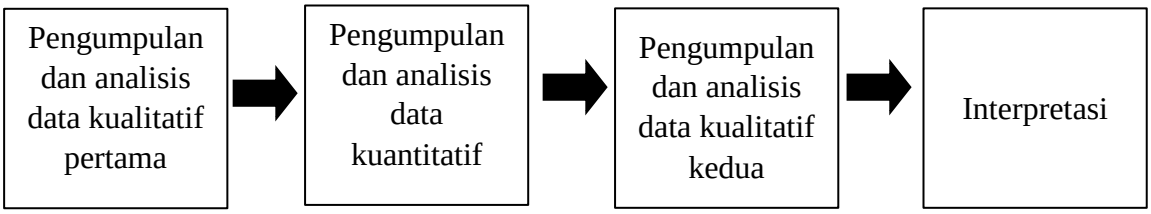
BAB III
METODE PENELITIAN

3.1 Metode dan Desain Penelitian

Penelitian ini adalah penelitian pengembangan dan penerapan STESDI untuk menganalisis miskonsepsi mahasiswa. Metode yang digunakan adalah penelitian dengan menggabungkan dua bentuk penelitian yaitu penelitian kuantitatif dan penelitian kualitatif. Menurut pendapat Sugiyono, menyatakan bahwa metode penelitian kombinasi (*mixed methods*) adalah suatu metode penelitian yang mengkombinasi atau menggabungkan antara metode kuantitatif dan metode kualitatif untuk digunakan bersama-sama dalam suatu kegiatan penelitian, sehingga diperoleh data yang lebih komprehensif, valid, reliabel dan obyektif. Gabungan antara metode kualitatif dan kuantitatif ini dapat berupa gabungan yang bermacam-macam. Namun, desain penelitian yang dianggap tepat untuk menggambarkan penelitian ini adalah *sequential exploratory design*. *Sequential exploratory design* yaitu mengumpulkan dan menganalisis data kualitatif, kemudian mengumpulkan dan menganalisis data kuantitatif (Creswell, 2010). Berikut merupakan diagram alur dalam penelitian.



Gambar 3. 1 Diagram Alur Penelitian *Sequential Exploratory Mixed*

Metode penelitian kombinasi *sequential exploratory mixed* pada penelitian ini, dicirikan dengan pengumpulan data dan analisis data kualitatif pada tahap pertama. Pada desain ini, peneliti menggunakan metode kualitatif untuk menggambarkan suatu fenomena yang terjadi di lapangan, kemudian hasil datanya sebagai bahan kajian untuk masuk pada tahap berikutnya yaitu tahap penelitian menggunakan metode kuantitatif. Tahap kedua merupakan pengumpulan serta analisis data kuantitatif, guna memperkuat hasil penelitian kualitatif yang dilakukan pada tahap pertama. Data kuantitatif yang digunakan dalam penelitian ini diperoleh dari hasil pengolahan data instrumen *six tier Earth Science diagnostic instrument* yang diberikan pada mahasiswa calon guru Fisika dari berbagai LPTK. Selanjutnya pada tahap ketiga dilakukan pengambilan data kualitatif dari hasil wawancara kepada mahasiswa yang telah mengerjakan instrumen *six tier Earth Science*

Wulan Cahya Kamila, 2020
SIX-TIER EARTH SCIENCE DIAGNOSTIC INSTRUMENT (STESDI) UNTUK MENGANALISIS TINGKAT MISKONSEPSI DAN PENYEBAB MISKONSEPSI PADA MAHASISWA
Universitas Pendidikan Indonesia | respository.upi.edu | perpustakaan.upi.edu

diagnostic instrument. Kemudian hasil dari ketiga tahap sebelumnya diinterpretasikan sehingga memperoleh kesimpulan yang diharapkan.

3.1.1 Model Produk

Model produk yang dihasilkan dalam penelitian pengembangan dan validasi ini adalah *Six Tier Earth Science Diagnostic Instrument* (STESDI) merupakan tes pilihan ganda bertingkat yang terdiri dari enam tingkat pada setiap butir soal, dengan pertanyaan sumber pilihan jawaban dan alasan pada tingkatan ketiga dan keenam. Tingkatan pertama (*first tier*) merupakan pertanyaan pilihan ganda tentang konsep ilmiah, tingkatan kedua (*second tier*) merupakan tingkat keyakinan mahasiswa dalam menjawab pertanyaan tingkat pertama, tingkatan ketiga (*third tier*) merupakan sumber pilihan jawaban dari pertanyaan tingkat pertama dengan dilengkapi oleh tingkat keyakinan jawaban, tingkatan keempat (*four tier*) merupakan pertanyaan pilihan ganda tentang alasan memilih jawaban pada tingkat pertama, tingkatan kelima (*five tier*) merupakan tingkat keyakinan mahasiswa dalam menjawab pertanyaan tingkat keempat, tingkatan keenam merupakan sumber pilihan jawaban dari pertanyaan tingkat keempat dengan dilengkapi oleh tingkat keyakinan jawaban. Tes diagnostik ini merupakan pengembangan yang dilakukan dengan alasan untuk menutupi kelemahan pada tes diagnostik *four-tier diagnostic test* yang dijelaskan pada bab I dan bab II sebelumnya.

3.1.2 Model Pengembangan

Model pengembangan dalam penelitian ini menggunakan model prosedural, yaitu model pengembangan yang bersifat deskriptif. Langkah-langkah sistematis yang dilakukan adalah dengan diawali pembuatan instrumen open ended, lalu melakukan validasi terhadap ahli, dan membuat produk berupa instrumen dalam mendiagnosis miskonsepsi dan penyebab miskonsepsi pada materi kebumian. Metode yang dilakukan adalah *sequential exploratory design* dengan mengumpulkan dan menganalisis data kualitatif dan mengumpulkan serta menganalisis data kuantitatif

3.2 Prosedur Penelitian

Prosedur penelitian yang dilakukan dalam penelitian ini dikelompokkan menjadi empat (4) tahapan kegiatan, yaitu tahap pengumpulan dan analisis data kualitatif pertama, tahap pengumpulan dan analisis data kuantitatif, tahap pengumpulan dan analisis data kualitatif kedua dan diakhiri dengan tahap interpretasi. Instrumen yang dimaksud dalam

penelitian ini adalah instrumen *six-tier earth science diagnostic test.* Keempat tahapan tersebut dirinci dalam penjelasan berikut:

3.2.1 Pengumpulan dan Analisis Data Kualitatif Pertama

Pada tahap ini, kegiatan dilakukan dengan menggali konsepsi mahasiswa pada materi Kebumian. Eksplorasi miskonsepsi dilakukan dengan mengkaji jurnal dan menentukan batasan isi materi yang berkaitan. Tahap ini dilakukan melalui beberapa tahap yaitu:

3.2.1.1 Analisis Batasan Materi

Kegiatan yang dilakukan pada tahapan ini adalah menentukan batasan isi materi dan mengidentifikasi konsep-konsep esensial yang terdapat dalam materi Kebumian yang telah ditentukan berdasarkan hasil kegiatan sebelumnya, yakni tahap analisis masalah. Tahap analisis batasan materi dilakukan dengan pengkajian RPS (Rencana Pembelajaran Semester) yang digunakan di salah satu universitas sebagai acuan yang digunakan untuk tolak ukur batasan materi yang akan digunakan.

3.2.1.2 Studi Literatur

Pada tahapan ini peneliti melakukan eksplorasi miskonsepsi dengan mengkaji jurnal yang berkaitan dengan materi Kebumian dengan tujuan untuk mengidentifikasi miskonsepsi serta konsep-konsep yang dapat terjadi miskonsepsi berdasarkan hasil penelitian sebelumnya. Informasi tentang miskonsepsi yang didapat dari kajian jurnal kemudian digunakan sebagai distraktor pada soal tingkat keempat (*four tier*). Selain berdasarkan data miskonsepsi dan jurnal, distraktor soal juga dikonstruksi dari data eksplorasi jawaban mahasiswa yang berpotensi mengalami miskonsepsi. Setelah melakukan eksplorasi miskonsepsi, dilakukan juga eksplorasi pada soal yang dapat mendeteksi pemahaman mahasiswa terhadap materi Kebumian, seperti ESDT (*Earth Science Diagnostic Test*) serta beberapa soal yang digunakan dalam jurnal untuk mendiagnosis pemahaman mahasiswa.

Dari hasil kajian pustaka dijumpai beberapa instrumen pendeteksi miskonsepsi yang telah dikembangkan oleh beberapa peneliti dengan

kelebihan dan kekurangannya. Penelitian ini akan mengembangkan instrumen pendeteksi miskonsepsi dan penyebabnya dengan mengadopsi instrumen yang dikembangkan oleh Caleon dan Subramaniam (2010b) dan juga instrumen yang dikembangkan oleh Gurel, Eryilmaz, dan McDermott (2017). Kedua instrumen ini digabungkan agar kelemahan yang didapat pada beberapa instrumen pendeteksi miskonsepsi dapat diatasi. Kelemahan-kelemahan dari beberapa instrumen tersebut telah dipaparkan pada latar belakang penelitian dan diuraikan dengan jelas pada kajian pustaka.

3.2.1.3 Pembuatan Kisi-Kisi

Pada tahap ini proses pengembangan produk dilakukan dengan menentukan indikator soal mengenai materi Kebumian yang sesuai dengan kurikulum yang berlaku. Indikator ini berfungsi sebagai pedoman dalam membuat soal sehingga dapat memenuhi validitas isi. Berdasarkan kisi-kisi ini, dihasilkan 17 butir soal dengan bentuk pilihan ganda terbuka atau *two-tier open-ended test*. Setiap soal terdiri dari inti soal dan alasan terbuka, setiap inti soal terdiri dari pokok pertanyaan dan empat pilihan jawaban (*option*). Alasan terbuka merupakan tempat kosong yang diisi mahasiswa dengan alasan mengapa peserta didik memilih salah satu pilihan jawaban dari empat pilihan jawaban yang disediakan.

Konsep Kebumian yang dipilih untuk dijadikan soal adalah struktur bumi (gunung api, jenis-jenis lempeng, pergerakan lempeng), efek rumah kaca, cuaca dan iklim (badai, curah hujan), bencana alam (banjir, longsor, gempa bumi, tsunami, dan likuifaksi). Alasan pengambilan konsep tersebut karena memiliki potensi terjadinya miskonsepsi setelah adanya kajian literatur penelitian sebelumnya.

3.2.1.4 Tes Pilihan Ganda Beralasan

Pada tahap ini proses pengembangan produk dilakukan dengan membuat *open ended two-tier test.* Soal pilihan ganda dengan distraktor soal yang dikonstruksi dari data eksplorasi jawaban mahasiswa dengan memberikan pertanyaan alasan, dimana alasan yang diberikan dalam uji terbuka. Distraktor pertanyaan alasan dibuat kosong dan diisi oleh

mahasiswa yang kemudian alasan akan divalidasi sehingga dapat diambil sebagai salah satu *tier* dalam *six-tier test*.

3.2.1.5 Penulisan Draf Soal

Pada tahap ini dilakukan transformasi dari tes *open ended two-tier test* menjadi tes diagnostik STESDI. Proses transformasi dilakukan dengan menempatkan dua atau tiga alasan terbaik dengan frekuensi paling tinggi dan teridentifikasi miskonsepsi akan dijadikan sebagai pilihan jawaban atau option pada tingkat ketiga. Alasan yang terpilih ini dikombinasikan dengan beberapa literatur mengenai miskonsepsi yang terjadi pada materi Kebumian. Tes diagnostik STESDI secara lengkap terdiri dari pilihan jawaban pada tingkat pertama, tingkat keyakinan memilih jawaban pada tingkat kedua, sumber penyebab pada tingkat ketiga, pilihan alasan pada tingkat keempat, tingkat keyakinan memilih alasan pada tingkat kelima serta sumber penyebab pemilihan alasan pada tingkatan keenam.

Transformasi berupa perubahan dari tes *open-ended two-tier* menjadi STESDI diawali dengan membuat format penampilan dari kedua bentuk tersebut. Untuk o*pen-ended* tes, bagiannya hanya terdiri dari dua yaitu pilihan jawaban pada tingkat atau tingkat pertama dan alasan terbuka yang akan diisi oleh peserta didik pada tingkat kedua. Sementara untuk format STESDI, seperti yang dijelaskan pada paragraf sebelumnya. Format dari kedua bentuk tersebut dapat diamati pada gambar di bawah ini.

Pertanyaan
Pilihan Jawaban (Tingkat I) A. B. C. D.
Tingkat Keyakinan Jawaban (Tingkat II) A. B. C. D.
Sumber Jawaban (Tingkat III) A. B. C. D.
Pilihan Alasan (Tingkat IV) A. B. C. D.
Tingkat Keyakinan Alasan (Tingkat V) A. B. C. D.
Sumber Alasan (Tingkat VI) A. B. C. D.

Pertanyaan
Pilihan Jawaban (Tingkat I) A. B. C. D.
Alasan (Tingkat II) ....

Gambar 3. 2 STESDI dan *Two-Tier Open-Ended Test*

Tahapan selanjutnya dari proses transformasi adalah perubahan dari *two-tier open ended* menjadi STESDI yang merupakan tabulasi konsep alternatif mahasiswa. Hasil tabulasi terhadap keseluruhan konsep alternatif sebanyak 11 mahasiswa. Untuk setiap butir soal akan dipilih sebanyak tiga atau empat alasan yang diberikan mahasiswa dengan frekuensi paling tinggi dan teridentifikasi miskonsepsi. Jika terdapat butir soal dengan konsep alternatif yang tidak mencukupi untuk dijadikan sebagai pilihan alasan pada tingkat ketiga maka akan ditambahkan dengan konsep alternatif yang diperoleh dari hasil kajian literatur. Baik konsep alternatif yang diperoleh oleh mahasiswa maupun dari hasil kajian literatur selanjutnya akan dianalisis lebih lanjut untuk memastikan bahwa konsep alternatif tersebut sesuai atau tidak dimasukkan sebagai alasan pada tingkat keempat. Jawaban yang dipilih sebagai alasan pada tingkat ketiga tetap mengacu pada kisi-kisi butir soal yang telah disusun lebih awal.

Setelah semua alasan untuk setiap butir soal berhasil diidentifikasi, langkah transformasi selanjutnya adalah dengan menggabungkan alasan tadi dengan pilihan jawaban pada tingkat pertama dan tingkat keyakinan pada tingkat kedua dan kelima serta pertanyaan sumber jawaban mahasiswa pada tingkat

ketiga dan keenam sehingga diperoleh perangkat tes diagnostik STESDI sebagai produk akhir dari proses pengembangan.

3.2.1.6 Validasi Kepada Ahli

Pada tahapan validasi dan uji coba (*Content validation and Piloting*), STESDI (*six tier earth science diagnostic instrument*) yang telah dibuat kemudian dikonsultasikan kepada para ahli (*judgment expert*) untuk mengevaluasi validitas isi serta konstruksi soal. Instrumen tes yang dikembangkan divalidasi oleh *judgment expert* (dosen ahli) sebanyak tiga (3) validator. Hasil validitas isi yang dilakukan oleh *judgment expert* kemudian dianalisis menggunakan analisis MINIFAC. Validitas instrumen tes juga terdapat catatan perbaikan dan saran untuk memperbaiki instrumen yang dikembangkan sehingga didapat instrumen yang valid. Selain validitas isi yang dilakukan oleh *judgment expert*, dilakukan pula validitas konstruksi menggunakan Rasch model pada software MINISTEP. Setelah soal dievaluasi dan divalidasi oleh ahli, kemudian dilakukan evaluasi dan perbaikan instrumen sehingga instrumen dapat dinyatakan valid. Uji validitas instrumen tes diagnostik STESDI dapat diketahui dengan menggunakan menu *output 10. Item (column): fit order* pada *software* MINISTEP. Kriteria pengukuran uji validitas instrumen yang diterima dapat dilihat dari hasil skor pada menu *Outfit mean square (MNSQ)* dengan skor yang diterima yaitu antara 0,5 sampai 1,5, dan *Outfit Z-standard (ZSTD)* dengan skor diterima antara -2,0 sampai dengan +2,0.

3.2.2 Pengumpulan dan Analisis Data Kuantitatif

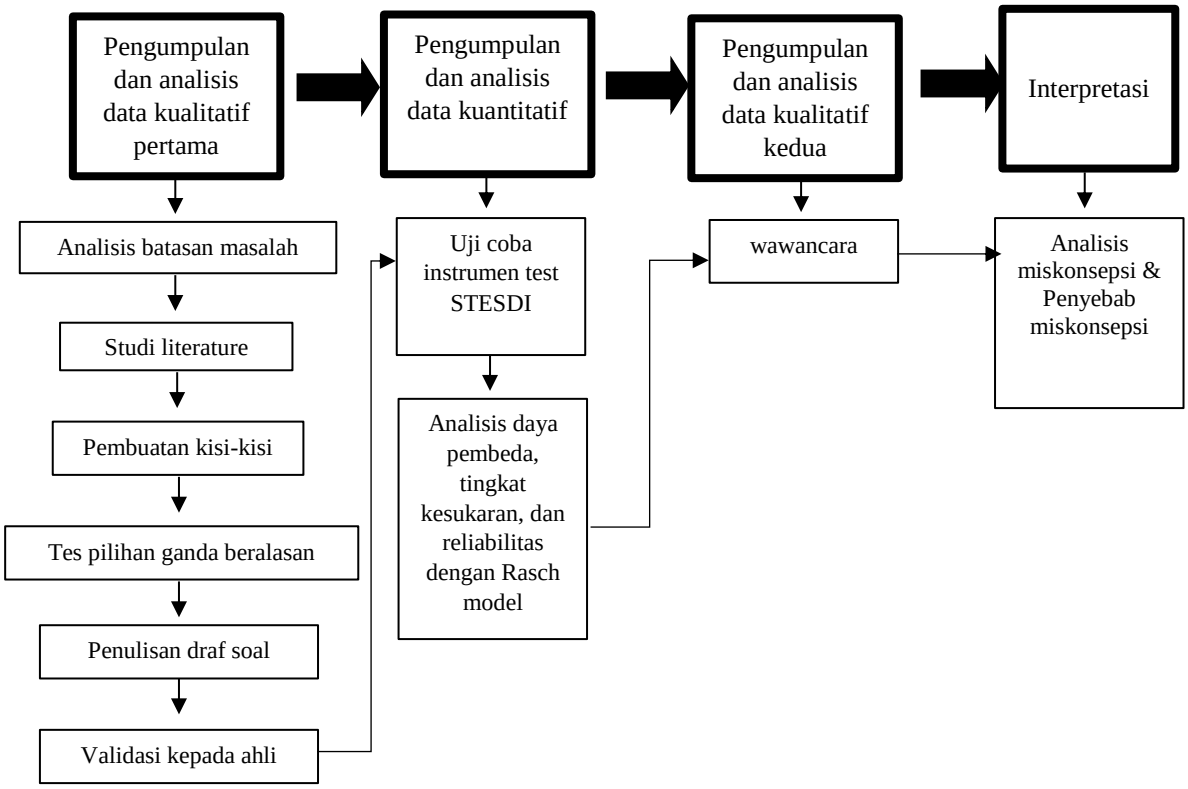
Pada tahap ini, peneliti menggunakan instrumen tes diagnostik dalam format *six-tier test* untuk mengetahui miskonsepsi dan penyebab miskonsepsi mahasiswa pada materi Kebumian. Pengumpulan data dilakukan dengan penyebaran instrumen tes pada skala besar yang dilakukan untuk melihat kinerja tes diagnostik *six-tier test* dalam menentukan miskonsepsi dan penyebab miskonsepsi yang terjadi pada mahasiswa terhadap materi Kebumian. Sampel yang digunakan tersebar pada beberapa mahasiswa di berbagai universitas yang telah maupun belum memperoleh mata kuliah Kebumian. Analisis instrumen dilakukan menggunakan Rasch model dengan melihat daya pembeda, reliabilitas, dan tingkat kesukaran instrumen.

3.2.3 Pengumpulan dan Analisis Data Kualitatif Kedua

Pada tahap ketiga ini, dilakukan dengan pengumpulan data kualitatif dari hasil wawancara kepada mahasiswa yang telah mengerjakan instrumen *six tier earth science diagnostic instrument*. Wawancara dilakukan pada 25% dari 110 responden yang telah mengerjakan instrumen tes STESDI (*six-tier earth science diagnostic instrument).*

3.2.4 Interpretasi

Kegiatan interpretasi merupakan tahap terakhir yang dilakukan untuk menilai kelebihan serta kekurangan yang ada pada instrumen tes diagnostik *six-tier test*. Tahap interpretasi dilakukan dengan menganalisis hasil dari uji coba instrumen pada tahap sebelumnya. Berikut merupakan diagram alur dalam penelitian:



Gambar 3. 3 Diagram Alur Penelitian

3.3 Lokasi dan Partisipan Penelitian

Target populasi penelitian ini adalah mahasiswa jurusan Fisika atau pendidikan Fisika, dimana diperoleh dari teknik pengambilan sampel *purposive random sampling*. Alasan pemilihan mahasiswa jurusan Fisika atau pendidikan Fisika sebagai target populasi penelitian disesuaikan dengan konsep materi yang akan diteliti, karena tes diagnostik STESDI yang akan dikembangkan mengenai materi Kebumian yang dipelajari oleh

mahasiswa Fisika atau pendidikan Fisika, sehingga detail partisipan penelitiannya adalah mahasiswa yang sudah maupun belum mengambil mata kuliah IPBA. Populasi terdiri dari enam universitas yang tersebar di Indonesia. Adapun subjek penelitian terdiri dari 110 mahasiswa. Mahasiswa berasal dari empat universitas, yaitu dua universitas negeri agama yang berlokasi di pulau Sumatera dan Jawa, dan dua universitas negeri di Indonesia bagian Barat (Sumatera dan Jawa)

3.4 Jenis Data

Data yang diperoleh, dianalisis dan digunakan dalam penelitian pengembangan ini terdiri dari 2 yaitu:

3.4.1 Data Kualitatif

Data kualitatif berupa studi literatur miskonsepsi materi Kebumian dari berbagai jurnal. Selain itu masukan dari validator dan review mengenai kualitas instrumen yang dikembangkan. Masukan dan saran yang diharapkan adalah berbagai hal yang berkaitan dengan konsep Kebumian dari miskonsepsi yang diidentifikasi, bahasa dan konstruksi serta ketepatan soal dalam mengidentifikasi miskonsepsi dan penyebab miskonsepsi pada materi Kebumian. Data kualitatif lain adalah tabulasi alasan yang diberikan mahasiswa, dimana alasan merupakan jawaban mahasiswa yang akan dianalisis untuk mengetahui alasan mana saja yang teridentifikasi miskonsepsi.

3.4.2 Data Kuantitatif

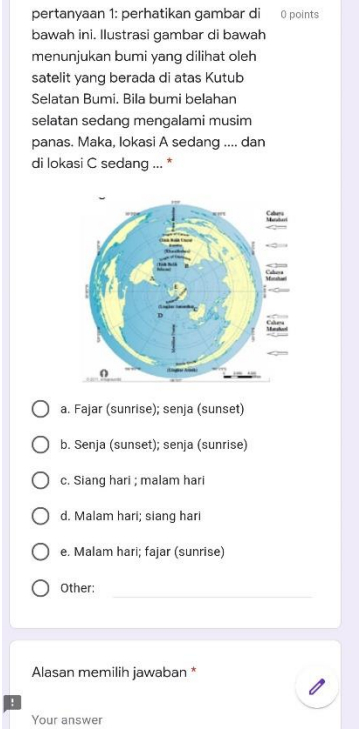
Data kuantitatif berupa tiga kelompok data yaitu hasil masukan validator, data uji coba instrumen dan data penerapan instrumen STESDI pada mahasiswa dengan tujuan mengidentifikasi miskonsepsi dan penyebab miskonsepsi pada materi Kebumian. Data dari hasil uji coba instrumen akan memberikan informasi mengenai validitas, reliabilitas, tingkat kesukaran dan daya pembeda. Sedangkan data hasil penggunaan instrumen STESDI digunakan untuk menjelaskan mengenai miskonsepsi dan penyebab miskonsepsi mahasiswa.

3.5 Instrumen Pengumpulan Data

Dalam menentukan kualitas instrumen tes yang dikembangkan, dilakukan estimasi validitas dan reliabilitas tes. Estimasi validitas dan reliabilitas dijelaskan sebagai berikut.

3.5.1 Instrumen Identifikasi Konsepsi Alternatif Mahasiswa (*Two Tier Open Ended Test*)

Instrumen t*wo tier open ended test* digunakan untuk mengumpulkan jawaban alternatif mahasiswa mengenai konsep Kebumian. Jawaban ditabulasi untuk kemudian dianalisis alasan-alasan mana saja yang teridentifikasi miskonsepsi. Jawaban mahasiswa yang teridentifikasi miskonsepsi nantinya akan dijadikan sebagai pilihan alasan pada *tier* keempat tes diagnostik STESDI dan dikombinasikan dengan temuan-temuan miskonsepsi hasil kajian literatur. Instrumen t*wo tier open ended test* terdiri dari 23 soal. Instrumen tersebut dapat dilihat pada Gambar 3.4 berikut (Lampiran 1)

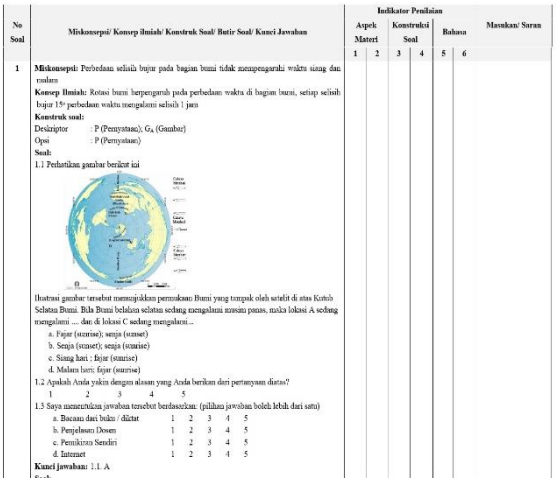


Gambar 3. 4 Instrumen *Two Tier Open Ended Test*

3.5.2 Instrumen Validasi Ahli

Pada penelitian ini digunakan validitas isi untuk instrumen yang dikembangkan maka dilakukan *expert judgment* dengan melibatkan ahli dalam bidang asesmen dan Fisika. Lembar validasi terdiri dari kolom tentang hal yang dikritisi dan masukan serta saran yang diberikan untuk setiap butir soal. Hasil rata-rata masukan dan saran yang diberikan *expert* juga akan dianalisis secara kuantitatif menggunakan menu data analisis uji *multi-rater* (*multi-faceted Rasch measurement*) pada *software* MINIFAC untuk dianalisis dengan Rasch. Hasil perhitungan dari analisis Rasch akan menjadi pertimbangan untuk menarik kesimpulan terhadap butir soal apakah valid atau tidak. Instrumen validasi terdiri

dari 17 butir soal. Instrumen tersebut dapat dilihat pada Gambar 3.5 berikut (lampiran 2)



No Soal	Miskonsepsi/ Konsep ilmiah/ Konstruk Soal/ Butir Soal/ Kunci Jawaban	Indikator Penilaian						Masukan/ Saran
		Aspek Materi		Konstruksi Soal		Bahasa		
		1	2	3	4	5	6	
1	**Miskonsepsi:** Perbedaan selisih bujur pada bagian bumi tidak mempengaruhi waktu siang dan malam **Konsep Ilmiah:** Rotasi bumi berpengaruh pada perbedaan waktu di bagian bumi, setiap selisih bujur 15° perbedaan waktu mengalami selisih 1 jam **Konstruk soal:** Deskriptor : P (Pernyataan); G_A (Gambar) Opsi : P (Pernyataan) **Soal:** 1.1 Perhatikan gambar berikut ini Ilustrasi gambar tersebut menunjukkan permukaan Bumi yang tampak oleh satelit di atas Kutub Selatan Bumi. Bila Bumi belahan selatan sedang mengalami musim panas, maka lokasi A sedang mengalami ... dan di lokasi C sedang mengalami ... a. Fajar (sunrise); senja (sunset) b. Senja (sunset); senja (sunrise) c. Siang hari ; fajar (sunrise) d. Malam hari; fajar (sunrise) 1.2 Apakah Anda yakin dengan alasan yang Anda berikan dari pertanyaan diatas? 1 2 3 4 5 1.3 Saya menentukan jawaban tersebut berdasarkan: (pilihan jawaban boleh lebih dari satu) a. Bacaan dari buku / diktat 1 2 3 4 5 b. Penjelasan Dosen 1 2 3 4 5 c. Pemikiran Sendiri 1 2 3 4 5 d. Internet 1 2 3 4 5 **Kunci jawaban:** 1.1. A							

Gambar 3. 5 Instrumen Validasi Soal kepada Ahli

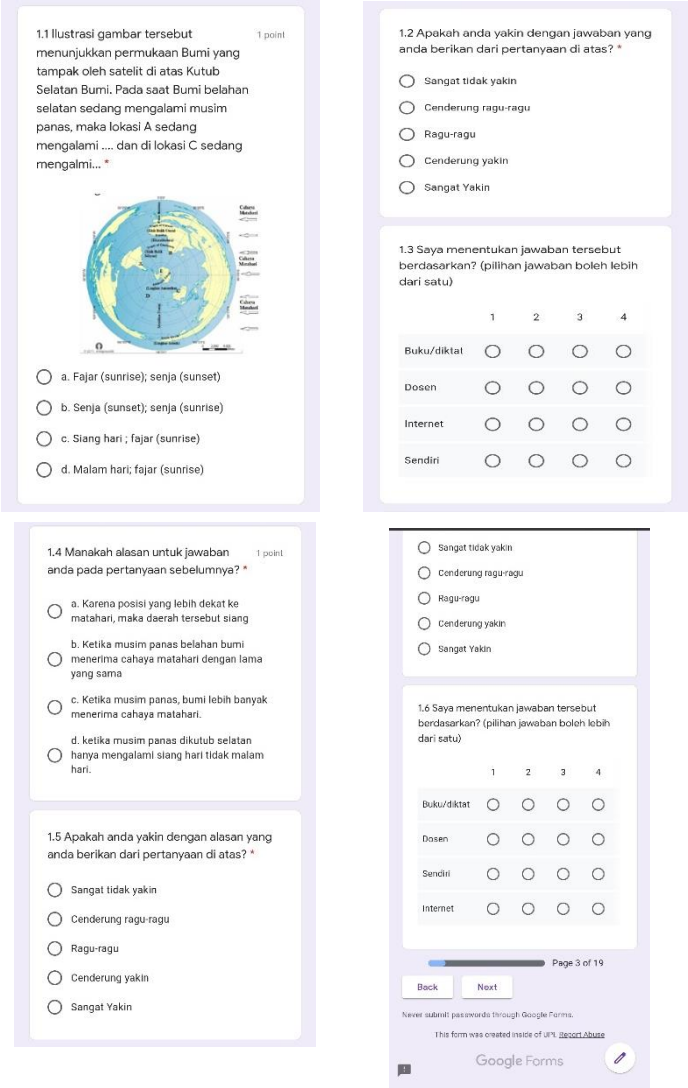
3.5.3 Instrumen Validitas Empiris

Instrumen untuk validasi empiris merupakan seperangkat tes diagnostik STESDI yang berjumlah 17 butir soal dari proses pengembangan yang dilakukan. Instrumen ini diperoleh dengan memasukkan konsepsi alternatif yang teridentifikasi miskonsepsi pada *tier* keempat pada tes diagnostik STESDI. Tes diagnostik ini merupakan rangkaian hasil revisian dari *expert judgment* dan transformasi dari instrumen *two tier open-ended* menjadi instrumen diagnostik STESDI.

3.5.4 Instrumen Identifikasi Miskonsepsi dan Penyebab Miskonsepsi Mahasiswa

Instrumen untuk identifikasi dan penyebab miskonsepsi merupakan seperangkat instrumen diagnostik STESDI yang terdiri dari 17 butir soal. Produk ini menjadi produk akhir dari pengembangan yang dilakukan setelah melewati beberapa tahapan pengujian kualitas diantaranya uji validitas, reliabilitas, tingkat kesukaran dan daya pembeda. Tes diagnostik STESDI yang telah dikembangkan kemudian digunakan untuk mengidentifikasi miskonsepsi dan penyebab miskonsepsi mahasiswa.

Tes diberikan pada mahasiswa dari enam universitas yang tersebar di Indonesia dengan menggunakan *Google Form*. Soal dapat diakses dengan link http://form.gle/f6jvAdqosZjuy8EH6. Contoh soal tersebut dapat dilihat pada Gambar 3.6 berikut.



Gambar 3. 6 Soal STESDI Menggunakan Format Google Form

3.6 Teknik Analisis Karakteristik Tes

Pengolahan data dimaksudkan untuk kepentingan penarikan kesimpulan. Pengolahan data meliputi kegiatan penentuan validitas dan reliabilitas dari tes diagnostik STESDI yang dikembangkan dan penentuan terhadap kunci determinasi pola jawaban yang diberikan mahasiswa. Pengujian validitas, reliabilitas, tingkat kesukaran dan daya pembeda dijelaskan sebagai berikut.

3.6.1 Uji Validitas

Validitas didefinisikan sebagai suatu kesesuaian, keberanian, dan kegunaan dari sebuah inferensi spesifik yang dibuat oleh peneliti berdasarkan data mereka kumpulkan (Fraenkel, Wallen, dan Hyun, 2011). Secara singkat, validitas adalah semua tentang penggambaran kesimpulan yang benar berdasarkan data yang diperoleh dari sebuah penilaian. Validasi dilakukan pada dua aspek yaitu *content-related evidence of validity* dan *criterion-related evidence of validity.*

Content-related evidence of validity ini mengacu pada konten dan format instrumen. Isi dan format tes harus sesuai dengan definisi variabel dan sampel subjek yang akan diukur. Oleh karena itu, validasi isi sebagian adalah masalah menentukan apakah isi instrumen adalah sampel yang memadai dari domain konten yang seharusnya diwakilinya. Aspek lain dari validasi konten adalah format instrumen, kejelasan pencetakan, kesesuaian bahasa, kejelasan arah, ukuran jenis dan lain sebagainya.

Criterion-related evidence of validity ini mengacu pada hubungan antar skor yang diperoleh dengan menggunakan instrumen dan skor yang diperoleh dengan menggunakan satu atau lebih instrumen lain. Kekuatan hubungan antara nilai tes dan kinerja kriteria memberikan kriteria terkait bukti validasi. Validasi dianalisis dengan pemodelan Rasch, dimana dengan pemodelan Rasch diperoleh beberapa kelebihan. Berdasarkan pemodelan Rasch dapat diprediksi skor untuk data yang hilang, dapat diidentifikasi response error, abilitas tidak bergantung pada jumlah jawaban benar serta adanya identifikasi tebakan (Sumintono dan Widhiarso, 2013).

3.6.1.1 Validasi Isi

Uji validitas yang dilakukan pada tahap ini adalah validasi isi. Validasi tes STESDI ditentukan dengan beberapa teknik secara kuantitatif dan kualitatif. Pertama-tama item dari tes diagnostik ini diperiksa dan dinilai oleh para ahli berdasarkan isi dan format di setiap tahap pengembangan tes, *open-ended,* pilihan ganda respon, dan pengujian instrumen tes yang dikembangkan untuk memperoleh *content-related evidence of validity.* Tabel spesifikasi disiapkan untuk konteks dan kasus yang diukur dengan tes dan tabel untuk pilihan kesalahpahaman diperiksa oleh para ahli untuk kecukupan dan kompatibilitasnya dengan item uji.

Termasuk didalamnya format, petunjuk arah, angka dan bahasa dinilai oleh para ahli.

Selanjutnya uji validitas dilakukan dengan menggunakan analisis uji *multi-rater* (*multi-faceted Rasch measurement*) pada *software* MINIFAC untuk dianalisis dengan Rasch. Langkah yang dilakukan untuk menghasilkan *output* uji *multi-rater*, adalah dengan membuat spesifikasi berkas data yang dipunyai dan jenis analisis yang dilakukan. File khusus tersebut tidak lain baris-baris perintah (*coding*) seperti halnya membuat satu program untuk tugas tertentu. *Coding* generik untuk analisis MINIFAC menggunakan program Notepad.

3.6.1.2 Validitas Empiris

Pada tahapan uji validasi ini, soal tes diagnostik STESDI diujicobakan. Setelah mendapatkan data hasil uji coba tersebut, selanjutnya uji validitas dilakukan dengan menggunakan menu *output 10. Item (column): fit order* pada *software* MINISTEP untuk dianalisis dengan Rasch. Pengukuran uji validitas instrumen ini dilihat dari hasil skor pada menu *Outfit mean square (MNSQ), Outfit Z-standard (ZSTD),* dan *Point Measure Correlation (Pt Mean Corr).* Menurut Boone, Staver, dan Yale (2014), kriteria tersebut digunakan untuk mengetahui tingkat kesesuaian butir (*item fit*) (Boone, Staver, dan Yale, 2014). Sedangkan penelitian yang dilakukan oleh Syahbudin, Mansor, Meerah, dan Muhammad (2018) menyatakan bahwa *Point Measure Correlation (Pt Mean Corr)* hanya digunakan untuk mengetahui daya pembeda dari suatu instrumen (Sabudin, Mansor, Meerah, dan Muhammad, 2018). Maka untuk mengukur kevalidan atau kesahihan suatu instrumen hanya menggunakan skor pada menu *Outfit mean square (MNSQ)* dan *Outfit Z-standard (ZSTD)* pada *software* MINISTEP. Mengetahui kevalidan dari tiap soal dapat dicari dengan interpretasi menurut Boone, Staver, dan Yale (2014) pada Tabel 3.1 sebagai berikut.

Tabel 3. 1

Interpretasi Nilai *Output* MNSQ dan ZSTD

Output item	**Skor**	**Keterangan**
Outfit MNSQ	0,5 < MNSQ < 1,5	Diterima

Output item	Skor	Keterangan
Outfit ZSTD	-2,0 < ZSTD < +2,0	Diterima

3.6.2 Uji Reliabilitas Instrumen

Reliabilitas instrumen mempunyai pengertian bahwa suatu instrumen cukup dapat dipercaya atau digunakan sebagai alat pengumpul data karena instrumen tersebut sudah baik dan tetap. Sebuah instrumen yang *reliabel* berapa kali digunakan dalam penelitian apapun hasilnya akan relatif sama sehingga hasil yang didapatkan dari instrumen tersebut dapat dipercaya. Uji reliabilitas pada penelitian ini menggunakan *software* MINISTEP 4.3.1. Perangkat lunak MINISTEP ini merupakan sebuah aplikasi pendukung untuk mengetahui kriteria dan kualitas suatu instrumen soal berdasarkan analisis Rasch. Salah satu menu *output* pada *software* ini yaitu *output 3.1 summary statistics* yang fungsinya yaitu untuk menampilkan beberapa nilai reliabilitas, diantaranya adalah *person reliability*, *item reliability*, dan *Cronbach alpha*. Nilai reliabilitas yang diambil dalam penelitian ini adalah i*tem reliability* menunjukan kualitas item tes, artinya kita dapat mengetahui kualitas sebuah instrumen STESDI yang digunakan dalam penelitian.

Interpretasi untuk nilai *item reliability* dapat dilihat pada Tabel 3.2 (Mohamad, Sulaiman, Sern, dan Salleh, 2015; Sumintono dan Widhiarso, 2015)

Tabel 3. 2

Interpretasi *Item Reliability*

Nilai *item reliability*	**Interpretasi**
0,94 < Nilai	Istimewa
0,91 ≤ Nilai ≤ 0,94	Bagus Sekali
0,80 ≤ Nilai ≤ 0,90	Bagus
0,67 ≤ Nilai ≤ 0,80	Cukup
Nilai < 0,67	Lemah

Nilai *item reliability* yang dipakai adalah *REAL RMSE* karena nilai ini merupakan kondisi terburuk reliabilitas batas bawah berdasarkan instrumen yang dipakai (Sumintono dan Widhiarso, 2015).

3.6.3 Tingkat Kesulitan

Tingkat kesukaran bertujuan untuk membedakan soal yang terlalu mudah, mudah, sulit, dan sangat sulit. Hal ini diperlukan untuk menghasilkan distribusi soal yang baik. Tingkat kesukaran tes diagnostik STESDI dapat diketahui dengan menggunakan *software* MINISTEP 4.3.1 pada menu *output Table 1 Variable (Wright) maps* dan *output Table 13 Item Measure* yang dianalisis dengan Rasch. *Output Variable (Wright) maps* pada *software* ini bertujuan untuk mendapat gambaran mengenai kekuatan setiap butir soal. Sesuai dengan penjelasan Boone dan Netemeyer (2017) bahwa *wright maps* mendukung para peneliti untuk mengukur kekuatan dan kelemahan instrumen, mendokumentasikan penilaian butir tes, membandingkan teori dengan data eksperimen, dan memberikan panduan kepada para peneliti. Sedangkan *Item Measure* digunakan untuk mengklasifikasikan tingkat kesulitan tiap butir soal berdasarkan nilai logit dan nilai Standar Deviasi (SD) yang didapat (Sumintono dan Widhiarso, 2015).

Klasifikasi kategori tingkat kesukaran pada instrumen STESDI ini dapat diinterpretasikan seperti pada Tabel 3.3 sebagai berikut.

Tabel 3. 3

Interpretasi Tingkat Kesukaran Soal

Tingkat Kesukaran (TK)	**Interpretasi**
0,50 < TK	Sangat Sukar
0,00 < TK ≤ 0,50	Sukar
-0,50 ≤ TK ≤ 0,00	Mudah
TK < -0,50	Sangat Mudah

Tingkat kesulitan butir dengan analisis Rasch pada dasarnya sama dengan taraf kesukaran teori tes klasik, yaitu perbandingan antara jumlah jawaban benar dengan jumlah soal yang diujikan. Perbedaannya, pada analisis Rasch nilai peluang diskalakan dengan memasukkan fungsi logaritma, sehingga dapat menghasilkan suatu pengukuran dengan interval yang sama karena terjadi transformasi data *odd ratio* dengan logaritma (Sumintono dan Widhiarso, 2015).

3.6.4 Daya Pembeda

Daya pembeda digunakan untuk mengetahui kemampuan soal dalam membedakan mahasiswa dengan kemampuan yang tinggi dan kemampuan yang rendah. Daya pembeda pada STESDI dapat diketahui menggunakan *software*

MINISTEP 4.3.1 dengan menu *output Table* 10 *Item Fit Order* yang dianalisis dengan Rasch.

Korelasi *point-measure* mengacu pada hubungan antara tingkat kesulitan item tiap individu dan tingkat kesulitan soal secara keseluruhan, dimana nilai satu menunjukan kemampuan mahasiswa yang tinggi menjawab soal dengan benar dan sebaliknya yang menunjukan korelasi sempurna menurut Rasch, nilai nol menunjukan tidak ada hubungan antara respon item dan item secara keseluruhan, sedangkan nilai negatif menunjukan masalah pada item soal karena sering mendapatkan skor rendah dibanding skor tinggi (Smiley, 2015).

Penggunaan *Point Measure Correlation (Pt Mean Corr)* seperti yang diungkapkan sebelumnya dapat memberikan informasi mengenai daya pembeda dari suatu instrumen soal yang digunakan untuk membedakan kemampuan mahasiswa. Untuk dapat mengetahui kategori daya pembeda tiap butir soal instrumen tes STESDI pada kolom *PT-MEASURE CORR.*, Smiley (2015) memberikan interpretasi untuk setiap nilai yang diberikan seperti yang terdapat pada Tabel 3.4 sebagai berikut.

Tabel 3. 4

Interpretasi nilai *Point Measure Correlation*

Pt Mean Corr	**Interpretasi**
0,40 < ID	Sangat baik
0,30 ≤ ID ≤ 0,40	Baik
0,20 ≤ ID < 0,30	Kurang baik
ID < 0,20	Jelek

3.7 Teknik Analisis Implementasi Tes

Tes diagnostik STESDI yang dikembangkan akan menghasilkan beberapa pola jawaban mahasiswa yang menunjukkan terjadinya miskonsepsi. Pengumpulan data dilakukan dengan cara melakukan tes diagnosis enam tingkatan (*six tier earth science diagnostic test*) yang disebar menggunakan *google form*. Setiap jawaban yang diberikan mahasiswa akan ditentukan berdasarkan jawaban yang mereka berikan untuk setiap tingkatan. Tingkatan pertama untuk pertanyaan utama, tingkatan kedua untuk tingkat keyakinan mahasiswa dalam menjawab tingkat pertama, tingkatan ketiga untuk sumber jawaban pada tingkatan pertama, tingkatan keempat untuk alasan memilih jawaban pada

tingkatan pertama dan tingkatan kelima untuk tingkatan keyakinan mahasiswa dalam menjawab tingkatan ketiga serta tingkatan keenam untuk sumber jawaban alasan pada tingkat keempat.

Teknik pemberian skor ini menggunakan angka 1 atau 0. Berdasarkan teknik pemberian skor yang dilakukan oleh Caleon dan Subramaniam (2010b), jika jawaban pada tingkat pertama (*first-tier*) atau alasan yang dipilih pada tingkat keempat (*four-tier*) benar, maka diberi skor 1 dan jika salah maka diberi nilai 0. Sedangkan untuk tingkat keyakinan (*confidence rating*) pada tingkat kedua (*second-tier*) dan tingkat kelima (*five-tier*) yaitu jika mahasiswa memberikan jawaban sangat tidak yakin maka diberi kode 1, jika memberi jawaban sangat yakin maka diberi kode 5, sedangkan untuk jawaban cenderung yakin diberi kode 4, jawaban cenderung tidak yakin diberi kode 2 serta untuk jawaban ragu-ragu diberi kode 3.

Setelah mahasiswa diberikan skor untuk tingkat pertama, tingkat kedua, tingkat ketiga serta tingkat kelima kemudian dilanjutkan dengan mengkategorikan kombinasi jawaban yang diberikan oleh mahasiswa. Kategori jawaban mahasiswa ini dibagi menjadi lima yaitu paham konsep atau *scientific conception* (SC), miskonsepsi (MSC), *false positif* (FP), *false negatif* (FN), dan kurangnya pemahaman atau *lack of knowledge* (LK) (Gurel, Eryilmaz, dan McDermott, 2015). Kategori ini dilihat berdasarkan jawaban dan tingkat keyakinan (*confidence rating*) yang diberikan oleh mahasiswa pada tingkat pertama, kedua, keempat serta kelima. Sedangkan tingkat ketiga dan keenam digunakan untuk mengetahui penyebab terjadinya miskonsepsi yang didapatkan dari jawaban mahasiswa terkait sumber yang mendasari mahasiswa memilih jawaban pada tingkatan-tingkatan sebelumnya.

Selain itu perhitungan data juga dilakukan dengan CDQ (*Confidence Discrimination Quotient*), dimana data yang digunakan untuk menganalisis miskonsepsi diperoleh dari hasil tes yang telah dikerjakan oleh mahasiswa pada uji lapangan akhir. Data yang diperoleh dianalisis untuk menentukan apakah mahasiswa dapat membedakan apa yang mereka pahami dan apa yang tidak mereka pahami (Fariyani, Rusilowati, dan Sugianto, 2015). Menurut Caleon dan Subramaniam (2010b) analisis data untuk menentukan CDQ (*Confidence Discrimination Quotient*) dapat digunakan persamaan

$$CDQ = \frac{(CFC - CFW)}{S} \quad \text{..... (2)}$$

Keterangan

CFC= rata-rata tingkat keyakinan mahasiswa yang menjawab dengan benar

CFW = rata-rata tingkat keyakinan mahasiswa yang menjawab dengan salah

S = standar deviasi tingkat keyakinan

Analisis CDQ (*Confidence Discrimination Quotient*) dibagi kedalam beberapa kelompok, diantaranya sebagai berikut:

1) Kelompok pilihan jawaban

Tabel 3. 5

Kelompok pilihan jawaban

No soal	**TK_{JB}**	**TK_{JS}**	**S_{JB-JS}**	**CDQ_1**

Keterangan:

TK_{JB} = tingkat keyakinan jawaban benar

TK_{JS} = tingkat keyakinan jawaban salah

S_{JB-JS} = standar deviasi tingkat keyakinan jawaban benar-jawaban salah

2) Kelompok alasan

Tabel 3. 6

Kelompok pilihan alasan

No soal	**TK_{AB}**	**TK_{AS}**	**S_{AB-AS}**	**CDQ_2**

Keterangan:

TK_{AB} = tingkat keyakinan alasan benar

TK_{AS} = tingkat keyakinan alasan salah

S_{AB-AS} = standar deviasi tingkat keyakinan alasan benar-alasan salah

3) Kelompok pilihan jawaban dan alasan

Tabel 3. 7

Kelompok pilihan jawaban benar-benar dan alasan salah-salah

No soal	**TK_{BB}**	**TK_{SS}**	**S_{BB-SS}**	**CDQ_3**

Keterangan:

TK_{BB} = tingkat keyakinan jawaban benar dan alasan benar

TK_{SS} = tingkat keyakinan jawaban salah dan alasan salah

S_{BB-SS} = standar deviasi tingkat keyakinan jawaban benar-alasan benar dan jawaban salah-alasan salah

Tabel 3. 8

Kelompok pilihan jawaban benar-salah dan alasan salah-benar

No soal	**TK_{BS}**	**TK_{SB}**	**S_{BS-SB}**	**CDQ_4**

Keterangan:

TK_{BS} = tingkat keyakinan jawaban benar dan alasan salah

TK_{SB} = tingkat keyakinan jawaban salah dan alasan benar

S_{BS-SB} = standar deviasi tingkat keyakinan jawaban benar-alasan benar dan jawaban salah-alasan salah

Kelompok miskonsepsi dari CDQ (*Confidence Discrimination Quotient*) yang diambil adalah kelompok pilihan jawaban dan alasan dimana pilihan jawaban benar alasan benar serta jawaban salah alasan salah. Tingkat miskonsepsi yang digunakan dalam penelitian ini ditinjau berdasarkan rata-rata tingkat keyakinan mahasiswa yang mengalami miskonsepsi pada setiap temuan miskonsepsi yang dianggap signifikan untuk tingkat kedua dan tingkat kelima.

Untuk penyebab terjadinya miskonsepsi, dapat diketahui dari tingkat ketiga dan tingkat keenam. Pada tingkat ketiga melihat sumber yang dipilih oleh mahasiswa untuk jawaban sedangkan tingkat keenam melihat sumber yang dipilih mahasiswa untuk alasan. Masing-masing sumber penyebab dilihat dari berapa banyak mahasiswa yang memilih sumber dan dihitung tingkat keyakinan untuk mahasiswa yang memilih sumber penyebab tersebut. Salah satu kategori sumber pilihan jawaban/alasan dapat menjadi penyebab miskonsepsi, apabila mahasiswa memilih jawaban salah dan tingkat keyakinan tinggi. Sedangkan apabila salah satu distraktor sumber pilihan jawaban/alasan dipilih mahasiswa dengan jawaban benar dan tingkat keyakinan tinggi maka tidak dapat dinyatakan sebagai penyebab miskonsepsi.

Tabel 3. 9

Analisis Sumber Penyebab Miskonsepsi

Jawaban Benar Alasan Benar – Jawaban Salah Alasan Salah				
No Soal	**CDQ**			
	Buku	**Dosen**	**Sendiri**	**Internet**

Jika kedua jawaban dipilih yaitu A dan B, maka kategori penyebab miskonsepsi pada mahasiswa adalah salah satu distraktor yang memiliki nilai negatif (jawaban salah namun tingkat keyakinan tinggi) yang lebih besar.