berhadapan sama besar dan bukan sudut siku-siku.
- Jumlah sudut yang berdekatan adalah 180° .
- Kedua diagonalnya tidak sama panjang dan saling berpotongan membagi diagonal-diagonal tersebut menjadi dua bagian yang sama panjang.
- Keliling jajargenjang adalah jumlah panjang keempat sisinya sedangkan luas jajargenjang adalah $L = a \times t$ dengan a sebagai alas dan t sebagai tinggi.



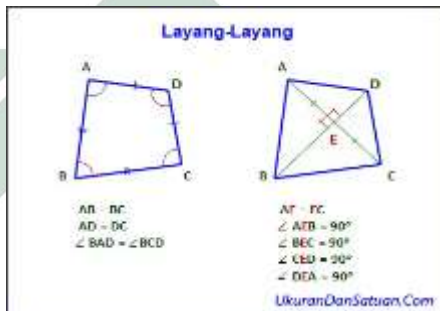
Gambar 2.5
Jajargenjang

6) Layang-layang

Layang-layang adalah segiempat yang salah satu diagonalnya membagi diagonal lain menjadi dua bagian yang sama panjang dan diagonal-diagonalnya saling tegak lurus. Adapun sifat dari layang-layang sebagai berikut:

- Memiliki dua pasang sisi yang berdampingan sama panjang dan sepasang sudut yang berhadapan sama besar.

- b) Kedua diagonalnya saling berpotongan tegak lurus dan membagi salah satu diagonal menjadi dua bagian yang sama panjang.
- c) Keliling layang-layang adalah jumlah keempat sisinya dan luas layang-layang adalah $L = \frac{1}{2} \times \overline{AC} \times \overline{BD} = \frac{1}{2} \times d_1 \times d_2$ dimana d_1 adalah diagonal satu dan d_2 adalah diagonal dua.



Gambar 2.6
Layang-Layang

- 7) Segiempat sembarang
Segiempat sembarang adalah segiempat yang tidak memiliki keteraturan khusus atau segiempat yang tidak termasuk dari salah satu segiempat istimewa yang dijelaskan di atas. Contoh untuk segiempat sembarang adalah sebagai berikut.



Gambar 2.7
Segiempat Sembarang

E. CRI (*Certainty of Response Index*)

Identifikasi miskonsepsi diartikan sebagai suatu cara yang dilakukan untuk mendeteksi belajar siswa yang diperkirakan mengalami kesalahan pemahaman konsep, dalam hal ini konsepsi siswa berbeda dengan para ahli.⁵⁶ Identifikasi yang dilakukan dalam penelitian ini dengan menggunakan tes diagnostik tertulis dengan metode *Certainty of Response Index* (CRI).⁵⁷ *Certainty of Response Index* (CRI) yang diperkenalkan oleh Saleem Hasan, Diola Bagayoko, dan Ella L. Kelley dalam jurnal mereka yang berjudul “*Misconceptions and The Certainty of Response Index (CRI)*” adalah sebuah cara untuk mengukur tingkat keyakinan atau kepastian responden dalam menjawab setiap pertanyaan (soal) yang diberikan.⁵⁸

CRI sering digunakan dalam ilmu sosial terutama dalam survei dimana responden diminta untuk memberikan tingkat kepastian yang dimilikinya dalam kemampuannya sendiri untuk memilih dan memanfaatkan pengetahuan, konsep, atau menjawab pertanyaan.⁵⁹

Certainty of Response Index (CRI) merupakan salah satu instrumen yang dikembangkan untuk mengidentifikasi miskonsepsi sekaligus mengklasifikasikan siswa dalam 3 kelompok yaitu: (1) siswa yang memahami konsep; (2) siswa yang mengalami miskonsepsi; (3) siswa yang tidak tahu konsep.⁶⁰ Berdasarkan penjelasan di atas, maka pengertian *Certainty of Response Index* (CRI) adalah sebuah cara untuk mengukur kesalahan seseorang dengan cara mengukur tingkat keyakinan atau kepastian dalam menjawab setiap pertanyaan yang diberikan. CRI biasanya didasarkan pada suatu skala, sebagai contoh skala enam (0 – 5) seperti pada Tabel 2.3 berikut.⁶¹

⁵⁶ Mega Teguh Budiarto, Rachmania Widya Ningrum, “Miskonsepsi Siswa SMP pada Materi Bangun Data Segiempat dan Alternatif Mengatasinya”, *Jurnal Ilmiah Pendidikan Matematika*, 1:5, 2016, hal. 61.

⁵⁷ Ibid.

⁵⁸ Yuyu R. Tayubi, “Identifikasi Miskonsepsi Pada Konsep-Konsep Fisika Menggunakan *Certainty of Response Index* (CRI)”. *Jurnal Mimbar Pendidikan*, 3:24, 2005, hal. 5.

⁵⁹ Saleem Hasan, et.al., “*Misconceptions and the Certainty of Response Index (CRI)*”. *Phys. Educ.*, 34:5, 1999, hal. 294.

⁶⁰ Muhammad Toni, dkk., “Analisis Kesalahan Siswa Menggunakan *Certainty of Response Index* (CRI) Termodifikasi pada Materi Pecahan”. *Jurnal pendidikan dan Pembelajaran*, 6:4, 2017, hal.3.

⁶¹ Tri Ade Mustaqim, dkk., “Identifikasi Miskonsepsi Siswa Dengan Menggunakan Metode (*Certainty of Response Index*) CRI Pada Konsep Fotosintesis dan Respirasi Tumbuhan”, *EDUSAINS*, 6:2, 2014, hal. 147.

Tabel 2.3
Skala CRI dan Kriteria

CRI	Kriteria	Kode
0	<i>Totally Guessed Answer</i> (Benar-Benar Tidak Tahu)	BBT
1	<i>Almost Guess</i> (Agak Tahu)	AT
2	<i>Not Sure</i> (Tidak Yakin)	TY
3	<i>Sure</i> (Yakin)	Y
4	<i>Almost Certain</i> (Agak Sangat Yakin)	ASY
5	<i>Certain</i> (Sangat Yakin)	SY

Sumber: Mustaqim, Zulfiani & Herlanti, 2014

Berdasarkan Tabel 2.3, angka 0 berarti tidak tahu konsep sama sekali dalam menjawab suatu pertanyaan (jawaban ditebak secara total), sementara angka 5 berarti kepercayaan diri yang penuh atas kebenaran akan konsep dalam menjawab suatu pertanyaan (soal) atau tidak ada unsur tebakan sama sekali.

Jika nilai 0 sampai 2 maka menunjukkan responden dalam menjawab dengan menebak, tanpa melihat apakah jawaban tersebut benar atau salah yang secara tidak langsung memperlihatkan ketidaktahuan responden akan konsep dalam menjawab suatu pertanyaan. Jika nilai 3 sampai 5 maka menunjukkan adanya tingkat kepercayaan diri pada responden dalam menjawab pertanyaan. Dalam hal ini apabila jawaban responden benar, ini dapat membuktikan bahwa tingkat keyakinan yang tinggi terhadap kebenaran konsepsi dapat teruji dengan baik. Akan tetapi, apabila jawaban responden salah maka menunjukkan terdapat suatu kekeliruan konsep yang dimilikinya dan dapat menjadi salah satu indikator terjadinya miskonsepsi.

Untuk membedakan antara tahu konsep, tidak tahu konsep dan miskonsepsi untuk responden baik secara individu maupun kelompok dapat dilihat pada Tabel 2.4.⁶²

⁶² Saleem Hasan, et.al., "Misconceptions and the Certainty of Response Index (CRI)", *Physics Education*, 34:5, 1999, hal. 296.

Tabel 2.4
Kriteria Jawaban CRI Kategori Rendah dan Tinggi

Kriteria Jawaban	CRI Rendah (< 2,5)	CRI Tinggi (> 2,5)
Jawaban benar	Jawaban benar tapi CRI rendah berarti tidak tahu konsep (kategori <i>lucky guess</i>)	Jawaban benar dan CRI tinggi berarti menguasai konsep dengan baik
Jawaban salah	Jawaban salah dan CRI rendah berarti tidak tahu konsep (kategori <i>lack of knowledge</i>)	Jawaban salah tapi CRI tinggi berarti terjadi miskonsepsi

Sumber: Saleem Hasan 1999

Pada Tabel 2.4 menunjukkan empat kemungkinan kombinasi dari sebuah jawaban yakni benar atau salah dan CRI dengan kategori tinggi atau rendah untuk tiap responden secara individu. Untuk setiap responden dan untuk suatu pertanyaan yang diberikan, jawaban benar dengan CRI berskala rendah menunjukkan tidak tahu konsep, dan jawaban benar dengan CRI berskala tinggi menunjukkan penguasaan konsep yang baik. Sedangkan jawaban salah dengan CRI rendah menunjukkan tidak tahu konsep dan jawaban salah dengan CRI berskala tinggi menunjukkan terjadinya miskonsepsi.

Penentuan kategori tingkat pemahaman siswa berdasarkan CRI dan alasan siswa terhadap pilihan jawaban didasarkan pada kategori tingkat pemahaman yang dimodifikasi oleh Hakim.⁶³ Adapun modifikasi kategori tingkat pemahaman siswa dapat dilihat pada Tabel 2.5.⁶⁴

⁶³ Aliefman Hakim., dkk., "Student Concept Understanding of Natural Product Chemistry in Primary and Secondary Metabolites Using the Data Collecting Technique of Modified CRI", *International Online Journal of Education Sciences*, 4:3, 2012, hal. 544-553.

⁶⁴ Tri Ade Mustaqim, dkk., "Identifikasi Miskonsepsi Siswa Dengan Menggunakan Metode (*Certainty of Response Index*) CRI Pada Konsep Fotosintesis dan Respirasi Tumbuhan", *EDUSAINS*, 6:2, 2014, hal. 148.

Tabel 2.5
Modifikasi Kategori Tingkatan Pemahaman Siswa

Jawaban	Alasan	Nilai CRI	Deskripsi	Kode
Benar	Benar	> 2,5	Memahami konsep dengan baik	PK
Benar	Benar	< 2,5	Memahami konsep tetapi kurang yakin	PKKY
Benar	Salah	> 2,5	Miskonsepsi	M
Benar	Salah	< 2,5	Tidak tahu konsep	TTK
Salah	Benar	> 2,5	Miskonsepsi	M
Salah	Benar	< 2,5	Tidak tahu konsep	TTK
Salah	Salah	> 2,5	Miskonsepsi	M
Salah	Salah	< 2,5	Tidak tahu konsep	TTK

Sumber: Mustaqim, Zulfiani & Herlanti, 2014

Dari Tabel 2.5 berikut menunjukkan kategori tingkat pemahaman siswa terbagi menjadi empat yaitu Paham Konsep (PK), Paham Konsep tetapi Kurang Yakin (PKKY), Tidak Tahu Konsep (TTK), dan Miskonsepsi (M). Tes diagnostik dengan menggunakan metode CRI digunakan untuk mengetahui miskonsepsi pada siswa namun tidak bisa mengungkapkan penyebab terjadinya miskonsepsi dan proses penalaran siswa.

F. Validitas

Validitas berasal dari kata *validity* yang mempunyai arti sejauh mana ketepatan dan kecermatan suatu alat ukur dalam melakukan fungsi ukurnya.⁶⁵ Istilah “validitas” dengan “valid” adalah sangat berbeda. “Validitas” merupakan sebuah kata benda, sedangkan “valid” merupakan kata sifat.⁶⁶ Suatu tes dikatakan memiliki validitas yang

⁶⁵ Widiastuti, *Tes dan Pengukuran Olahraga*, (Jakarta: Rajawali Pers.), 2015, hal.8.

⁶⁶ Arikunto, *Dasar-Dasar Evaluasi Pendidikan (Cetakan kesepuluh)*, (Jakarta: Bumi Angkasa), 1993, hal. 56-57.

tinggi apabila alat tersebut menjalankan fungsi ukur secara tepat atau memberikan hasil ukur yang sesuai dengan maksud dilakukannya pengukuran tersebut.⁶⁷ Sejalan dengan pendapat Suryabrata, validitas tes pada dasarnya menunjuk pada derajat fungsi pengukuran suatu tes atau derajat kecermatan ukurnya suatu tes.⁶⁸ Validitas berkenaan dengan ketepatan alat penilaian terhadap apa yang seharusnya dinilai sehingga benar-benar menilai apa yang akan dinilai.

Tes memiliki validitas yang tinggi jika hasilnya sesuai dengan kriteria, dalam arti memiliki kesejajaran antara tes dan kriteria.⁶⁹ Sebuah tes dikatakan valid apabila tes tersebut digunakan untuk mengukur apa yang seharusnya diukur. Validitas tidak berlaku secara umum dikarenakan bergantung pada tujuan penelitian dan situasi. Berdasarkan beberapa penjelasan di atas maka pengertian validitas adalah suatu ukuran yang menunjukkan tingkat kevalidan atau kesahihan suatu tes.

Menurut Arikunto, ada 4 macam validitas yang berasal dari dasar pembagian jenis di atas, yaitu:⁷⁰

a. Validitas isi (*content validity*)

Validitas isi merupakan validitas yang diestimasi lewat pengujian terhadap isi tes dengan analisis rasional atau lewat *professional judgment*. Validitas isi ini harus memuat isi yang relevan dan tidak keluar dari batasan tujuan ukur. Menurut Purwanto, validitas isi (*content validity*) adalah pengujian validitas dilakukan atas isinya untuk memastikan apakah butir tes hasil belajar mengukur secara tepat keadaan yang ingin diukur.⁷¹ Sedangkan menurut Arikunto, tes memiliki validitas isi apabila mengukur tujuan khusus tertentu yang sejajar dengan materi atau isi pelajaran yang diinginkan.⁷² Sebuah tes dikatakan memiliki validitas isi apabila mengukur tujuan khusus tertentu yang berkaitan dengan materi. Dari penjelasan di atas, dapat diartikan validitas isi adalah sebuah tes yang mengukur secara tepat antara tujuan khusus dengan materi.

⁶⁷ Zulkifli Matondang, "Validitas dan Reliabilitas Suatu Instrumen Penelitian", *Jurnal Tabularasa PPS UNIMED*, 6:1, 2009, hal. 89.

⁶⁸ Sumadi Suryabrata, *Pengembangan Alat Ukur Psikologis*, (Yogyakarta: Andi, 2000), hal. 41.

⁶⁹ Arikunto, *Dasar-Dasar Evaluasi pendidikan*, (Jakarta: Bumi Angkasa), 1999, hal.65.

⁷⁰ Ibid.

⁷¹ Elina Aenny Zahra, *Analisis Soal Ulangan Akhir Semester I Kelas X SMAN Banyumas Mata Pelajaran Matematika Tahun Pelajaran 2011/2012*, Diakses dari http://www.eprints.walisongo.ac.id/953/6/083511033_Coverdll.pdf pada 24 Juni 2019.

⁷² Arikunto. *Penelitian Tindakan*. (Yogyakarta: Aditya Media, 2010), hal.67.

b. Validitas konstruksi (*construct validity*)

Secara etimologis, kata “konstruk” mengandung arti susunan, kerangka atau rekaan. Validitas konstruk (*construct validity*) berkaitan dengan konstruksi atau konsep bidang ilmu yang akan diuji validitas alat ukurnya. Validitas konstruk berkenaan dengan kesanggupan alat ukur untuk mengukur konsep kemampuan terhadap materi yang akan diukur sehingga perlunya penjabaran dari setiap konsep dikembangkan menjadi beberapa indikator. Dengan adanya indikator pada setiap konsep maka dapat memudahkan dalam menetapkan cara pengukuran.

c. Validitas “ada sekarang” (*concurrent validity*)

Validitas ini dikenal dengan sebutan validitas bandingan, dikarenakan perlunya kriterium atau alat banding luar berupa alat ukur lain yang serupa dalam menetapkan tingkat kevalidannya. Apabila hasil dari alat ukur yang telah dibakukan memiliki tingkat kesesuaian dengan hasil dari alat ukur lain yang serupa maka tes tersebut terbukti memenuhi *concurrent validity*.

d. Validitas prediksi (*predictive validity*)

Memprediksi artinya meramal, dan meramal selalu mengenai hal yang akan datang. Sebuah tes dikatakan memiliki validitas prediksi atau validitas ramalan apabila mempunyai kemampuan untuk meramalkan apa yang akan terjadi pada masa yang akan datang. Dalam validitas ini lebih mengutamakan kriteria dari pada isi tes, apakah alat ukur tersebut dapat digunakan untuk meramalkan perilaku tertentu atau kriteria tertentu yang diharapkan.

Untuk menghitung koefisien korelasi antara skor butir dengan skor total instrumen digunakan rumus statistika yang sesuai dengan jenis skor butir dari instrumen tersebut.⁷³ Jika skor butir kontinum maka untuk menghitung koefisien korelasi antara skor butir dengan skor total instrumen digunakan koefisien korelasi *product moment*.⁷⁴

$$r_{xy} = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{\{N \sum X^2 - (\sum X)^2\} \{N \sum Y^2 - (\sum Y)^2\}}}$$

Keterangan:

r_{xy} = Koefisien korelasi antara variabel X dan variabel Y

$\sum X$ = Jumlah rerata nilai X

$\sum Y$ = Jumlah rerata nilai Y

⁷³ Ali Hamzah, *Evaluasi Pembelajaran Matematika*, (Jakarta: PT. Rajagrafindo Persada, 2014), hal. 220.

⁷⁴ Ibid.

N = Banyaknya responden

Namun jika skor butir dikotomi (0,1) maka untuk menghitung koefisien korelasi antara skor butir dengan skor total instrumen digunakan koefisien korelasi poin biserial (r_{pbis}) yang menggunakan rumus sebagai berikut.⁷⁵

$$r_{pbis} = \frac{X_i - X_t}{S_t} \sqrt{\frac{p_i}{q_i}}$$

Keterangan:

r_{pbis} = Koefisien korelasi poin biserial antara skor butir soal nomor dengan skor total

X_i = Rata-rata skor total responden yang menjawab benar butir soal nomor i ,

X_t = Rata-rata skor total semua responden

S_t = Standar deviasi skor total semua responden

p_i = Proporsi jawaban yang benar untuk butir soal nomor i

q_i = Proporsi jawaban yang salah untuk butir soal nomor i

Untuk mengetahui valid atau tidaknya butir soal, maka hasil perhitungan r_{xy} dapat dikorelasikan dengan r_{tabel} .⁷⁶ Apabila $r_{xy} > r_{tabel}$, maka butir soal dikatakan valid, sebaliknya jika $r_{xy} < r_{tabel}$, maka butir soal dikatakan tidak valid.⁷⁷ Sedangkan untuk menginterpretasikan tingkat validitas, maka koefisien korelasi dikategorikan pada Tabel 2.6.⁷⁸

Tabel 2.6
Kriteria Validitas Instrumen Tes

Nilai r	Interpretasi
0,81 – 1,00	Sangat tinggi
0,61 – 0,80	Tinggi
0,41 – 0,60	Cukup
0,21 – 0,40	Rendah
0,00 – 0,20	Sangat rendah

Sumber: Suharsimi Arikunto, 1991

⁷⁵ Ibid, hal. 221.

⁷⁶ Ibid, hal. 222.

⁷⁷ Ibid.

⁷⁸ Arikunto, *Metodelogi Penelitian*, (Jakarta: Rineka Cipta, 1991).

G. Reliabilitas

Reliabilitas adalah serangkaian pengukuran atau serangkaian alat ukur yang memiliki konsistensi bila pengukuran yang dilakukan dengan alat ukur itu dilakukan secara berulang.⁷⁹ Menurut Anastasia dan Urbina, reliabilitas adalah sesuatu yang merujuk pada konsistensi skor yang dicapai oleh orang yang sama ketika mereka diuji ulang dengan tes yang sama pada kesempatan yang berbeda, atau dengan seperangkat butir-butir ekuivalen (*equivalent items*) yang berbeda, atau di bawah kondisi pengujian yang berbeda.⁸⁰ Azwar menyatakan bahwa reliabilitas merupakan salah-satu ciri atau karakter utama instrumen pengukuran yang baik.⁸¹ Dari beberapa pendapat para pakar, maka uji reliabilitas adalah sebuah tes yang apabila diujikan berkali-kali mendapatkan hasil yang relatif tetap.

Menurut Arikunto, ada beberapa metode untuk mengetahui reliabilitas antara lain.⁸²

a. Metode bentuk paralel (*equivalent*)

Tes paralel atau tes ekuivalen adalah dua buah tes yang mempunyai kesamaan tujuan, tingkat kesukaran dan susunan, tetapi butir-butir soalnya berbeda. Dalam istilah bahasa Inggris disebut *alternative-forms method (parallel forms)*. Dengan metode bentuk paralel ini, dua buah tes yang paralel, misalnya tes Matematika Seri A yang akan dicari reliabilitasnya dan tes Seri B diteskan kepada sekelompok siswa yang sama, kemudian hasilnya dikorelasikan.

Koefisien korelasi dari kedua hasil tes inilah yang menunjukkan koefisien reliabilitas tes Seri A. jika koefisiennya tinggi maka tes tersebut sudah reliabel dan dapat digunakan sebagai alat pengetes yang terandalkan. Kelemahan dari metode ini adalah bahwa pengetes pekerjaannya berat karena harus menyusun dua seri tes. Lagi pula harus tersedia waktu yang lama untuk mencobakan dua kali tes.

b. Metode tes ulang (*test-retest method*)

Metode tes ini dilakukan orang untuk menghindari penyusunan dua seri tes. Dalam menggunakan tes atau metode ini pengetes hanya memiliki satu seri tes tetapi dicobakan dua kali. Oleh karena tesnya

⁷⁹ Sugiyono, *Metode Penelitian Bisnis*, (Bandung: Alfabeta, 2008).

⁸⁰ Anne Anastasia, Susana Urbina, *Psychological Testing*, (New Jersey: Prentice-Hall Inc, 1997).

⁸¹ Saifuddin Azwar, *Sikap Manusia Teori dan Pengukurannya*, (Yogyakarta: Pustaka Pelajar, 2003), hal. 176.

⁸² Arikunto, *Dasar-Dasar Evaluasi pendidikan*, (Jakarta: Bumi Angkasa, 1999), hal.85.

hanya satu dan dicobakan dua kali, maka metode ini dapat disebut dengan *single-test-double-trial-method*. Kemudian hasil dari kedua kali tes tersebut dihitung korelasinya.

Untuk tes yang banyak mengungkap pengetahuan (ingatan) dan pemahaman, cara ini kurang mengena karena tercoba akan masih ingat akan butir-butir soalnya. Oleh karena itu tenggang waktu antara pemberian tes pertama dengan kedua menjadi permasalahan tersendiri. Jika tenggang waktu terlalu sempit, siswa masih banyak ingat materi. Sebaliknya kalau tenggang waktu terlalu lama, maka faktor-faktor atau kondisi tes sudah akan berbeda, dan siswa sendiri barangkali sudah mempelajari sesuatu. Tentu saja faktor-faktor ini akan berpengaruh pula terhadap reliabilitas.

Pada umumnya hasil tes yang kedua cenderung lebih baik dari pada hasil tes pertama. Hal ini tidak mengapa karena pengetes harus sadar akan adanya *practice effect* dan *carry-over effect*. Yang penting adalah adanya kesejajaran hasil atau ketetapan hasil yang ditunjukkan oleh koefisien korelasi yang tinggi. Walaupun nampaknya skornya naik, akan tetapi kenaikannya dialami oleh semua siswa. Metode ini juga disebut *self-correlation method* (korelasi diri sendiri) karena mengkorelasikan hasil dari tes yang sama.

c. Metode belah dua atau *split-half method*

Kelemahan menggunakan metode dua-tes dua kali percobaan dan satu-tes dua kali percobaan di atasi dengan metode ketiga ini yaitu metode belah dua. Dalam metode ini pengetes hanya menggunakan sebuah tes dan dicobakan satu kali. Oleh karena itu disebut juga dengan *single-test-single-trial method*.

Berbeda dengan metode pertama dan kedua yang setelah ditemukan koefisien langsung diartikan itulah koefisien reliabilitas, maka dengan metode ketiga ini tidak dapat demikian. Pada waktu membelah dua dan mengkorelasikan dua belahan, baru diketahui separuh tes. Untuk mengetahui reliabilitas seluruh tes harus menggunakan rumus Sperman-Bown. Banyak pemakai metode ini salah membelah hasil tes pada waktu menganalisis. Yang mereka lakukan adalah mengelompokkan hasil separuh subyek peserta tes dan separuh yang lain kemudian hasil kedua kelompok ini dikorelasikan. Yang benar adalah membelah item atau butir soal. Tidak akan keliru kiranya bagi pemakai metode ini harus ingat bahwa banyaknya butir soal harus genap agar dapat dibelah.

Ada dua cara membelah butir soal ini yaitu: (1) membelah atas item-item genap dan item-item ganjil yang selanjutnya disebut belahan

ganjil-genap dan (2) membelah atas item-item awal dan item-item akhir yaitu separuh jumlah pada nomor-nomor awal dan separuh pada nomor-nomor akhir yang selanjutnya disebut belahan awal-akhir.

Ada beberapa prosedur untuk menghitung indeks reliabilitas tes, di antaranya melalui pendekatan tes ulang (*test-retest*), pendekatan paralel, dan pendekatan konsistensi internal. Di antara pendekatan konsistensi internal adalah metode *Kuder-Richardson 20* (KR.20) dan *Alpha Cronbach*.⁸³ Menurut Nitko, *Kuder-Richardson 20* (KR.20) digunakan untuk menghitung nilai reliabilitas tes dalam bentuk tes objektif yang hanya menggunakan skor dikotomi, yaitu bila benar = 1 dan salah = 0 seperti pada tes pilihan ganda, sedangkan koefisien *Alpha Cronbach* digunakan untuk menghitung nilai reliabilitas tes dalam bentuk uraian atau skala.⁸⁴ Reliabilitas tes dapat ditentukan dengan menggunakan persamaan *Kuder-Richardson* (K-R.20), rumusnya sebagai berikut.⁸⁵

$$r_{11} = \left(\frac{n}{n-1} \right) \left(\frac{S^2 - \sum pq}{S^2} \right)$$

Keterangan:

r_{11} = Reliabilitas suatu tes

p = Proporsi subjek yang menjawab soal dengan benar

q = Proporsi subjek yang menjawab soal dengan salah

n = Banyaknya soal

S = Standar deviasi

Adapun tolak ukur untuk menginterpretasikan derajat reliabilitas instrumen yang diperoleh sesuai dengan Tabel 2.7.⁸⁶

⁸³ Idrus Alwi, "Kriteria Empirik Dalam Menentukan Ukuran Sampel Pada Pengujian Hipotesis Statistika dan Analisis Butir", *Jurnal Formatif*, 2:2, 2019, hal. 144.

⁸⁴ Anthoni J. Nitko, *Educational Test and Measurement An Introduction*, (New York: Harcourt Brace Jovanovich, 1983), hal. 395.

⁸⁵ Arikunto, *Dasar-Dasar Evaluasi Pendidikan Edisi 2*, (Jakarta: Bumi Aksara, 2016), hal. 115.

⁸⁶ Arikunto, *Prosedur Penelitian Suatu Praktek*, (Jakarta: Bima Aksara, 2003).

Tabel 2.7
Interpretasi Reliabilitas

Koefisien Korelasi (r)	Kriteria Reliabilitas
$0,81 < r \leq 1,00$	Sangat tinggi
$0,61 < r \leq 0,80$	Tinggi
$0,41 < r \leq 0,60$	Cukup
$0,21 < r \leq 0,40$	Rendah
$0,00 < r \leq 0,20$	Sangat rendah

Sumber: Suharsimi Arikunto, 2003

H. Kualitas Butir Soal

1. Daya Pembeda Soal

Menurut Arikunto, daya pembeda soal adalah kemampuan suatu soal untuk membedakan antara siswa yang berkemampuan tinggi dengan siswa yang berkemampuan rendah.⁸⁷ Daya pembeda butir soal adalah kemampuan suatu butir soal untuk membedakan kelompok dalam aspek yang diukur sesuai dengan perbedaan yang ada dalam kelompok itu.⁸⁸ Sejalan dengan pendapat Sudijono yang menyatakan daya pembeda item merupakan suatu butir item tes hasil belajar untuk dapat membedakan antara *testee* (responden) yang berkemampuan tinggi dengan *testee* (responden) yang berkemampuan rendah.⁸⁹ Daya pembeda soal yang merupakan suatu dasar pegangan untuk menyusun sebuah butir soal yang mampu memberikan pencerminan adanya perbedaan kemampuan individu siswa. Berdasarkan beberapa penjelasan di atas, maka dapat disimpulkan daya pembeda soal adalah kemampuan suatu butir soal yang dapat membedakan antara siswa yang telah menguasai materi dengan siswa yang belum menguasai materi.

Untuk menentukan daya pembeda (nilai D) dibedakan antara kelompok kecil (kurang dari 100) dan kelompok besar (100 orang ke atas).⁹⁰ Untuk kelompok kecil *testee* dibagi dua sama besar, 50%

⁸⁷ Ibid, hal. 211.

⁸⁸ Widyanuklida, "Analisis Tingkat Kesukaran dan Daya Pembeda Butir Soal Ujian Pelatihan Radiografi Tingkat I", *Repo-nkmbatan*, 16:1, 2017, hal. 3.

⁸⁹ Anas Sudijono, *Pengantar Evaluasi Pendidikan*, (Jakarta: PT. Grafindo Pustaka, 2012), hal.385.

⁹⁰ Arikunto, *Dasar-Dasar Evaluasi Pendidikan Edisi 2*, (Jakarta: Bumi Aksara, 2016), hal. 227.

kelompok atas dan 50% kelompok bawah. Seluruh pengikut tes, dideretkan mulai dari skor teratas sampai terbawah lalu dibagi dua. Sedangkan untuk kelompok besar, mengingat waktu dan biaya untuk menganalisis biasanya hanya diambil kedua kutubnya saja yaitu 27% skor teratas sebagai kelompok atas (J_A) dan 27% skor terbawah sebagai kelompok bawah (J_B).⁹¹ Rumus untuk menentukan indeks deskriminasi adalah sebagai berikut.

$$D = \frac{B_A}{J_A} - \frac{B_B}{J_B} = P_A - P_B$$

Keterangan:

J_A = Banyaknya peserta kelompok atas

J_B = Banyaknya peserta kelompok bawah

B_A = Banyaknya peserta kelompok atas yang menjawab soal dengan benar

B_B = Banyaknya peserta kelompok bawah yang menjawab soal dengan benar

Butir-butir soal yang baik adalah butir-butir soal yang mempunyai indeks deskriminasi 0,4 – 0,7.⁹² Adapun klasifikasi interpretasi daya pembeda dapat dilihat pada Tabel 2.8 di bawah ini.⁹³

Tabel 2.8
Klasifikasi Interpretasi Daya Pembeda

Nilai Daya Pembeda (D_p)	Interpretasi
$D_p < 0,00$	Sangat Jelek
$0,00 \leq D_p < 0,20$	Jelek
$0,20 \leq D_p < 0,40$	Cukup
$0,40 \leq D_p < 0,70$	Baik
$0,70 \leq D_p < 1,00$	Sangat Baik

Sumber: Suharsimi Arikunto, 2009

⁹¹ Ibid.

⁹² Ibid, hal. 232.

⁹³ Arikunto, *Prosedur Penelitian Suatu Pendekatan Praktik*, (Jakarta: Rineka Cipta, 2009), hal. 218.

2. Tingkat Kesukaran

Tingkat kesukaran adalah bilangan yang menunjukkan sukar atau mudahnya sesuatu soal.⁹⁴ Menurut Arikunto, indeks kesukaran adalah suatu instrumen untuk mengetahui tingkat kesulitan tiap butir soal.⁹⁵ Arifin menyatakan bahwa tingkat kesukaran soal adalah persentase atau proporsi dari peserta tes untuk menjawab benar suatu butir soal.⁹⁶ Besarnya tingkat kesukaran berkisar 0,00 – 1,00.⁹⁷ Semakin besar tingkat kesukaran yang diperoleh sebesar 1,00 berarti semakin mudah butir soal yang dijawab benar oleh siswa dan harus direvisi, sedangkan semakin kecil tingkat kesukaran yang diperoleh sebesar 0,00 berarti semakin sukar butir soal yang tidak bisa dijawab benar oleh siswa. Berdasarkan beberapa penjelasan di atas, maka dapat disimpulkan tingkat kesukaran adalah sebuah bilangan yang menunjukkan sukar atau mudahnya suatu butir soal.

Tingkat kesukaran butir soal memiliki 2 kegunaan, yakni bagi guru dan bagi pengujian dan pengajaran. Kegunaan bagi guru, antara lain: 1) sebagai pengenalan konsep terhadap pembelajaran ulang dan memberi masukan pada siswa tentang hasil belajar; 2) memperoleh informasi tentang penekanan kurikulum atau mengecek terhadap butir soal yang bias, sedangkan kegunaan bagi pengujian dan pengajaran, antara lain: 1) pengenalan yang diperlukan untuk diajarkan ulang; 2) mengecek terhadap kelebihan dan kelemahan pada kurikulum sekolah; 3) memberi masukan kepada siswa; 4) mengecek kemungkinan adanya butir soal yang bias.⁹⁸ Adapun rumus tentang tingkat kesukaran suatu soal sebagai berikut.⁹⁹

$$P = \frac{B}{J}$$

Keterangan:

P = Taraf kesukaran

⁹⁴ Arikunto, *Prosedur Penelitian: Suatu Pendekatan Praktik Edisi Revisi IV*, (Yogyakarta: Rineka Cipta, 1999), hal. 207.

⁹⁵ Ibid.

⁹⁶ Zaenal Arifin, "Kriteria Instrumen Dalam Suatu Penelitian", *Jurnal THEOREMS (The Original Research of Mathematics)*, 2:1, 2017, hal. 31.

⁹⁷ Ibid.

⁹⁸ Ibid, hal. 32.

⁹⁹ Ali Hamzah, *Evaluasi Pembelajaran Matematika*, (Jakarta: Rajawali Pers, 2014), hal. 246.

B = Banyak subjek yang menjawab benar

J = Banyak subjek yang mengikuti tes

Tolak ukur yang menginterpretasikan tingkat kesukaran tiap butir soal digunakan kriteria pada Tabel 2.9.¹⁰⁰

Tabel 2.9
Klasifikasi Interpretasi Tingkat Kesukaran

Nilai Tingkat Kesukaran (P)	Interpretasi
$P = 0,00$	Sangat sukar
$0,00 < P \leq 0,30$	Sukar
$0,30 < P \leq 0,70$	Sedang
$0,70 < P < 1,00$	Mudah
$P = 1,00$	Sangat mudah

Sumber: Ali Hamzah, 2014

3. Pengecoh atau Distraktor

Pengecoh atau *distractor* yang juga dikenal dengan istilah penyesat atau penggoda adalah pilihan jawaban yang bukan merupakan kunci jawaban.¹⁰¹ Tes bentuk pilihan ganda terstruktur atas item permasalahan yang ditanyakan atau pokok soal dan *option* atau sejumlah pilihan jawaban.¹⁰² *Option* sendiri terbagi menjadi 2 yakni kunci jawaban dan pengecoh, berarti dalam sekian pilihan jawaban hanya terdapat satu jawaban yang tepat disebut kunci jawaban dan kemungkinan jawaban lain adalah jawaban yang tidak tepat atau disebut pengecoh atau distraktor. Butir soal yang baik, pengecohnya akan dipilih oleh semua siswa yang menjawab salah. Sedangkan butir soal yang kurang baik, pengecohnya akan dipilih tidak merata oleh siswa.

Seorang yang membuat soal pilihan ganda, terkadang ia tidak mudah untuk membuat pengecohnya dikarenakan kecilnya angka daya pembeda soal membuat kurang berfungsinya distraktor.¹⁰³ Salah satu tujuan analisis distraktor adalah untuk memeriksa berapa banyak siswa

¹⁰⁰ Ibid.

¹⁰¹ Ngilman Purwanto, *Prinsip-Prinsip dan Teknik Evaluasi Pengajaran*, (Bandung: PT. Remaja Rosdakarya, 2009), hal. 120.

¹⁰² Zaenal Arifin, Op., Cit, hal. 33.

¹⁰³ Supranata, *Analisis, Validitas, Reliabilitas, dan Interpretasi Hasil Tes Implementasi Kurikulum 2004*, (Bandung: PT. Remaja Rosdakarya, 2009), hal. 49.

pada kelompok atas dan kelompok bawah yang memilih tiap-tiap jawaban pada tes pilihan ganda.¹⁰⁴ Dengan demikian, dalam membuat distraktor perlunya memuat dua komponen yang perlu diperhatikan.

Pertama, apabila tidak ada siswa yang memilih distraktor maka distraktor tersebut tidak berfungsi. Apabila distraktor menyatakan jelas sebagai jawaban tidak benar namun tidak ada siswa yang memilihnya maka distraktor tersebut perlu direvisi. Kedua, apabila distraktor dipilih oleh siswa yang berkemampuan bawah, maka distraktor tersebut baik.

Menurut Arikunto, sebuah pengecoh dapat dikatakan berfungsi dengan baik apabila pengecoh tersebut mempunyai daya tarik yang besar bagi pengikut-pengikut tes yang kurang memahami konsep atau kurang menguasai bahan.¹⁰⁵ Berdasarkan penjelasan di atas maka dapat disimpulkan bahwa pengecoh atau distraktor adalah pilihan jawaban yang bukan kunci jawaban dan dipilih minimal 5% dari peserta tes. Indeks pengecoh butir soal dapat dihitung dengan menggunakan rumus sebagai berikut.¹⁰⁶

$$IP = \frac{P}{(N-B)/(n-1)} \times 100\%$$

Keterangan:

IP = Indeks pengecoh

P = Jumlah siswa yang memilih pengecoh

N = Jumlah siswa yang mengikuti tes

B = Jumlah siswa yang menjawab benar

n = Jumlah alternatif jawaban (*options*)

Suatu distraktor yang efektif harus dipilih oleh beberapa peserta tes atau minimal dipilih oleh 5% peserta tes.¹⁰⁷ Pengecoh dikatakan berfungsi efektif apabila banyak dipilih oleh siswa yang berkemampuan bawah dan sebaliknya apabila banyak dipilih oleh siswa yang berkemampuan tinggi maka pengecoh tersebut tidak berfungsi. Sebagaimana pada Tabel 2.10 sebagai berikut.¹⁰⁸

¹⁰⁴ Zaenal Arifin, Op., Cit, hal 33.

¹⁰⁵ Arikunto, *Prosedur Penelitian: Suatu Pendekatan Praktek Edisi Revisi IV*, (Jakarta: PT. Rineka Cipta, 1997), hal. 225.

¹⁰⁶ Zainal Arifin, *Evaluasi Pembelajaran*, (Bandung: PT. Remaja Rosdakarya, 2013), hal. 279.

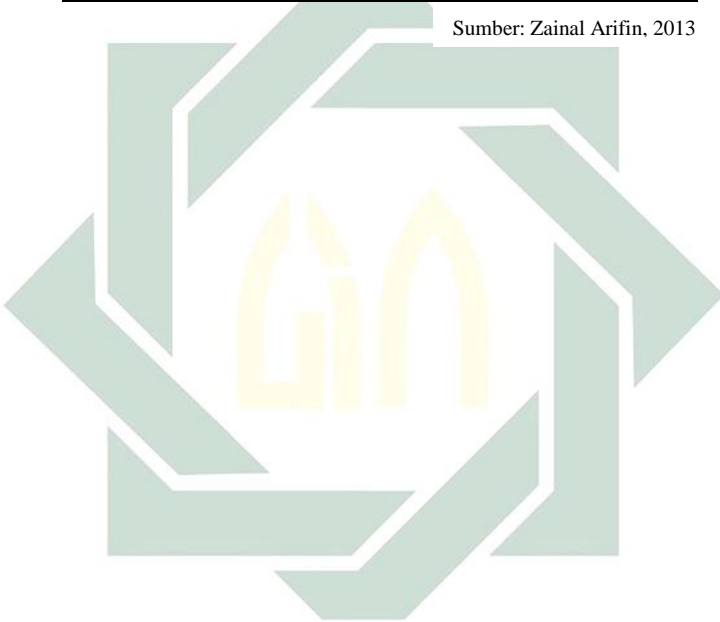
¹⁰⁷ Basuki-Hariyanto, *Asesemen Pembelajaran*, (Bandung: PT. Remaja Rosdakarya, 2014), hal. 144.

¹⁰⁸ Zainal Arifin, Op., Cit, hal. 279.

Tabel 2.10
Klasifikasi Interpretasi Indeks Pengecoh

Nilai Indeks Pengecoh (<i>IP</i>)	Interpretasi
Lebih dari 200%	Sangat jelek
0% – 25% atau 176% – 200%	Jelek
26% – 50% atau 151% – 175%	Cukup baik
51% – 75% atau 126% – 150%	Baik
76% – 125%	Sangat baik

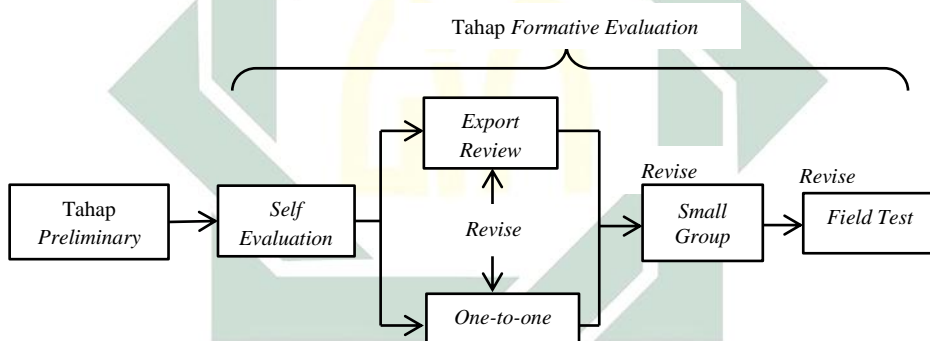
Sumber: Zainal Arifin, 2013



BAB III METODE PENELITIAN

A. Model Penelitian dan Pengembangan

Penelitian ini merupakan penelitian dan pengembangan yang bertujuan untuk mengembangkan instrumen tes diagnostik berbentuk pilihan ganda tiga tingkat (*three-tier*) untuk mengidentifikasi miskonsepsi pada materi geometri yang valid dan reliabel dalam pembelajaran matematika. Model pengembangan yang digunakan adalah model pengembangan *tipe formative research Tessmer* yang terdiri dari dua tahapan yaitu tahap *preliminary* dan tahap *formative evaluation* yang meliputi *self evaluation*, *prototyping*, *small group* dan *field test*. Prosedur penelitian yang dikemukakan oleh Tessmer dapat dilihat pada Gambar 3.1.¹⁰⁹



Gambar 3.1
Alur Tahapan Tessmer

B. Prosedur Penelitian dan Pengembangan

Pada pengembangan instrumen tes ini menggunakan model Tessmer yang terbagi menjadi dua tahapan. Adapun tahapan dalam penelitian ini sebagai berikut.

¹⁰⁹ Martin Tessmer, *Planning and Conducting Formative Evaluations: Improving the Quality of Educational and Training*, (London: Kogan Page, 1996).

1. Tahap *Preliminary* (Tahap Persiapan)

Tahap ini merupakan tahapan penentuan tempat dan subjek dalam penelitian. Namun sebelum itu melakukan studi kepustakaan mengenai miskonsepsi, tes diagnostik pilihan ganda *three tier*, serta mengenai konsep geometri. Studi kepustakaan dapat diambil dari buku, jurnal atau penelitian-penelitian yang relevan. Selain itu juga mengumpulkan informasi yang dibutuhkan pada saat penelitian, seperti mengatur jadwal penelitian dan prosedur kerjasama dengan guru matematika yang akan dijadikan lokasi penelitian.

Setelah mendapatkan beberapa informasi yang berkaitan dengan penelitian maka akan dilaksanakan kegiatan penentuan lokasi dan subjek uji coba dengan cara menghubungi beberapa pihak sekolah yaitu kepala sekolah dan guru matematika, kemudian melakukan wawancara kepada guru matematika mengenai pembelajaran matematika yang berkurikulum 2013, kondisi belajar siswa dalam menyelesaikan soal matematika berdasarkan kemampuan pemahaman konsep dan faktor ketidaktuntasan mereka.

2. Tahap *Formative Evaluation* (Tahap Penilaian Formatif)

Pada tahapan ini terdapat beberapa langkah yang harus dilakukan, antara lain:

a. *Self Evaluation* (Evaluasi Diri)

1) Analisis

Tahap ini merupakan tahap awal dalam penelitian pengembangan. Dalam hal ini peneliti akan melakukan beberapa hal di antaranya:

a) Analisis kurikulum

Peneliti melakukan telaah kurikulum pada mata pelajaran matematika, tantangan, dan literatur yang mendukung sehingga diperoleh tes diagnostik pilihan ganda *three-tier* yang dapat mengidentifikasi miskonsepsi siswa.

b) Analisis siswa

Peneliti menganalisis siswa yang akan menjadi subjek dalam penelitian baik itu secara karakter dari siswa maupun kompetensi akademik siswa secara keseluruhan.

c) Analisis materi

Peneliti mengidentifikasi, menyusun, dan merinci secara berurutan pokok bahasan atau materi yang berkaitan dengan geometri. Dengan analisis ini dapat mempermudah peneliti dalam mengembangkan tes diagnostik pilihan ganda *three-tier*.

2) Desain

Kegiatan yang dilakukan pada tahap ini yaitu peneliti menentukan jumlah butir soal yang akan dikembangkan, membuat kisi-kisi soal pada tes diagnostik, membuat butir soal tes diagnostik pilihan ganda *three-tier* beserta kunci jawabannya. Setelah membuat desain tes diagnostik pilihan ganda *three-tier* yang dikembangkan maka dikonsultasikan ke dosen pembimbing. Hasil dari desain produk ini disebut sebagai *prototype*. Pada masing-masing *prototype* fokus pada beberapa aspek penilaian yaitu konten, konstruk, dan bahasa. Penjelasan dapat dilihat pada Tabel 3.1.

Tabel 3.1
Aspek Penilaian yang Menjadi Fokus *Prototype*

No.	Aspek Penilaian	Kategori
1.	Konten	Instrumen yang dibuat sesuai dengan indikator yang dibuat
		Instrumen yang diukur sesuai dengan kompetensi keterampilan yang akan diukur
		Isi atau materi yang ditanyakan sesuai dengan kompetensi keterampilan yang akan diukur
2.	Konstruk	Menggunakan petunjuk pengerjaan soal yang jelas
		Mengembangkan kemampuan pemahaman konsep siswa
		Terdapat rubrik penilaian
3.	Bahasa	Instrumen yang dibuat menggunakan bahasa Indonesia

		sesuai EYD
		Menggunakan bahasa yang sederhana, mudah dipahami dan sesuai dengan taraf berpikir siswa SMP
		Soal tidak mengandung penafisiran ganda atau salah pengertian

b. Prototyping

Pada tahap ini, produk yang sudah didesain akan dievaluasi dengan diujicobakan dalam tiga kelompok yaitu:

1) *Expert Review* (Ulasan Para Ahli)

Tahap ini hasil *prototype* pertama akan diberikan kepada para pakar atau para ahli yang mana akan dicermati, dinilai, dan dievaluasi. Para ahli atau validator akan menelaah konten, konstruk, dan bahasa pada masing-masing *prototype* I. Pada penelitian ini terdapat tiga validator yaitu dua dosen pendidikan matematika dan guru bidang matematika. Validator memberikan saran-saran dan tanggapan yang digunakan sebagai bahan untuk merevisi desain tes diagnostik pilihan ganda *three-tier* yang disusun oleh peneliti. Saran dan tanggapan yang diberikan oleh validator ditulis pada lembar validasi yang sudah disiapkan oleh peneliti dan menyatakan bahwa tes diagnostik pilihan ganda *three-tier* tersebut telah valid. Adapun *prototype* I yang akan diberikan kepada para ahli terlampir pada Lampiran 4.1.

2) *One-to-one* (Perorangan)

Hasil dari revisi pada tahap *expert review* akan diujicobakan pada 3 siswa non subjek penelitian untuk menjawab tes diagnostik pilihan ganda *three-tier* yang telah direvisi. Ketiga siswa ini adalah siswa yang memiliki kategori kemampuan tinggi, sedang, dan rendah, kemudian peneliti meminta komentar dari ketiga siswa tersebut. Komentar yang didapat digunakan untuk merevisi desain tes diagnostik pilihan ganda *three-tier* yang telah dibuat oleh peneliti. Hasil

revisi pada tahap ini disebut sebagai *prototype* II yang terlampir pada Lampiran 4.2.

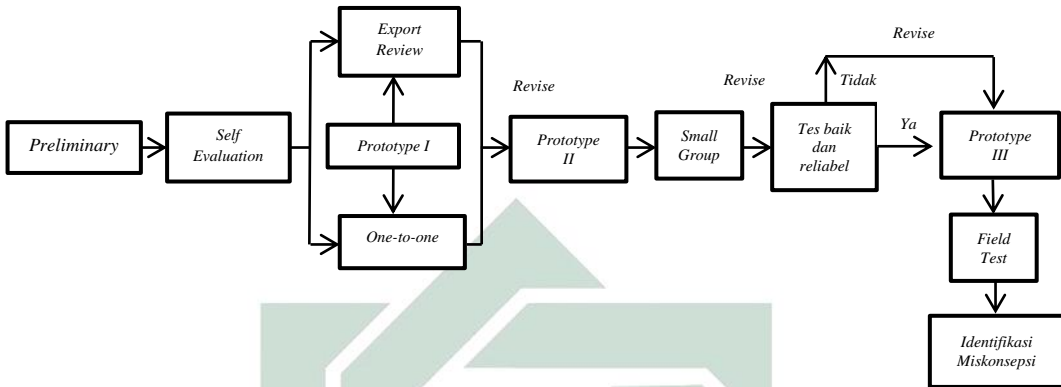
3) *Small Group* (Kelompok Kecil)

Hasil revisi dari tahap *one-to-one* dan kesulitan yang dialami pada uji coba *prototype* I akan digunakan sebagai bahan revisi *prototype* II. Hasil dari *prototype* II akan diujicobakan kepada 6 siswa non subjek penelitian yang terdiri dari dua siswa yang memiliki kemampuan tinggi, dua siswa yang memiliki kemampuan sedang, dan dua siswa yang memiliki kemampuan rendah. Siswa-siswa tersebut dimintai komentar dan tanggapan terhadap tes diagnostik pilihan ganda *three-tier* yang diujikan. Berdasarkan tanggapan dan komentar siswa tersebut digunakan sebagai tambahan untuk merevisi desain tes diagnostik pilihan ganda *three-tier* ke tahap berikutnya. Pada tahap ini peneliti mengharapkan dapat menghasilkan tes diagnostik pilihan ganda *three-tier* yang mampu mengidentifikasi miskonsepsi. Hasil revisi pada tahap *one-to-one* diberi nama *prototype* III yang terlampir pada Lampiran 4.3.

c. *Field Test* (Uji Coba Lapangan)

Pada tahap ini, komentar atau masukan yang didapatkan dari uji coba *small group* digunakan untuk merevisi desain *prototype* III. Hasil revisi akan diuji cobakan ke subjek penelitian. Uji coba pada tahap ini merupakan *field test* atau uji coba lapangan. Uji coba lapangan ini akan diberikan kepada siswa kelas IX SMPN 13 Surabaya.

Adapun desain pengembangan pada penelitian ini adalah sebagai berikut.



Gambar 3.2
Alur Tahapan Pengembangan Tesser

C. Uji Coba Produk

1. Subjek Uji Coba

Subjek dalam penelitian pengembangan ini adalah siswa kelas IX-C SMPN 13 Surabaya. Teknik pengambilan sampel pada penelitian ini adalah *simple random sampling* atau teknik acak sederhana.

2. Jenis Data

Data adalah suatu bahan yang berisi keterangan tentang suatu objek penelitian yang diperoleh dari tempat penelitian.¹¹⁰ Adapun data dalam penelitian ini:

a. Data Hasil Validasi Ahli Terhadap Tes Diagnostik Pilihan Ganda *Three-Tier* (Lembar Validasi)

Data hasil validator terhadap tes diagnostik pilihan ganda *three-tier* berupa data pernyataan mengenai kevalidan tes diagnostik pilihan ganda *three-tier* dari segi kesesuaian tes dengan materi. Sumber data dalam penelitian ini berupa tiga validator yang berkompeten.

¹¹⁰ Burhan Mungin, *Metodologi Penelitian Kuantitatif: Komunikasi, Ekonomi, dan Kebijakan Publik serta Ilmu-ilmu Sosial Lainnya*, (Jakarta: Kencana Prenada Media Group, 2011), hal. 129.

b. Data Hasil Uji coba lapangan

Data hasil uji coba lapangan pada penelitian ini berupa data tentang validitas, reliabilitas, daya pembeda soal, dan tes diagnostik pilihan ganda *three-tier* untuk mengidentifikasi miskonsepsi pada materi geometri yang telah dikembangkan. Data tersebut dapat bertujuan memberikan kesimpulan bahwa tes diagnostik pilihan ganda *three-tier* yang dikembangkan telah memenuhi syarat tes yang baik.

D. Teknik Pengumpulan Data

Metode pengumpulan data merupakan strategi untuk mendapatkan data yang diperlukan dalam sebuah penelitian. Teknik pengumpulan data yang dilakukan pada penelitian ini adalah sebagai berikut:

1. Metode Angket

Pada penelitian ini angket yang digunakan adalah lembar validasi tes diagnostik pilihan ganda *three-tier*. Lembar validasi yang digunakan adalah lembar validasi yang diadopsi dari penelitian yang dilakukan oleh Wardah untuk aspek konten, konstruksi, dan bahasa.

2. Metode Tes Diagnostik Pilihan Ganda *Three-Tier*

Tes merupakan alat untuk melakukan pengukuran yaitu mengumpulkan informasi karakteristik dari suatu objek. Tes yang diberikan pada penelitian ini berupa soal-soal tes diagnostik pilihan ganda *three-tier* pada pokok bahasan segiempat. Tes ini diberikan kepada siswa kelas IX-C SMPN 13 Surabaya. Hasil dari tes diagnostik pada penelitian ini digunakan untuk mengidentifikasi miskonsepsi siswa pada materi geometri tentang segiempat.

E. Instrumen Pengumpulan Data

Instrumen pengumpulan data merupakan alat yang memiliki fungsi untuk mengumpulkan data yang akan diteliti. Adapun instrumen pengumpulan data yang digunakan dalam penelitian ini antara lain:

1. Lembar Angket Validasi

Lembar angket validasi pada penelitian ini adalah lembar validasi instrumen yang disusun untuk menilai aspek konten, konstruk, dan bahasa. Lembar validasi ditujukan untuk menggali kualitas tes diagnostik pilihan ganda *three-tier* untuk

mengidentifikasi miskonsepsi siswa pada materi geometri yang telah dikembangkan dari validator. Adapun lembar validasi terlampir pada Lampiran 4.4.

2. Lembar Tes Diagnostik Pilihan Ganda *Three-Tier*

Lembar tes diagnostik pilihan ganda *three-tier* dalam penelitian ini menggunakan soal-soal matematika kelas VII SMP/MTs/Sederajat dengan pokok bahasan segiempat sebanyak 10 butir soal. Tes yang diujikan ini dalam bentuk pilihan ganda tiga tingkat yang disesuaikan dengan indikator yang dibuat peneliti guna mengidentifikasi miskonsepsi siswa pada materi geometri tentang segiempat. Adapun soal tes diagnostik *three-tier* yang akan diujikan terlampir pada Lampiran 4.3.

F. Teknik Analisis Data

Dari data yang diperoleh maka akan dilakukan analisis data sebagai berikut:

1. Analisis Data Lembar Validasi

Pada tahap ini menganalisis hasil penilaian validator terhadap lembar validasi tes diagnostik pilihan ganda *three-tier* yang dibuat oleh peneliti. Analisis data hasil validasi lembar tes diagnostik pilihan ganda *three-tier* untuk mengidentifikasi miskonsepsi siswa pada materi geometri dengan mencari rata-rata dari setiap kategori, rata-rata setiap aspek dalam lembar validasi, dan rata-rata total penilaian validator. Adapun rumus-rumus yang digunakan dalam menganalisis data hasil validasi tes diagnostik *three-tier*.

a. Mencari rata-rata setiap kategori dari semua validator.¹¹¹

$$RK_i = \frac{\sum_{j=1}^n V_{ji}}{n}$$

Keterangan:

RK_i = Rata-rata kategori ke- i

V_{ji} = Skor hasil penilaian validator ke- j untuk kategori ke- i

n = Banyaknya validator

¹¹¹ Siti Khabibah, Disertasi: “*Pengembangan Perangkat Pembelajaran Matematika dengan Soal Terbuka Untuk Meningkatkan Kreatifitas Siswa Sekolah Dasar*”. (Surabaya: Unesa, 2006 tidak dipublikasikan), hal. 34.

- b. Mencari rata-rata setiap aspek dari semua validator

$$RA_i = \frac{\sum_{j=1}^n RK_{ji}}{n}$$

Keterangan:

RA_i = Rata-rata aspek ke- i

RK_{ji} = Rata-rata aspek ke- j untuk kategori ke- i

n = Banyaknya kategori dalam aspek ke- i

- c. Mencari rata-rata total validitas

$$RTV = \frac{\sum_{i=1}^n RA_i}{n}$$

Keterangan:

RTV = Rata-rata total validitas

RA_i = Rata-rata aspek ke- i

n = Banyaknya aspek

Khabibah menyatakan bahwa untuk menentukan kategori kevalidan suatu tes dapat diperoleh dengan mencocokkan rata-rata total (\bar{x}) dengan kategori kevalidan perangkat pembelajaran sebagaimana ditunjukkan dalam tabel berikut.¹¹²

Tabel 3.2

Kriteria Penilaian Kevalidan Perangkat Pembelajaran

Interval Soal	Kategori Kevalidan
$4 \leq RTV \leq 5$	Sangat valid
$3 \leq RTV < 4$	Valid
$2 \leq RTV < 3$	Kurang valid
$1 \leq RTV < 2$	Tidak valid

Sumber: Siti Khabibah, 2006

Keterangan:

RTV = Rata-rata total hasil penilaian validator terhadap instrumen yang dikembangkan

2. Analisis Soal Tes Diagnostik Pilihan Ganda *Three-Tier*

- a. Uji Validitas Empiris

Dalam menganalisis validitas butir soal dalam penelitian ini menggunakan koefisien korelasi antara skor butir dengan skor total tes dengan menggunakan rumus koefisien korelasi poin biserial sebagai berikut.¹¹³

¹¹² Ibid.

¹¹³ Ali Hamzah, *Evaluasi Pembelajaran Matematika*, (Jakarta: PT. Rajagrafindo Persada, 2014), hal. 221.

$$r_{pbis} = \frac{X_i - X_t}{S_t} \sqrt{\frac{p_i}{q_i}}$$

Keterangan:

r_{pbis} = Koefisien korelasi poin biserial antara skor butir soal nomor i dengan skor total

X_i = Rata-rata skor total responden yang menjawab benar butir soal nomor i ,

X_t = Rata-rata skor total semua responden

S_t = Standar deviasi skor total semua responden

p_i = Proporsi jawaban yang benar untuk butir soal nomor i

q_i = Proporsi jawaban yang salah untuk butir soal nomor i

Interpretasi dari koefisien korelasi yang diperoleh dapat dilihat pada Tabel 3.3.¹¹⁴

Tabel 3.3
Interpretasi Koefisien Korelasi

Koefisien Korelasi (r)	Tafsiran
$0,40 \leq r < 1,00$	Soal baik
$0,30 \leq r < 0,40$	Terima dan diperbaiki
$0,20 \leq r < 0,30$	Soal diperbaiki
$0,19 \leq r < 0,00$	Soal ditolak

Sumber: Suharsimi Arikunto, 2006

Untuk menafsirkan harga koefisien korelasi yang diperoleh dapat menggunakan dua cara, yakni: (1) dengan melihat nilai r dan diinterpretasikan seperti soal baik, soal diterima dan diperbaiki, soal diperbaiki, serta soal ditolak; (2) dengan berkonsultasi tabel kritis r *product moment* sehingga dapat diketahui signifikan atau tidaknya korelasi tersebut. Adapun yang digunakan dalam penelitian ini, yaitu tes diagnostik pilihan ganda *three-tirer* dikatakan valid apabila nilai r lebih dari sama dengan 0,400.¹¹⁵ Cara menghitung validitas dalam penelitian ini, yaitu menggunakan rumus koefisien korelasi poin biserial

¹¹⁴ Arikunto, *Prosedur Penelitian: Suatu Pendekatan Praktik*, (Jakarta: Rineka Cipta, 2006), hal. 283.

¹¹⁵ Ibid.

dengan berbantuan program *SPSS Statistic Versi 16.0*. Adapun kategori nilai korelasi pengujian sebagai berikut:

- 1) Jika nilai korelasi berada di atas 0,400, maka tes diagnostik pilihan ganda *three-tier* dikatakan valid.
- 2) Jika nilai korelasi di bawah 0,400, maka tes diagnostik pilihan ganda *three-tier* dikatakan tidak valid.

b. Uji Reliabilitas

Uji reliabilitas dilakukan untuk mengetahui apakah tes diagnostik ini dapat konsisten jika instrumen yang dihasilkan dipakai berulang kali.¹¹⁶ Pada penelitian ini uji reliabilitas diukur dengan menggunakan rumus persamaan K-R.20. Rumus tersebut adalah sebagai berikut.¹¹⁷

$$r_{11} = \left(\frac{n}{n-1} \right) \left(\frac{S^2 - \sum pq}{S^2} \right)$$

Keterangan:

r_{11} = Reliabilitas tes secara keseluruhan

p = Proporsi subjek yang menjawab item dengan benar

q = Proporsi subjek yang menjawab item dengan salah
($q = 1 - p$)

$\sum pq$ = Jumlah hasil perkalian antara p dan q

n = Banyaknya item

S = Standar deviasi dari tes (standar deviasi adalah akar varians)

Sebuah tes dikatakan baik jika bersifat *reliabel*. Tes dikatakan *reliabel* jika tes tersebut diujikan berkali-kali dengan dua atau lebih penilai kepada subjek yang sama maka hasilnya relatif sama. Menurut Arikunto, untuk interpretasi koefisien reliabilitas dapat menggunakan pedoman kriteria penafsiran koefisien reliabilitas sebagaimana pada Tabel 3.4.¹¹⁸

¹¹⁶ Zainal Arifin, *Evaluasi Pembelajaran Cet. V*, (Bandung: PT. Remaja Rosdakarya, 2013), hal. 259.

¹¹⁷ Arikunto, *Dasar-Dasar Evaluasi Pendidikan Edisi 2*, (Jakarta: Bumi Aksara, 2016), hal. 115.

¹¹⁸ Arikunto, *Prosedur Penelitian Suatu Pendekatan Praktik*, (Jakarta: Rineka Cipta, 2009), hal. 276.

Tabel 3.4
Kategori Penafsiran Koefisien Reliabilitas

Koefisien Reliabilitas (r_{11})	Tafsiran
$0,80 \leq r_{11} < 1,00$	Reliabilitas sangat tinggi
$0,60 \leq r_{11} < 0,80$	Reliabilitas tinggi
$0,40 \leq r_{11} < 0,60$	Reliabilitas sedang (cukup)
$0,20 \leq r_{11} < 0,40$	Reliabilitas rendah
$r_{11} \leq 0,20$	Reliabilitas sangat rendah

Sumber: Suharsimi Arikunto, 2006

Dalam penelitian ini, uji reliabilitas menggunakan rumus *Kuder-Richardson* (K-R).20 dengan berbantuan program *SPSS Statistic Versi 16.0*. Untuk kriteria koefisien reliabilitas menggunakan batasan 0,600.¹¹⁹ Oleh karena itu, kriteria tes diagnostik pilihan ganda *three-tier* dapat dikatakan reliabel apabila nilai korelasinya di atas atau sama dengan 0,600.

c. Daya Pembeda

Daya pembeda soal merupakan kemampuan suatu soal yang dapat membedakan antara siswa yang telah menguasai materi dengan siswa yang belum menguasai materi. Daya pembeda soal mengkaji butir-butir soal dengan tujuan untuk mengetahui kesanggupan soal dalam membedakan siswa yang tergolong mampu (tinggi prestasinya) dengan siswa yang tergolong kurang atau lemah prestasinya.¹²⁰

Butir-butir soal dikatakan baik jika butir soal tersebut dapat membedakan kemampuan siswa dalam memahami suatu konsep. Untuk menentukan indeks diskriminasi dapat

¹¹⁹ Domenic Vincent Cichetti, et.al., "Rating Scales, Scales of Measurement, Issues of Reliability Resolving Some Critical Issues for Clinicians and Research", *The Journal of Nervous and Mental Disease* (2006), 194, hal. 557-564.

¹²⁰ Nana Sudjana, *Penilaian Hasil Proses Belajar Mengajar Cet. XIII*, (Bandung: PT. Remaja Rosdakarya, 2009), hal. 141.

menggunakan rumus yang diungkapkan oleh Arikunto sebagai berikut.¹²¹

$$D = \frac{B_A}{J_A} - \frac{B_B}{J_B} = P_A - P_B$$

Keterangan:

D = Indeks daya pembeda butir soal

J = Jumlah peserta tes

J_A = Banyaknya peserta kelompok atas

J_B = Banyaknya peserta kelompok bawah

B_A = Banyaknya peserta kelompok atas yang menjawab soal dengan benar

B_B = Banyaknya peserta kelompok bawah yang menjawab soal dengan benar

P_A = Proporsi peserta kelompok atas yang menjawab benar

P_B = Proporsi peserta kelompok bawah yang menjawab benar

Untuk mengetahui interpretasi daya pembeda butir soal dapat dilihat pada Tabel 3.5.¹²²

Tabel 3.5
Klasifikasi Interpretasi Daya Pembeda

Indeks Daya Pembeda (D_p)	Tafsiran
$0,70 \leq D_p < 1,00$	Baik Sekali
$0,40 \leq D_p < 0,70$	Baik
$0,20 \leq D_p < 0,40$	Cukup
$0,00 \leq D_p < 0,20$	Jelek

Sumber: Suharsimi Arikunto, 2006

Dalam penelitian ini menggunakan kriteria indeks diskriminasi butir soal dengan batasan 0,400. Dengan demikian, kriteria indeks diskriminasi butir soal dikatakan baik apabila nilai indeksnya berada di atas atau sama dengan 0,400

¹²¹ Suwanto, *Pengembangan Tes Diagnostik dalam Pembelajaran*, (Yogyakarta: Pustaka Pelajar, 2013), hal. 109.

¹²² Arikunto, *Prosedur Penelitian Suatu Pendekatan Praktik*, (Jakarta: Rineka Cipta, 2006), hal. 218.

d. Tingkat Kesukaran

Tingkat kesukaran butir soal merupakan salah satu indikator kualitas butir soal yang menunjukkan butir soal tersebut termasuk sukar, sedang, atau mudah. Suatu butir soal dikatakan mudah apabila sebagian besar siswa dapat menjawabnya dengan benar dan suatu butir soal dikatakan sukar apabila sebagian besar siswa tidak dapat menjawab dengan benar.¹²³ Tingkat kesukaran dihitung melalui indeks kesukaran yaitu angka yang menunjukkan proporsi siswa yang menjawab benar pada setiap butir soal. Semakin tinggi angka indeks kesukaran semakin mudah soal tersebut, sebaliknya semakin kecil indeks kesukaran semakin sukar soal tersebut.¹²⁴ Adapun rumus untuk menghitung indeks kesukaran butir soal adalah sebagai berikut.¹²⁵

$$P = \frac{B}{J}$$

Keterangan:

P = Taraf kesukaran

B = Banyak subjek yang menjawab benar

J = Banyak subjek yang mengikuti tes

Tolak ukur untuk menginterpretasikan taraf kesukaran pada setiap butir soal dapat dilihat pada Tabel 3.6 di bawah ini.

Tabel 3.6
Klasifikasi Interpretasi Tingkat Kesukaran

Nilai Tingkat Kesukaran (P)	Interpretasi
$P = 0$	Sangat sukar
$0,00 < P \leq 0,30$	Sukar
$0,30 < P \leq 0,70$	Sedang
$0,70 < P < 1,00$	Mudah
$P = 1,00$	Sangat mudah

Sumber: Ali Hamzah, 2014

¹²³ Ali Hamzah, *Evaluasi Pembelajaran Matematika*, (Jakarta: Rajawali Pers, 2014), hal. 244.

¹²⁴ Ibid.

¹²⁵ Ibid, hal. 246.

Dalam penelitian ini. kriteria tingkat kesukaran yang digunakan antara 0,31 – 0,70.¹²⁶ Jadi, kriteria butir soal tes diagnostik pilihan ganda *three-tier* memiliki tingkat kesukaran yang baik antara 0,31 – 0,70.

e. Pengecoh atau Distraktor

Pengecoh atau *distractor* merupakan pilihan jawaban yang bukan termasuk kunci jawaban. Misalnya, pada soal objektif pilihan ganda empat pilihan a, b, c, d dan kunci jawabannya adalah c maka a, b, d merupakan pengecoh.¹²⁷ Pada butir soal yang baik pengecohnya akan dipilih merata oleh siswa yang menjawab salah, sebaliknya butir soal yang kurang baik pengecohnya akan dipilih secara tidak merata. Untuk indeks pengecoh dapat menggunakan rumus di bawah ini.¹²⁸

$$IP = \frac{P}{(N-B)/(n-1)} \times 100\%$$

Keterangan:

IP = Indeks pengecoh

P = Jumlah siswa yang memilih pengecoh

N = Jumlah siswa yang mengikuti tes

B = Jumlah siswa yang menjawab benar

n = Jumlah alternatif jawaban (*options*)

Tabel 3.7

Klasifikasi Interpretasi Indeks Pengecoh

Nilai Indeks Pengecoh (<i>IP</i>)	Interpretasi
Lebih dari 200%	Sangat jelek
0% – 25% atau 176% – 200%	Jelek
26% – 50% atau 151% – 175%	Cukup baik
51% – 75% atau 126% – 150%	Baik
76% – 125%	Sangat baik

Sumber: Zainal Arifin, 2013

Suatu distraktor yang efektif harus dipilih oleh beberapa peserta tes atau minimal dipilih oleh 5% peserta

¹²⁶ Ali Hamzah, *Evaluasi Pembelajaran Matematika*, (Jakarta: Rajawali Pers, 2014), hal. 244.

¹²⁷ Purwanto, *Evaluasi Hasil Belajar*, (Yohyakarta: Pustaka Pelajar, 2014), hal. 75.

¹²⁸ Zainal Arifin, *Evaluasi Pembelajaran*, (Bandung: PT. Remaja Rosdakarya, 2013), hal. 279.

tes.¹²⁹ Dalam penelitian ini, pengecoh dikatakan baik apabila dipilih paling sedikit 2 dari subjek penelitian. Apabila pengecoh dipilih secara merata maka pengecoh tersebut baik, sedangkan apabila pengecoh dipilih lebih banyak oleh siswa yang berkemampuan atas maka pengecoh tersebut perlu direvisi atau diperbaiki.

3. Analisis Data Hasil Tes Diagnostik Pilihan Ganda *Three-tier*
Langkah-langkah dalam menganalisis data hasil tes diagnostik *three-tier* antara lain:
 - a. Merekapitulasi setiap jawaban responden dan mengubahnya menjadi skor

Dalam soal tes diagnostik *three-tier* terdiri dari tiga *tier* atau tiga tingkatan. Tingkatan pertama berisi pilihan jawaban dengan kriteria penilaian seperti Tabel 3.8. Pada *-tier* atau tingkatan kedua berisi sejumlah pilihan alasan untuk jawaban yang dipilih pada *tier* pertama dengan kriteria seperti pada Tabel 3.9 serta menambahkan kolom CRI (*Certainty of Response Index*) pada tingkatan ketiga dengan skala 1 sampai 4 dengan kategori tertentu, apabila interpretasi skala 1 dan 2 maka CRI tingkat rendah dan apabila interpretasi skala 3 dan 4 maka CRI tingkat tinggi dapat dilihat pada Tabel 3.10.¹³⁰

Tabel 3.8
Kriteria Penilaian Pilihan Jawaban

Bentuk Soal	Nilai	Keterangan
Pilihan ganda	1	Jawaban benar
	0	Jawaban salah

Tabel 3.9
Kriteria Penilaian Pilihan Alasan

Bentuk Soal	Nilai	Keterangan
Pilihan ganda	1	Alasan benar
	0	Alasan salah

¹²⁹ Basuki-Hariyanto, *Asemen Pembelajaran*, (Bandung: PT. Remaja Rosdakarya, 2014), hal. 144.

¹³⁰ Nursiwin, "*Menggali Miskonsepsi Siswa Pada Materi Perhitungan Kimia Menggunakan Certainty of Response Index*". (Artikel Universitas Tanjungpura, Pontianak, 2014).

Tabel 3.10
Interpretasi Skala CRI

CRI	Kriteria	Tingkat Kategori
1	Sangat tidak yakin	Rendah
2	Tidak yakin	Rendah
3	Yakin	Tinggi
4	Sangat yakin	Tinggi

Sumber: Nursiwin, 2014

Untuk menganalisis miskonsepsi siswa, yang harus dilakukan pertama kali adalah memberikan skor pada masing-masing jawaban siswa. Adapun kriteria penilaian pilihan jawaban dan alasan pada tes diagnostik pilihan ganda *three-tier* adalah sebagai berikut:

- 1) Jika siswa memilih pilihan jawaban dan pilihan alasan benar maka butir soal mendapatkan skor 1
 - 2) Jika siswa memilih pilihan jawaban salah dan pilihan alasan benar maka butir soal mendapatkan skor 0
 - 3) Jika siswa memilih pilihan jawaban benar dan pilihan alasan salah maka butir soal mendapatkan skor 0
 - 4) Jika siswa memilih pilihan jawaban salah dan pilihan alasan salah maka butir soal mendapatkan skor 0
- b. Menginterpretasikan jawaban responden

Setelah merekapitulasi data hasil tes diagnostik *three-tier* dan mengubahnya menjadi skor akan dilakukan interpretasi jawaban-jawaban dalam kategori paham, tidak paham, dan miskonsepsi seperti pada Tabel 3.11.¹³¹

¹³¹Tri Ade Mustaqim, dkk., "Identifikasi Miskonsepsi Siswa Dengan Menggunakan Metode (*Certainty of Response Index*) CRI Pada Konsep Fotosintesis dan Respirasi Tumbuhan", *EDUSAINS*, 6:2, 2014, hal. 148.

Tabel 3.11
Interpretasi Hasil

Tingkat Pertama	Tingkat Kedua	Tingkat Kategori CRI	Deskripsi
Benar	Benar	Tinggi	Paham
Benar	Benar	Rendah	Tidak paham
Benar	Salah	Tinggi	Miskonsepsi
Benar	Salah	Rendah	Tidak paham
Salah	Benar	Tinggi	Miskonsepsi
Salah	Benar	Rendah	Tidak paham
Salah	Salah	Tinggi	Miskonsepsi
Salah	Salah	Rendah	Tidak paham

Sumber: Mustaqim, Zulfiani & Herlanti, 2014

- c. Menghitung persentase dan memasukkan ke dalam kategori tingkat miskonsepsi

Melakukan perhitungan persentase terhadap hasil penilaian pada setiap kategori dengan menggunakan rumus berikut ini.¹³²

$$P = \frac{f}{N} \times 100\%$$

Keterangan:

P = Persentase siswa tiap kategori

f = Jumlah siswa tiap kategori

N = Jumlah seluruh siswa yang menjadi subjek penelitian

Pada tahap berikutnya melakukan rekapitulasi dan memasukkan kategori yang diperoleh siswa dari perhitungan persentase sebelumnya dengan kategori tingkat miskonsepsi dapat dilihat pada Tabel 3.12.¹³³

Tabel 3.12
Kriteria Penilaian Tingkat Miskonsepsi

Persentase	Kategori
$0\% \leq P < 30\%$	Rendah
$30\% \leq P < 60\%$	Sedang
$60\% \leq P < 100\%$	Tinggi

Sumber: Anas Sudijono, 2009

¹³² Anas Sudijono, *Pengantar Statistik Pendidikan*, (Jakarta: PT. Grafindo Persada, 2010).

¹³³ Anas Sudijono, *Pengantar Evaluasi Pendidikan*, (Jakarta: Rajawali Press, 2009).

BAB IV HASIL PENELITIAN DAN PEMBAHASAN

A. Deskripsi Data

1. Deskripsi Data Kevalidan Tes Diagnostik Pilihan Ganda *Three-Tier* Untuk Mengidentifikasi Miskonsepsi Pada Materi Geometri

a. Data Validasi Para Ahli (*Expert Judgment*)

Tes diagnostik *three-tier* berupa soal pilihan ganda *three-tier* untuk mengidentifikasi miskonsepsi siswa yang akan divalidasi oleh 3 ahli (*expert judgment*), yakni 2 dosen pendidikan matematika dari UIN Sunan Ampel Surabaya dan 1 guru matematika dari SMP Negeri 13 Surabaya. Dalam penelitian ini, validasi para ahli meliputi konten, konstruk dan bahasa. Tujuan dilakukan validasi para ahli dalam penelitian ini adalah menghasilkan butir soal yang valid, baik dari segi isi, materi maupun bahasa. Selain itu, validasi para ahli juga menghasilkan produk yang berkualitas baik dari segi format penilaian, skala penilaian dan kriteria penilaian. Oleh sebab itu, validasi para ahli dalam penelitian ini bertujuan untuk memperbaiki soal tes diagnostik pilihan ganda *three-tier*, kisi-kisi soal tes diagnostik pilihan ganda *three-tier*, kunci jawaban beserta pedoman penskoran. Berikut nama validator beserta kode validator.

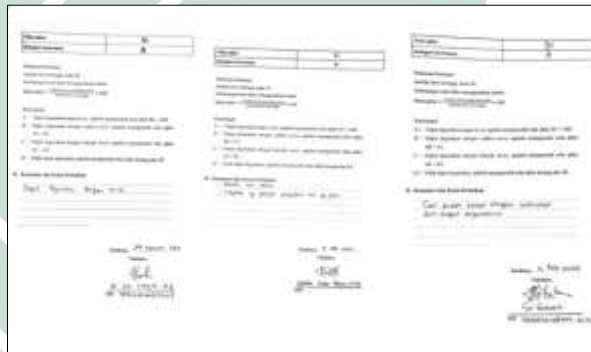
Tabel 4.1
Daftar Nama Validator Tes Diagnostik Pilihan Ganda
***Three-Tier* Beserta Kodenya**

No.	Nama Validator	Kode Validator
1.	Dr. Siti Lailiyah, M.Si	V1
2.	Novita Vindri Harini, M.Pd	V2
3.	Sri Suharti, S.Pd	V3

Para ahli tersebut memberikan penilaian untuk menentukan apakah tes diagnostik pilihan ganda *three-tier* untuk mengidentifikasi miskonsepsi pada materi geometri yang telah dikembangkan ini layak digunakan, layak digunakan dengan perbaikan, atau tidak layak digunakan untuk aspek yang telah

ditentukan yaitu dari aspek konten, aspek konstruk, dan aspek penggunaan bahasa.

Dalam penelitian ini, proses validasi para ahli dilaksanakan selama 3 minggu dengan validator yang berkompeten dan mengerti tentang pengembangan tes diagnostik pilihan ganda *three-tier* untuk mengidentifikasi miskonsepsi siswa pada materi geometri serta mampu memberikan komentar/saran untuk memperbaiki dan menyempurnakan tes diagnostik yang telah dikembangkan. Adapun bukti komentar/saran dari validator ahli dapat dilihat pada Gambar 4.1 berikut.



Gambar 4.1
Bukti Komentar/Saran Validator Ahli

Untuk hasil validasi menurut penilaian para ahli terlampir pada Lampiran 4.5, sedangkan penilaian hasil validasi penilaian menurut para ahli secara keseluruhan sebagai berikut.

Tabel 4.2
Data Hasil Validasi Para Ahli Tes Diagnostik Pilihan Ganda *Three-Tier*

No.	Aspek Penilaian	Kategori	V1	V2	V3
1.	Konten	Soal tes diagnostik yang dibuat sesuai dengan indikator pembelajaran.	4	3	4

		Soal tes diagnostik yang dibuat sesuai dengan materi pembelajaran geometri tentang segiempat di SMP/MTs/Sederajat.	4	3	4
		Butir-butir soal tes diagnostik dibuat untuk mengidentifikasi miskonsepsi siswa pada materi segiempat.	2	4	4
2.	Konstruksi	Soal tes diagnostik tidak memberikan petunjuk ke arah jawaban yang benar.	4	3	4
		Gambar atau tabel jelas dan berfungsi.	3	3	4
		Menggunakan petunjuk pengerjaan soal secara jelas dan runtut.	4	3	4
		Waktu pengerjaan soal adalah 90 menit.	3	3	3
		Pilihan jawaban yang disajikan homogen dan logis dari segi materi.	4	3	2
		Pilihan jawaban tidak mengandung pernyataan “semua jawaban benar” atau “semua jawaban salah”	4	4	4
		Pilihan alasan yang disajikan homogen	3	3	4

		dan logis dari segi materi.			
		Pilihan alasan bersesuaian dengan pilihan jawaban.	4	4	4
3.	Bahasa	Menggunakan Bahasa Indonesia yang baik dan benar.	4	3	4
		Menggunakan bahasa yang sesuai dengan taraf berpikir siswa SMP.	4	4	4
		Tidak menggunakan kata atau ungkapan yang menimbulkan salah paham atau penafsiran ganda.	4	3	4

Berdasarkan Tabel 4.2 dapat diketahui bahwa tes diagnostik yang dikembangkan melewati proses revisi dan validasi sebanyak 3 kali oleh dua dosen pendidikan matematika dan guru matematika. Validasi para ahli menilai 3 aspek dalam penelitian yaitu konten, konstruksi, bahasa. Validator para ahli memberikan penilaian berbeda-beda pada setiap aspek sehingga pada aspek konten diperoleh dengan jumlah skor 32, aspek konstruksi diperoleh jumlah skor 84, sedangkan untuk aspek bahasa diperoleh jumlah skor 34.

Selain penilaian kelayakan terhadap tes diagnostik pilihan ganda *three-tier* untuk mengidentifikasi miskonsepsi pada materi geometri di atas, validator juga memberikan komentar atau saran yang akan dijadikan sebagai bahan revisi tes diagnostik tersebut.

b. Data Validasi Empirik

Validitas empiris digunakan apabila sudah melakukan uji coba lapangan. Hasil uji coba lapangan terhadap kevalidan empirik tes diagnostik *three-tier* dapat disajikan dalam Tabel 4.3 berikut.

Tabel 4.3
Data Hasil Validasi Empirik Tes Diagnostik Pilihan Ganda
Three-Tier

No.	Kode Siswa	Skor Tiap Butir Soal									
		1	2	3	4	5	6	7	8	9	10
1	SB-1	1	1	0	0	1	0	1	0	1	0
2	SB-2	0	0	0	1	0	0	0	0	0	1
3	SB-3	0	1	0	1	1	1	1	1	1	1
4	SB-4	0	0	0	1	0	1	0	1	0	0
5	SB-5	0	1	0	1	0	1	1	1	0	0
6	SB-6	0	1	0	1	1	1	1	1	1	1
7	SB-7	0	0	0	0	0	1	1	0	0	0
8	SB-8	0	0	0	0	0	1	1	1	1	1
9	SB-9	1	0	0	0	0	1	1	1	1	1
10	SB-10	0	1	0	1	0	1	1	0	1	0
11	SB-11	1	1	0	0	0	0	1	1	1	0
12	SB-12	0	1	0	1	0	0	0	0	0	0
13	SB-13	0	1	0	1	0	0	0	0	0	0
14	SB-14	0	0	0	1	0	0	1	0	0	0
15	SB-15	1	1	0	1	0	0	0	1	1	1
16	SB-16	0	0	0	1	0	0	0	0	1	0
17	SB-17	0	1	0	1	0	1	1	1	1	0
18	SB-18	0	0	0	1	0	0	0	1	0	0
19	SB-19	0	1	0	0	1	0	0	1	1	1
20	SB-20	0	1	0	0	0	1	0	1	0	0
21	SB-21	0	1	0	0	0	0	0	0	0	0
22	SB-22	0	1	0	0	0	0	1	0	0	0
23	SB-23	1	0	0	0	0	1	1	1	1	0
24	SB-24	0	0	0	0	0	0	1	0	1	0
25	SB-25	0	0	0	0	0	0	1	0	0	0
26	SB-26	0	0	0	0	0	0	0	0	1	0
27	SB-27	0	0	0	1	0	0	1	0	0	0
28	SB-28	0	1	0	0	1	0	0	0	0	0
29	SB-29	0	0	0	1	0	0	0	0	1	0
30	SB-30	0	1	0	0	0	0	1	0	1	0
31	SB-31	0	1	0	1	0	0	1	0	0	1
32	SB-32	0	0	0	0	0	0	0	1	0	0

33	SB-33	0	0	0	0	0	0	0	1	0	0
34	SB-34	0	0	0	1	0	1	1	0	1	0
35	SB-35	0	1	0	1	0	1	1	1	1	0
36	SB-36	1	1	0	1	0	0	1	0	1	0
37	SB-37	0	0	0	0	0	0	0	0	1	0
38	SB-38	0	0	0	1	0	0	0	0	0	0
39	SB-39	0	0	0	1	0	0	0	0	0	0
40	SB-40	1	1	0	1	0	0	1	0	1	0
r_{hitung}		0,44	0,57	0	0,23	0,41	0,58	0,55	0,53	0,60	0,52

Pada Tabel 4.3 dapat dilihat bahwa data hasil validasi empirik berupa jawaban dari 40 subjek penelitian menghasilkan nilai r_{hitung} berbeda-beda pada setiap butir soal dengan menggunakan rumus korelasi biserial berbantuan *SPSS Statistic Versi 16.0*. Pada setiap butir soal mendapatkan penilaian tertentu yakni siswa mendapat skor 1 apabila jawaban benar dan alasan benar sedangkan siswa mendapat skor 0 apabila jawaban salah atau alasan salah atau tidak memberikan jawaban. Validitas empiris bertujuan untuk menentukan butir soal tes diagnostik pilihan ganda *three-tier* yang dikembangkan valid dan tidak valid terhadap hasil uji coba lapangan.

2. Deskripsi Data Reliabilitas Tes Diagnostik Pilihan Ganda *Three-Tier* Untuk Mengidentifikasi Miskonsepsi Pada Materi Geometri

Data reliabilitas ini diperoleh dari data hasil tes diagnostik pilihan ganda *three-tier*. Hasil uji reliabilitas terhadap tes diagnostik pilihan ganda *three-tier* yang dikembangkan disajikan dalam Tabel 4.4 sebagai berikut.

Tabel 4.4
Data Hasil Reliabilitas Tes Diagnostik Pilihan Ganda *Three-Tier*

No.	Kode Siswa	Skor Tiap Butir Soal										Skor Total
		1	2	3	4	5	6	7	8	9	10	
1	SB-1	1	1	0	0	1	0	1	0	1	0	5
2	SB-2	0	0	0	1	0	0	0	0	0	1	2
3	SB-3	0	1	0	1	1	1	1	1	1	1	8
4	SB-4	0	0	0	1	0	1	0	1	0	0	3
5	SB-5	0	1	0	1	0	1	1	1	0	0	5

6	SB-6	0	1	0	1	1	1	1	1	1	1	8
7	SB-7	0	0	0	0	0	1	1	0	0	0	2
8	SB-8	0	0	0	0	0	1	1	1	1	1	5
9	SB-9	1	0	0	0	0	1	1	1	1	1	6
10	SB-10	0	1	0	1	0	1	1	0	1	0	5
11	SB-11	1	1	0	0	0	0	1	1	1	0	5
12	SB-12	0	1	0	1	0	0	0	0	0	0	2
13	SB-13	0	1	0	1	0	0	0	0	0	0	2
14	SB-14	0	0	0	1	0	0	1	0	0	0	2
15	SB-15	1	1	0	1	0	0	0	1	1	1	6
16	SB-16	0	0	0	1	0	0	0	0	1	0	2
17	SB-17	0	1	0	1	0	1	1	1	1	0	6
18	SB-18	0	0	0	1	0	0	0	1	0	0	2
19	SB-19	0	1	0	0	1	0	0	1	1	1	5
20	SB-20	0	1	0	0	0	1	0	1	0	0	3
21	SB-21	0	1	0	0	0	0	0	0	0	0	1
22	SB-22	0	1	0	0	0	0	1	0	0	0	2
23	SB-23	1	0	0	0	0	1	1	1	1	0	5
24	SB-24	0	0	0	0	0	0	1	0	1	0	2
25	SB-25	0	0	0	0	0	0	1	0	0	0	1
26	SB-26	0	0	0	0	0	0	0	0	1	0	1
27	SB-27	0	0	0	1	0	0	1	0	0	0	2
28	SB-28	0	1	0	0	1	0	0	0	0	0	2
29	SB-29	0	0	0	1	0	0	0	0	1	0	2
30	SB-30	0	1	0	0	0	0	1	0	1	0	3
31	SB-31	0	1	0	1	0	0	1	0	0	1	4
32	SB-32	0	0	0	0	0	0	0	1	0	0	1
33	SB-33	0	0	0	0	0	0	0	1	0	0	1
34	SB-34	0	0	0	1	0	1	1	0	1	0	4
35	SB-35	0	1	0	1	0	1	1	1	1	0	6
36	SB-36	1	1	0	1	0	0	1	0	1	0	5
37	SB-37	0	0	0	0	0	0	0	0	1	0	1
38	SB-38	0	0	0	1	0	0	0	0	0	0	1
39	SB-39	0	0	0	1	0	0	0	0	0	0	1
40	SB-40	1	1	0	1	0	0	1	0	1	0	5

Dari Tabel 4.4 dapat dilihat bahwa data hasil reliabilitas berupa jawaban dari 40 subjek penelitian menghasilkan skor total berbeda-beda pada setiap butir soal dengan menggunakan rumus KR.20 berbantuan *SPSS Statistic Versi 16.0*. Pada setiap butir soal mendapatkan penilaian tertentu yakni siswa mendapat skor 1 apabila jawaban benar dan alasan benar sedangkan siswa mendapat skor 0 apabila jawaban salah atau alasan salah atau tidak memberikan jawaban. Data tersebut digunakan sebagai bahan untuk mengetahui kualitas soal atau konsistensi soal tes diagnostik pilihan ganda *three-tier*.

3. Deskripsi Data Kualitas Butir Soal Tes Diagnostik Pilihan Ganda *Three-Tier* Untuk Mengidentifikasi Miskonsepsi Pada Materi Geometri

a. Data Daya Pembeda

Daya pembeda soal digunakan untuk mengetahui kemampuan suatu butir soal yang dapat membedakan antara siswa yang memiliki kemampuan tinggi dengan siswa yang memiliki kemampuan rendah. Daya pembeda soal dalam penelitian ini bertujuan untuk mengetahui seberapa jauh setiap butir soal dapat membedakan antara siswa yang telah menguasai materi dengan siswa yang belum menguasai materi. Indeks yang digunakan dalam membedakan antara siswa yang berkemampuan tinggi dengan siswa yang berkemampuan rendah adalah indeks diskriminasi. Indeks ini ditetapkan dari selisih proporsi yang menjawab dari masing-masing kelompok yakni kelompok atas dan kelompok bawah. Adapun hasil daya pembeda soal tes diagnostik pilihan ganda *three-tier* dapat dilihat pada Tabel 4.5.

Tabel 4.5
Data Hasil Daya Pembeda Soal Tes Diagnostik Pilihan Ganda
Three-Tier

No.	Kode Siswa	Skor Tiap Butir Soal									
		1	2	3	4	5	6	7	8	9	10
1	SB-3	0	1	0	1	1	1	1	1	1	1
2	SB-6	0	1	0	1	1	1	1	1	1	1
3	SB-9	1	0	0	0	0	1	1	1	1	1
4	SB-15	1	1	0	1	0	0	0	1	1	1
5	SB-17	0	1	0	1	0	1	1	1	1	0
6	SB-35	0	1	0	1	0	1	1	1	1	0

7	SB-1	1	1	0	0	1	0	1	0	1	0
8	SB-5	0	1	0	1	0	1	1	1	0	0
9	SB-8	0	0	0	0	0	1	1	1	1	1
10	SB-10	0	1	0	1	0	1	1	0	1	0
11	SB-11	1	1	0	0	0	0	1	1	1	0
12	SB-19	0	1	0	0	1	0	0	1	1	1
13	SB-23	1	0	0	0	0	1	1	1	1	0
14	SB-36	1	1	0	1	0	0	1	0	1	0
15	SB-40	1	1	0	1	0	0	1	0	1	0
16	SB-31	0	1	0	1	0	0	1	0	0	1
17	SB-34	0	0	0	1	0	1	1	0	1	0
18	SB-4	0	0	0	1	0	1	0	1	0	0
19	SB-20	0	1	0	0	0	1	0	1	0	0
20	SB-30	0	1	0	0	0	0	1	0	1	0
Kelompok Atas		0,35	0,75	0	0,6	0,2	0,6	0,8	0,65	0,8	0,35
21	SB-2	0	0	0	1	0	0	0	0	0	1
22	SB-7	0	0	0	0	0	1	1	0	0	0
23	SB-12	0	1	0	1	0	0	0	0	0	0
24	SB-13	0	1	0	1	0	0	0	0	0	0
25	SB-14	0	0	0	1	0	0	1	0	0	00
26	SB-16	0	0	0	1	0	0	0	0	1	0
27	SB-18	0	0	0	1	0	0	0	1	0	0
28	SB-22	0	1	0	0	0	0	1	0	0	0
29	SB-24	0	0	0	0	0	0	1	0	1	0
30	SB-27	0	0	0	1	0	0	1	0	0	0
31	SB-28	0	1	0	0	1	0	0	0	0	0
32	SB-39	0	0	0	1	0	0	0	0	1	0
33	SB-21	0	1	0	0	0	0	0	0	0	0
34	SB-25	0	0	0	0	0	0	1	0	0	0
35	SB-26	0	0	0	0	0	0	0	0	1	0
36	SB-32	0	0	0	0	0	0	0	1	0	0
37	SB-33	0	0	0	0	0	0	0	1	0	0
38	SB-37	0	0	0	0	0	0	0	0	1	0
39	SB-38	0	0	0	1	0	0	0	0	0	0
40	SB-39	0	0	0	1	0	0	0	0	0	0
Kelompok Bawah		0	0,25	0	0,5	0,05	0,05	0,3	0,15	0,25	0,05
Daya Pembeda		0,35	0,5	0	0,1	0,15	0,55	0,5	0,5	0,55	0,3

Tabel 4.5 menunjukkan bahwa data hasil daya pembeda terbagi menjadi dua kelompok yakni kelompok atas (kelompok yang memiliki kemampuan tinggi) dan kelompok bawah (kelompok yang memiliki kemampuan rendah). Data tersebut didapatkan dengan mengurutkan skor siswa tertinggi ke skor siswa terendah dengan menggunakan rumus pada BAB III berbantuan *SPSS Statistic Versi 16.0*. Pada setiap butir soal mendapatkan penilaian tertentu yakni siswa mendapat skor 1 apabila jawaban benar dan alasan benar sedangkan siswa mendapat skor 0 apabila jawaban salah atau alasan salah atau tidak memberikan jawaban.

b. Data Tingkat Kesukaran

Data tingkat kesukaran diperoleh setelah melakukan uji lapangan. Tingkat kesukaran digunakan untuk mengetahui apakah butir soal termasuk butir soal yang mudah, sedang, atau sukar. Butir soal yang baik adalah butir soal yang tidak terlalu mudah atau tidak terlalu sulit. Tingkat kesukaran dihitung melalui indeks kesukaran atau *difficulty index* yaitu angka yang menunjukkan proporsi siswa yang menjawab benar soal tersebut. Hasil dari tingkat kesukaran pada tes diagnostik pilihan ganda *three-tier* disajikan dalam Tabel 4.6 di bawah ini.

Tabel 4.6

**Data Hasil Tingkat Kesukaran Tes Diagnostik Pilihan Ganda
*Three-Tier***

No.	Kode Siswa	Skor Tiap Butir Soal									
		1	2	3	4	5	6	7	8	9	10
1	SB-1	1	1	0	0	1	0	1	0	1	0
2	SB-2	0	0	0	1	0	0	0	0	0	1
3	SB-3	0	1	0	1	1	1	1	1	1	1
4	SB-4	0	0	0	1	0	1	0	1	0	0
5	SB-5	0	1	0	1	0	1	1	1	0	0
6	SB-6	0	1	0	1	1	1	1	1	1	1
7	SB-7	0	0	0	0	0	1	1	0	0	0
8	SB-8	0	0	0	0	0	1	1	1	1	1
9	SB-9	1	0	0	0	0	1	1	1	1	1
10	SB-10	0	1	0	1	0	1	1	0	1	0
11	SB-11	1	1	0	0	0	0	1	1	1	0

12	SB-12	0	1	0	1	0	0	0	0	0	0
13	SB-13	0	1	0	1	0	0	0	0	0	0
14	SB-14	0	0	0	1	0	0	1	0	0	0
15	SB-15	1	1	0	1	0	0	0	1	1	1
16	SB-16	0	0	0	1	0	0	0	0	1	0
17	SB-17	0	1	0	1	0	1	1	1	1	0
18	SB-18	0	0	0	1	0	0	0	1	0	0
19	SB-19	0	1	0	0	1	0	0	1	1	1
20	SB-20	0	1	0	0	0	1	0	1	0	0
21	SB-21	0	1	0	0	0	0	0	0	0	0
22	SB-22	0	1	0	0	0	0	1	0	0	0
23	SB-23	1	0	0	0	0	1	1	1	1	0
24	SB-24	0	0	0	0	0	0	1	0	1	0
25	SB-25	0	0	0	0	0	0	1	0	0	0
26	SB-26	0	0	0	0	0	0	0	0	1	0
27	SB-27	0	0	0	1	0	0	1	0	0	0
28	SB-28	0	1	0	0	1	0	0	0	0	0
29	SB-29	0	0	0	1	0	0	0	0	1	0
30	SB-30	0	1	0	0	0	0	1	0	1	0
31	SB-31	0	1	0	1	0	0	1	0	0	1
32	SB-32	0	0	0	0	0	0	0	1	0	0
33	SB-33	0	0	0	0	0	0	0	1	0	0
34	SB-34	0	0	0	1	0	1	1	0	1	0
35	SB-35	0	1	0	1	0	1	1	1	1	0
36	SB-36	1	1	0	1	0	0	1	0	1	0
37	SB-37	0	0	0	0	0	0	0	0	1	0
38	SB-38	0	0	0	1	0	0	0	0	0	0
39	SB-39	0	0	0	1	0	0	0	0	0	0
40	SB-40	1	1	0	1	0	0	1	0	1	0
<i>B</i>		7	22	0	22	5	13	22	16	21	8
<i>P</i>		0,1 75	0, 55	0	0,5 5	0,1 25	0, 32 5	0, 55	0, 40	0, 52 5	0, 20

Keterangan:

B = Banyaknya Subjek yang Menjawab Benar

P = *Difficulty Index*

Berdasarkan Tabel 4.6 menunjukkan bahwa data hasil tingkat kesukaran memiliki *difficultuty index* (indeks kesukaran) dan jumlah subjek menjawab benar yang berbeda-beda pada tiap butir soal. Pada setiap butir soal mendapatkan penilaian tertentu yakni siswa mendapat skor 1 apabila jawaban benar dan alasan benar sedangkan siswa mendapat skor 0 apabila jawaban salah atau alasan salah atau tidak memberikan jawaban. Data hasil tingkat kesukaran tersebut didapatkan dengan menggunakan rumus pada BAB III berbantuan *Microsoft excel*.

c. Data Pengecoh atau Distraktor

Data pengecoh atau distraktor diperoleh setelah melakukan uji lapangan. Dalam penelitian ini pengecoh atau distraktor terbagi menjadi dua bagian, yakni pengecoh pilihan jawaban dan pengecoh pilihan alasan. Pengecoh atau distraktor tidak sekedar pelengkap opsi jawaban, namun sengaja dibuat untuk menyesatkan atau mengecoh siswa dalam menjawab soal tersebut. Hasil dari pengecoh pilihan jawaban tes diagnostik pilihan ganda *three-tier* dapat dilihat pada Tabel 4.7 di bawah ini.

Tabel 4.7
Data Hasil Pengecoh Pilihan Jawaban Tes Diagnostik
Pilihan Ganda *Three-Tier*

Kode Siswa	Pilihan Jawaban Tiap Butir Soal									
	1	2	3	4	5	6	7	8	9	10
SB-1	C	A	B	A	D	B	D	B	C	B
SB-2	B	D	B	C	0	C	B	A	C	A
SB-3	B	C	D	0	D	B	D	0	A	C
SB-4	B	B	B	C	0	B	D	0	D	0
SB-5	0	C	B	C	D	0	D	0	C	C
SB-6	0	D	0	C	D	B	D	C	A	A
SB-7	0	C	C	C	D	B	D	B	B	C
SB-8	A	D	B	B	0	B	D	D	0	C
SB-9	C	B	B	0	0	0	D	0	A	0
SB-10	A	C	B	0	A	B	D	C	C	B
SB-11	C	D	0	C	0	D	D	0	0	B

SB-12	D	D	0	A	0	0	C	0	A	A
SB-13	A	0	B	C	A	C	C	0	A	0
SB-14	A	B	0	0	0	0	D	C	0	B
SB-15	C	0	B	D	A	B	B	B	D	0
SB-16	0	0	0	C	0	A	D	C	0	C
SB-17	0	0	B	C	C	B	D	0	0	C
SB-18	0	A	B	D	D	A	A	0	B	0
SB-19	0	0	0	0	D	D	D	B	0	0
SB-20	0	A	B	A	D	0	C	B	0	0
SB-21	0	0	B	0	D	D	A	C	0	0
SB-22	A	B	B	C	C	B	D	C	0	B
SB-23	C	A	B	C	C	0	D	B	B	0
SB-24	0	0	0	D	C	0	D	C	C	B
SB-25	D	0	B	C	B	0	D	0	B	D
SB-26	0	B	B	0	C	C	D	C	0	C
SB-27	0	A	0	0	0	C	D	C	D	C
SB-28	0	0	0	D	D	C	B	A	0	C
SB-29	0	0	B	0	B	0	B	D	C	0
SB-30	0	0	0	B	0	C	D	C	0	0
SB-31	0	A	B	D	0	C	D	C	C	D
SB-32	0	0	0	0	0	0	D	B	0	C
SB-33	0	0	0	0	B	B	D	B	D	C
SB-34	A	A	0	0	0	B	D	C	C	B
SB-35	0	0	B	0	0	D	D	B	0	B
SB-36	A	0	B	A	0	0	D	0	D	0
SB-37	A	0	0	0	0	0	D	0	C	0
SB-38	D	0	B	0	0	B	D	0	C	0
SB-39	A	A	B	C	B	C	0	B	C	0
SB-40	0	0	B	B	0	C	D	0	0	0

Keterangan: Warna biru = Kunci jawaban

Dari Tabel 4.7 didapatkan data hasil pengecoh jawaban dengan subjek penelitian berjumlah 40 siswa. Data pengecoh jawaban berada pada tingkat pertama dalam tes diagnostik pilihan ganda *three-tier* yang mana pada tingkat tersebut siswa harus memilih jawaban yang benar dan apabila siswa tidak menjawab maka mendapatkan nilai 0. Distraktor yang baik harus sejenis atau dibuat semirip mungkin dengan kunci jawaban sedangkan distraktor yang buruk adalah distraktor yang sama sekali tidak dipilih oleh siswa. Dari data hasil pengecoh jawaban tersebut dapat dilihat bahwa hampir keseluruhan butir soal tes diagnostik yang dikembangkan pengecoh jawaban dapat berfungsi dengan baik. Sedangkan hasil dari pengecoh pilihan alasan tes diagnostik pilihan ganda *three-tier* dapat dilihat pada Tabel 4.8 berikut ini.

Tabel 4.8
Data Hasil Pengecoh Pilihan Alasan Tes Diagnostik
Pilihan Ganda *Three-Tier*

Kode Siswa	Pilihan Alasan Tiap Butir Soal									
	1	2	3	4	5	6	7	8	9	10
SB-1	a	d	a	b	d	d	a	b	b	d
SB-2	a	d	a	b	c	d	d	a	b	d
SB-3	0	0	c	b	d	c	c	a	b	d
SB-4	a	d	a	c	c	0	a	a	b	b
SB-5	d	d	d	b	b	c	c	a	a	d
SB-6	a	0	c	b	d	0	c	a	b	d
SB-7	a	a	a	d	b	0	c	0	b	d
SB-8	b	c	c	d	c	c	0	0	b	d
SB-9	0	c	d	d	c	c	a	a	b	d
SB-10	0	d	a	0	c	c	0	b	b	0
SB-11	a	0	0	a	c	d	c	a	b	d
SB-12	0	d	c	b	c	c	c	b	b	a
SB-13	c	d	d	c	a	d	a	b	b	d
SB-14	b	c	c	b	a	c	0	b	b	0
SB-15	0	d	a	0	c	d	0	a	b	d
SB-16	a	0	0	b	c	0	a	c	b	b

SB-17	0	d	c	0	c	0	c	a	b	b
SB-18	0	b	c	b	c	b	d	d	b	b
SB-19	a	d	0	0	d	c	0	0	b	0
SB-20	a	0	c	c	d	c	c	a	0	0
SB-21	0	d	0	d	c	b	d	b	a	b
SB-22	c	0	a	a	c	c	c	0	b	c
SB-23	d	b	a	b	c	0	0	a	b	d
SB-24	a	0	a	b	0	0	0	b	b	b
SB-25	c	a	d	b	a	a	c	0	b	0
SB-26	0	c	a	b	c	c	a	b	b	c
SB-27	a	d	0	0	c	0	c	0	b	b
SB-28	a	0	a	b	d	c	c	c	b	b
SB-29	0	d	0	b	b	a	b	0	b	b
SB-30	d	d	d	0	c	d	0	c	b	b
SB-31	0	c	a	b	c	c	c	0	b	b
SB-32	0	c	a	b	c	0	a	c	b	d
SB-33	a	c	0	a	c	c	0	a	b	d
SB-34	c	c	c	b	a	0	0	0	b	d
SB-35	0	d	0	0	c	0	a	a	b	0
SB-36	0	0	d	b	0	c	b	0	b	0
SB-37	c	c	a	0	0	c	0	d	b	d
SB-38	c	0	c	0	b	c	a	0	b	d
SB-39	b	c	a	b	c	0	0	d	c	d
SB-40	0	0	a	a	c	0	a	0	0	0

Keterangan: Warna biru = Kunci jawaban

Berdasarkan Tabel 4.8 menunjukkan data hasil pengecoh alasan dengan subjek penelitian berjumlah 40 siswa. Data pengecoh alasan berada pada tingkat kedua dalam tes diagnostik pilihan ganda *three-tier* yang mana pada tingkat tersebut siswa harus memilih alasan yang benar dan apabila siswa tidak menjawab maka mendapatkan nilai 0. Distraktor yang baik harus sejenis atau dibuat semirip mungkin dengan kunci jawaban sedangkan distraktor yang buruk adalah

distraktor yang sama sekali tidak dipilih oleh siswa. Dari data hasil pengecoh alasan tersebut dapat dilihat bahwa beberapa pengecoh alasan pada setiap butir soal tes diagnostik yang dikembangkan dapat berfungsi dengan baik.

4. Deskripsi Data Hasil Tes Diagnostik Pilihan Ganda *Three-Tier* Untuk Mengidentifikasi Miskonsepsi Pada Materi Geometri

Data hasil tes diagnostik pilihan ganda *three-tier* yang dikembangkan ini diperoleh setelah melakukan uji lapangan. Dalam penelitian ini hasil tes diagnostik yang dikembangkan terdapat tiga tingkatan. Tingkatan pertama berupa pilihan jawaban, tingkatan kedua berupa pilihan alasan, dan tingkatan ketiga berupa kolom CRI yang menunjukkan seberapa yakin siswa dalam menjawab butir soal tersebut dengan skala 4. Adapun data hasil tes diagnostik pilihan ganda *three-tier* dapat dilihat pada Lampiran 4.6.

Pada data hasil tes diagnostik pilihan ganda *three-tier* terdapat 3 kolom dalam setiap butir soal yaitu kolom jawaban (J), kolom alasan (A), serta kolom CRI yang mana dalam setiap kolom berisikan jawaban dan alasan yang dipilih siswa dan tingkat keyakinan siswa dalam menjawab butir soal tersebut. Penilaian skor 1 diberikan apabila jawaban benar atau alasan benar sedangkan siswa mendapat skor 0 apabila jawaban salah atau alasan salah atau tidak memberikan jawaban. Untuk penilaian tingkat keyakinan apabila siswa memilih CRI dengan skala 1 atau 2 maka termasuk kategori rendah sedangkan apabila siswa memilih CRI dengan skala 3 atau 4 maka termasuk kategori tinggi.

B. Analisis Data

1. Analisis Data Kevalidan Tes Diagnostik *Three-Tier* Untuk Mengidentifikasi Miskonsepsi Pada Materi Geometri

a. Analisis Data Validasi Para Ahli (*Expert Judgments*)

Sebuah tes diagnostik dapat diuji cobakan ketika sudah divalidasi oleh validator. Oleh karena itu, tes diagnostik pilihan ganda *three-tier* untuk mengidentifikasi miskonsepsi siswa pada materi bangun datar segiempat yang dikembangkan harus divalidasi terlebih dahulu supaya dapat diuji cobakan di lapangan. Berdasarkan deskripsi data validitas, diperoleh informasi penilaian validator terhadap empat belas aspek yang dinilai dalam tes diagnostik pilihan ganda *three-tier* untuk mengidentifikasi miskonsepsi pada materi geometri yang

dikembangkan. Berdasarkan data Tabel 4.9, maka dapat dianalisis sebagai berikut.

Tabel 4.9
Hasil Analisis Data Validasi Para Ahli

No.	Aspek Penilaian	Kategori	RK	RA
1	Konten	Soal tes diagnostik yang dibuat sesuai dengan indikator pembelajaran.	3,67	3,56
		Soal tes diagnostik yang dibuat sesuai dengan materi pembelajaran geometri tentang segiempat di SMP/MTs/Sederajat.	3,67	
		Butir-butir soal tes diagnostik dibuat untuk mengidentifikasi miskonsepsi siswa pada materi segiempat.	3,33	
2	Konstruksi	Soal tes diagnostik tidak memberikan petunjuk ke arah jawaban yang benar.	3,67	3,50
		Gambar atau tabel jelas dan berfungsi.	3,33	
		Menggunakan petunjuk pengerjaan soal secara jelas dan runtut.	3,67	
		Waktu pengerjaan soal adalah 90 menit.	3,00	
		Pilihan jawaban yang disajikan homogen dan logis dari segi materi.	3,00	
		Pilihan jawaban tidak mengandung pernyataan “semua jawaban benar” atau “semua jawaban salah”	4,00	
		Pilihan alasan yang disajikan homogen dan logis dari segi materi.	3,33	

		Pilihan alasan bersesuaian dengan pilihan jawaban.	4,00	
3	Bahasa	Menggunakan Bahasa Indonesia yang baik dan benar.	3,67	3,78
		Menggunakan bahasa yang sesuai dengan taraf berpikir siswa SMP.	4,00	
		Tidak menggunakan kata atau ungkapan yang menimbulkan salah paham atau penafsiran ganda.	3,67	
Rata-Rata Total Validitas (RTV) Tes Diagnostik <i>Three-Tier</i>			3,61	

Keterangan:

RK : Rata-rata tiap kategori

RA : Rata-rata tiap aspek

Berdasarkan hasil analisis data pada Tabel 4.3 diperoleh nilai rata-rata dari setiap aspek penilaian validasi tes diagnostik adalah sebagai berikut: 1) segi konten memperoleh rata-rata 3,56; 2) segi konstruksi memperoleh rata-rata 3,50; 3) segi bahasa memperoleh rata-rata 3,78, sehingga nilai rata-rata total validitas tes diagnostik dari ketiga validator adalah 3,61. Berdasarkan kategori kevalidan yang telah ditetapkan pada BAB III dalam Tabel 3.2 serta mencocokkan rata-rata total validitas yang didapat maka tes diagnostik pilihan ganda *three-tier* yang dikembangkan oleh peneliti ini termasuk dalam kategori valid.

b. Analisis Data Validasi Empirik

Uji validitas empiris digunakan untuk mengetahui valid tidaknya butir-butir soal tes. Butir soal yang valid mengartikan butir soal tersebut baik untuk digunakan tes sedangkan butir soal yang tidak valid mengartikan butir soal tersebut perlu diperbaiki atau ditolak. Pada penelitian ini untuk menguji validitas empiris menggunakan rumus koefisien korelasi poin biserial dengan berbantuan program *SPSS Statistic Versi 16.0*. Kriteria butir soal tes diagnostik pilihan ganda *three-tier* dikatakan valid apabila nilai korelasi berada di atas 0,400, sebaliknya jika nilai korelasi di bawah 0,400, maka tes

diagnostik tersebut dikatakan tidak valid, sehingga dapat diperoleh hasil analisis validitas empiris sebagaimana dalam Tabel 4.10.

Tabel 4.10
Hasil Analisis Validasi Empirik Butir Soal Tes Diagnostik
Pilihan Ganda *Three-Tier*

Butir Soal Nomor	Koefisien Korelasi	Kategori
1	0,44	Valid
2	0,57	Valid
3	0	Tidak valid
4	0,23	Tidak valid
5	0,41	Valid
6	0,58	Valid
7	0,55	Valid
8	0,53	Valid
9	0,60	Valid
10	0,52	Valid

Dari hasil analisis tersebut menunjukkan bahwa terdapat 8 butir soal yang valid yaitu butir soal nomor 1, 2, 5, 6, 7, 8, 9, 10 yang artinya kedelapan soal tersebut dapat digunakan untuk tes diagnostik pilihan ganda *three-tier* sedangkan butir soal yang tidak valid terdapat 2 butir soal yaitu soal nomor 3 dan 4 artinya kedua butir soal tersebut perlu diperbaiki atau tidak digunakan. Adapun untuk perhitungan secara lebih jelasnya dapat dilihat pada Lampiran 4.7 dan Lampiran 4.8.

2. Analisis Data Reliabilitas Tes Diagnostik *Three-Tier* Untuk Mengidentifikasi Miskonsepsi Pada Materi Geometri

Uji reliabilitas berdasarkan uji coba lapangan (*field test*) yang melibatkan kelas IX SMP Negeri 13 Surabaya dengan jumlah subjek lapangan sebanyak 40 siswa menghasilkan analisis reliabilitas tes diagnostik pilihan ganda *three-tier* dengan menggunakan rumus KR.20 dan berbantuan program *SPSS Statistic Versi 16.0*. Uji reliabilitas adalah pengujian suatu alat ukur untuk mengetahui sejauh mana pengukuran dapat memberikan hasil yang sama apabila alat ukur digunakan dalam satu kali tes dengan subjek

penelitian yang berbeda-beda. Sebelum melakukan uji reliabilitas, perlu dilihat terlebih dahulu apakah data yang ada sudah valid ataukah ada yang harus dikeluarkan. Hasil dari uji reliabilitas dengan menggunakan *SPSS Statistic Versi 16.0* adalah sebagai berikut.

Tabel 4.11
Hasil Analisis Reliabilitas Tes Diagnostik Pilihan Ganda *Three-Tier*

Case Processing Summary

		N	%
Cases	Valid	40	100.0
	Excluded ^a	0	.0
	Total	40	100.0

a. Listwise deletion based on all variables in the procedure.

Reliability Statistics

Cronbach's Alpha	Part 1	Value	.638
		N of Items	5 ^a
	Part 2	Value	.630
		N of Items	5 ^b
		Total N of Items	10
Correlation Between Forms			.612
Spearman-Brown Coefficient	Equal Length		.675
	Unequal Length		.675
Guttman Split-Half Coefficient			.641

a. The items are: X1, X2, X3, X4, X5.

b. The items are: X6, X7, X8, X9, X10.

Dari Tabel 4.11, dapat dilihat bahwa data yang valid berjumlah 40 dan tidak ada yang harus dikeluarkan. Selanjutnya, untuk menentukan suatu tes dikatakan reliabel atau tidak, dapat menggunakan batasan 0,600.¹³⁴ Dari tabel *Reliability Statistics* di atas menunjukkan bahwa nilai korelasi *Guttman Split-Half Coefficient* adalah sebesar 0,641. Artinya, reliabilitas tes diagnostik rata-rata memiliki reliabilitas yang baik. Untuk lebih jelasnya hasil dari analisis reliabilitas tes diagnostik *three-tier* dapat dilihat pada Lampiran 4.9.

3. Analisis Data Kualitas Butir Soal Tes Diagnostik *Three-Tier* Untuk Mengidentifikasi Miskonsepsi Pada Materi Geometri

a. Daya Pembeda

Butir-butir soal pada tes diagnostik pilihan ganda *three-tier* dapat dikatakan baik apabila butir-butir soal tersebut memiliki daya pembeda paling kecil adalah 0,200. Dalam hal ini menunjukkan bahwa butir soal memiliki daya pembeda minimal cukup. Daya pembeda pada butir soal yang dikembangkan diperoleh dari hasil tes diagnostik *three-tier* yang telah diuji coba lapangan (*field test*). Hasil analisis daya pembeda soal dalam tes diagnostik *three-tier* untuk mengidentifikasi miskonsepsi pada materi geometri adalah sebagai berikut.

Tabel 4.12
Hasil Analisis Daya Pembeda Soal Tes Diagnostik
Pilihan Ganda *Three-Tier*

Nomor Butir Soal	Indeks Daya Pembeda	Kategori
1	0,35	Cukup baik
2	0,50	Baik
3	0	Sangat jelek
4	0,1	Jelek
5	0,15	Jelek
6	0,55	Baik
7	0,50	Baik
8	0,50	Baik
9	0,55	Baik
10	0,30	Cukup baik

¹³⁴ Duwi Priyatno, *SPSS 22 Pengolah Data Terpraktis*, (Yogyakarta: Andi, 2014), hal. 64.

Tabel 4.12 menunjukkan bahwa daya pembeda butir soal tes diagnostik pilihan ganda *three-tier* yakni terdapat 1 butir soal yang memiliki daya pembeda soal sangat jelek dengan indeks korelasi 0,00, sebanyak 2 butir soal yang memiliki daya pembeda soal jelek dengan indeks korelasi rata-rata sebesar 0,125, sebanyak 2 butir soal yang memiliki daya pembeda soal cukup baik dengan indeks korelasi rata-rata sebesar 0,125 serta sebanyak 5 butir soal yang memiliki daya pembeda soal baik dengan indeks korelasi rata-rata sebesar 0,52. Untuk lebih jelasnya hasil data daya pembeda butir soal disajikan pada Lampiran 4.10. Dari hasil tersebut, butir soal yang memiliki daya pembeda soal yang cukup baik hendaknya dapat digunakan kembali pada tes yang akan datang atau dihimpun dalam kumpulan bank soal, sedangkan butir soal yang memiliki daya pembeda masih rendah maka sebaiknya ditelusuri untuk kemudian diperbaiki atau dibuang dan tidak digunakan dalam tes.

b. Tingkat Kesukaran

Butir soal tes diagnostik pilihan ganda *three-tier* dikatakan baik apabila butir soal tes tersebut memiliki tingkat kesukaran antara 0,31 – 0,70. Dalam hal ini menunjukkan bahwa butir soal tersebut tidak terlalu mudah dan tidak terlalu susah. Tingkat kesukaran didapat dari hasil pekerjaan subjek penelitian pada uji coba lapangan atau *field test*. Hasil analisis tingkat kesukaran pada tes diagnostik pilihan ganda *three-tier* untuk mengidentifikasi miskonsepsi dapat dilihat pada Tabel 4.13 berikut.

Tabel 4.13
Hasil Analisis Tingkat Kesukaran Tes Diagnostik
Pilihan Ganda *Three-Tier*

Nomor Butir Soal	Indeks Kesukaran	Kategori
1	0,175	Sukar
2	0,55	Sedang
3	0	Sangat sukar
4	0,55	Sedang
5	0,125	Sukar
6	0,325	Sedang

7	0,55	Sedang
8	0,40	Sedang
9	0,525	Sedang
10	0,20	Sukar

Berdasarkan Tabel 4.13 menunjukkan bahwa sebagian besar butir soal tes diagnostik pilihan ganda *three-tier* dapat dilihat bahwa sebanyak 6 butir soal dengan indeks korelasi rata-rata sebesar 0,48 berada pada kategori soal tingkat sedang, terdapat 3 butir soal dengan indeks korelasi rata-rata sebesar 0,17 yang termasuk kategori soal tingkat sukar dan hanya ada satu butir soal dengan indeks korelasi rata 0,00 yang termasuk kategori soal tingkat sangat sukar. Untuk lebih jelasnya hasil data tingkat kesukaran butir soal disajikan pada Lampiran 4.11.

Dalam kaitannya dengan hasil analisis tersebut dapat dilakukan beberapa tindakan lanjutan seperti 1) butir soal yang termasuk kategori sedang sebaiknya dapat digunakan kembali pada tes berikutnya; 2) butir soal yang termasuk kategori sukar atau sangat sukar ada 3 kemungkinan tindak lanjut yaitu dibuang dan tidak digunakan pada tes, ditelusuri kembali untuk dapat mengetahui penyebab banyaknya siswa yang tidak bisa menjawab pada butir soal tersebut, serta dapat digunakan pada tes kemampuan tinggi.

c. Pengecoh atau Distraktor

Tingkat pertama pada tes diagnostik pilihan ganda *three-tier* yakni berupa soal dengan empat pilihan jawaban, satu pilihan sebagai kunci jawaban dan tiga pilihan sebagai pengecoh atau distraktor. Sedangkan pada tingkat kedua pada tes diagnostik pilihan ganda *three-tier* berupa empat pilihan alasan dalam memilih jawaban, satu pilihan alasan sebagai kunci jawaban dan tiga pilihan alasan lainnya sebagai pengecoh atau distraktor. Pengecoh dibuat supaya dapat menarik perhatian siswa untuk menguji ketelitian mereka dalam memilih jawaban yang benar.

Kriteria pengecoh yang baik adalah apabila pengecoh tersebut dipilih oleh paling sedikit 5% dari semua siswa yang mengikuti tes. Pengecoh dikatakan berfungsi dengan baik apabila pengecoh tersebut memiliki daya tarik bagi siswa yang kurang menguasai materi. Jumlah pilihan pada tes diagnostik

pilihan ganda *three-tier* yang dikembangkan adalah sebanyak 40 opsi dari 10 butir soal pada pilihan jawaban dan 40 opsi dari pilihan alasan. Adapun deskripsi hasil analisis pengecoh atau distraktor pada tingkat pertama dapat dilihat pada Tabel 4.14 dan pada tingkat kedua dapat dilihat pada Tabel 4.15.

Tabel 4.14
Hasil Analisis Pengecoh Jawaban Tes Diagnostik Pilihan
Ganda *Three-Tier*

Nomor Butir Soal	Alternatif Jawaban				Total	Keterangan
	A	B	C	D		
1	9	3	*	3	15	Semua pengecoh berfungsi baik (dipilih $\geq 5\%$ dari peserta)
2	8	5	*	5	18	Semua pengecoh berfungsi baik (dipilih $\geq 5\%$ dari peserta)
3	0	24	1	*	25	Alternatif A dan C tidak berfungsi baik (dipilih $< 5\%$ dari peserta)
4	4	3	*	5	12	Semua pengecoh berfungsi baik (dipilih $\geq 5\%$ dari peserta)
5	3	4	5	*	12	Semua pengecoh berfungsi baik (dipilih $\geq 5\%$ dari peserta)
6	2	*	9	3	14	Semua pengecoh berfungsi baik (dipilih $\geq 5\%$ dari peserta)
7	3	3	3	*	9	Semua pengecoh berfungsi baik (dipilih $\geq 5\%$ dari peserta)
8	3	*	9	3	15	Semua pengecoh berfungsi baik (dipilih $\geq 5\%$ dari peserta)
9	5	4	*	5	14	Semua pengecoh berfungsi baik (dipilih

						≥ 5% dari peserta)
10	*	8	11	2	21	Semua pengecoh berfungsi baik (dipilih ≥ 5% dari peserta)

Keterangan :

* = Kunci jawaban

Tabel 4.15
Hasil Analisis Pengecoh Alasan Tes Diagnostik Pilihan
Ganda *Three-Tier*

Nomor Butir Soal	Alternatif Jawaban				Total	Keterangan
	A	B	C	D		
1	*	3	6	3	12	Semua pengecoh berfungsi baik (dipilih ≥ 5% dari peserta)
2	2	2	10	*	14	Semua pengecoh berfungsi baik (dipilih ≥ 5% dari peserta)
3	*	0	10	6	16	Alternatif B tidak berfungsi baik (dipilih < 5% dari peserta)
4	4	*	3	3	10	Semua pengecoh berfungsi baik (dipilih ≥ 5% dari peserta)
5	4	4	4	*	12	Semua pengecoh berfungsi baik (dipilih ≥ 5% dari peserta)
6	2	2	*	6	10	Semua pengecoh berfungsi baik (dipilih ≥ 5% dari peserta)
7	10	2	*	3	15	Semua pengecoh berfungsi baik (dipilih ≥ 5% dari peserta)
8	*	8	4	3	15	Semua pengecoh berfungsi baik (dipilih ≥ 5% dari peserta)

9	2	*	1	0	3	Alternatif C dan D tidak berfungsi baik (dipilih < 5% dari peserta)
10	1	11	2	*	14	Alternatif A tidak berfungsi baik (dipilih < 5% dari peserta)

Keterangan :

* = Kunci jawaban

Berdasarkan hasil analisis Tabel 4.14 dan Tabel 4.15 diketahui bahwa dari 40 pengecoh jawaban pada tingkat pertama ada 38 pengecoh yang dapat dipakai atau telah berfungsi baik dengan persentase 95% dikarenakan pengecoh tersebut dipilih paling sedikit 5% dari semua siswa yang mengikuti tes dan 2 pengecoh yang tidak berfungsi baik yakni pengecoh jawaban A dan C pada butir soal nomor 3 dengan persentase 5% dikarenakan pengecoh tersebut dipilih kurang dari 5% oleh siswa yang mengikuti tes. Sedangkan pada tingkat kedua yaitu pengecoh alasan menunjukkan bahwasanya dari 40 pengecoh alasan terdapat 37 pengecoh yang dapat dipakai atau telah berfungsi baik dengan persentase 92,5% dikarenakan pengecoh tersebut dipilih paling sedikit 5% dari siswa yang mengikuti tes dan 3 pengecoh alasan yang tidak berfungsi baik yakni pengecoh alasan pada butir soal nomor 9 dan 10 dengan persentase 7,5%, hal ini dikarenakan pengecoh tersebut dipilih kurang dari 5% oleh siswa yang mengikuti tes. Untuk lebih jelasnya hasil data pengecoh atau distraktor tes diagnostik pilihan ganda *three-tier* dapat disajikan pada Lampiran 4.12.

4. Analisis Data Hasil Tes Diagnostik Pilihan Ganda *Three-Tier* Untuk Mengidentifikasi Miskonsepsi Pada Materi Geometri

Tes yang digunakan dalam penelitian ini adalah tes diagnostik pilihan ganda tiga tingkat (*three-tier*). Pada tahap ini siswa mengerjakan tes pilihan ganda *three-tier* dengan memilih salah satu jawaban dan alasan yang tepat dari 4 pilihan jawaban pada tingkat pertama serta 4 pilihan alasan pada tingkat kedua, sedangkan pada tingkat ketiga mereka pun diharuskan untuk memilih tingkat keyakinan dalam menjawab soal tersebut.

Jumlah butir soal yang digunakan pada tahap uji coba lapangan (*field test*) sebanyak 10 soal. Skor yang didapatkan dilihat dari jumlah soal yang dijawab benar oleh siswa, dan persentase jumlah jawaban benar diperoleh dari jumlah jawaban benar dibagi banyak soal dikalikan seratus persen. Dari penelitian ini dihasilkan data jumlah jawaban benar siswa yang dapat dilihat pada Tabel 4.16 sebagai berikut.

Tabel 4.16
Data Jumlah Jawaban Benar yang Diperoleh Siswa

Kode Siswa	Jumlah Benar	Persentase Skor (%)
SB-1	5	50
SB-2	2	20
SB-3	8	80
SB-4	3	30
SB-5	5	50
SB-6	8	80
SB-7	2	20
SB-8	5	50
SB-9	6	60
SB-10	5	50
SB-11	5	50
SB-12	2	20
SB-13	2	20
SB-14	2	20
SB-15	6	60
SB-16	2	20
SB-17	6	60
SB-18	2	20
SB-19	5	50
SB-20	3	30
SB-21	1	10
SB-22	2	20
SB-23	5	50
SB-24	2	20
SB-25	1	10
SB-26	1	10
SB-27	2	20
SB-28	2	20

SB-29	2	20
SB-30	3	30
SB-31	4	40
SB-32	1	10
SB-33	1	10
SB-34	4	40
SB-35	6	60
SB-36	5	50
SB-37	1	10
SB-38	1	10
SB-39	1	10
SB-40	5	50

Berdasarkan Tabel 4.16 terlihat bahwa kemampuan siswa dalam menjawab soal tes diagnostik pilihan ganda *three-tier* berbeda-beda. Siswa yang menjawab benar paling rendah 1 butir soal sebanyak 8 orang dengan persentase mencapai 10%, sedangkan siswa yang menjawab benar paling tinggi sebanyak 8 butir soal sebanyak 2 orang dengan persentase 80%.

Siswa yang menjawab benar tertinggi belum dapat dikatakan memahami konsep, begitu juga dengan siswa yang menjawab benar terendah belum dapat dikatakan tidak memahami konsep, karenanya perlu dilihat apakah alasan yang mereka pilih pada setiap jawaban benar atau salah serta seberapa tingkat keyakinan mereka dalam menjawab soal, sehingga akan terlihat apakah siswa tersebut memahami konsep, miskonsepsi, atau tidak memahami konsep. Apabila setiap jawaban benar baik dari pilihan jawaban maupun pilihan alasan maka diberi skor 1, dan sebaliknya apabila jawaban salah baik dari pilihan jawaban maupun pilihan alasan maka diberi skor 0. Untuk lebih jelasnya hasil tes diagnostik *three-tier* untuk mengidentifikasi miskonsepsi pada materi geometri dapat disajikan pada Lampiran 4.13.

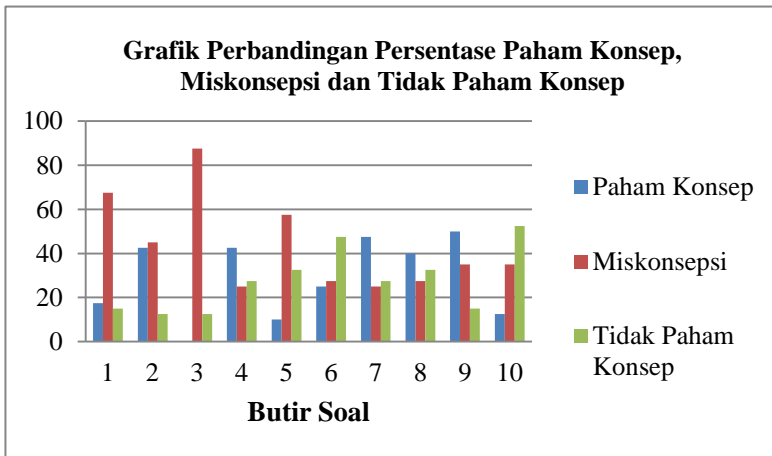
Persentase siswa yang paham konsep diambil dari jawaban dan alasan benar dengan tingkat keyakinan tinggi, persentase siswa yang mengalami miskonsepsi diambil dari jawaban benar-alasan salah/jawaban salah-alasan benar/jawaban salah-alasan salah dengan tingkat keyakinan tinggi, sedangkan persentase siswa yang tidak paham konsep diambil dari jawaban benar-alasan

benar/jawaban benar-alasan salah/jawaban salah-alasan
 benar/jawaban salah-alasan salah dengan tingkat keyakinan rendah.
 Lebih jelasnya dapat dilihat pada Tabel 4.17 berikut ini.

Tabel 4.17
Persentase Paham Konsep, Miskonsepsi, dan Tidak Paham
Konsep Berdasarkan Butir Soal

Butir Soal Nomor	Kriteria					
	Paham Konsep		Miskonsepsi		Tidak Paham Konsep	
	Frekuensi	%	Frekuensi	%	Frekuensi	%
1	7	17,5	27	67,5	6	15
2	17	42,5	18	45	5	12,5
3	0	0	35	87,5	5	12,5
4	17	42,5	10	25	11	27,5
5	4	10	23	57,5	13	32,5
6	10	25	11	27,5	19	47,5
7	19	47,5	10	25	11	27,5
8	16	40	11	27,5	13	32,5
9	20	50	14	35	6	15
10	5	12,5	14	35	21	52,5
Jumlah	115	287,5	173	432,5	110	275
Rata-rata	11,5	28,75	17,3	43,25	11	27,5

Berdasarkan Tabel 4.17 menunjukkan bahwa persentase siswa yang paham konsep dari hasil penelitian ini sebesar 28,75%, persentase siswa yang mengalami miskonsepsi sebesar 43,25%, sedangkan persentase siswa yang tidak paham konsep sebesar 27,5%. Dalam tabel tersebut terlihat bahwa persentase siswa yang paham konsep, miskonsepsi, dan tidak paham konsep pada setiap butir soal sangat bervariasi. Oleh karena itu, untuk lebih jelasnya dapat dilihat pada grafik berikut ini.



Gambar 4.2 Grafik Perbandingan Persentase Paham Konsep, Miskonsepsi, dan Tidak Paham Konsep

Dari Gambar 4.2 terlihat persentase miskonsepsi pada setiap butir soal sangat berbeda-beda. Untuk lebih jelasnya dalam melihat persentase miskonsepsi yang terjadi pada siswa, maka dapat dilihat dari perhitungan persentase miskonsepsi yang berdasarkan indikator pencapaian kompetensi dan indikator soal pada Tabel 4.18 berikut.

Tabel 4.18

Perhitungan Persentase Miskonsepsi Berdasarkan Indikator Pencapaian Kompetensi dan Indikator Soal

No.	Indikator Pencapaian Kompetensi	Indikator Soal	Ranah Kognitif						Tingkat Miskonsepsi
			C 1	C 2	C 3	C 4	C 5	C 6	
1.	Menjelaskan definisi bangun datar segiempat untuk mengidentifikasi miskonsepsi definisi bangun datar segiempat	Disajikan beberapa gambar bangun datar kemudian siswa diminta untuk menjelaskan bangun datar segiempat		✓					67,5%
2.	Menyebutkan sifat-sifat dari	Disajikan berbagai macam bangun	✓						45%

	berbagai jenis segiempat (persegi, persegi panjang, belahketupat, jajargenjang, trapesium, dan layang-layang) untuk mengidentifikasi miskonsepsi sifat-sifat bangun datar segiempat	datar segiempat kemudian siswa dimintai untuk menentukan bangun persegi panjang							
3.		Disajikan beberapa pernyataan mengenai sifat-sifat bangun datar segiempat, kemudian siswa diminta untuk menentukan pernyataan yang sesuai dengan sifat-sifat dari bangun datar jajargenjang	✓						87,5%
4.		Disajikan ilustrasi mengenai bangun datar, siswa diminta untuk menentukan ilustrasi yang berbentuk jajargenjang	✓						25%
Rata-rata persentase									52,5%
5.	Menentukan rumus keliling dan luas untuk berbagai jenis segiempat (persegi, persegi panjang, belahketupat, jajargenjang, trapesium, dan layang-layang) untuk	Diberikan dua gambar bangun datar yang berbeda, siswa diminta untuk menentukan luas bangun datar yang tidak diarsir			✓				57,5%
6.		Siswa diminta untuk menentukan luas layang-layang			✓				27,5%
7.		Siswa diminta menentukan salah			✓				25%

	mengidentifikasi miskonsepsi rumus keliling dan luas bangun datar segiempat	satu diagonal yang belum diketahui pada bangun datar belah ketupat							
8.		Siswa diminta untuk menentukan luas trapesium sama kaki dengan tidak diketahui tinggi trapesium tersebut			✓				35%
Rata-rata persentase									36,25%
9.	Menyelesaikan masalah kontekstual yang berkaitan dengan luas dan keliling bangun datar segiempat	Disajikan ilustrasi kebun berbentuk bangun datar, siswa diminta untuk menentukan hasil penjualan dari kebun tersebut			✓				27,5%
10.	untuk mengidentifikasi miskonsepsi penyelesaian masalah kontekstual pada bangun datar segiempat	Siswa menentukan keliling tanah dengan tidak diketahui lebar tanah tersebut			✓				35%
Rata-rata persentase									31,25%
Rata-rata persentase miskonsepsi keseluruhan									43,25%

Pada Tabel 4.18 dapat dilihat persentase miskonsepsi yang terjadi paling tinggi pada indikator pencapaian kompetensi yaitu kemampuan menjelaskan sifat-sifat dari berbagai jenis segiempat dengan indikator soal sifat-sifat dari bangun datar jajargenjang yang mencapai 87,5%. Hal ini dikarenakan banyaknya siswa yang belum memahami konsep ilmiah jajargenjang terutama pada sifat bangun datar tersebut yang sebagian besar siswa menyebutkan jajargenjang memiliki sepasang sisi yang miring. Dari tabel tersebut dapat diketahui pula bahwa butir soal yang memiliki nilai persentase

miskonsepsi tinggi yaitu sebesar 67,5% adalah butir soal nomor 1 dengan indikator pencapaian kompetensi yaitu kemampuan menjelaskan definisi bangun datar segiempat dengan indikator soal menentukan beberapa gambar segiempat menunjukkan bahwasanya kebanyakan siswa mengalami kesulitan terutama dalam mengungkapkan definisi bangun datar segiempat, yang mana menganggap bahwa segiempat selalu dalam bentuk beraturan saja. Di samping itu, terdapat butir soal yang memiliki nilai persentase miskonsepsi rendah dengan rata-rata sebesar 31,25% terjadi pada indikator pencapaian kompetensi yaitu kemampuan menyelesaikan masalah kontekstual yang berkaitan dengan luas dan keliling bangun datar segiempat menunjukkan bahwasanya ada beberapa siswa yang masih mengalami kelemahan dalam menerapkan hubungan antara rumus bangun datar yang digunakan dengan permasalahan kontekstual. Dari analisis hasil tes diagnostik pilihan ganda *three-tier* yang telah dilakukan, maka dapat diperoleh kesimpulan bahwa miskonsepsi sering terjadi pada indikator kemampuan siswa dalam menjelaskan konsep definisi bangun datar segiempat dan menyebutkan sifat-sifat jajargenjang.

C. Revisi Produk

Validator tidak hanya menilai kelayakan suatu tes diagnostik pilihan ganda *three-tier*, melainkan validator juga memberikan komentar atau saran yang dijadikan bahan untuk merevisi tes diagnostik pilihan ganda *three-tier* sehingga layak untuk digunakan. Berikut hasil revisi tes diagnostik pilihan ganda *three-tier* untuk mengidentifikasi miskonsepsi pada materi geometri yang dikembangkan.

1. Kisi-Kisi Soal Tes Diagnostik Pilihan Ganda *Three-Tier*

Kisi-kisi soal tes diagnostik pilihan ganda *three-tier* yang dikembangkan peneliti diambil dari buku paket matematika kurikulum 2013 edisi revisi pada materi segiempat dan literatur yang mendukung. Yang selanjutnya, memperoleh komentar dan saran dari para validator. Adapun komentar dan saran validator dapat disajikan pada Tabel 4.19.

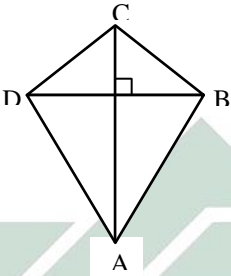
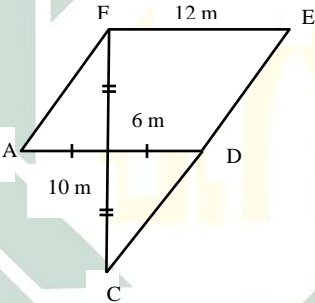
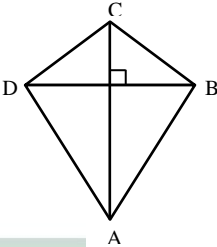
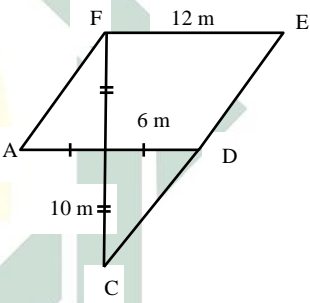
Tabel 4.19
Hasil Revisi Kisi-Kisi Soal Tes Diagnostik
Pilihan Ganda *Three-Tier*

No.	Komentar/Saran Validator	
	Sebelum Revisi	Setelah Revisi
1.	Tidak tercantum indikator miskonsepsi pada indikator pencapaian kompetensi	Dijelaskan indikator miskonsepsi pada indikator pencapaian kompetensi
2.	Penetapan dimensi kognitifnya tidak sesuai dengan indikator pencapaian kompetensi	Penetapan dimensi kognitif pada setiap butir soal disempurnakan dengan menerapkan taksonomi Bloom yang sesuai dengan indikator pencapaian kompetensi
3.	Penulisan tata bahasa belum sempurna	Penulisan tata bahasa dalam kisi-kisi soal disempurnakan dengan menggunakan EYD yang benar

2. Soal Tes Diagnostik Pilihan Ganda *Three-Tier*

Tabel 4.20
Hasil Revisi Soal Tes Diagnostik Pilihan Ganda *Three-Tier*

No.	Komentar/Saran Validator	
	Sebelum Revisi	Setelah Revisi
1.	Gambar pada butir soal nomor 6 dan 8 tidak jelas dan letaknya membingungkan.	Gambar pada butir soal nomor 6 dan 8 diperjelas dan menempatkannya dengan tepat.

	<p>Soal nomor 6 :</p>  <p>Soal nomor 8 :</p> 	<p>Soal nomor 6 :</p>  <p>Soal nomor 8 :</p> 
2.	Waktu pengerjaan soal dibatasi dengan waktu yang sudah ditentukan.	Waktu pengerjaan soal tidak perlu dibatasi dengan membiarkan siswa mengerjakan sampai selesai.
3.	<p>Pilihan alasan pada butir soal nomor 3 kurang tepat.</p> <p>Soal nomor 3 : Pilihan Alasan: a. Jajargenjang merupakan bangun datar dengan kedua pasang sisi yang berhadapan kongruen b. Jajargenjang merupakan</p>	<p>Pilihan alasan pada butir soal nomor 3 diperbaiki dengan memberikan alasan yang tepat dan sesuai dengan konsep pada materi segiempat.</p> <p>Soal nomor 3 : Pilihan Alasan: a. Jajargenjang merupakan</p>

	<p>bangun datar dengan tepat sepasang sisi yang berhadapan sejajar serta sepasang sisi yang miring</p> <p>c. Jajargenjang merupakan bangun datar dengan semua sisi yang kongruen serta sudut yang sama besar</p> <p>d. Jajargenjang merupakan bangun datar dengan kedua pasang sisi berhadapan yang sejajar serta sudut yang sama besar</p>	<p>bangun datar dengan pasang sisi yang berhadapan sejajar</p> <p>b. Jajargenjang merupakan bangun datar yang memiliki tepat sepasang sisi yang berhadapan sejajar serta sepasang sisi yang miring</p> <p>c. Jajargenjang merupakan bangun datar dengan semua sisi yang kongruen serta sudut yang sama besar</p> <p>d. Jajargenjang merupakan bangun datar dengan kedua pasang sisi berhadapan sejajar serta sudut yang sama besar</p>
4.	<p>Penulisan tata bahasa belum sempurna.</p> <p>Soal nomor 7 : Panjang salah satu diagonal (d_1) belah ketupat adalah 24 cm. Jika luas belah ketupat 120 cm^2. Maka panjang diagonal satunya (d_2) pada bangun tersebut adalah ...</p>	<p>Penulisan tata bahasa pada soal tes diagnostik pilihan ganda <i>three-tier</i> disempurnakan dengan menggunakan EYD yang benar.</p> <p>Soal nomor 7 : Panjang salah satu diagonal (d_1) belah ketupat adalah 24 cm. Jika luas belah ketupat 120 cm^2 maka panjang diagonal lainnya (d_2) pada bangun tersebut adalah ...</p>

3. Kunci Jawaban dan Pedoman Penskoran

Yang perlu direvisi dari kunci jawaban dan pedoman penskoran adalah penulisan tata bahasa yang digunakan perlu diperbaiki dan

memperjelas kembali penilaian dalam setiap tingkatan pada tes diagnostik pilihan ganda *three-tier*. Adapun hasil revisi dari kunci jawaban dan pedoman penskoran tes diagnostik pilihan ganda *three-tier* dapat dilihat pada Tabel 4.21.

Tabel 4.21
Hasil Revisi Kunci Jawaban dan Pedoman Penskoran Tes Diagnostik Pilihan Ganda *Three-Tier*

No.	Komentar/Saran Validator	
	Sebelum Revisi	Setelah Revisi
1.	<p>Pada kunci jawaban butir soal nomor 6 tidak ada jawabannya.</p> <p>Penyelesaian soal nomor 6 : Diket : Keliling = 66 cm, AB = 20 cm, BD = 24 cm Ditanya : Luas layang-layang ABCD Jawab : Dikarenakan AD = AB dan BC = CD maka Keliling = AB + BC + CD + AD $66 = 20 + 20 + CD + CD$ $66 = 40 + 2 CD$ $26 = 2 CD$ $13 = CD$ CD = BC = 13 cm $OC = \sqrt{BC^2 - OB^2}$ $= \sqrt{13^2 - 12^2}$ $= \sqrt{25} = 5 \text{ cm}$ $OA = \sqrt{AB^2 - OB^2}$ $= \sqrt{20^2 - 12^2}$ $= \sqrt{256} = 16 \text{ cm}$ AC = OC + OA = 5 + 16 = 21 cm Luas layang-layang ABCD $= \frac{1}{2} \times AC \times BD$ $= \frac{1}{2} \times 21 \times 20 = 210$</p>	<p>Kunci jawaban pada butir soal nomor 6 diperbaiki dengan mengecek kembali penulisan maupun perhitungan.</p> <p>Penyelesaian soal nomor 6 : Diket : Keliling = 66 cm, AB = 20 cm, BD = 24 cm Ditanya : Luas layang-layang ABCD Jawab : Dikarenakan AD = AB dan BC = CD maka Keliling = AB + BC + CD + AD $66 = 20 + 20 + CD + CD$ $66 = 40 + 2 CD$ $26 = 2 CD$ $13 = CD$ CD = BC = 13 cm $OC = \sqrt{BC^2 - OB^2} = \sqrt{13^2 - 12^2}$ $= \sqrt{25} = 5 \text{ cm}$ $OA = \sqrt{AB^2 - OB^2} = \sqrt{20^2 - 12^2}$ $= \sqrt{256} = 16 \text{ cm}$ AC = OC + OA = 5 + 16 = 21 cm Luas layang-layang ABCD $= \frac{1}{2} \times AC \times BD$ $= \frac{1}{2} \times 21 \times 24 = 252$</p>

	Jadi, luas layang-layang ABCD adalah 210 cm^2	Jadi, luas layang-layang ABCD adalah 252 cm^2
2.	<p>Pedoman penskoran pada bagian penentuan skor belum jelas.</p> <p>Dengan ketentuan :</p> <ol style="list-style-type: none"> 1) Skor 1 apabila jawaban benar atau alasan benar 2) Skor 0 apabila jawaban salah atau alasan salah atau tidak memberikan jawaban 	<p>Pedoman penskoran pada bagian penentuan skor diperbaiki dan memperjelas penilaian.</p> <p>Dengan ketentuan :</p> <ol style="list-style-type: none"> 1) Skor 1 diberikan apabila jawaban benar dan alasan benar 2) Skor 0 diberikan apabila jawaban salah atau alasan salah atau tidak memberikan jawaban

D. Kajian Produk Akhir

Berdasarkan hasil pengembangan dan uji coba produk yang telah dilakukan di lapangan, maka dapat dikemukakan berbagai kajian terkait produk akhir tes diagnostik *three-tier* yang dihasilkan. Berikut temuan yang diperoleh:

1. Temuan Terkait Tes Diagnostik Pilihan Ganda *Three-Tier* Untuk Mengidentifikasi Miskonsepsi Pada Materi Geometri

Kisi-kisi soal, soal tes diagnostik pilihan ganda *three-tier*, kunci jawaban dan pedoman penskoran yang dikembangkan digunakan untuk mengidentifikasi miskonsepsi pada materi geometri. Kisi-kisi soal, soal tes diagnostik pilihan ganda *three-tier*, kunci jawaban dan pedoman penskoran yang telah disusun sebelumnya telah divalidasi oleh validator. Berdasarkan analisis data hasil validasi para ahli, didapatkan bahwa kisi-kisi soal, soal tes diagnostik pilihan ganda *three-tier*, kunci jawaban, dan pedoman penskoran telah dinyatakan “valid” dengan total rata-rata validasi sebesar 3,61. Seperti yang dikemukakan di BAB II, bahwa perangkat pembelajaran dinyatakan valid jika rata-rata nilai yang didapatkan dari validator termasuk dalam kategori interval skor “valid”. Hampir semua aspek memperoleh rata-rata skor dalam rentang $3 \leq \text{RTV} < 4$ yang berarti menunjukkan rata-rata total hasil penilaian validator terhadap tes yang dikembangkan termasuk kategori valid.

Selain dinyatakan valid, kisi-kisi soal, soal tes diagnostik pilihan ganda *three-tier*, kunci jawaban dan pedoman penskoran

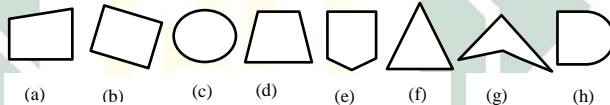
yang dikembangkan mendapatkan penilaian “B” oleh ketiga validator yang berarti dapat digunakan dengan sedikit revisi. Setelah dinyatakan valid dan layak maka dapat diuji cobakan kepada subjek penelitian yang telah ditetapkan oleh peneliti.

Berdasarkan penyajian dan analisis data kevalidan, reliabilitas dan kualitas butir soal tes diagnostik pilihan ganda *three-tier* terdapat temuan menarik pada masing-masing butir soal tes diagnostik *three-tier* yang dikembangkan. Temuan yang dimaksudkan terkait dengan kevalidan, reliabilitas, dan kualitas butir soal yang meliputi daya pembeda, tingkat kesukaran, dan keefektifan pengecoh. Berikut temuan terkait pada setiap butir soal:

a. Butir soal nomor 1

Setelah melalui tahap revisi, berikut produk akhir terkait tes diagnostik *three-tier* untuk mengidentifikasi miskonsepsi pada materi geometri pada butir soal nomor 1:

“Perhatikan gambar bangun datar berikut. Mana saja yang merupakan bangun datar segiempat?”



Pilihan jawaban:

- A. (b)
- B. (a), (b), (d)
- C. (a), (b), (d), (g)
- D. (a), (b), (d), (e), (f), (g)

Pilihan alasan:

- a. *Bangun datar memiliki 4 sisi dan 4 titik sudut*
- b. *Bangun datar memiliki 4 titik sudut yang sama besar*
- c. *Bangun datar memiliki 4 sisi yang sama panjang*
- d. *Bangun datar memiliki 2 sisi yang berhadapan sama panjang dan 4 titik sudut siku-siku”*

Butir soal nomor 1 memiliki validitas dengan koefisien korelasi sebesar 0,44. Validitas berkenaan dengan ketepatan alat penilai terhadap apa yang seharusnya dinilai sehingga benar-benar menilai apa yang akan dinilai. Validitas tes pada dasarnya menunjuk pada derajat fungsi pengukuran suatu tes

atau derajat kecermatan ukurnya suatu tes.¹³⁵ Butir soal dikatakan valid apabila nilai koefisien relasinya lebih dari sama dengan 0,400. Hasil dari butir soal nomor 1, koefisien korelasi lebih dari kisaran yang ditentukan. Hal ini menunjukkan bahwa butir soal nomor 1 dapat dikatakan valid dan layak untuk diuji coba.

Korelasi reliabilitas pada butir soal nomor 1 dengan menggunakan rumus KR.20 dan berbantuan program *SPSS Statistic Versi 16.0* sebesar 0,578. Arifin menyatakan reliabilitas adalah tingkat atau derajat konsistensi dari sebuah tes, reliabilitas tes berkenaan dengan pertanyaan apakah suatu tes yang diteliti dan dapat dipercaya sesuai dengan kriteria yang telah ditetapkan.¹³⁶ Sebuah tes dikatakan baik jika bersifat reliabel. Tes dikatakan reliabel jika tes tersebut diujikan berkali-kali dengan dua atau lebih penilai kepada subjek yang sama maka hasilnya relatif sama. Nilai korelasi reliabilitas yang baik yaitu mempunyai nilai di atas atau sama dengan 0,600. Hasil nilai dari reliabilitas pada butir soal nomor 1 menunjukkan butir soal tersebut reliabel dan layak digunakan dalam tes.

Selanjutnya kualitas butir soal nomor 1 meliputi daya pembeda soal, tingkat kesukaran dan keefektifan pengecoh. Butir soal nomor 1 mempunyai daya pembeda soal sebesar 0,35. Daya pembeda soal merupakan kemampuan suatu butir soal yang dapat membedakan antara siswa yang telah menguasai materi dengan siswa yang belum menguasai materi. Semakin tinggi daya pembeda suatu item, semakin baik suatu item dalam membedakan kemampuan penempuh tes.¹³⁷ Daya pembeda soal yang baik berkisar antara $0,40 \leq D_p < 0,70$. Hasil dari butir soal nomor 1, daya pembeda soal berada pada kategori cukup baik sebagaimana telah ditetapkan pada BAB III dalam Tabel 3.5. Hal ini menunjukkan bahwa butir soal nomor 1 cukup memadai dalam membedakan antara siswa yang sudah memahami materi dengan siswa yang belum memahami materi.

¹³⁵ Sumadi Suryabrata, *Pengembangan Alat Ukur Psikologis*, (Yogyakarta: Andi, 2000), hal. 41.

¹³⁶ Zainal Arifin, *Evaluasi Pembelajaran*, (Bandung: PT. Rosdakarya, 2013), hal. 258.

¹³⁷ Zainal Arifin, *Evaluasi Pembelajaran*, (Bandung: PT. Remaja Rosdakarya, 2009).

Tingkat kesukaran butir soal nomor 1 memiliki nilai sebesar 0,175 dengan kategori sukar. Suatu soal hendaknya tidak terlalu sukar dan tidak terlalu mudah. Jika suatu soal memiliki tingkat kesukaran seimbang (proporsional), maka dapat dikatakan bahwa soal tersebut baik.¹³⁸ Tingkat kesukaran yang baik yaitu memiliki rentang $0,30 < P \leq 0,70$. Hasil dari butir soal nomor 1, tingkat kesukaran kurang dari kisaran yang ditentukan. Hal tersebut menunjukkan bahwa siswa dengan kemampuan tinggi merasa kesulitan dalam mengerjakan dan sebaliknya siswa dengan kemampuan rendah merasa mudah menjawab butir soal nomor 1. Dilihat dari perolehan menjawab, sebanyak 7 siswa yang menjawab benar pada butir soal nomor 1 sehingga kemungkinan siswa yang menjawab kebanyakan adalah siswa dengan kemampuan rendah.

Keefektifan pengecoh jawaban dan alasan. pada butir soal nomor 1 dipilih lebih dari 5% dari peserta tes. Sebuah pengecoh dapat dikatakan berfungsi dengan baik apabila pengecoh tersebut mempunyai daya tarik yang besar bagi pengikut-pengikut tes yang kurang memahami konsep atau kurang menguasai bahan.¹³⁹ Pengecoh yang efektif dikatakan baik apabila pengecoh tersebut dipilih paling sedikit 2 dari siswa yang mengikuti tes. Hasil dari butir soal nomor 1, keefektifan pengecoh jawaban dan alasan lebih dari kisaran yang ditentukan sehingga menunjukkan pengecoh jawaban dan alasan tersebut berfungsi dengan baik.

Untuk mendapatkan karakteristik butir soal dengan kriteria baik, hendaknya memenuhi 3 kriteria persyaratan yaitu memiliki koefisien korelasi validitas yang baik, koefisien reliabilitas yang baik, dan kualitas butir soal yang baik. Karakteristik butir soal nomor 1 menunjukkan kriteria yang baik.

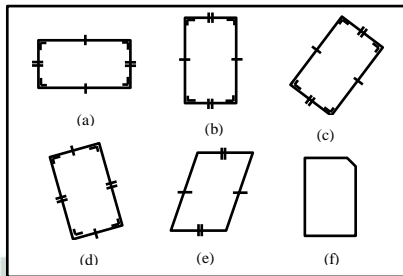
b. Butir soal nomor 2

Setelah melalui tahap revisi, berikut produk akhir terkait tes diagnostik *three-tier* untuk mengidentifikasi miskonsepsi pada materi geometri pada butir soal nomor 2:

¹³⁸ Zainal Arifin, *Evaluasi Pembelajaran*, (Bandung: PT. Remaja Rosdakarya, 2009).

¹³⁹ Arikunto, *Prosedur Penelitian: Suatu Pendekatan Praktek Edisi RevisiIV*, (Jakarta: PT. Rineka Cipta, 1997), hal. 225.

“Manakah yang termasuk persegi panjang?”



Pilihan jawaban:

- A. (a), (b), (e)
- B. (a), (b), (c)
- C. (a), (b), (c), (d)
- D. (a), (b), (d), (e), (f)

Pilihan alasan:

- a. Persegi panjang memiliki 4 sisi sejajar dan sama panjang
- b. Persegi panjang memiliki 4 titik sudut yang sama besar
- c. Persegi panjang memiliki sisi yang berhadapan sejajar dan sama panjang serta sudut yang sama besar
- d. Persegi panjang memiliki dua pasang sisi yang sejajar dan sama panjang serta setiap sudutnya siku-siku”

Butir soal nomor 2 memiliki validitas dengan koefisien korelasi sebesar 0,57. Validitas berkenaan dengan ketepatan alat penilai terhadap apa yang seharusnya dinilai sehingga benar-benar menilai apa yang akan dinilai. Validitas tes pada dasarnya menunjuk pada derajat fungsi pengukuran suatu tes atau derajat kecermatan ukurnya suatu tes.¹⁴⁰ Butir soal dikatakan valid apabila nilai koefisien relasinya lebih dari sama dengan 0,400. Hasil dari butir soal nomor 2, koefisien korelasinya lebih dari kisaran yang ditentukan. Hal ini menunjukkan bahwa butir soal nomor 2 dapat dikatakan valid dan layak untuk diuji coba.

Korelasi reliabilitas pada butir soal nomor 2 dengan menggunakan rumus KR.20 dan berbantuan program SPSS

¹⁴⁰ Sumadi Suryabrata, *Pengembangan Alat Ukur Psikologis*, (Yogyakarta: Andi, 2000), hal. 41.

Statistic Versi 16.0 sebesar 0,670. Arifin menyatakan reliabilitas adalah tingkat atau derajat konsistensi dari sebuah tes, reliabilitas tes berkenaan dengan pertanyaan apakah suatu tes yang diteliti dan dapat dipercaya sesuai dengan kriteria yang telah ditetapkan.¹⁴¹ Sebuah tes dikatakan baik jika bersifat reliabel. Tes dikatakan reliabel jika tes tersebut diujikan berkali-kali dengan dua atau lebih penilai kepada subjek yang sama maka hasilnya relatif sama. Nilai korelasi reliabilitas yang baik yaitu mempunyai nilai di atas atau sama dengan 0,600. Hasil nilai dari reliabilitas pada butir soal nomor 2 menunjukkan butir soal tersebut reliabel dan layak digunakan dalam tes.

Selanjutnya kualitas butir soal nomor 2 meliputi daya pembeda soal, tingkat kesukaran dan keefektifan pengecoh. Butir soal nomor 2 mempunyai daya pembeda soal sebesar 0,50. Daya pembeda soal merupakan kemampuan suatu butir soal yang dapat membedakan antara siswa yang telah menguasai materi dengan siswa yang belum menguasai materi. Semakin tinggi daya pembeda suatu item, semakin baik suatu item dalam membedakan kemampuan penempuh tes.¹⁴² Daya pembeda soal yang baik berkisar antara $0,40 \leq D_p < 0,70$. Hasil dari butir soal nomor 2, daya pembeda soal berada kisaran yang ditentukan. Dengan demikian, butir soal nomor 2 dapat membedakan antara siswa yang sudah memahami materi dengan siswa yang belum memahami materi.

Tingkat kesukaran butir soal nomor 2 memiliki nilai sebesar 0,55. Suatu soal hendaknya tidak terlalu sukar dan tidak terlalu mudah. Jika suatu soal memiliki tingkat kesukaran seimbang (proporsional), maka dapat dikatakan bahwa soal tersebut baik.¹⁴³ Tingkat kesukaran yang baik yaitu memiliki rentang $0,30 < P \leq 0,70$. Hasil dari butir soal nomor 2, tingkat kesukaran berada pada rentang yang ditentukan. Siswa dengan kemampuan tinggi akan merasa mudah dalam mengerjakan dan sebaliknya siswa dengan kemampuan rendah akan sulit menjawab butir soal nomor 2.

¹⁴¹ Zainal Arifin, *Evaluasi Pembelajaran*, (Bandung: PT. Rosdakarya, 2013), hal. 258.

¹⁴² Zainal Arifin, *Evaluasi Pembelajaran*, (Bandung: PT. Remaja Rosdakarya, 2009).

¹⁴³ Zainal Arifin, *Evaluasi Pembelajaran*, (Bandung: PT. Remaja Rosdakarya, 2009).

Keefektifan pengecoh jawaban dan alasan. pada butir soal nomor 2 dipilih lebih dari 5% dari peserta tes. Sebuah pengecoh dapat dikatakan berfungsi dengan baik apabila pengecoh tersebut mempunyai daya tarik yang besar bagi pengikut-pengikut tes yang kurang memahami konsep atau kurang menguasai bahan.¹⁴⁴ Pengecoh yang efektif dikatakan baik apabila pengecoh tersebut dipilih paling sedikit 2 dari siswa yang mengikuti tes. Hasil dari butir soal nomor 2, keefektifan pengecoh jawaban dan alasan lebih dari kisaran yang ditentukan sehingga menunjukkan pengecoh jawaban dan alasan tersebut berfungsi dengan baik.

Untuk mendapatkan karakteristik butir soal dengan kriteria baik, hendaknya memenuhi 3 kriteria persyaratan yaitu memiliki koefisien korelasi validitas yang baik, koefisien reliabilitas yang baik, dan kualitas butir soal yang baik. Karakteristik butir soal nomor 2 menunjukkan kriteria yang baik.

c. Butir soal nomor 3

Setelah melalui tahap revisi, berikut produk akhir terkait tes diagnostik *three-tier* untuk mengidentifikasi miskonsepsi pada materi geometri pada butir soal nomor 3:

“Perhatikan pernyataan-pernyataan di bawah ini :

- (1) *Sisi-sisi yang berhadapan sejajar dan sama panjang.*
 - (2) *Setiap sudutnya siku-siku.*
 - (3) *Sudut yang berhadapan sama besar.*
 - (4) *Jumlah dua besar sudut yang berdekatan adalah 90° .*
 - (5) *Kedua diagonalnya saling berpotongan dan membagi diagonal-diagonal tersebut menjadi dua bagian sama panjang.*
 - (6) *Sepasang sisinya miring.*
 - (7) *Jumlah dua besar sudut yang berdekatan adalah 180° .*
- Manakah yang merupakan sifat-sifat dari jajargenjang?*

Pilihan jawaban:

- A. *(1), (2), (4), (5)*
- B. *(1), (3), (5), (6)*
- C. *(1), (3), (5), (6)*

¹⁴⁴ Arikunto, *Prosedur Penelitian: Suatu Pendekatan Praktek Edisi RevisiIV*, (Jakarta: PT. Rineka Cipta, 1997), hal. 225.

D. (1), (3), (5), (7)

Pilihan alasan:

- a. *Jajargenjang merupakan bangun datar dengan pasang sisi yang berhadapan sejajar*
- b. *Jajargenjang merupakan bangun datar yang memiliki tepat sepasang sisi yang berhadapan sejajar serta sepasang sisi yang miring*
- c. *Jajargenjang merupakan bangun datar dengan semua sisi yang kongruen serta sudut yang sama besar*
- d. *Jajargenjang merupakan bangun datar dengan kedua pasang sisi berhadapan sejajar serta sudut yang sama besar”*

Butir soal nomor 3 memiliki validitas dengan koefisien korelasi sebesar 0,00. Validitas berkenaan dengan ketepatan alat penilai terhadap apa yang seharusnya dinilai sehingga benar-benar menilai apa yang akan dinilai. Validitas tes pada dasarnya menunjuk pada derajat fungsi pengukuran suatu tes atau derajat kecermatan ukurnya suatu tes.¹⁴⁵ Butir soal dikatakan valid apabila nilai koefisien relasinya lebih dari sama dengan 0,400. Hasil dari butir soal nomor 3, koefisien korelasinya kurang dari kisaran yang ditentukan. Hal ini menunjukkan bahwa butir soal nomor 3 dapat dikatakan tidak valid dan tidak layak untuk diuji coba.

Korelasi reliabilitas pada butir soal nomor 3 dengan menggunakan rumus KR.20 dan berbantuan program *SPSS Statistic Versi 16.0* sebesar 0,610. Arifin menyatakan reliabilitas adalah tingkat atau derajat konsistensi dari sebuah tes, reliabilitas tes berkenaan dengan pertanyaan apakah suatu tes yang diteliti dan dapat dipercaya sesuai dengan kriteria yang telah ditetapkan.¹⁴⁶ Sebuah tes dikatakan baik jika bersifat reliabel. Tes dikatakan reliabel jika tes tersebut diujikan berkali-kali dengan dua atau lebih penilai kepada subjek yang sama maka hasilnya relatif sama. Nilai korelasi reliabilitas yang baik yaitu mempunyai nilai di atas atau sama dengan 0,600. Hasil nilai dari reliabilitas pada butir soal nomor 3

¹⁴⁵ Sumadi Suryabrata, *Pengembangan Alat Ukur Psikologis*, (Yogyakarta: Andi, 2000), hal. 41.

¹⁴⁶ Zainal Arifin, *Evaluasi Pembelajaran*, (Bandung: PT. Rosdakarya, 2013), hal. 258.

menunjukkan butir soal tersebut reliabel dan apabila diujikan berkali-kali maka mendapatkan nilai yang relatif sama .

Selanjutnya kualitas butir soal nomor 3 meliputi daya pembeda soal, tingkat kesukaran dan keefektifan pengecoh. Butir soal nomor 3 mempunyai daya pembeda soal sebesar 0,00. Daya pembeda soal merupakan kemampuan suatu butir soal yang dapat membedakan antara siswa yang telah menguasai materi dengan siswa yang belum menguasai materi. Semakin tinggi daya pembeda suatu item, semakin baik suatu item dalam membedakan kemampuan penempuh tes.¹⁴⁷ Daya pembeda soal yang baik berkisar antara $0,40 \leq D_p < 0,70$. Hasil dari butir soal nomor 3, daya pembeda soal kurang dari kisaran yang ditentukan. Hal tersebut menunjukkan bahwa butir soal nomor 3 tidak dapat membedakan antara siswa yang sudah memahami materi dengan siswa yang belum memahami materi.

Tingkat kesukaran butir soal nomor 3 memiliki nilai sebesar 0,00. Suatu soal hendaknya tidak terlalu sukar dan tidak terlalu mudah. Jika suatu soal memiliki tingkat kesukaran seimbang (proporsional), maka dapat dikatakan bahwa soal tersebut baik.¹⁴⁸ Tingkat kesukaran yang baik yaitu memiliki rentang $0,30 < P \leq 0,70$. Hasil dari butir soal nomor 3, tingkat kesukaran kurang dari kisaran yang ditentukan. Hal tersebut menunjukkan bahwa siswa dengan kemampuan tinggi merasa kesulitan dalam mengerjakan dan sebaliknya siswa dengan kemampuan rendah merasa mudah menjawab butir soal nomor 3. Dilihat dari perolehan menjawab, tidak ada siswa yang menjawab benar pada butir soal nomor 3 sehingga menunjukkan butir soal tersebut semakin sukar untuk diselesaikan bagi siswa yang berkemampuan tinggi maupun siswa yang berkemampuan rendah.

Keefektifan pengecoh jawaban pada butir soal nomor 3, alternatif A dan C dipilih kurang dari 5% dari peserta tes sedangkan pengecoh alasan dengan alternatif B juga dipilih kurang dari 5% dari peserta tes. Sebuah pengecoh dapat dikatakan berfungsi dengan baik apabila pengecoh tersebut mempunyai daya tarik yang besar bagi pengikut-pengikut tes

¹⁴⁷ Zainal Arifin, *Evaluasi Pembelajaran*, (Bandung: PT. Remaja Rosdakarya, 2009).

¹⁴⁸ Zainal Arifin, *Evaluasi Pembelajaran*, (Bandung: PT. Remaja Rosdakarya, 2009).

yang kurang memahami konsep atau kurang menguasai bahan.¹⁴⁹ Pengecoh yang efektif dikatakan baik apabila pengecoh tersebut dipilih paling sedikit 2 dari siswa yang mengikuti tes. Hasil dari butir soal nomor 3, keefektifan pengecoh jawaban dan alasan kurang dari kisaran yang ditentukan sehingga menunjukkan pengecoh jawaban dan alasan tersebut berfungsi tidak baik.

Untuk mendapatkan karakteristik butir soal dengan kriteria baik, hendaknya memenuhi 3 kriteria persyaratan yaitu memiliki koefisien korelasi validitas yang baik, koefisien reliabilitas yang baik, dan kualitas butir soal yang baik. Karakteristik butir soal nomor 3 menunjukkan kriteria yang tidak baik.

d. Butir soal nomor 4

Setelah melalui tahap revisi, berikut produk akhir terkait tes diagnostik *three-tier* untuk mengidentifikasi miskonsepsi pada materi geometri pada butir soal nomor 4:



“Benda-benda tersebut berbentuk bangun datar. Manakah yang termasuk jajargenjang?”

Pilihan jawaban:

- A. Kue, kertas lipat, jam dinding, ketupat*
- B. Layang-layang, ketupat, papan, bingkai foto*
- C. Kertas lipat, kue, bingkai foto, ketupat*
- D. Jam dinding, kertas lipat, DVD, bingkai foto*

Pilihan alasan:

- a. Benda tersebut memiliki bentuk seperti persegi, belah ketupat, dan layang-layang yang termasuk dalam himpunan bangun jajargenjang*

¹⁴⁹ Arikunto, *Prosedur Penelitian: Suatu Pendekatan Praktek Edisi RevisiIV*, (Jakarta: PT. Rineka Cipta, 1997), hal. 225.

- b. Benda tersebut memiliki bentuk seperti persegi, persegi panjang, dan belah ketupat yang termasuk dalam himpunan bangun jajargenjang
- c. Benda tersebut memiliki bentuk seperti lingkaran, persegi, dan layang-layang yang termasuk dalam himpunan bangun jajargenjang
- d. Benda tersebut memiliki bentuk seperti persegi, persegi panjang, dan layang-layang yang termasuk dalam himpunan bangun jajargenjang”

Butir soal nomor 4 memiliki validitas dengan koefisien korelasi sebesar 0,23. Validitas berkenaan dengan ketepatan alat penilai terhadap apa yang seharusnya dinilai sehingga benar-benar menilai apa yang akan dinilai. Validitas tes pada dasarnya menunjuk pada derajat fungsi pengukuran suatu tes atau derajat kecermatan ukurnya suatu tes.¹⁵⁰ Butir soal dikatakan valid apabila nilai koefisien relasinya lebih dari sama dengan 0,400. Hasil dari butir soal nomor 4, koefisien korelasinya kurang dari kisaran yang ditentukan. Hal ini menunjukkan bahwa butir soal nomor 4 dapat dikatakan tidak valid dan layak untuk diuji coba apabila direvisi terlebih dahulu.

Korelasi reliabilitas pada butir soal nomor 4 dengan menggunakan rumus KR.20 dan berbantuan program *SPSS Statistic Versi 16.0* sebesar 0,653. Arifin menyatakan reliabilitas adalah tingkat atau derajat konsistensi dari sebuah tes, reliabilitas tes berkenaan dengan pertanyaan apakah suatu tes yang diteliti dan dapat dipercaya sesuai dengan kriteria yang telah ditetapkan.¹⁵¹ Sebuah tes dikatakan baik jika bersifat reliabel. Tes dikatakan reliabel jika tes tersebut diujikan berkali-kali dengan dua atau lebih penilai kepada subjek yang sama maka hasilnya relatif sama. Nilai korelasi reliabilitas yang baik yaitu mempunyai nilai di atas atau sama dengan 0,600. Hasil nilai dari reliabilitas pada butir soal nomor 4 menunjukkan butir soal tersebut reliabel dan layak digunakan dalam tes.

¹⁵⁰ Sumadi Suryabrata, *Pengembangan Alat Ukur Psikologis*, (Yogyakarta: Andi, 2000), hal. 41.

¹⁵¹ Zainal Arifin, *Evaluasi Pembelajaran*, (Bandung: PT. Rosdakarya, 2013), hal. 258.

Selanjutnya kualitas butir soal nomor 4 meliputi daya pembeda soal, tingkat kesukaran dan keefektifan pengecoh. Butir soal nomor 4 mempunyai daya pembeda soal sebesar 0,10. Daya pembeda soal merupakan kemampuan suatu butir soal yang dapat membedakan antara siswa yang telah menguasai materi dengan siswa yang belum menguasai materi. Semakin tinggi daya pembeda suatu item, semakin baik suatu item dalam membedakan kemampuan penempuh tes.¹⁵² Daya pembeda soal yang baik berkisar antara $0,40 \leq D_p < 0,70$. Hasil dari butir soal nomor 4, daya pembeda soal berada pada kategori jelek sebagaimana telah ditetapkan pada BAB III dalam Tabel 3.5. Hal ini menunjukkan bahwa butir soal nomor 4 tidak dapat membedakan antara siswa yang sudah memahami materi dengan siswa yang belum memahami materi.

Tingkat kesukaran butir soal nomor 4 memiliki nilai sebesar 0,55. Suatu soal hendaknya tidak terlalu sukar dan tidak terlalu mudah. Jika suatu soal memiliki tingkat kesukaran seimbang (proporsional), maka dapat dikatakan bahwa soal tersebut baik.¹⁵³ Tingkat kesukaran yang baik yaitu memiliki rentang $0,30 < P \leq 0,70$. Hasil dari butir soal nomor 4, tingkat kesukaran berada pada rentang yang ditentukan. Siswa dengan kemampuan tinggi akan merasa mudah dalam mengerjakan dan sebaliknya siswa dengan kemampuan rendah akan sulit menjawab butir soal nomor 4.

Keefektifan pengecoh jawaban dan alasan pada butir soal nomor 4 dipilih lebih dari 5% dari peserta tes. Sebuah pengecoh dapat dikatakan berfungsi dengan baik apabila pengecoh tersebut mempunyai daya tarik yang besar bagi pengikut-pengikut tes yang kurang memahami konsep atau kurang menguasai bahan.¹⁵⁴ Pengecoh yang efektif dikatakan baik apabila pengecoh tersebut dipilih paling sedikit 2 dari siswa yang mengikuti tes. Hasil dari butir soal nomor 4, keefektifan pengecoh jawaban dan alasan lebih dari kisaran yang ditentukan sehingga menunjukkan pengecoh jawaban dan alasan tersebut berfungsi dengan baik.

¹⁵² Zainal Arifin, *Evaluasi Pembelajaran*, (Bandung: PT. Remaja Rosdakarya, 2009).

¹⁵³ Zainal Arifin, *Evaluasi Pembelajaran*, (Bandung: PT. Remaja Rosdakarya, 2009).

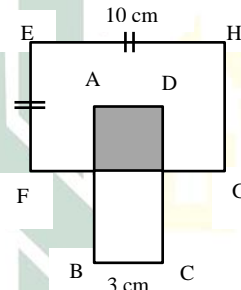
¹⁵⁴ Arikunto, *Prosedur Penelitian: Suatu Pendekatan Praktek Edisi RevisiIV*, (Jakarta: PT. Rineka Cipta, 1997), hal. 225.

Untuk mendapatkan karakteristik butir soal dengan kriteria baik, hendaknya memenuhi 3 kriteria persyaratan yaitu memiliki koefisien korelasi validitas yang baik, koefisien reliabilitas yang baik, dan kualitas butir soal yang baik. Karakteristik butir soal nomor 4 menunjukkan kriteria yang baik.

e. Butir soal nomor 5

Setelah melalui tahap revisi, berikut produk akhir terkait tes diagnostik *three-tier* untuk mengidentifikasi miskonsepsi pada materi geometri pada butir soal nomor 5:

"Bidang EFGH adalah persegi dan bidang ABCD adalah persegi panjang. Jika panjang AB = 14 cm dan luas daerah yang diarsir adalah 28 cm² maka luas daerah yang tidak diarsir adalah ..."



Pilihan jawaban:

- A. 142 cm²
- B. 128 cm²
- C. 114 cm²
- D. 86 cm²

Pilihan alasan:

- a. *Luas daerah tidak diarsir = Luas total + Luas daerah diarsir*
- b. *Luas daerah tidak diarsir = Luas total + (2 × Luas daerah diarsir)*
- c. *Luas daerah tidak diarsir = Luas total – Luas daerah diarsir*
- d. *Luas daerah tidak diarsir = Luas total – (2 × Luas daerah diarsir)"*

Butir soal nomor 5 memiliki validitas dengan koefisien korelasi sebesar 0,41. Validitas berkenaan dengan ketepatan alat penilai terhadap apa yang seharusnya dinilai sehingga benar-benar menilai apa yang akan dinilai. Validitas tes pada dasarnya menunjuk pada derajat fungsi pengukuran suatu tes atau derajat kecermatan ukurnya suatu tes.¹⁵⁵ Butir soal dikatakan valid apabila nilai koefisien relasinya lebih dari sama dengan 0,400. Hasil dari butir soal nomor 5, koefisien korelasinya lebih dari kisaran yang ditentukan. Hal ini menunjukkan bahwa butir soal nomor 5 dapat dikatakan valid dan layak untuk diuji coba.

Korelasi reliabilitas pada butir soal nomor 5 dengan menggunakan rumus KR.20 dan berbantuan program *SPSS Statistic Versi 16.0* sebesar 0,679. Arifin menyatakan reliabilitas adalah tingkat atau derajat konsistensi dari sebuah tes, reliabilitas tes berkenaan dengan pertanyaan apakah suatu tes yang diteliti dan dapat dipercaya sesuai dengan kriteria yang telah ditetapkan.¹⁵⁶ Sebuah tes dikatakan baik jika bersifat reliabel. Tes dikatakan reliabel jika tes tersebut diujikan berkali-kali dengan dua atau lebih penilai kepada subjek yang sama maka hasilnya relatif sama. Nilai korelasi reliabilitas yang baik yaitu mempunyai nilai di atas atau sama dengan 0,600. Hasil nilai dari reliabilitas pada butir soal nomor 5 menunjukkan butir soal tersebut reliabel dan layak digunakan dalam tes.

Selanjutnya kualitas butir soal nomor 5 meliputi daya pembeda soal, tingkat kesukaran dan keefektifan pengecoh. Butir soal nomor 5 mempunyai daya pembeda soal sebesar 0,15. Daya pembeda soal merupakan kemampuan suatu butir soal yang dapat membedakan antara siswa yang telah menguasai materi dengan siswa yang belum menguasai materi. Semakin tinggi daya pembeda suatu item, semakin baik suatu item dalam membedakan kemampuan penempuh tes.¹⁵⁷ Daya pembeda soal yang baik berkisar antara $0,40 \leq D_p < 0,70$. Hasil dari butir soal nomor 5, daya pembeda soal berada pada

¹⁵⁵ Sumadi Suryabrata, *Pengembangan Alat Ukur Psikologis*, (Yogyakarta: Andi, 2000), hal. 41.

¹⁵⁶ Zainal Arifin, *Evaluasi Pembelajaran*, (Bandung: PT. Rosdakarya, 2013), hal. 258.

¹⁵⁷ Zainal Arifin, *Evaluasi Pembelajaran*, (Bandung: PT. Remaja Rosdakarya, 2009).

kategori jelek sebagaimana telah ditetapkan pada BAB III dalam Tabel 3.5. Hal ini menunjukkan bahwa butir soal nomor 5 tidak dapat membedakan antara siswa yang sudah memahami materi dengan siswa yang belum memahami materi.

Tingkat kesukaran butir soal nomor 5 memiliki nilai sebesar 0,125. Suatu soal hendaknya tidak terlalu sukar dan tidak terlalu mudah. Jika suatu soal memiliki tingkat kesukaran seimbang (proporsional), maka dapat dikatakan bahwa soal tersebut baik.¹⁵⁸ Tingkat kesukaran yang baik yaitu memiliki rentang $0,30 < P \leq 0,70$. Hasil dari butir soal nomor 5, tingkat kesukaran kurang dari kisaran yang ditentukan. Hal tersebut menunjukkan bahwa siswa dengan kemampuan tinggi merasa kesulitan dalam mengerjakan dan sebaliknya siswa dengan kemampuan rendah merasa mudah dalam menjawab butir soal nomor 5. Dilihat dari perolehan menjawab, sebanyak 6 siswa yang menjawab benar pada butir soal nomor 5 sehingga kemungkinan siswa yang menjawab kebanyakan adalah siswa dengan kemampuan rendah.

Keefektifan pengecoh jawaban dan alasan pada butir soal nomor 5 dipilih lebih dari 5% dari peserta tes. Sebuah pengecoh dapat dikatakan berfungsi dengan baik apabila pengecoh tersebut mempunyai daya tarik yang besar bagi pengikut-pengikut tes yang kurang memahami konsep atau kurang menguasai bahan.¹⁵⁹ Pengecoh yang efektif dikatakan baik apabila pengecoh tersebut dipilih paling sedikit 2 dari siswa yang mengikuti tes. Hasil dari butir soal nomor 5, keefektifan pengecoh jawaban dan alasan lebih dari kisaran yang ditentukan sehingga menunjukkan pengecoh jawaban dan alasan tersebut berfungsi dengan baik.

Untuk mendapatkan karakteristik butir soal dengan kriteria baik, hendaknya memenuhi 3 kriteria persyaratan yaitu memiliki koefisien korelasi validitas yang baik, koefisien reliabilitas yang baik, dan kualitas butir soal yang baik. Karakteristik butir soal nomor 5 menunjukkan kriteria yang baik.

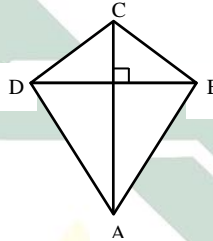
¹⁵⁸ Zainal Arifin, *Evaluasi Pembelajaran*, (Bandung: PT. Remaja Rosdakarya, 2009).

¹⁵⁹ Arikunto, *Prosedur Penelitian: Suatu Pendekatan Praktek Edisi Revisi IV*, (Jakarta: PT. Rineka Cipta, 1997), hal. 225.

f. Butir soal nomor 6

Setelah melalui tahap revisi, berikut produk akhir terkait tes diagnostik *three-tier* untuk mengidentifikasi miskonsepsi pada materi geometri pada butir soal nomor 6:

“Dari gambar layang-layang berikut diketahui kelilingnya 66 cm, $AB = 20$ cm, titik O adalah perpotongan antara d_1 dengan d_2 dan $BD = 24$ cm. Luas layang-layang $ABCD$ adalah ...



Pilihan jawaban:

- A. 210 cm^2
- B. 252 cm^2
- C. 270 cm^2
- D. 310 cm^2

Pilihan alasan:

- a. Panjang d_1 adalah 23 cm
- b. Panjang d_2 adalah 22 cm
- c. Luas = $\frac{1}{2} \times d_2 \times d_1$
- d. Luas = $\frac{1}{2} \times (d_1 + d_2)$ ”

Butir soal nomor 6 memiliki validitas dengan koefisien korelasi sebesar 0,58. Validitas berkenaan dengan ketepatan alat penilai terhadap apa yang seharusnya dinilai sehingga benar-benar menilai apa yang akan dinilai. Validitas tes pada dasarnya menunjuk pada derajat fungsi pengukuran suatu tes atau derajat kecermatan ukurnya suatu tes.¹⁶⁰ Butir soal dikatakan valid apabila nilai koefisien relasinya lebih dari sama dengan 0,400. Hasil dari butir soal nomor 6, koefisien korelasinya lebih dari kisaran yang ditentukan. Hal ini

¹⁶⁰ Sumadi Suryabrata, *Pengembangan Alat Ukur Psikologis*, (Yogyakarta: Andi, 2000), hal. 41.

menunjukkan bahwa butir soal nomor 6 dapat dikatakan valid dan layak untuk diuji coba.

Korelasi reliabilitas pada butir soal nomor 6 dengan menggunakan rumus KR.20 dan berbantuan program *SPSS Statistic Versi 16.0* sebesar 0,641. Arifin menyatakan reliabilitas adalah tingkat atau derajat konsistensi dari sebuah tes, reliabilitas tes berkenaan dengan pertanyaan apakah suatu tes yang diteliti dan dapat dipercaya sesuai dengan kriteria yang telah ditetapkan.¹⁶¹ Sebuah tes dikatakan baik jika bersifat reliabel. Tes dikatakan reliabel jika tes tersebut diujikan berkali-kali dengan dua atau lebih penilai kepada subjek yang sama maka hasilnya relatif sama. Nilai korelasi reliabilitas yang baik yaitu mempunyai nilai di atas atau sama dengan 0,600. Hasil nilai dari reliabilitas pada butir soal nomor 6 menunjukkan butir soal tersebut reliabel dan layak digunakan dalam tes.

Selanjutnya kualitas butir soal nomor 6 meliputi daya pembeda soal, tingkat kesukaran dan keefektifan pengecoh. Butir soal nomor 6 mempunyai daya pembeda soal sebesar 0,55. Daya pembeda soal merupakan kemampuan suatu butir soal yang dapat membedakan antara siswa yang telah menguasai materi dengan siswa yang belum menguasai materi. Semakin tinggi daya pembeda suatu item, semakin baik suatu item dalam membedakan kemampuan penempuh tes.¹⁶² Daya pembeda soal yang baik berkisar antara $0,40 \leq D_p < 0,70$. Hasil dari butir soal nomor 6, daya pembeda soal berada kisaran yang ditentukan. Dengan demikian, butir soal nomor 6 dapat membedakan antara siswa yang sudah memahami materi dengan siswa yang belum memahami materi.

Tingkat kesukaran butir soal nomor 6 memiliki nilai sebesar 0,325. Suatu soal hendaknya tidak terlalu sukar dan tidak terlalu mudah. Jika suatu soal memiliki tingkat kesukaran seimbang (proporsional), maka dapat dikatakan bahwa soal tersebut baik.¹⁶³ Tingkat kesukaran yang baik yaitu memiliki rentang $0,30 < P \leq 0,70$. Hasil dari butir soal nomor 6, tingkat

¹⁶¹ Zainal Arifin, *Evaluasi Pembelajaran*, (Bandung: PT. Rosdakarya, 2013), hal. 258.

¹⁶² Zainal Arifin, *Evaluasi Pembelajaran*, (Bandung: PT. Remaja Rosdakarya, 2009).

¹⁶³ Zainal Arifin, *Evaluasi Pembelajaran*, (Bandung: PT. Remaja Rosdakarya, 2009).

kesukaran berada pada rentang yang ditentukan. Siswa dengan kemampuan tinggi akan merasa mudah dalam mengerjakan dan sebaliknya siswa dengan kemampuan rendah akan sulit menjawab butir soal nomor 6.

Keefektifan pengecoh jawaban dan alasan. pada butir soal nomor 6 dipilih lebih dari 5% dari peserta tes. Sebuah pengecoh dapat dikatakan berfungsi dengan baik apabila pengecoh tersebut mempunyai daya tarik yang besar bagi pengikut-pengikut tes yang kurang memahami konsep atau kurang menguasai bahan.¹⁶⁴ Pengecoh yang efektif dikatakan baik apabila pengecoh tersebut dipilih paling sedikit 2 dari siswa yang mengikuti tes. Hasil dari butir soal nomor 6, keefektifan pengecoh jawaban dan alasan lebih dari kisaran yang ditentukan sehingga menunjukkan pengecoh jawaban dan alasan tersebut berfungsi dengan baik.

Untuk mendapatkan karakteristik butir soal dengan kriteria baik, hendaknya memenuhi 3 kriteria persyaratan yaitu memiliki koefisien korelasi validitas yang baik, koefisien reliabilitas yang baik, dan kualitas butir soal yang baik. Karakteristik butir soal nomor 6 menunjukkan kriteria yang baik.

g. Butir soal nomor 7

Setelah melalui tahap revisi, berikut produk akhir terkait tes diagnostik *three-tier* untuk mengidentifikasi miskonsepsi pada materi geometri pada butir soal nomor 7:

“Panjang salah satu diagonal (d_1) belah ketupat adalah 24 cm. Jika luas belah ketupat 120 cm^2 maka panjang diagonal lainnya (d_2) pada bangun tersebut adalah ...

Pilihan jawaban:

- A. 30 cm
- B. 24 cm
- C. 20 cm
- D. 10 cm

Pilihan alasan:

- a. $d_2 = d_1$
- b. $d_2 > d_1$

¹⁶⁴ Arikunto, *Prosedur Penelitian: Suatu Pendekatan Praktek Edisi RevisiIV*, (Jakarta: PT. Rineka Cipta, 1997), hal. 225.

$$c. d_2 = \frac{2 \times Luas}{d_1}$$

$$d. d_2 = \sqrt{2 \times Luas}$$

Butir soal nomor 7 memiliki validitas dengan koefisien korelasi sebesar 0,55. Validitas berkenaan dengan ketepatan alat penilai terhadap apa yang seharusnya dinilai sehingga benar-benar menilai apa yang akan dinilai. Validitas tes pada dasarnya menunjuk pada derajat fungsi pengukuran suatu tes atau derajat kecermatan ukurnya suatu tes.¹⁶⁵ Butir soal dikatakan valid apabila nilai koefisien relasinya lebih dari sama dengan 0,400. Hasil dari butir soal nomor 7, koefisien korelasinya lebih dari kisaran yang ditentukan. Hal ini menunjukkan bahwa butir soal nomor 7 dapat dikatakan valid dan layak untuk diuji coba.

Korelasi reliabilitas pada butir soal nomor 7 dengan menggunakan rumus KR.20 dan berbantuan program *SPSS Statistic Versi 16.0* sebesar 0,657. Arifin menyatakan reliabilitas adalah tingkat atau derajat konsistensi dari sebuah tes, reliabilitas tes berkenaan dengan pertanyaan apakah suatu tes yang diteliti dan dapat dipercaya sesuai dengan kriteria yang telah ditetapkan.¹⁶⁶ Sebuah tes dikatakan baik jika bersifat reliabel. Tes dikatakan reliabel jika tes tersebut diujikan berkali-kali dengan dua atau lebih penilai kepada subjek yang sama maka hasilnya relatif sama. Nilai korelasi reliabilitas yang baik yaitu mempunyai nilai di atas atau sama dengan 0,600. Hasil nilai dari reliabilitas pada butir soal nomor 7 menunjukkan butir soal tersebut reliabel dan layak digunakan dalam tes.

Selanjutnya kualitas butir soal nomor 7 meliputi daya pembeda soal, tingkat kesukaran dan keefektifan pengecoh. Butir soal nomor 7 mempunyai daya pembeda soal sebesar 0,50. Daya pembeda soal merupakan kemampuan suatu butir soal yang dapat membedakan antara siswa yang telah menguasai materi dengan siswa yang belum menguasai materi. Semakin tinggi daya pembeda suatu item, semakin baik suatu

¹⁶⁵ Sumadi Suryabrata, *Pengembangan Alat Ukur Psikologis*, (Yogyakarta: Andi, 2000), hal. 41.

¹⁶⁶ Zainal Arifin, *Evaluasi Pembelajaran*, (Bandung: PT. Rosdakarya, 2013), hal. 258.

item dalam membedakan kemampuan penempuh tes.¹⁶⁷ Daya pembeda soal yang baik berkisar antara $0,40 \leq D_p < 0,70$. Hasil dari butir soal nomor 7, daya pembeda soal berada kisaran yang ditentukan. Dengan demikian, butir soal nomor 7 dapat membedakan antara siswa yang sudah memahami materi dengan siswa yang belum memahami materi.

Tingkat kesukaran butir soal nomor 7 memiliki nilai sebesar 0,55. Suatu soal hendaknya tidak terlalu sukar dan tidak terlalu mudah. Jika suatu soal memiliki tingkat kesukaran seimbang (proporsional), maka dapat dikatakan bahwa soal tersebut baik.¹⁶⁸ Tingkat kesukaran yang baik yaitu memiliki rentang $0,30 < P \leq 0,70$. Hasil dari butir soal nomor 7, tingkat kesukaran berada pada rentang yang ditentukan. Siswa dengan kemampuan tinggi akan merasa mudah dalam mengerjakan dan sebaliknya siswa dengan kemampuan rendah akan sulit menjawab butir soal nomor 7.

Keefektifan pengecoh jawaban dan alasan, pada butir soal nomor 7 dipilih lebih dari 5% dari peserta tes. Sebuah pengecoh dapat dikatakan berfungsi dengan baik apabila pengecoh tersebut mempunyai daya tarik yang besar bagi pengikut-pengikut tes yang kurang memahami konsep atau kurang menguasai bahan.¹⁶⁹ Pengecoh yang efektif dikatakan baik apabila pengecoh tersebut dipilih paling sedikit 2 dari siswa yang mengikuti tes. Hasil dari butir soal nomor 7, keefektifan pengecoh jawaban dan alasan lebih dari kisaran yang ditentukan sehingga menunjukkan pengecoh jawaban dan alasan tersebut berfungsi dengan baik.

Untuk mendapatkan karakteristik butir soal dengan kriteria baik, hendaknya memenuhi 3 kriteria persyaratan yaitu memiliki koefisien korelasi validitas yang baik, koefisien reliabilitas yang baik, dan kualitas butir soal yang baik. Karakteristik butir soal nomor 7 menunjukkan kriteria yang baik.

¹⁶⁷ Zainal Arifin, *Evaluasi Pembelajaran*, (Bandung: PT. Remaja Rosdakarya, 2009).

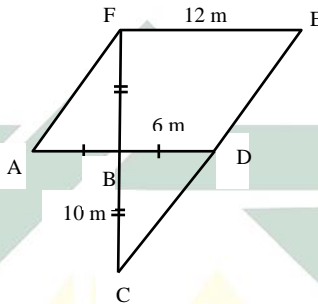
¹⁶⁸ Zainal Arifin, *Evaluasi Pembelajaran*, (Bandung: PT. Remaja Rosdakarya, 2009).

¹⁶⁹ Arikunto, *Prosedur Penelitian: Suatu Pendekatan Praktek Edisi RevisiIV*, (Jakarta: PT. Rineka Cipta, 1997), hal. 225.

h. Butir soal nomor 8

Setelah melalui tahap revisi, berikut produk akhir terkait tes diagnostik *three-tier* untuk mengidentifikasi miskonsepsi pada materi geometri pada butir soal nomor 8:

“Pak Ali mempunyai kebun dengan bentuk seperti pada gambar.



Kebun tersebut akan dijual dengan harga Rp 200.000,00/m². Hasil penjualan kebun Pak Ali adalah ...

Pilihan jawaban:

- A. Rp 28.800.000,00
- B. Rp 30.000.000,00
- C. Rp 36.000.000,00
- D. Rp 57.600.000,00

Pilihan alasan:

- a. $\text{Harga tanah} = \text{Harga jual} \times \text{Luas kebun}$
- b. $\text{Harga tanah} = \frac{\text{Harga jual}}{\text{Luas kebun}}$
- c. Luas kebun adalah 100 m²
- d. Luas kebun adalah 130 m²

Butir soal nomor 8 memiliki validitas dengan koefisien korelasi sebesar 0,53. Validitas berkenaan dengan ketepatan alat penilai terhadap apa yang seharusnya dinilai sehingga benar-benar menilai apa yang akan dinilai. Validitas tes pada dasarnya menunjuk pada derajat fungsi pengukuran suatu tes atau derajat kecermatan ukurnya suatu tes.¹⁷⁰ Butir soal dikatakan valid apabila nilai koefisien relasinya lebih dari sama dengan 0,400. Hasil dari butir soal nomor 8, koefisien

¹⁷⁰ Sumadi Suryabrata, *Pengembangan Alat Ukur Psikologis*, (Yogyakarta: Andi, 2000), hal. 41.

korelasinya lebih dari kisaran yang ditentukan. Hal ini menunjukkan bahwa butir soal nomor 8 dapat dikatakan valid dan layak untuk diuji coba.

Korelasi reliabilitas pada butir soal nomor 8 dengan menggunakan rumus KR.20 dan berbantuan program *SPSS Statistic Versi 16.0* sebesar 0,661. Arifin menyatakan reliabilitas adalah tingkat atau derajat konsistensi dari sebuah tes, reliabilitas tes berkenaan dengan pertanyaan apakah suatu tes yang diteliti dan dapat dipercaya sesuai dengan kriteria yang telah ditetapkan.¹⁷¹ Sebuah tes dikatakan baik jika bersifat reliabel. Tes dikatakan reliabel jika tes tersebut diujikan berkali-kali dengan dua atau lebih penilai kepada subjek yang sama maka hasilnya relatif sama. Nilai korelasi reliabilitas yang baik yaitu mempunyai nilai di atas atau sama dengan 0,600. Hasil nilai dari reliabilitas pada butir soal nomor 8 menunjukkan butir soal tersebut reliabel dan layak digunakan dalam tes.

Selanjutnya kualitas butir soal nomor 8 meliputi daya pembeda soal, tingkat kesukaran dan keefektifan pengecoh. Butir soal nomor 8 mempunyai daya pembeda soal sebesar 0,50. Daya pembeda soal merupakan kemampuan suatu butir soal yang dapat membedakan antara siswa yang telah menguasai materi dengan siswa yang belum menguasai materi. Semakin tinggi daya pembeda suatu item, semakin baik suatu item dalam membedakan kemampuan penempuh tes.¹⁷² Daya pembeda soal yang baik berkisar antara $0,40 \leq D_p < 0,70$. Hasil dari butir soal nomor 8, daya pembeda soal berada kisaran yang ditentukan. Dengan demikian, butir soal nomor 8 dapat membedakan antara siswa yang sudah memahami materi dengan siswa yang belum memahami materi.

Tingkat kesukaran butir soal nomor 8 memiliki nilai sebesar 0,40. Suatu soal hendaknya tidak terlalu sukar dan tidak terlalu mudah. Jika suatu soal memiliki tingkat kesukaran seimbang (proporsional), maka dapat dikatakan bahwa soal tersebut baik.¹⁷³ Tingkat kesukaran yang baik yaitu memiliki

¹⁷¹ Zainal Arifin, *Evaluasi Pembelajaran*, (Bandung: PT. Rosdakarya, 2013), hal. 258.

¹⁷² Zainal Arifin, *Evaluasi Pembelajaran*, (Bandung: PT. Remaja Rosdakarya, 2009).

¹⁷³ Zainal Arifin, *Evaluasi Pembelajaran*, (Bandung: PT. Remaja Rosdakarya, 2009).

rentang $0,30 < P \leq 0,70$. Hasil dari butir soal nomor 8, tingkat kesukaran berada pada rentang yang ditentukan. Siswa dengan kemampuan tinggi akan merasa mudah dalam mengerjakan dan sebaliknya siswa dengan kemampuan rendah akan sulit menjawab butir soal nomor 8.

Keefektifan pengecoh jawaban dan alasan. pada butir soal nomor 8 dipilih lebih dari 5% dari peserta tes. Sebuah pengecoh dapat dikatakan berfungsi dengan baik apabila pengecoh tersebut mempunyai daya tarik yang besar bagi pengikut-pengikut tes yang kurang memahami konsep atau kurang menguasai bahan.¹⁷⁴ Pengecoh yang efektif dikatakan baik apabila pengecoh tersebut dipilih paling sedikit 2 dari siswa yang mengikuti tes. Hasil dari butir soal nomor 8, keefektifan pengecoh jawaban dan alasan lebih dari kisaran yang ditentukan sehingga menunjukkan pengecoh jawaban dan alasan tersebut berfungsi dengan baik.

Untuk mendapatkan karakteristik butir soal dengan kriteria baik, hendaknya memenuhi 3 kriteria persyaratan yaitu memiliki koefisien korelasi validitas yang baik, koefisien reliabilitas yang baik, dan kualitas butir soal yang baik. Karakteristik butir soal nomor 8 menunjukkan kriteria yang baik.

i. Butir soal nomor 9

Setelah melalui tahap revisi, berikut produk akhir terkait tes diagnostik *three-tier* untuk mengidentifikasi miskonsepsi pada materi geometri pada butir soal nomor 9:

“Bu Nur memiliki sebidang tanah berbentuk persegi panjang dengan panjang 10 m dan lebar x m. Jika luasnya adalah 60 m^2 maka keliling tanah tersebut adalah ...

Pilihan jawaban:

- A. 54 m^2
- B. 46 m^2
- C. 32 m^2
- D. 28 m^2

Pilihan alasan:

- a. *Lebar = Luas \times Panjang*

¹⁷⁴ Arikunto, *Prosedur Penelitian: Suatu Pendekatan Praktek Edisi RevisiIV*, (Jakarta: PT. Rineka Cipta, 1997), hal. 225.

$$b. \text{ Lebar} = \frac{\text{Luas}}{\text{Panjang}}$$

$$c. \text{ Keliling} = 2 \times \text{Panjang} \times \text{Lebar}$$

$$d. \text{ Keliling} = 4 \times (\text{Panjang} + \text{Lebar})''$$

Butir soal nomor 9 memiliki validitas dengan koefisien korelasi sebesar 0,60. Validitas berkenaan dengan ketepatan alat penilai terhadap apa yang seharusnya dinilai sehingga benar-benar menilai apa yang akan dinilai. Validitas tes pada dasarnya menunjuk pada derajat fungsi pengukuran suatu tes atau derajat kecermatan ukurnya suatu tes.¹⁷⁵ Butir soal dikatakan valid apabila nilai koefisien relasinya lebih dari sama dengan 0,400. Hasil dari butir soal nomor 9, koefisien korelasinya lebih dari kisaran yang ditentukan. Hal ini menunjukkan bahwa butir soal nomor 9 dapat dikatakan valid dan layak untuk diuji coba.

Korelasi reliabilitas pada butir soal nomor 9 dengan menggunakan rumus KR.20 dan berbantuan program *SPSS Statistic Versi 16.0* sebesar 0,637. Arifin menyatakan reliabilitas adalah tingkat atau derajat konsistensi dari sebuah tes, reliabilitas tes berkenaan dengan pertanyaan apakah suatu tes yang diteliti dan dapat dipercaya sesuai dengan kriteria yang telah ditetapkan.¹⁷⁶ Sebuah tes dikatakan baik jika bersifat reliabel. Tes dikatakan reliabel jika tes tersebut diujikan berkali-kali dengan dua atau lebih penilai kepada subjek yang sama maka hasilnya relatif sama. Nilai korelasi reliabilitas yang baik yaitu mempunyai nilai di atas atau sama dengan 0,600. Hasil nilai dari reliabilitas pada butir soal nomor 9 menunjukkan butir soal tersebut reliabel dan layak digunakan dalam tes.

Selanjutnya kualitas butir soal nomor 9 meliputi daya pembeda soal, tingkat kesukaran dan keefektifan pengecoh. Butir soal nomor 9 mempunyai daya pembeda soal sebesar 0,55. Daya pembeda soal merupakan kemampuan suatu butir soal yang dapat membedakan antara siswa yang telah menguasai materi dengan siswa yang belum menguasai materi. Semakin tinggi daya pembeda suatu item, semakin baik suatu

¹⁷⁵ Sumadi Suryabrata, *Pengembangan Alat Ukur Psikologis*, (Yogyakarta: Andi, 2000), hal. 41.

¹⁷⁶ Zainal Arifin, *Evaluasi Pembelajaran*, (Bandung: PT. Rosdakarya, 2013), hal. 258.

item dalam membedakan kemampuan penempuh tes.¹⁷⁷ Daya pembeda soal yang baik berkisar antara $0,40 \leq D_p < 0,70$. Hasil dari butir soal nomor 9, daya pembeda soal berada kisaran yang ditentukan. Dengan demikian, butir soal nomor 9 dapat membedakan antara siswa yang sudah memahami materi dengan siswa yang belum memahami materi.

Tingkat kesukaran butir soal nomor 9 memiliki nilai sebesar 0,525. Suatu soal hendaknya tidak terlalu sukar dan tidak terlalu mudah. Jika suatu soal memiliki tingkat kesukaran seimbang (proporsional), maka dapat dikatakan bahwa soal tersebut baik.¹⁷⁸ Tingkat kesukaran yang baik yaitu memiliki rentang $0,30 < P \leq 0,70$. Hasil dari butir soal nomor 9, tingkat kesukaran berada pada rentang yang ditentukan. Siswa dengan kemampuan tinggi akan merasa mudah dalam mengerjakan dan sebaliknya siswa dengan kemampuan rendah akan sulit menjawab butir soal nomor 9.

Keefektifan pengecoh jawaban pada butir soal nomor 9 dipilih lebih dari 5% dari peserta tes tetapi pada pengecoh alasan dengan alternatif C dan D dipilih kurang dari 5% dari peserta tes. Sebuah pengecoh dapat dikatakan berfungsi dengan baik apabila pengecoh tersebut mempunyai daya tarik yang besar bagi pengikut-pengikut tes yang kurang memahami konsep atau kurang menguasai bahan.¹⁷⁹ Pengecoh yang efektif dikatakan baik apabila pengecoh tersebut dipilih paling sedikit 2 dari siswa yang mengikuti tes. Hasil dari butir soal nomor 9, keefektifan pengecoh alasan kurang dari kisaran yang ditentukan sehingga menunjukkan pengecoh alasan tersebut berfungsi tidak baik.

Untuk mendapatkan karakteristik butir soal dengan kriteria baik, hendaknya memenuhi 3 kriteria persyaratan yaitu memiliki koefisien korelasi validitas yang baik, koefisien reliabilitas yang baik, dan kualitas butir soal yang baik. Karakteristik butir soal nomor 9 menunjukkan kriteria yang cukup baik.

¹⁷⁷ Zainal Arifin, *Evaluasi Pembelajaran*, (Bandung: PT. Remaja Rosdakarya, 2009).

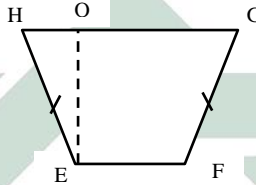
¹⁷⁸ Zainal Arifin, *Evaluasi Pembelajaran*, (Bandung: PT. Remaja Rosdakarya, 2009).

¹⁷⁹ Arikunto, *Prosedur Penelitian: Suatu Pendekatan Praktek Edisi RevisiIV*, (Jakarta: PT. Rineka Cipta, 1997), hal. 225.

j. Butir soal nomor 10

Setelah melalui tahap revisi, berikut produk akhir terkait tes diagnostik *three-tier* untuk mengidentifikasi miskonsepsi pada materi geometri pada butir soal nomor 10:

“Perhatikan gambar trapesium sama kaki tersebut. Jika diketahui $GH = 30$ cm, $EF = 20$ cm dan kelilingnya adalah 76 cm, maka luas trapesium tersebut adalah ...



Pilihan jawaban:

- A. 300 cm^2
- B. 475 cm^2
- C. 650 cm^2
- D. 825 cm^2

Pilihan alasan:

- a. $\text{Luas} = a \times t$
- b. $\text{Luas} = \frac{a \times t}{2}$
- c. $\text{Luas} = \text{Jumlah dua sisi sejajar} \times t$
- d. $\text{Luas} = \frac{1}{2} \times \text{Jumlah dua sisi sejajar} \times t$

Butir soal nomor 10 memiliki validitas dengan koefisien korelasi sebesar 0,52. Validitas berkenaan dengan ketepatan alat penilai terhadap apa yang seharusnya dinilai sehingga benar-benar menilai apa yang akan dinilai. Validitas tes pada dasarnya menunjuk pada derajat fungsi pengukuran suatu tes atau derajat kecermatan ukurnya suatu tes.¹⁸⁰ Butir soal dikatakan valid apabila nilai koefisien relasinya lebih dari sama dengan 0,400. Hasil dari butir soal nomor 10, koefisien korelasinya lebih dari kisaran yang ditentukan. Hal ini menunjukkan bahwa butir soal nomor 10 dapat dikatakan valid dan layak untuk diuji coba.

¹⁸⁰ Sumadi Suryabrata, *Pengembangan Alat Ukur Psikologis*, (Yogyakarta: Andi, 2000), hal. 41.

Korelasi reliabilitas pada butir soal nomor 10 dengan menggunakan rumus KR.20 dan berbantuan program *SPSS Statistic Versi 16.0* sebesar 0,656. Arifin menyatakan reliabilitas adalah tingkat atau derajat konsistensi dari sebuah tes, reliabilitas tes berkenaan dengan pertanyaan apakah suatu tes yang diteliti dan dapat dipercaya sesuai dengan kriteria yang telah ditetapkan.¹⁸¹ Sebuah tes dikatakan baik jika bersifat reliabel. Tes dikatakan reliabel jika tes tersebut diujikan berkali-kali dengan dua atau lebih penilai kepada subjek yang sama maka hasilnya relatif sama. Nilai korelasi reliabilitas yang baik yaitu mempunyai nilai di atas atau sama dengan 0,600. Hasil nilai dari reliabilitas pada butir soal nomor 10 menunjukkan butir soal tersebut reliabel dan layak digunakan dalam tes.

Selanjutnya kualitas butir soal nomor 10 meliputi daya pembeda soal, tingkat kesukaran dan keefektifan pengecoh. Butir soal nomor 10 mempunyai daya pembeda soal sebesar 0,30. Daya pembeda soal merupakan kemampuan suatu butir soal yang dapat membedakan antara siswa yang telah menguasai materi dengan siswa yang belum menguasai materi. Semakin tinggi daya pembeda suatu item, semakin baik suatu item dalam membedakan kemampuan penempuh tes.¹⁸² Daya pembeda soal yang baik berkisar antara $0,40 \leq D_p < 0,70$. Hasil dari butir soal nomor 10, daya pembeda soal berada pada kategori cukup baik sebagaimana telah ditetapkan pada BAB III dalam Tabel 3.5. Hal ini menunjukkan bahwa butir soal nomor 10 cukup memadai dalam membedakan antara siswa yang sudah memahami materi dengan siswa yang belum memahami materi.

Tingkat kesukaran butir soal nomor 10 memiliki nilai sebesar 0,20. Suatu soal hendaknya tidak terlalu sukar dan tidak terlalu mudah. Jika suatu soal memiliki tingkat kesukaran seimbang (proporsional), maka dapat dikatakan bahwa soal tersebut baik.¹⁸³ Tingkat kesukaran yang baik yaitu memiliki rentang $0,30 < P \leq 0,70$. Hasil dari butir soal nomor 10,

¹⁸¹ Zainal Arifin, *Evaluasi Pembelajaran*, (Bandung: PT. Rosdakarya, 2013), hal. 258.

¹⁸² Zainal Arifin, *Evaluasi Pembelajaran*, (Bandung: PT. Remaja Rosdakarya, 2009).

¹⁸³ Zainal Arifin, *Evaluasi Pembelajaran*, (Bandung: PT. Remaja Rosdakarya, 2009).

tingkat kesukaran kurang dari kisaran yang ditentukan. Hal tersebut menunjukkan bahwa siswa dengan kemampuan tinggi merasa kesulitan dalam mengerjakan dan sebaliknya siswa dengan kemampuan rendah merasa mudah dalam menjawab butir soal nomor 10. Dilihat dari perolehan menjawab, sebanyak 8 siswa yang menjawab benar pada butir soal nomor 10 sehingga kemungkinan siswa yang menjawab kebanyakan adalah siswa dengan kemampuan rendah.

Keefektifan pengecoh jawaban pada butir soal nomor 9 dipilih lebih dari 5% dari peserta tes tetapi pada pengecoh alasan dengan alternatif A dipilih kurang dari 5% dari peserta tes. Sebuah pengecoh dapat dikatakan berfungsi dengan baik apabila pengecoh tersebut mempunyai daya tarik yang besar bagi pengikut-pengikut tes yang kurang memahami konsep atau kurang menguasai bahan.¹⁸⁴ Pengecoh yang efektif dikatakan baik apabila pengecoh tersebut dipilih paling sedikit 2 dari siswa yang mengikuti tes. Hasil dari butir soal nomor 10, keefektifan pengecoh alasan kurang dari kisaran yang ditentukan sehingga menunjukkan pengecoh alasan tersebut berfungsi tidak baik.

Untuk mendapatkan karakteristik butir soal dengan kriteria baik, hendaknya memenuhi 3 kriteria persyaratan yaitu memiliki koefisien korelasi validitas yang baik, koefisien reliabilitas yang baik, dan kualitas butir soal yang baik. Karakteristik butir soal nomor 10 menunjukkan kriteria yang cukup baik.

2. Temuan Terkait Hasil Tes Diagnostik Pilihan Ganda *Three-Tier* Untuk Mengidentifikasi Miskonsepsi Pada Materi Geometri

Berdasarkan hasil analisis data yang diperoleh, terlihat berbagai miskonsepsi yang terjadi pada siswa. Hal ini membuktikan bahwa penanaman konsep pada siswa sangat penting terutama pada materi bangun datar segiempat. Sebagian besar pendidik hanya mengenalkan konsep bukan memahamkan konsep. Peneliti berusaha mengidentifikasi miskonsepsi apa saja yang terjadi pada materi bangun datar segiempat melalui tes diagnostik pilihan ganda *three-tier* yang telah dikembangkan dan telah diuji kevalidan, reliabilitas,

¹⁸⁴ Arikunto, *Prosedur Penelitian: Suatu Pendekatan Praktek Edisi RevisiIV*, (Jakarta: PT. Rineka Cipta, 1997), hal. 225.

dan kualitas butir soalnya. Berikut ini merupakan pembahasan mengenai persentase serta penyebab dari temuan miskonsepsi yang dialami oleh siswa.

Indikator pencapaian kompetensi pertama yakni menjelaskan definisi bangun datar segiempat, menghasilkan persentase miskonsepsi tinggi sebesar 67,5% dikarenakan siswa banyak mengalami kesulitan dalam mengungkapkan definisi bangun datar segiempat. Dilihat dari permasalahan tersebut, siswa mengalami miskonsepsi teoritikal yakni bentuk miskonsepsi yang didasarkan atas kesalahan dalam mempelajari fakta-fakta atau kejadian-kejadian dalam sistem yang terorganisir.¹⁸⁵ Hal ini sejalan dengan salah satu indikator miskonsepsi yang dijelaskan Ainiyah dalam jurnalnya berupa siswa kurang mampu dalam mendefinisikan konsep bangun datar.¹⁸⁶ Sebagian besar siswa salah memahami konsep saat mendefinisikan segiempat yang terlihat dalam hasil tes diagnostik pilihan ganda *three-tier* yang mana mereka menganggap bahwa segiempat selalu dalam bentuk beraturan saja.

Indikator pencapaian kompetensi kedua yaitu menyebutkan sifat-sifat dari berbagai jenis segiempat menghasilkan persentase miskonsepsi dengan rata-rata sebesar 52,5%. Pada butir soal nomor 3 siswa mengalami miskonsepsi tinggi pada indikator sub pokok bahasan sifat bangun datar jajargenjang, menghasilkan persentase miskonsepsi berkategori tinggi sebesar 87,5% dikarenakan sebagian besar siswa belum memahami konsep ilmiah jajargenjang. Dalam permasalahan tersebut sejalan dengan pemikiran Salirawati yang mengatakan bahwa siswa tersebut termasuk kedalam miskonsepsi teoritikal yakni siswa mendefinisikan jajargenjang yang tidak sesuai dengan pengertiannya.¹⁸⁷ Sebagian besar siswa mengalami kesulitan dalam menyebutkan sifat bangun datar jajargenjang yang mana mereka menyebutkan jajargenjang adalah bangun datar yang memiliki sepasang sisi miring. Namun pada butir soal nomor 4, siswa mengalami miskonsepsi berkategori rendah sebesar 25%. Dalam keadaan tersebut, siswa termasuk kedalam miskonsepsi

¹⁸⁵ Das Salirawati, Disertasi: "Pengembangan Instrumen Pendeteksi Miskonsepsi Kimia pada Peserta Didik SMA", (Yogyakarta: PPs-UNY, 2011), hal. 36.

¹⁸⁶ Lutfia Afifatul Ainiyah, "*Identifikasi Miskonsepsi Siswa Dalam Materi Geometri Pada Pembelajaran Matematika Siswa Kelas VIII SMP Negeri 1 Punggalen*", diakses dari journal.student.uny.ac.id, pada tanggal 14 Desember 2019.

¹⁸⁷ Das Salirawati, Op., Cit, hal. 36.

klasifikasional yakni miskonsepsi yang didasarkan atas kesalahan klasifikasi fakta-fakta ke dalam bagan-bagan yang terorganisir.¹⁸⁸ Siswa melakukan kesalahan dalam menentukan benda-benda yang berbentuk lingkaran dan layang-layang termasuk jajargenjang atau menganggap benda-benda yang berbentuk persegi, persegi panjang, dan belah ketupat adalah bukan termasuk jajargenjang.

Pada indikator pencapaian kompetensi ketiga yaitu menentukan rumus keliling dan luas dari berbagai bangun datar segiempat, menghasilkan persentase miskonsepsi berkategori sedang dengan rata-rata sebesar 36,25%. Pada indikator butir soal nomor 5 siswa mengalami miskonsepsi sebesar 57,5% dikarenakan siswa kurang paham dalam mempresentasikan setiap soal dalam bentuk gambar. Dalam hal ini siswa tersebut termasuk kedalam miskonsepsi korelasional yakni siswa tidak dapat menjelaskan hubungan antar bangun datar.¹⁸⁹ Siswa mengalami kelemahan dalam menerapkan hubungan antar rumus bangun datar yang digunakan dengan permasalahan yang terdapat dalam soal. Begitu pun dengan indikator pencapaian kompetensi keempat yaitu menyelesaikan masalah kontekstual yang berkaitan dengan luas dan keliling bangun datar segiempat, menghasilkan persentase miskonsepsi berkategori sedang dengan rata-rata sebesar 31,25%. Siswa kesulitan dalam merepresentasikan permasalahan yang berkaitan dengan kehidupan sehari-hari, sebagaimana sejalan dengan pemikiran Salirawati yang mengatakan bahwa siswa tersebut termasuk kedalam miskonsepsi korelasional yakni bentuk miskonsepsi yang didasarkan atas kesalahan mengenai kejadian-kejadian khusus yang saling berhubungan, atau observasi-observasi yang terdiri atas dugaan-dugaan terutama berbentuk formulasi prinsip-prinsip umum.¹⁹⁰ Dalam hal ini, siswa mengalami kelemahan dalam menerapkan hubungan antar rumus bangun datar yang digunakan dengan permasalahan kontekstual yang terdapat dalam soal serta kesulitan dalam merepresentasikan soal ke dalam bentuk jawaban.

¹⁸⁸ Ibid.

¹⁸⁹ Lutfia Afifatul Ainiah, Op., Cit. hal. 19.

¹⁹⁰ Das Salirawati, Op., Cit, hal. 36.

BAB V

PENUTUP

A. Simpulan

Berdasarkan hasil penelitian pengembangan tes diagnostik pilihan ganda *three-tier* untuk mengidentifikasi miskonsepsi pada materi geometri dapat diambil kesimpulan sebagai berikut:

1. Hasil analisis validitas dari tes diagnostik pilihan ganda *three-tier* menunjukkan bahwa dari kesepuluh butir soal yang dikembangkan terdapat 8 dari 10 butir soal yang memiliki tingkat kevalidan yang baik dan 2 butir soal yang perlu direvisi kembali karena memiliki tingkat kevalidan yang rendah.
2. Reliabilitas tes diagnostik pilihan ganda *three-tier* yang dikembangkan dinyatakan reliabel dengan rata-rata kategori tinggi yaitu 0,641.
3. Kualitas 10 butir soal tes diagnostik pilihan ganda *three-tier* yang dikembangkan yang meliputi daya pembeda soal, tingkat kesukaran dan keefektifan pengecoh termasuk kategori baik, hanya butir soal nomor 3 dengan daya pembeda, tingkat kesukaran dan keefektifan pengecoh yang mempunyai kategori tidak baik, butir soal nomor 4 dan 5 dengan daya pembeda soal yang tidak baik, serta butir soal nomor 9 dan 10 dengan keefektifan pengecoh alasan yang tidak berfungsi dengan baik.
4. Tes diagnostik pilihan ganda *three-tier* yang dikembangkan menunjukkan bahwa secara umum siswa mengalami jenis miskonsepsi teoritikal pada indikator mendefinisikan bangun datar segiempat dan jajargenjang.

B. Saran

Berdasarkan kesimpulan di atas maka peneliti memberikan beberapa saran sebagai berikut:

1. Produk yang berupa tes diagnostik pilihan ganda *three-tier* untuk mengidentifikasi miskonsepsi pada materi geometri, dapat dijadikan contoh atau acuan bagi guru matematika Sekolah Menengah Pertama (SMP) untuk membuat tes diagnostik yang bertujuan mengidentifikasi miskonsepsi yang terjadi pada siswa. Bagi peneliti yang akan mengembangkan penelitian ini, disarankan melakukan penelitian lanjutan mengenai pengaruh tingkat kesukaran butir soal tes diagnostik

pilihan ganda *three-tier* terhadap miskonsepsi yang terjadi pada siswa.

2. Tes diagnostik pilihan ganda *three-tier* ini hendaknya diuji cobakan pada kapasitas kelas besar atau sekolah lain sehingga diperoleh tes diagnostik yang lebih baik.
3. Bagi peneliti yang ingin mengembangkan penelitian ini, disarankan bentuk soal yang dikembangkan nantinya lebih bervariasi tidak hanya pada bentuk soal pilihan ganda saja.



DAFTAR PUSTAKA

- Ainiyah, Lutfia Afifatul. *Identifikasi Miskonsepsi Siswa Dalam Materi Geometri Pada Pembelajaran Matematika Siswa Kelas VIII SMP Negeri 1 Punggalen*. Diakses pada 14 Desember, 2019; journal.student.uny.ac.id; internet.
- Anne, Anastasia., Susana Urbina. 1997. *Psychological Testing*. New Jersey: Prentice-Hall Inc.
- Arifin, Zaenal. 2017. "Kriteria Instrumen Dalam Suatu Penelitian", *Jurnal THEOREMS (The Original Research of Mathematics)*. Vol. 2 No. 1.
- Arifin, Zainal. 2013. *Evaluasi Pembelajaran*. Bandung: PT. Remaja Rosdakarya.
- Arikunto, Suharsimi. 2016. *Dasar-Dasar Evaluasi Pendidikan Edisi 2*. Jakarta: Bumi Aksara.
- Arikunto, Suharsimi. 2010. *Penelitian Tindakan*. Yogyakarta: Aditya Media.
- Arikunto, Suharsimi. 2006. *Prosedur Penelitian Suatu Pendekatan Praktik*. Jakarta: PT. Rineka Cipta.
- Arikunto, Suharsimi. 2003. *Prosedur Penelitian Suatu Praktek*. Jakarta: Bima Aksara.
- Arikunto, Suharsimi. 1999. *Prosedur Penelitian: Suatu Pendekatan Praktik Edisi Revisi IV*. Yogyakarta: Rineka Cipta.
- Arikunto, Suharsimi. 1993. *Dasar-Dasar Evaluasi Pendidikan Cetakan 10*. Jakarta: Bumi Aksara.
- Arikunto, Suharsimi. 1991. *Metode Penelitian*. Jakarta: Rineka Cipta.
- Arslan, Harika Ozge., Ceyhan Cigdemoglu & Christine Mosley. 2012. "A Three-Tier Diagnostic Test to Assess Pre-Service Teacher's

Misconceptions About Global Warming, Greenhouse Effect, Ozone Layer Depletion, and Acid Rain”. *International Journal of Science Education*. Vol. 34 No. 11.

- Basuki, Novila Rahmad. “*Analisis Kesulitan Siswa SMK Pada Materi Pokok Geometri dan Alternatif Pemecahannya*”. Prosiding Seminar Nasional Pendidikan Matematika Universitas Muhammadiyah Surakarta, 2012.
- Budiarto, Mega Teguh. “*Pembelajaran Geometri dan Berpikir Geometri*”. Prosiding Seminar Nasional Matematika FMIPA ITS Surabaya, Surabaya, 2000.
- Budiarto, Mega Teguh., Rudianto Artiono. 2019. “Geometri dan Permasalahan Dalam Pembelajarannya (Suatu Penelitian Meta Analisis)”, *Jurnal Megister Pendidikan Matematika*. Vo. 1 No. 1.
- Budiarto, Mega Teguh., Rachmania Widya Ningrum. 2016. “Miskonsepsi Siswa SMP Pada Materi Bangun Datar Segiempat dan Alternatif Mengatasinya”, *Jurnal Ilmiah Pendidikan Matematika*. Vol. 1 No. 5.
- Bruckner, Leo John., Ernest O. Melby. 1981. *Diagnostic and Remedial Teaching*. Boston: Houghton Mifflin Company.
- Cahyani, Fatmawati Nur Indah. Skripsi: “*Analisis Miskonsepsi Siswa Materi Bangun Datar Segiempat Dibedakan dari Gaya Kognitif Siswa*”. Surabaya: UIN Sunan Ampel Surabaya, 2018.
- Cichetti, Domenic Vincent., Richard Bronen, and Peter Tyrer. 2006. “Rating Scales, Scales of Measurement, Issues of Reliability Resolving Some Critical Issues for Clinicians and Research”. *The Journal of Nervous and Mental Disease*, 194, 557-564.
- Dahar, Ratna Wilis. 2011. *Teori-Teori Belajar & Pembelajaran*. Bandung: Erlangga.

- Dindar, Ayla Cetin., Omer Geban. "Developmet of a Three-Tier Test to Assess High School Students Understanding of Acids and Bases", *Procedia Sosial and Behavioral Sciences* 15 Turkey, 2011.
- Ebel, Robbert L., David A. Fribie. 1991. *Essentials of Educational Measurment*. New Jersey : Prentice-Hall Inc.
- Farida, Anisatul. "Analisis Miskonsepsi Siswa Terhadap Simbol dan Istilah Matematika pada Konsep Hubungan Bangun Datar Segiempat Melalui Permainan dengan Alat Peraga (SD Muhammadiyah 1 Surakarta)". *Prosiding Konferensi Nasional Penelitian Matematika dan Pembelajarannya*, Yogyakarta, 2016.
- Hamzah, Ali. 2014. *Evaluasi Pembelajaran Matematika*. Jakarta: Rajawali Press.
- Hariwijaya, Muhammad., Sutan Surya. 2012. *Adventure in Math Tes IQ Matematika*. Jakarta: Oryza.
- Hasan, Saleem., Diola Bagayoko, and Ella L. Kelley. 1999. "Misconceptions and the Certainty of Response Index (CRI)". *Physics Education Research Amerika Journal of Physics*. Vol. 34 No. 5.
- Hakim, Aliefman,. Liliasari, dan Asep Kadarohman. 2012. "Student Concept Understanding of Natural Product Chemistry in Primary and Secondary Metabolites Using the Data Collecting Technique of Modified CRI". *International Online Journal of EducationSciences*. Vol.4No.3.
- Joel, Klammer. 1998. *An Overview of Techniques for Identifying, Acknowledging and Overcoming Alternate Conceptions in Physics Education. 1997/98 Klingenstein Project Report*. (Teachers College-Columbia University).
- Khabibah, Siti. Disertasi: "Pengembangan Perangkat Pembelajaran Matematika dengan Soal Terbuka Untuk Meningkatkan

Kreatifitas Siswa Sekolah Dasar". (Surabaya: Unesa, tidak dipublikasikan, 2006).

- Kirbulut, Zubeyde Demet. 2014. "Using Three-Tier Diagnostic Test to Assess Student's Misconceptions of States of Matter". *Eurasia Journal of Mathematics, Science & Technology Education*. Vol. 10 No. 5.
- Kristiani, Kornelia Devi., Tantri Mayasari, dan Erawan Kurniadi. "Pengembangan Asesmen Keterampilan Berpikir Kreatif Siswa SMP Negeri 5 Madiun Pada Materi Cahaya dan Alat Optik". Prosiding Seminar Nasional Fisika dan Aplikasinya (SNFA), Madiun, 2017.
- Kuncoro, Achmad., Engkos, dan Riduan. 2011. *Cara Menggunakan dan Memakai Path Analysis*. Bandung: CV Alfabeta.
- Lestari, Karunia Eka., Yudhanegara, Muhammad Ridwan. 2015. *Penelitian Pendidikan Matematika Cet. I*. Bandung: PT. Rafika Aditama.
- Lin, Robert L., Noman E. Gronlund. 2008. *Measurement and Assessment in Teaching*. New York: Mac Milian Publishing Co.
- Majid, Abdul. Tesis: "Pengembangan Modul Matematika pada Materi Garis dan Sudut Setting Pembelajaran Kooperatif Tipe Think Pair Share (TPS) untuk Siswa Kelas VII SMP". (Makassar: UNM, 2014).
- Makmun, Abin Syamsuddin. 2009. *Psikologi Kependidikan Perangkat Sistem Pengajaran Modul*. Bandung: Remaja Rosdakarya.
- Matondang, Zulkifli. 2009. "Validitas dan Reliabilitas Suatu Instrumen Penelitian". *Jurnal Tabularasa PPS UNIMED*. Vol. 6 No. 1.
- Mubarak, Syarifatul., Endang Susilaningsih, dan Edy Cahyono. 2016. "Pengembangan Tes Diagnostik Three Tier Multiple Choice Untuk Mengidentifikasi Miskonsepsi Peserta Didik Kelas XI". *Journal of Innovative Science Education*. Vol. 5 No. 2.

- Mungin, Burhan. 2011. *Metodologi Penelitian Kuantitatif: Komunikasi, Ekonomi, dan Kebijakan Publik serta Ilmu-ilmu Sosial Lainnya*. Jakarta: Kencana Prenada Media Group.
- Mustaqim, Tri Ade., Zulfiani, dan Yanti Herlanti. 2014. "Identifikasi Miskonsepsi Siswa dengan Menggunakan Metode *Certainty of Response Index* (CRI) Pada Konsep Fotosintesis dan Respirasi Tumbuhan". *EDUSAINS*. Vol. 6 No. 2.
- Ngilman, Malikatun., Endah Budi Rahaju. 2017. "Identifikasi Tahap Pemahaman Geometri Siswa Berdasarkan Teori Van Hiele Ditinjau dari Perbedaan Gender pada Materi Persegipanjang Kelas VIII". *Jurnal Ilmiah Pendidikan Matematika*. Vol. 2 No. 6.
- Nursiwin, "Menggali Miskonsepsi Siswa Pada Materi Perhitungan Kimia Menggunakan *Certainty of Response Index*". Artikel Universitas Tanjungpura, Pontianak, 2014.
- Salirawati, Das. Disertasi: "*Pengembangan Instrumen Pendeteksi Miskonsepsi Kimia pada Peserta Didik SMA*". (Yogyakarta: PPs-UNY, 2011).
- Savira, Intan., Sri Wardani, Harjito, dan Any Noorhayati. 2019. "Desain Instrumen Tes Three Tiers Multiple Choice Untuk Analisis Miskonsepsi Siswa Terkait Larutan Penyangga". *Jurnal Inovasi Pendidikan Kimia*. Vol. 13 No. 1.
- Setiawati, Gusti Ayu Dewi., Ida Bagus Ari Sanjaya, dan Ni Wayan Ekayanti. 2014. "Identifikasi Miskonsepsi dalam Materi Fotosintesis dan Respirasi Tumbuhan pada Siswa Kelas IX SMP di Kota Denpasar". *Jurnal Bakti Saraswati (BJS)*. Vol. 3 No. 2.
- Sudijono, Anas. 2012. *Pengantar Evaluasi Pendidikan*. Jakarta: PT. Grafindo Pustaka.

- Sudjana, Nana. 2009. *Penilaian Hasil Proses Belajar Mengajar Cet. XIII*. Bandung: PT. Remaja Rosdakarya.
- Sugiyono. 2008. *Metode Penelitian Bisnis*. Bandung: Pusat Bahasa Depdiknas.
- Sulaiha. Tesis: “*Profil Berpikir Geometri Siswa MTs pada Materi Bangun Segiempat Ditinjau dari Gaya Kognitif*”. Surabaya: Program Pascasarjana (PPs) Universitas Negeri Surabaya, 2016.
- Suparno, Paul. 2013. *Miskonsepsi dan Perubahan Konsep Pendidikan Fisika*. Jakarta : PT. Grasindo.
- Suryabrata, Sumadi. 2000. *Pengembangan Alat Ukur Psikologis*. Yogyakarta: Andi.
- Suwarto. 2013. *Pengembangan Tes Diagnostik Dalam Pembelajaran*. Yogyakarta: Pustaka Pelajar.
- Syahrul, Dimas Ardiansyah., Woro Setyarsih. 2015. “Identifikasi Miskonsepsi dan Penyebab Miskonsepsi Siswa dengan Three-Tier Diagnostic Test Pada Materi Dinamika Rotasi”. *Jurnal Inovasi Pendidikan Fisika (JIPF)*. Vol. 4 No. 3.
- Taslidere, Erdal. 2016. “Development and Use of a Three-Tier Diagnostic Test to Assess High School Student’s Misconceptions About the Photoelectric Effect”. *Journal Research in Science & Technological Education*. Vol. 34 No. 2.
- Tayubi, Yuyu R. 2005. “Identifikasi Miskonsepsi Pada Konsep-Konsep Fisika Menggunakan Certainty of Response Index (CRI)”. *Jurnal Mimbar Pendidikan*. Vol. 3 No. 24.
- Tessmer, Martin. 1996. *Planning and Conducting Formative Evaluations: Improving the Quality of Educational and Training*. London: Kogan Page.

- Toni, Muhammad., Zubaidah R., dan Ahmad Yani T. 2017. “Analisis Kesalahan Siswa Menggunakan Certainty of Response Index (CRI) Termodifikasi pada Materi Pecahan”. *Jurnal pendidikan dan Pembelajaran*. Vol. 6 No. 4.
- Treagust, David F. 1988. “Development and Use of Diagnostic Test to Evaluate Student’s Misconceptions in Science”. *International Journal of Science Education*. Vol. 10 No. 2.
- Widiastuti. 2015. *Tes dan Pengukuran Olahraga*. Jakarta: Rajawali Press.
- Widyanuklida. 2017. “Analisis Tingkat Kesukaran dan Daya Pembeda Butir Soal Ujian Pelatihan Radiografi Tingkat I”. *Reponkmbatan*. Vol. 16 No. 1.
- Wulandari, Chirstine. 2017. “Menanamkan Konsep Bentuk Geometri”. *Jurnal Pengabdian Masyarakat Ipteks*. Vol. 3 No. 1

