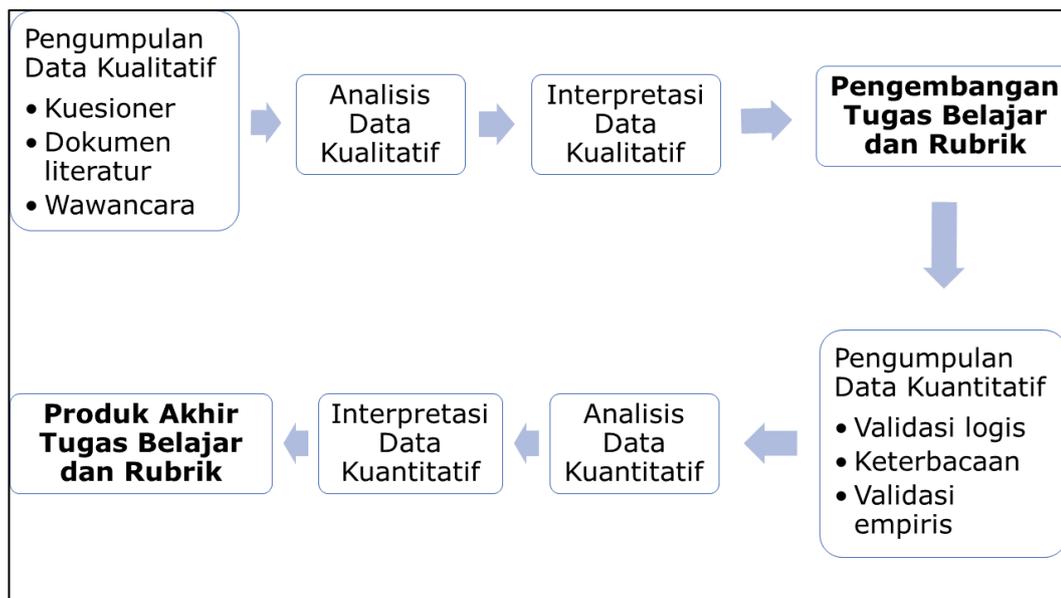


## BAB III

### METODE PENELITIAN

#### 3.1 Desain Penelitian

Penelitian ini difokuskan pada rancang bangun tugas belajar topik gelombang cahaya berbasis kerangka pembelajaran tiga dimensi dengan teknologi *mobile web*. Metode penelitian yang digunakan adalah *sequential mixed methods* yaitu melibatkan pengumpulan data secara kualitatif dan kuantitatif, penggabungan dua bentuk data tersebut serta menggunakan rancangan dengan melibatkan asumsi-asumsi dan kerangka teoritis (Creswell & Creswell, 2018). Pada penelitian ini, model pengembangan yang digunakan adalah *Exploratory Design: Instrument Development Model*, model ini merupakan desain penelitian untuk mengembangkan instrumen dengan mengeksplorasi menggunakan penelitian kualitatif yang diikuti dengan penelitian kuantitatif. Pada penelitian kualitatif, diperoleh data kualitatif untuk mengembangkan asesmen dengan melibatkan ahli, kemudian data tersebut dianalisis dan diperoleh kesimpulan konstruksi pengembangan asesmen. Selanjutnya dilakukan penelitian kuantitatif sehingga diperoleh data kuantitatif hasil uji coba asesmen yang telah dikembangkan. Data tersebut dianalisis kemudian diinterpretasikan sehingga menghasilkan produk berupa tugas belajar topik gelombang cahaya berbasis kerangka pembelajaran tiga dimensi yang dapat diakses secara *online* melalui teknologi *mobile web* seperti *smartphone*, laptop, komputer, serta perangkat *mobile* lainnya. Adaptasi *Exploratory Design: Instrument Development Model* yang digunakan pada penelitian ini disajikan pada Gambar 3.1.



Gambar 3.1 Adaptasi *Exploratory Design: Instrument Development Model* dalam Pengembangan Tugas Belajar Topik Cahaya Berbasis Kerangka Pembelajaran Tiga Dimensi

### 3.2 Partisipan Penelitian

Dalam penelitian ini melibatkan beberapa partisipan, yaitu dosen ahli, guru, dan peserta didik. Dosen ahli yang terlibat dalam penelitian ini berjumlah empat orang beserta tiga guru untuk melakukan validasi logis baik dari segi konstruk dan konten. Selain itu, terdapat juga 20 orang guru dan 150 peserta didik dari SMAN 1 Jatiluhur di Kabupaten Purwakarta yang akan menjadi partisipan sebagai responden terkait produk tugas belajar topik gelombang cahaya berbasis kerangka pembelajaran tiga dimensi dengan teknologi *mobile web* yang diuji coba. Hal tersebut menyesuaikan dengan kriteria partisipan untuk analisis *Rasch model* pada tingkatan kepercayaan 95% (Linacre, 1994; Sumintono & Widhiarso, 2015). Partisipan tersebut memiliki karakteristik sebagai berikut: 1) Peserta didik berusia antara 15-18 tahun; 2) Memiliki dan terampil menggunakan perangkat *mobile*.

### 3.3 Populasi dan Sampel

Populasi pada penelitian ini yaitu seluruh peserta didik kelas XI semester 2 dari SMAN 1 Jatiluhur di Kabupaten Purwakarta pada tahun ajaran 2021/2022. Dalam penelitian ini pengambilan sampel ditentukan dengan *convenience sampling*, yaitu salah satu jenis teknik *non random sampling* dimana anggota populasi target yang memenuhi kriteria praktis tertentu, seperti aksesibilitas mudah,

ketersediaan pada waktu tertentu, atau kesediaan untuk berpartisipasi dalam penelitian tertentu (Creswell & Creswell, 2018). Melalui teknik sampling tersebut dipilih peserta didik dalam lima kelas yang sedang mempelajari materi gelombang cahaya. Selain itu partisipan juga ditentukan sesuai dengan kesediaan pihak sekolah dan peserta didik untuk mengikuti penelitian ini.

### **3.4 Instrumen Penelitian**

Penelitian ini menggunakan lima jenis instrumen yaitu lembar kuesioner pembelajaran tiga dimensi dan lembar wawancara terstruktur untuk data kualitatif, kemudian untuk data kuantitatif digunakan lembar validasi logis, lembar keterbacaan serta tugas belajar yang dikembangkan. Peran kelima instrumen penelitian tersebut dalam mengembangkan tugas belajar topik cahaya berbasis kerangka pembelajaran tiga dimensi dapat dijelaskan sebagai berikut.

#### **3.4.1 Lembar Kuesioner Pembelajaran Tiga Dimensi**

Kuesioner pembelajaran tiga dimensi diberikan kepada guru dan peserta didik SMA di Purwakarta untuk memperoleh data kebutuhan pembelajaran. Kuesioner pembelajaran tiga dimensi menggunakan *google form* yang terdiri dari tiga bagian. Bagian pertama berisi karakteristik responden. Bagian kedua berisi tentang pengetahuan awal responden terkait pembelajaran tiga dimensi. Bagian ketiga membahas tentang tugas belajar yang digunakan di sekolah. Kuesioner bertujuan untuk menganalisis kebutuhan sebagai dasar dalam mengembangkan tugas belajar gelombang cahaya berbasis kerangka pembelajaran tiga dimensi. Lembar kuesioner pembelajaran tiga dimensi selengkapnya dapat dilihat di Lampiran A.1. dan Lampiran A.2.

#### **3.4.2 Lembar Wawancara Terstruktur**

Wawancara terhadap lima guru Fisika dan lima peserta didik XI SMAN 1 Jatiluhur ini dilakukan dengan pertanyaan-pertanyaan sebagai berikut:

- 1) Apa yang Anda ketahui terkait pembelajaran tiga dimensi (*three-dimensional learning*)?
- 2) Hal apa saja yang penting dalam pembelajaran Fisika masa kini?
- 3) Apakah tugas belajar yang diberikan sudah berbasis teknologi *mobile web*?
- 4) Bagaimana pengembangan tugas belajar fisika yang harus dilakukan?

Lembar wawancara selengkapnya dapat dilihat di Lampiran A.3. dan Lampiran A.4.

### 3.4.3 Lembar Validasi Logis

Pada penelitian ini digunakan validitas logis untuk instrumen yang dikembangkan maka dilakukan *expert judgment* dengan melibatkan ahli dalam bidang asesmen dan Fisika. Lembar validasi menggunakan *Google Form* untuk memberikan penilaian pada tiap tugas belajar. Lembar validasi tugas belajar secara daring yang dikembangkan menggunakan *website* dapat diakses melalui link <https://bit.ly/LembarValidasiTugasGelCahaya> agar konten digital dalam tugas belajar seperti video dan simulasi dapat ditampilkan dengan baik. Validator akan memeriksa kesesuaian butir tugas belajar berdasarkan delapan aspek dalam *Learning Object Review Instrument (LORI) 2.0* (Nesbit dkk, 2009). Tabel 3.1 menjelaskan kriteria yang diadaptasi dari LORI 2.0 pada lembar validasi logis tugas belajar.

Tabel 3.1

Kualitas Tugas Belajar Berdasarkan LORI 2.0

| No | Aspek                                                                                                                                                                         | Skala* |   |   |   |   |
|----|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------|---|---|---|---|
|    |                                                                                                                                                                               | 1      | 2 | 3 | 4 | 5 |
| 1  | <b>Kualitas Konten</b><br>Konten pada tugas belajar sudah sesuai dengan konsep fisika. Ide-ide yang disampaikan dapat digunakan kembali dengan konteks yang berbeda.          |        |   |   |   |   |
| 2  | <b>Kesesuaian Tujuan Pembelajaran</b><br>Tugas belajar sudah sesuai dengan <i>learning objective</i> .                                                                        |        |   |   |   |   |
| 3  | <b>Respon dan Adaptasi</b><br>Konten pada tugas belajar dapat beradaptasi dan memberikan umpan balik yang dapat digunakan oleh peserta didik atau model pembelajaran berbeda. |        |   |   |   |   |
| 4  | <b>Motivasi</b><br>Tugas belajar dapat memberi motivasi dan menarik perhatian peserta didik dalam menggunakan bahan ajar tersebut.                                            |        |   |   |   |   |

| No                                                                                | Aspek                                                                                                                                               | Skala* |   |   |   |   |
|-----------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------|--------|---|---|---|---|
|                                                                                   |                                                                                                                                                     | 1      | 2 | 3 | 4 | 5 |
| 5                                                                                 | <b>Desain Penyajian</b><br>Desain dari visual dan audio memuat informasi yang dapat meningkatkan pembelajaran dan proses mental yang efisien.       |        |   |   |   |   |
| 6                                                                                 | <b>Kemudahan Penggunaan</b><br>Tugas belajar memiliki navigasi yang baik dan antarmuka yang intuitif serta mudah untuk digunakan oleh pengguna.     |        |   |   |   |   |
| 7                                                                                 | <b>Aksesibilitas</b><br>Desain tugas belajar dapat dikontrol dan format penyajian mengakomodasi berbagai karakter atau keadaan peserta didik.       |        |   |   |   |   |
| 8                                                                                 | <b>Pemenuhan Standar</b><br>Tugas belajar memenuhi suatu standar yang berlaku dan dapat dioperasikan dalam platform yang umum dan banyak digunakan. |        |   |   |   |   |
| *) 1: Tidak Sesuai; 2: Kurang Sesuai; 3: Agak Sesuai; 4: Sesuai; 5: Sangat Sesuai |                                                                                                                                                     |        |   |   |   |   |

(diadaptasi dari Nesbit, Belfer, & Leacock, 2009)

Penilaian diberikan untuk setiap butir tugas belajar dengan memilih kriteria penilaian yang paling menggambarkan kualitas butir soal yang dikembangkan, apakah setiap butir tugas belajar **Tidak Sesuai, Kurang Sesuai, Cukup Sesuai, Sesuai atau Sangat Sesuai** dengan ketiga kaidah penulisan yang dimaksud.

#### 3.4.4 Lembar Keterbacaan

Pada lembar keterbacaan ini bertujuan untuk melihat keterbacaan, jawaban, dan refleksi kemudahan atau kesulitan prototipe dipahami dan dijawab oleh peserta didik. Instrumen menggunakan *Google Form* untuk memberikan penilaian pada tiap butir tugas belajar, meliputi keterbacaan tugas belajar beserta alasannya. Keterbacaan tugas belajar terdiri dari 5 kategori yaitu **Tidak Jelas, Kurang Jelas,**

**Cukup Jelas, Jelas dan Sangat Jelas.** Kemudian peserta didik memberikan alasan terkait keterbacaan tugas belajar yang dipilih.

Peserta didik dapat mengakses instrumen secara daring yang dikembangkan menggunakan *website* melalui link <https://bit.ly/UjiTerbatasTugasBelajar1>, <https://bit.ly/UjiTerbatasTugasBelajar2>, <https://bit.ly/UjiTerbatasTugasBelajar3> agar video dan simulasi dapat ditampilkan dengan baik.

### **3.4.5 Tugas Belajar Topik Gelombang Cahaya Berbasis Kerangka Pembelajaran Tiga Dimensi yang Dikembangkan**

Tugas belajar topik gelombang cahaya berbasis kerangka pembelajaran tiga dimensi juga digunakan sebagai salah satu instrumen penelitian. Instrumen ini merupakan seperangkat lembar kerja yang terdiri dari tiga sub materi yaitu interferensi, difraksi dan polarisasi cahaya. Setiap tugas belajar memiliki 16 butir pertanyaan uraian dengan stimulus berupa video dan simulasi dari proses pengembangan yang dilakukan. Instrumen ini dikembangkan berdasarkan kerangka pembelajaran tiga dimensi dan kurikulum terkait materi gelombang cahaya. Jawaban peserta didik dianalisis menggunakan rubrik yang disusun mengacu pada *performance expectation* (PE) dan *learning objective* (LO).

### **3.5 Prosedur Penelitian**

Prosedur penelitian yang dilakukan pada penelitian ini melalui dua fase utama yaitu fase kualitatif kemudian selanjutnya fase kuantitatif, pada penelitian ini tujuan utamanya adalah menghasilkan produk berupa tugas belajar dan rubrik. Penelitian *Exploratory Design: Instrument Development Model* dengan latar belakang kualitatif yang kuat dimana peneliti mengumpulkan data kualitatif pada fase pertama, menganalisis hasil, dan kemudian menggunakan hasil untuk merencanakan fase pengembangan produk.

Hasil data kualitatif berisi aspek dari *three-dimensional learning* yang akan dinilai serta konstruksi dari instrumen. Fase pengembangan produk mengkonstruksi spesifikasi produk yang dikembangkan dan menguji hasil konstruksi dengan meminta masukan kepada dosen ahli dan guru. Data kuantitatif membantu dalam menjelaskan lebih lanjut hasil dari data kualitatif. Pada penelitian ini data kuantitatif diperlukan untuk menggali lebih jauh terkait implementasi tugas belajar termasuk kritik dan saran untuk memperbaiki tugas belajar yang dikembangkan.

Berdasarkan tahapan deskripsi kegiatan pada *Exploratory Design: Instrument Development Model*, pada bagian pertama penelitian pengembangan ini akan dilakukan beberapa studi kualitatif yang hasilnya digunakan untuk menyusun *blueprint* yang akan memandu proses pengembangan prototipe selanjutnya. Studi kualitatif ini mencakup analisis kebutuhan pembelajaran dari kuesioner pembelajaran tiga dimensi, hasil wawancara, dokumen kerangka pembelajaran tiga dimensi dan kurikulum terkait materi gelombang cahaya, serta studi dari penelitian lain yang relevan terkait karakteristik tugas belajar yang diberikan. Tahap yang dilakukan yaitu membuat *performance expectation* (PE) dan *learning objective* (LO) tentang materi gelombang cahaya yang disesuaikan dengan kurikulum yang berlaku. Selanjutnya pemetaan praktik saintifik (*Scientific Practices*), konsep lintas bidang (*crosscutting concepts*) dan gagasan inti disiplin ilmu (*Disciplinary Core Ideas*) *Disciplinary Core Ideas*) pada materi gelombang cahaya serta penulisan batasan materi.

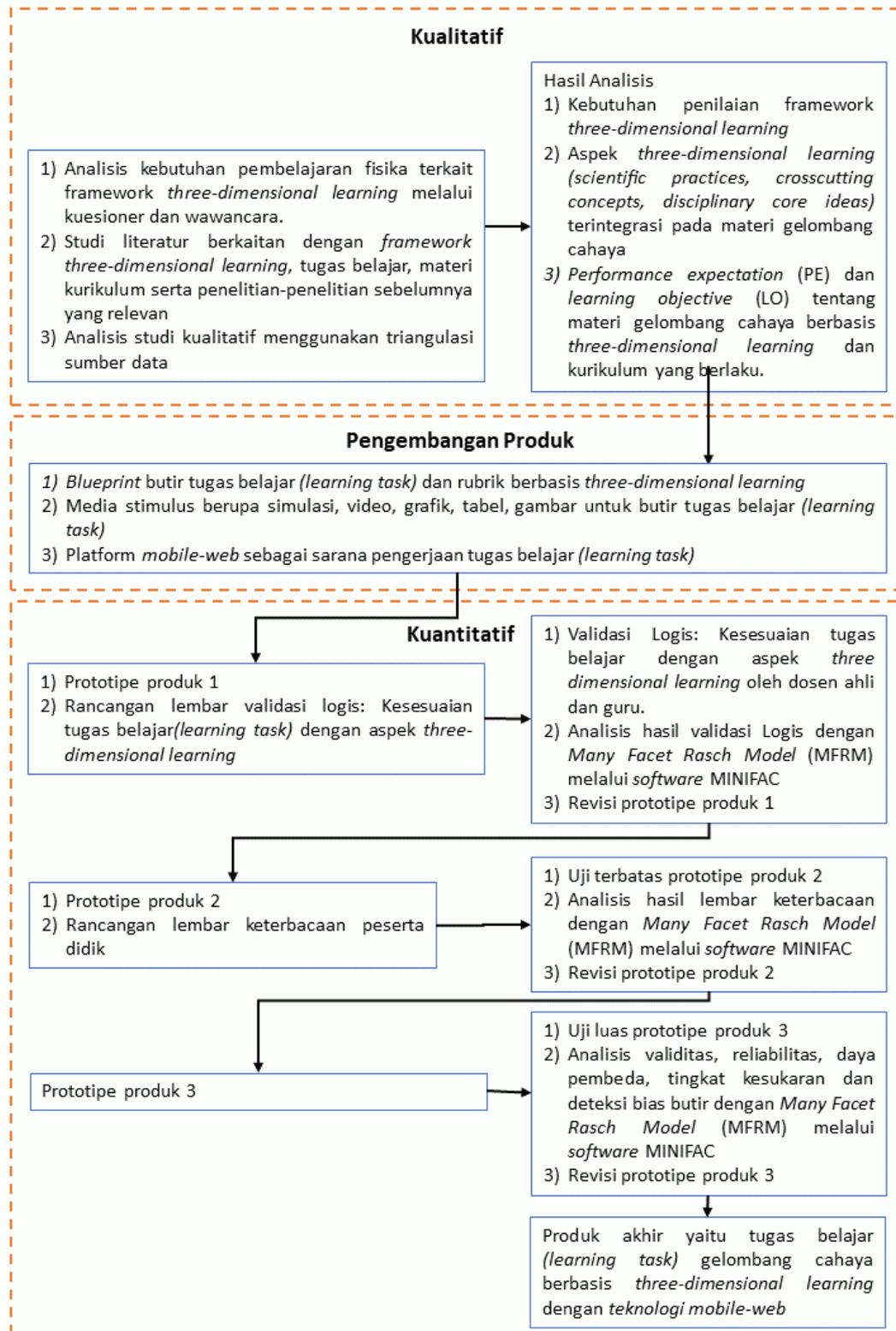
Pada tahap pengembangan produk, dilakukan pembuatan *blueprint* tugas belajar. *Blueprint* yang dihasilkan dari studi kualitatif dibuat menjadi Prototipe 1 dengan mempertimbangkan pula kepraktisan, efektivitas dan efisiensi media yang digunakan. Selanjutnya penyusunan simulasi menggunakan *Physics Education Technology* (PhET) serta menyusun grafik, tabel, gambar pendukung dan video pendukung. Prototipe 1 ini yang akan dievaluasi melalui validasi ahli.

Pada tahapan studi kuantitatif, tugas belajar berbasis kerangka pembelajaran tiga dimensi yang telah dibuat kemudian dikonsultasikan kepada para ahli (*judgment experts*) untuk mengevaluasi validitas isi serta konstruksi soal. Tugas belajar yang dikembangkan divalidasi oleh *judgment experts* sebanyak tujuh validator. Validasi ini mengukur kualitas tugas belajar berdasarkan delapan aspek dalam *Learning Object Review Instrument* (LORI) 2.0 (Nesbit dkk, 2009) menggunakan *rating scale* skala 1 sampai 5. Hasil validasi logis kemudian dianalisis menggunakan analisis *Many Facet Rasch Measurement* (MFRM) pada *software* MINIFAC. Validitas logis juga mengandung catatan perbaikan dan saran untuk memperbaiki instrumen yang dikembangkan sehingga didapat instrumen yang valid. Setelah soal divalidasi oleh ahli, kemudian dilakukan evaluasi dan

perbaikan tugas belajar sehingga dapat dinyatakan valid. Hasil evaluasi digunakan untuk mengembangkan Prototipe 2.

Uji terbatas pada Prototipe 2 digunakan untuk mengevaluasi proses penggunaannya dalam pembelajaran dan mengidentifikasi kesulitan-kesulitan peserta didik yang diakibatkan dari kekurangan prototipe. Pada tahapan analisis pengujian ini, peneliti akan melakukan analisis *Many Facet Rasch Measurement* (MFRM) pada *software* MINIFAC untuk mendapatkan informasi yang dibutuhkan. Setelah data dikumpulkan dan dianalisis, hasilnya digunakan untuk memperbaiki Prototipe 2 yang dibuat menjadi Prototipe 3.

Pada tahap terakhir, Protipe 3 diujicobakan secara luas ke lebih banyak peserta didik. Pengujian ini ditujukan untuk menilai secara empiris kualitas tugas belajar meliputi validitas, reliabilitas, daya pembeda, tingkat kesukaran dan deteksi bias butir. Hasil dari pengujian Prototipe 3 dianalisis menggunakan *Many Facet Rasch Measurement* (MFRM) pada *software* MINIFAC untuk diimplementasikan ke dalam revisi akhir produk. Setelahnya, proses dilanjutkan ke dalam evaluasi secara keseluruhan untuk menghasilkan simpulan, saran dan rekomendasi baik untuk kebutuhan praktis, kebutuhan teoritis, maupun kebutuhan riset lanjutan. Ketiga tahapan besar di atas terangkum dalam Gambar 3.2.



Gambar 3.2 Prosedur Penelitian

### 3.6 Analisis Data Hasil Penelitian

Terdapat dua jenis data yang diperoleh dalam penelitian ini, yaitu data kualitatif dan data kuantitatif. Untuk masing-masing jenis data, dianalisis menggunakan metode yang berbeda sebagai berikut ini.

#### 3.6.1 Analisis Data Kualitatif

Data yang didapatkan dari penelitian kualitatif terdiri dari data kuesioner, wawancara, dokumen kerangka pembelajaran tiga dimensi dan kurikulum. Data kualitatif tersebut dianalisis menggunakan triangulasi sumber data. Triangulasi data dapat digambarkan sebagai penggunaan beberapa sumber data untuk mendapatkan pandangan yang berbeda tentang situasi dalam satu penelitian (Cohen dkk., 2017). Triangulasi menggambarkan penggunaan beberapa sumber data dalam penelitian yang sama untuk tujuan validasi.

#### 3.6.2 Analisis Data Kuantitatif

Data yang didapatkan dari penelitian kuantitatif terdiri dari data hasil validasi logis, lembar keterbacaan pada uji terbatas, serta validasi empiris pada uji luas. Data tersebut kemudian diolah menggunakan *Many Facet Rasch Measurement* (MFRM). Georg Rasch (1960) mengembangkan satu model analisis dari teori respon butir atau *Item Response Theory* (IRT), biasa disebut 1PL (satu parameter logis) (Olsen, 2003). *Many Facet Rasch Measurement* (MFRM) merupakan metode pengujian dengan menggunakan skala peringkat yang dilakukan beberapa validator sehingga diperoleh penilaian secara mendasar dari metode penilaian (Sumintono & Widhiarso, 2014; Zahir & Sumintono, 2017). MFRM memiliki kelebihan yaitu setiap hasil validasi dapat ditunjukkan berdasarkan cara penilai menggunakan skala peringkat sehingga ahli dapat mendefinisikan tentang skala peringkatnya (Bond & Fox, 2015; Boone dkk., 2014; Engelhard, 2013; Zahir & Sumintono, 2017).

Pada MFRM terdapat beberapa nilai yang dimunculkan dalam tabel hasil data. *Item fit* menjelaskan apakah butir soal berfungsi normal melakukan pengukuran atau tidak (Sumintono & Widhiarso, 2015). Nilai *infit* menitikberatkan pada responden (*person*) yang kemampuannya dekat dengan kesulitan item, sementara itu nilai *outfit* tidak menitikberatkan siapapun sehingga lebih sensitif kepada pengaruh jawaban yang tidak wajar (Bond & Fox, 2015). Nilai MNSQ

menunjukkan ukuran keacakan, yaitu jumlah penyimpangan dalam sistem pengukuran (Sumintono & Widhiarso, 2015). Nilai ZSTD menunjukkan seberapa besar kemungkinan terjadinya penyimpangan (Bond & Fox, 2015). Nilai *Pt Measure Corr* merupakan hubungan antara kesulitan masing-masing *item* dan kesulitan instrumen secara keseluruhan (Smiley, 2015).

Untuk melakukan pengolahan data menggunakan *Many Facet Rasch Measurement* (MFRM) ini dibantu dengan program Minifac yang dikembangkan oleh Winstep. Prinsip pada pengolahan ini adalah data mentah dari *Microsoft Excel* diubah menjadi sebuah program pengkodean khusus untuk dianalisis dengan *inter-rater* (Zahir & Sumintono, 2017). Langkah-langkah pengkodean yang dilakukan untuk melakukan analisis MFRM dapat dilihat pada Lampiran B.8.

Kualitas tugas belajar yang dikembangkan meliputi aspek validitas dan reliabilitas. Untuk karakteristik tugas belajar dilihat dari unidimensionalitas, tingkat kesukaran, daya pembeda, deteksi bias butir tugas belajar. Berikut penjelasan tahap analisis secara lebih rinci.

### 3.6.2.1. Uji Unidimensionalitas

Uji unidimensionalitas (*unidimensionality*) instrumen adalah ukuran penting untuk mengevaluasi apakah instrumen yang dikembangkan mampu mengukur apa yang seharusnya diukur berdasarkan tujuan yang diharapkan (Sumintono & Widhiarso, 2015). Uji unidimensionalitas dilakukan pada tahap validasi logis, analisis keterbacaan (uji terbatas) dan validasi empiris (uji luas). Untuk nilai validitas pada tingkat instrumen ditentukan dengan melihat unidimensional untuk nilai *variance explained by Rasch measures*. Kriteria *unidimensionality* pada *Rasch model* dijelaskan pada Tabel 3.2.

Tabel 3.2

#### Kriteria *Unidimensionality*

| <i>Variance explained by Rasch measures</i> | <b>Interpretasi</b> |
|---------------------------------------------|---------------------|
| $20\% \leq X \leq 40\%$                     | Cukup               |
| $40\% < X \leq 60\%$                        | Baik                |
| $X > 60\%$                                  | Sangat Baik         |

(Sumintono & Widhiarso, 2015)

### 3.6.2.2. Uji Validitas

Uji validitas bertujuan untuk mengukur ketepatan suatu instrumen yang digunakan dalam penelitian agar penelitian dapat dilaksanakan sesuai tujuan yang telah ditetapkan. Uji validitas dilakukan pada tahap validasi logis dan validasi empiris (uji luas). Person dan item dinyatakan valid berdasarkan *Rasch Model* menurut Sumintono & Widhiarso (2015) apabila memenuhi minimal dua dari tiga kriteria berikut.

- 1) Nilai *Outfit Mean Square* (MNSQ), untuk menguji konsistensi jawaban dengan tingkat kesulitan butir pernyataan;
- 2) Nilai *Outfit Z-Standard* (ZSTD), untuk mendeskripsikan *how much* hasil measure merupakan butir outfit, tidak mengukur atau terlalu mudah, atau terlalu sulit;
- 3) Nilai *Point Measure Correlation* (*Pt Measure Corr*), untuk mendeskripsikan *how good* butir pernyataan dipahami, direspon beda, atau membingungkan dengan item lainnya.

Untuk butir dinyatakan *fit* berdasarkan kriteria penerimaan *Outfit MnSq*, *Outfit ZStd* dan *PT Measure Correlation* yang dapat dilihat pada Tabel 3.3.

Tabel 3.3

Kriteria Penerimaan *Outfit Mnsq*, *Outfit Zstd* dan *PT Measure Correlation*

| Nilai                         | Kriteria                              |
|-------------------------------|---------------------------------------|
| <i>Outfit MnSq</i>            | $0,5 < MnSq < 1,5$                    |
| <i>Outfit ZStd</i>            | $-2,0 < ZStd < +2,0$                  |
| <i>PT Measure Correlation</i> | $0,4 < PT Measure Correlation < 0,85$ |

### 3.6.2.3. Uji Reliabilitas

Reliabilitas berhubungan dengan kestabilan atau konsistensi dari hasil pengukuran suatu instrumen (Silalahi, 2017). Suatu instrumen asesmen dikatakan mempunyai reliabilitas atau taraf kepercayaan yang tinggi jika instrumen asesmen tersebut memberikan hasil yang tetap (Arikunto, 2016). Hasil pengukuran itu harus tetap sama jika pengukurannya diberikan pada subjek yang sama meskipun oleh orang yang berbeda, waktu yang berbeda dan tempat yang berbeda pula. Hasil asesmen tidak terpengaruh oleh pelaku, situasi dan kondisi. Menurut Sumintono &

Widhiarso (2015), uji reliabilitas dengan menggunakan *Rasch Model* terdiri dari reliabilitas inter-rater (*inter-rater reliability*) serta reliabilitas person dan item (*person and item reliability*) sebagai berikut.

1) *Inter-rater Reliability*

Reliabilitas inter-rater (*inter-rater reliability*) menyatakan probabilitas persetujuan validator atau rater terhadap suatu instrumen. Reliabilitas inter-rater digunakan pada tahap validasi logis dan analisis keterbacaan (uji terbatas). Probabilitas persetujuan validator atau *inter-rater agreement* diperoleh dengan membandingkan nilai persetujuan nyata (*exact agreement*) dan nilai persetujuan yang diharapkan (*expected agreement*).

2) *Person and Item Reliability*

Reliabilitas person dan item (*person and item reliability*) menunjukkan keajegan responden dan item instrumen dalam melakukan pengukuran. Reliabilitas person digunakan pada tahap validasi logis dan analisis keterbacaan (uji terbatas). Kemudian reliabilitas item digunakan pada tahap validasi empiris (uji luas). Nilai *person reliability* dan *item reliability* dalam *Rasch model* memiliki kriteria yang terdapat pada Tabel 3.4.

Tabel 3.4

Kriteria *Person Reliability* dan *Item Reliability*

| <i>Person reliability</i> dan <i>item reliability</i> | Interpretasi |
|-------------------------------------------------------|--------------|
| $0,94 \leq \text{nilai}$                              | Istimewa     |
| $0,91 \leq \text{nilai} < 0,94$                       | Sangat Baik  |
| $0,81 \leq \text{nilai} < 0,90$                       | Baik         |
| $0,67 \leq \text{nilai} < 0,80$                       | Cukup        |
| $\text{nilai} < 0,67$                                 | Lemah        |

(Sumintono & Widhiarso, 2015)

#### 3.6.2.4. Tingkat Kesukaran

Tingkat kesukaran menyatakan seberapa sukar atau mudahnya suatu soal. Secara kuantitatif, tingkat kesulitan dinyatakan dengan indeks kesukaran (Arikunto, 2016). Tingkat kesulitan butir pada *Rasch model* pada dasarnya sama dengan taraf kesukaran teori klasik, yaitu perbandingan antara jumlah jawaban benar dengan jumlah soal yang diujikan (*odd-ratio*). Hanya saja yang membedakan adalah nilai

peluang itu kemudian diskalakan dengan memasukkan fungsi logaritma. Hasil estimasi logit dari *odd-ratio* inilah yang disebut logit atau *item measure*. Nilai logit yang tinggi menunjukkan item tersebut sukar (Sumintono & Widhiarso, 2015). Kriteria tingkat kesukaran dalam *Rasch model* terdapat pada Tabel 3.5.

Tabel 3.5  
Kriteria Tingkat Kesukaran

| Tingkat Kesukaran (TK) | Interpretasi |
|------------------------|--------------|
| $X > 0,50$             | Sangat Sukar |
| $0,00 < X < 0,50$      | Sukar        |
| $-0,50 < X < 0,00$     | Mudah        |
| $X < -0,50$            | Sangat Mudah |

### 3.6.2.5. Daya Pembeda

Daya pembeda soal adalah kemampuan suatu soal untuk membedakan antara peserta didik yang berkemampuan tinggi (pandai) dengan peserta didik yang berkemampuan rendah (tidak pandai) (Arikunto, 2016). Angka yang menunjukkan besarnya daya pembeda disebut indeks diskriminasi. Daya diskriminasi Rasch atau nilai korelasi skor butir dan skor Rasch (*Point Measure Correlation*) pada prinsipnya sama dengan daya diskriminasi item yang diukur dengan pendekatan teori klasik. Hanya saja jika pada teori klasik komputasinya menggunakan skor mentah, untuk *Rasch model* yang digunakan adalah skor *Point Measure Correlation*.

Nilai *Point Measure Correlation* (*Pt Measure Corr.*) yang tinggi mengindikasikan bahwa peserta didik dengan abilitas rendah menjawab butir dengan salah dan peserta didik dengan abilitas tinggi menjawab butir dengan benar. Sementara nilai *Point Measure Correlation* negatif mengindikasikan butir soal yang menyesatkan karena peserta didik dengan kemampuan rendah mampu menjawab butir dengan benar dan peserta didik dengan kemampuan tinggi justru menjawab salah. Soal-soal dengan nilai korelasi negatif harus diperiksa untuk melihat apakah kunci jawaban salah, perlu direvisi, atau dihapus dari tugas belajar (Smiley, 2015).

Seperti pada teori klasik, nilai korelasi skor butir dan skor Rasch yang ideal adalah positif serta tidak mendekati nol. Alagumalai, Curtis, & Hungi (2005) mengklasifikasikan daya pembeda berdasarkan nilai *Point Measure Correlation* dalam *Rasch model* seperti yang ditunjukkan pada Tabel 3.6.

Tabel 3.6

## Kriteria Daya Pembeda

| <b>Daya Pembeda</b> | <b>Interpretasi</b>          |
|---------------------|------------------------------|
| >0,40               | Sangat Baik                  |
| 0,30–0,39           | Baik                         |
| 0,20-0,29           | Cukup                        |
| 0,00-0,19           | Kurang                       |
| <0,00               | Mebutuhkan pemeriksaan butir |

(Alagumalai, Curtis, & Hungi, 2005)

### 3.6.2.6. Deteksi Bias

Butir instrumen pengukuran dapat bersifat bias, yaitu ketika sebuah butir lebih memihak pada salah satu individu dengan karakteristik tertentu. Salah satu contohnya adalah bias terhadap jenis kelamin atau bias gender. Butir yang cenderung bias dalam pengukuran mengalami keberfungsian butir diferensial atau *differential item functioning* (DIF). *Rasch model* menyediakan menu untuk memfasilitasi deteksi adanya butir-butir yang bias. Suatu butir dinyatakan bias gender apabila memiliki nilai probabilitas lebih kecil dari 0,05 atau 5% untuk jenis kelamin tertentu (Sumintono & Widhiarso, 2015).