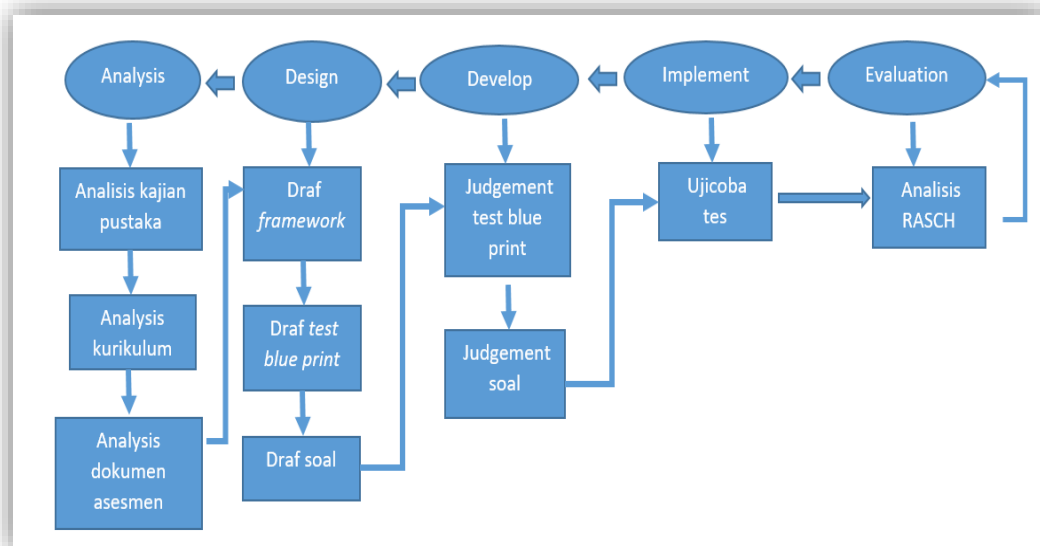


BAB III METODE PENELITIAN

3.1 Desain Penelitian

Metode yang digunakan dalam penelitian ini merupakan metode penelitian pengembangan atau *Research and Development* (R&D) atau sering juga disebut metode pengembangan yaitu metode penelitian yang digunakan untuk menghasilkan produk tertentu. Desain penelitian pengembangan mengadopsi dari rancangan pengembangan model *Analysis, Design, Development, Implementation, Evaluation* (ADDIE) (Branch, 2009). Model ADDIE dalam mendesain sistem instruksional pendekatan sistem. Esensi dari pendekatan sistem adalah untuk mengatur langkah-langkah ke dalam unsur logis, kemudian menggunakan *output* dari setiap langkah sebagai *input* pada langkah berikutnya (Januszewski & Molenda, 2008). Model instruksional ADDIE pada penelitian dapat dilihat pada Gambar 3.1, sebagai berikut.



Gambar 3. 1 Bagan Alur Pengembangan Instrumen tes.

Model pengembangan yang digunakan dalam penelitian ini adalah model ADDIE (*Analysis, Design, Development, Implementation, and Evaluation*). Alasan dalam penelitian pengembangan ini menggunakan model ADDIE karena memiliki tahapan-tahapan yang sistematis dan mudah dipelajari. Tahapan model ADDIE

terdapat revisi pada setiap tahapannya, sehingga model ADDIE cocok digunakan untuk pengembangan instrumen tes tertulis agar menghasilkan instrumen tes yang valid dan praktis. tahapan model pengembangan ADDIE yang di adaptasi dari Robert Marine Branch (2009). Model ini memiliki kesempatan untuk revisi dan evaluasi secara berulang-ulang pada setiap tahapan, bertujuan agar menghasilkan produk yang valid dan praktis. berdasarkan hal tersebut model pengembangan ADDIE dirasa cocok digunakan untuk mengembangkan suatu produk. Model pengembangan ADDIE hanya sampai tahap evaluasi, dimana evaluasi diperoleh dari komentar dan saran para validator dan peserta didik sebagai acuan perbaikan produk yang dikembangkan. Berikut penjelasan prosedur Pengembangan dari tahap pengembangan ADDIE yang akan peneliti lakukan.

Prosedur pengembangan ADDIE meliputi 5 tahapan yaitu: *Analyze* (Analisis), *Design* (Perancangan), *Development* (pengembangan), *implementation* (Implementasi), *Evaluate* (Evaluasi). Model ADDIE biasa digunakan dalam penelitian pengembangan karena memiliki tahapan-tahapan yang sistematis. Tahapan model ADDIE hanya sampai tahap evaluasi tanpa ada tahap penyebaran, sehingga dalam penelitian ini sesuai dengan tahapan yang telah ditetapkan. berikut penjelasan dari 5 tahapan tersebut.

1. Tahap Analisis (*Analyze*)

Kegiatan utama tahap analysis adalah menganalisis perlunya pengembangan instrumen tes tertulis baru dan menganalisis syarat-syarat kelayakan pengembangan instrumen tes baru. kegiatan yang dilakukan berupa analisis kajian pustaka, analysis kurikulum dan analysis dokumen asesmen. Berikut ini penjelasan dari masing-masing kegiatan analisis:

a. Analisis kajian pustaka

Kajian pustaka adalah kegiatan yang meliputi mencari, membaca, dan menelaah laporan-laporan penelitian dan bahan pustaka yang memuat teori-teori yang relevan dengan penelitian yang akan dilakukan. Kajian pustaka dalam suatu penelitian ilmiah merupakan satu bagian penting dari keseluruhan langkah-langkah metode penelitian.

b. Analisis kurikulum

Berdasarkan analisis kurikulum dan silabus mata pelajaran Biologi kurikulum 2013 didapatkan pokok bahasan yang akan dikembangkan menjadi soal. pokok bahasan materi Biologi yaitu materi virus dan akan dikembangkan pada konten pandemik COVID-19. Pokok bahasan materi yang telah ditentukan dapat digunakan untuk menentukan wacana soal yang akan kita kembangkan, dengan melihat kompetensi dasar dari pokok bahasan yang sudah ditentukan.

c. Analisis dokumen asesmen

Analisis asesmen tes tertulis pada soal tentang keterampilan menjelaskan fenomena secara ilmiah mata pelajaran Biologi di SMA yang merujuk pada *framework* PISA 2018. Data kualitatif diperoleh melalui karakteristik asesmen tes tertulis tentang keterampilan menjelaskan fenomena secara ilmiah dari Menganalisis soal-soal yang terdapat pada lembar kerja peserta didik (LKPD), Ulangan Harian (UH), penilaian tengah semester (PTS), dan penilaian akhir semester (PAS) sehubungan dengan karakteristik prasyaratnya yang dikelompokkan dengan kriteria soal “tidak sesuai”, “mendekati sesuai” dan “sesuai”.

2. Tahap Desain (*Design*)

Tahapan desain meliputi beberapa perencanaan pengembangan instrumen tes diantaranya meliputi beberapa kegiatan sebagai berikut: 1) Mengembangkan *draf framework* yang mengintegrasikan *framework* PISA 2018, indikator pada Kompetensi Dasar (KD biologi Sekolah Menengah Atas (SMA) materi virus dengan indikator keterampilan menjelaskan keterampilan fenomena secara ilmiah pada konten pandemik COVID-19 yang akan dikembangkan; (2) Membuat *draf test blueprint* yang merupakan kelanjutan setelah *draf framework* PISA setelah selesai dibuat yang terdiri dari uraian indikator keterampilan menjelaskan fenomena secara ilmiah, konten materi virus kelas X kurikulum 2013, konten pandemik virus, indikator pencapaian kompetensi, dan nomor soal; (3) Menghasilkan *draf soal* yang merupakan kelanjutan setelah *draf test blueprint* selesai dibuat. Diawali dengan menyusun soal keterampilan menjelaskan

fenomena secara ilmiah yang dibuat peneliti yang memiliki beberapa komponen yang dinilai akan divalidasi oleh ahli, kemudian menentukan 20% soal *linking* untuk selanjutnya membentuk draf soal. Draft soal yang dibuat berupa soal paralel tiga paket, yang masing-masing paket soal terdiri dari 30 butir soal.

3. Tahap *Development* (Pengembangan)

Pengembangan dalam Model ADDIE berisi kegiatan realisasi rancangan produk dalam hal ini adalah instrumen tes. Langkah pengembangan dalam penelitian ini meliputi kegiatan validasi isi dan validasi soal kepada *judgement* untuk *test blue print* dan soal ke ahli/*expert* lalu merevisi soal berdasarkan uji coba. Dalam tahap desain telah disusun kerangka konseptual pengembangan instrumen tes. Dalam tahap pengembangan kerangka konseptual tersebut direalisasikan dalam bentuk produk pengembangan instrumen tes yang siap diimplementasikan sesuai dengan tujuan. Dalam melakukan langkah pengembangan instrumen tes, ada dua tujuan penting yang perlu dicapai antara lain adalah : 1) Memproduksi atau merevisi instrumen tes yang akan digunakan untuk mencapai tujuan pembelajaran yang telah dirumuskan, 2) Memilih instrumen tes terbaik yang akan digunakan untuk mencapai tujuan pembelajaran.

4. Tahap Implementasi (*Implementation*)

Pada tahapan implementasi dalam penelitian ini merupakan tahapan untuk mengimplementasikan rancangan instrumen soal yang telah dikembangkan pada situasi yang nyata dikelas. Selama implementasi, rancangan instrumen tes yang telah dikembangkan diterapkan pada kondisi yang sebenarnya. Materi instrumen tes yang telah dikembangkan dilakukan uji butir soal disekolah pada siswa kelas XI yang dilakukan secara daring dengan menggunakan *Google formulir*.

5. Tahap Evaluasi (*Evaluation*)

Pada tahap ini, melakukan analisis butir soal menggunakan Model *RASCH* pada program aplikasi *Winstep* versi 3.73. untuk mengukur kualitas soal, dan untuk mengetahui profil awal siswa jenjang SMA dalam kompetensi abad ke-21: keterampilan menjelaskan fenomena secara ilmiah pada konten pandemik COVID-19 berdasarkan *proportion correct* dari *item* yang dihasilkan. Kemudian peneliti melakukan revisi terakhir terhadap instrumen soal yang dikembangkan di

lakukan analisis RASCH, menilai kualitas butir soal, menentukan soal yang ditolak, direvisi dan digunakan berdasarkan analisis, serta mendeskripsikan capaian siswa berdasarkan *proportion correct*. dan analisis RASCH untuk menganalisis dan menentukan kualitas butir soal serta menentukan *profil* capaian siswa berdasarkan *proportion correct*.

3.2 Subyek Penelitian

Subyek dalam penelitian ini adalah Sekolah Menengah Atas (SMA) baik negeri maupun swasta di Kota Bandung yang ditentukan dengan teknik *stratified random sampling*, teknik tersebut digunakan untuk populasi yang tidak homogen, teknik ini juga digunakan bila populasi mempunyai anggota atau unsur yang berstrata (tingkat). Teknik ini digunakan karena memungkinkan para peneliti untuk mendapatkan sampel populasi yang paling mewakili seluruh populasi yang diteliti.

Subyek penelitian diambil berdasarkan hasil Ujian Nasional di Kota Bandung dengan kategori tinggi, sedang, dan rendah. Pemilihan sekolah dilakukan dengan nilai UN di Kota Bandung berdasarkan dengan kriteria sesuai kriteria Badan Standar Nasional Pendidikan (BNSP) sebagai berikut Baik: $70 < \text{Rerata UN}$; Sedang: $60 \leq \text{Rerata UN} \leq 70$; Kurang: $\text{Rerata UN} < 60$ yang merujuk kepada laporan penelitian Wulan *et al.*, (2019). Kemudian setelah dilakukan *random* terpilih 8 sekolah yaitu 3 sekolah kategori rata-rata UN baik, 3 sekolah dengan kategor UN cukup, dan 2 sekolah dengan kategori UN kurang. Daftar sekolah yang dijadikan subyek penelitian dapat dilihat pada tabel 3.1.

Subyek penelitian ada dua, pada tahap pertama subyek penelitian adalah sekolah yang terdiri dari enam SMA yang ada di kota Bandung yaitu dua sekolah negeri dan empat sekolah swasta, pada tahap kedua subyek penelitian adalah siswa. Subyek ini merupakan peserta yang akan diujikan pada hasil pengembangan instrumen tes tulis. Sekolah yang menjadi partisipan berasal dari enam Sekolah Menengah Atas (SMA) dengan dua sekolah negeri dan empat sekolah swasta.

Tabel 3.1 Daftar dan kode Sekolah yang dijadikan Subyek penelitian.

NO	NAMA SEKOLAH	STATUS SEKOLAH	NILAI RERATA UN	KATEGORI
1	SMAB01Sw	Swasta	84.15	BAIK
2	SMAB02Sw	Swasta	71,48	

NO	NAMA SEKOLAH	STATUS SEKOLAH	NILAI RERATA UN	KATEGORI
3	SMAB02Sw	Swasta	72.65	CUKUP
4	SMAC01N	Swasta	68,43	
5	SMAC02N	Negeri	64.71	
6	SMAC01Sw	Negeri	57.77	KURANG
7	SMAK02Sw	Swasta	52.02	
8	SMAK02Sw	Swasta	48.43	

Sumber : Puspendik, 2019

Subyek penelitian dari delapan Sekolah Menengah Atas (SMA) Kota Bandung hanya dari lima sekolah saja. Tiga sekolah lain yang berkategori baik tidak memberikan izin penelitian, sedangkan satu sekolah negeri dan setengah partisipan sekolah swasta yang sudah menyatakan kesediaan uji coba soal sedikit tertunda dalam pelaksanaannya. Instrumen soal terdiri dari paket 1,2, dan 3 untuk setiap kelas XI yang dipilih secara acak untuk kelompok A, B, dan C dengan masing-masing paket terdiri dari 30 soal kemudian dibagi menjadi dua menjadi setiap paket terdiri dari 15 soal. siswa yang berpartisipasi dalam uji coba soal berasal dari empat sekolah berjumlah 138 siswa. Subyek uji coba soal dapat dilihat pada tabel 3.2 berikut ini.

Tabel 3.2 Subyek penelitian uji coba soal

Kode Sekolah	Status	Kategori	Jumlah Siswa Kelas XI			Total
			Paket 1	Paket 2	Paket 3	
SMAC02N	Negeri	Cukup	12	9	12	33
SMAC03SW	Swasta	Cukup	8	12	21	41
SMAK01Sw	Swasta	Kurang	15	17	22	54
SMAK02Sw	swasta	Kurang	4	3	3	10
Jumlah Soal						138

3.3 Definisi Operasional

Untuk menghindari kesalahpahaman dalam penelitian yang akan dilakukan, berikut dipaparkan definisi operasional yang memuat tentang variabel penelitian yang terdiri atas variabel independen dan variabel dependen. Variabel independen pada penelitian ini adalah pengembangan instrumen tes tertulis. Sementara itu, variabel dependen penelitian ini meliputi keterampilan menjelaskan fenomena secara ilmiah. Berikut definisi operasional pada penelitian ini:

- a. Instrumen tes tertulis yang dikembangkan dalam penelitian ini berupa pilihan ganda dengan empat opsi jawaban yang terdiri dari 3 paket soal pada konten

pandemik COVID-19. Instrumen tes tertulis ini telah melalui beberapa tahap pengembangan hingga akhirnya layak untuk diimplementasikan pada tahap uji coba di lapangan (138 responden). instrumen tes tertulis ini dikembangkan untuk mengukur keterampilan menjelaskan fenomena secara ilmiah yang dimiliki oleh siswa SMA kelas XI pada konten pandemik COVID-19.

- b. *Test blueprint* disebut juga dengan *table specification* atau kisi – kisi soal yang disusun berdasarkan indikator soal keterampilan menjelaskan fenomena secara ilmiah sebelum instrumen tes tersebut diimplementasikan pada tahap uji coba di lapangan.
- c. Keterampilan menjelaskan fenomena secara ilmiah merupakan salahsatu kompetensi literasi sains yang diukur meliputi indikator : mengingat dan menerapkan pengetahuan ilmiah yang sesuai dengan situasi tertentu, menggunakan gambaran serta model dengan jelas, membuat dan memberikan prediksi yang sesuai, menawarkan hipotesis yang bersifat menjelaskan. dan menjelaskan keterlibatan potensial pengetahuan ilmiah bagi masyarakat. kelima indikator keterampilan menjelaskan fenomena secara ilmiah ini masing-masing diwakili oleh beberapa butir soal.
- d. Konten pandemik COVID-19 adalah konten tentang spesifik virus pandemik COVID-19 yang disebabkan oleh *Coronavirus* jenis baru yang diberi nama SARS-CoV-2 yang merupakan permasalahan yang yang kontekstual yang sedang terjadi di masyarakat, yang nantinya akan di sajikan dalam tes tertulis yang dikembangkan.

3.4 Instrumen Penelitian

Instrumen penelitian merupakan suatu alat yang digunakan untuk mengumpulkan data dan menjawab pertanyaan dalam penelitian. Instrumen penelitian yang digunakan untuk mengumpulkan data adalah tabel analisis soal menjelaskan keterampilan fenomena secara ilmiah, tabel analisis kriteria soal keterampilan menjelaskan fenomena secara ilmiah, *draft framework*, lembar validasi instrumen untuk *test blueprint* dan soal, dan Instrumen Tes Tertulis (*test blueprint* dan Butir Soal).

Tabel 3.3 Kesesuaian pertanyaan penelitian dan instrumen yang digunakan.

No	Pertanyaan penelitian yang diharapkan	Instrumen yang digunakan.
1.	Bagaimana asesmen tes tertulis tentang keterampilan menjelaskan fenomena secara ilmiah pada konten Virus yang terdapat pada soal LKPD, UH, PTS dan PAS yang digunakan di SMA?	Tabel analisis soal keterampilan menjelaskan fenomena secara ilmiah.
2	Bagaimana kriteria asesmen tes tertulis tentang keterampilan menjelaskan fenomena secara ilmiah dari soal LKPD, UH, PTS, dan PAS sehubungan dengan karakteristik prasyaratnya yang digunakan di SMA?	Tabel analisis kriteria soal keterampilan menjelaskan fenomena secara ilmiah.
3.	Bagaimanakah desain instrumen tes tertulis untuk mengukur keterampilan menjelaskan fenomena secara ilmiah siswa SMA pada konten pandemik COVID-19.	<i>draft framework</i> <i>draft blueprint</i> <i>draft soal</i>
4.	Bagaimanakah kualitas <i>test blueprint</i> (kiri-kisi) dan karakteristik instrumen butir soal yang dikembangkan berdasarkan analisis pemodelan RASCH untuk mengukur keterampilan menjelaskan fenomena secara ilmiah siswa SMA kelas XI pada konten pandemik COVID-19.	Lembar validasi tes <i>Blueprint</i> Instrumen tes tertulis (<i>Blueprint</i> dan Butir soal)
5	Bagaimanakah profil awal (<i>pilot profile</i>) keterampilan abad ke-21 siswa dalam keterampilan menjelaskan fenomena secara ilmiah pada konten pandemik COVID-19, pada jenjang SMA berdasarkan <i>proportion correct</i> (PC) dari item yang dihasilkan dari pengerjaan instrumen soal yang di kembangkan?	Instrumen Soal

5. Tabel analisis soal keterampilan menjelaskan fenomena secara ilmiah.

Tabel analisis diperlukan untuk mengumpulkan soal penilaian tentang keterampilan mengevaluasi dan merancang penyelidikan ilmiah yang terdapat pada soal UH, PTS, dan PAS. Pengumpulan data dilakukan melalui teknik dokumentasi soal, dokumentasi soal bertujuan untuk memperoleh data keterampilan menjelaskan fenomena secara ilmiah pada soal pokok bahasan Biologi dengan menggunakan pengkodean pada semua soal. Adapun soal yang digunakan pada penelitian ini sebagai berikut.

Pengkodean Soal : Sekolah (SMA) + Kategori + Nomor Sekolah + Status + LKPD /UH/PTS/PAS + Bentuk soal (Pilihan Ganda/essay) + Nomor Soal

a. Tabel analisis soal keterampilan menjelaskan fenomena secara ilmiah.

Tabel analisis soal keterampilan menjelaskan fenomena secara ilmiah digunakan untuk melihat penggunaan soal dengan indikator keterampilan menjelaskan fenomena secara ilmiah. Tabel 3.3 yang digunakan sebagai berikut.

Tabel 3.4 data hasil analisis pokok uji kesesuaian soal dengan kriteria.

No. Soal	Bentuk soal	Indikator Keterampilan	Hasil analisis			Kesesuaian soal	
			Tidak sesuai	Soal mendekati sesuai	sesuai	Sesuai untuk kode indikator	Kesesuaian dengan indikator pencapaian keterampilan menjelaskan fenomena secara ilmiah
Jumlah Soal							

b. Tabel analisis kriteria soal keterampilan menjelaskan fenomena secara ilmiah.

Tabel analisis soal menggunakan *framework* PISA yang terintegrasi dengan kurikulum 2013. Soal yang dianalisis memiliki soal PTS dan PAS menjadi soal yang tidak sesuai, soal yang mendekati sesuai, dan soal yang sesuai dengan keterampilan menjelaskan fenomena secara ilmiah. Berikut Tabel 3.5 adalah kode dan indikator dari keterampilan menjelaskan fenomena secara ilmiah.

Tabel 3.5 Tabel pengkodean indikator keterampilan menjelaskan fenomena secara ilmiah

Kompetensi ilmiah	Indikator	Kode Indikator
Keterampilan menjelaskan fenomena secara ilmiah.	a. Mengingat dan menerapkan pengetahuan ilmiah yang sesuai dengan situasi tertentu.	MPI01
	b. Mengidentifikasi, menggunakan, dan membuat model dan gambaran sederhana untuk menjelaskan fenomena ilmiah yang dijumpai dalam kehidupan sehari-hari.	MGSM02
	c. Membuat dan memberikan prediksi yang sesuai	MMP03
	d. Menawarkan hipotesis yang bersifat menjelaskan.	MHP04
	e. Menjelaskan keterlibatan potensial pengetahuan ilmiah bagi masyarakat.	MIM05

Tabel yang digunakan untuk melihat persentase dari jumlah soal yang dianalisis menjadi soal yang tidak sesuai, soal yang mendekati sesuai, dan soal yang sesuai dengan indikator keterampilan menjelaskan fenomena secara ilmiah seperti pada Tabel 3.3. Tabel rekap soal yang mendekati sesuai dan sesuai indikator keterampilan menjelaskan fenomena secara ilmiah seperti pada Tabel 3.4 berikut ini

Tabel 3. 6 Data hasil analisis pokok uji di SMA untuk soal mendekati sesuai dan sesuai kriteria.

No	Indikator	Kode Soal	Kode indikator Keterampilan	Jumlah soal sesuai keterampilan	Kode soal	Persentase dari soal keseluruhan
JUMLAH TOTAL						

Tabel berikutnya adalah tabel rekap data soal LKPD, UH, PTS, dan PAS yang digunakan untuk melihat persentase dari jumlah soal yang dianalisis menjadi soal yang tidak sesuai, soal mendekati sesuai, dan soal yang sesuai dengan indikator keterampilan menjelaskan fenomena secara ilmiah. selanjutnya menganalisis soal LKPD, UH, PTS, dan PAS. dengan Tabel 3.5 sebagai berikut.

Tabel 3.7 Rekap data soal LKPD, UH, PTS, dan PAS

No	Kode Soal	Soal

6. *Draf framework*

Draf framework merupakan draf yang disusun sebagai kerangka untuk mengembangkan *blueprint* soal keterampilan menjelaskan fenomena secara ilmiah konten virus dan juga konten pandemik COVID-19. merupakan *framework* indikator-indikator keterampilan menjelaskan keterampilan fenomena secara ilmiah pada konten pandemik COVID-19 yang akan diintegrasikan dengan indikator pada Kompetensi Dasar (KD) Biologi di Sekolah Menengah Atas (SMA) pada KD 3.4 Menganalisis struktur, replikasi, dan peran virus dalam kehidupan dan aspek konteks (personal/lokal/nasional/global).

7. Lembar Validasi *test blueprint*.

Lembar validasi *test blueprint* digunakan sebagai kisi-kisi instrumen untuk

mengungkap keterampilan menjelaskan fenomena secara ilmiah yang akan divalidasi oleh ahli. Hasil penilaian digunakan sebagai bahan perbaikan *test blueprint* yang merupakan pengembangan dari menjelaskan fenomena secara ilmiah pada konten pandemik COVID-19 dengan materi spesifik virus dan pandemik COVID-19. Lembar validasi ini akan digunakan untuk memperoleh data kualitas dari *draft test blueprint* yang dinilai dengan beberapa komponen oleh ahli. Instrumen yang dinilai ini disusun menggunakan skala tanggapan sangat kurang, kurang, baik, dan sangat baik. Adapun lembar validasi dapat dilihat pada lampiran 7 Berisi kriteria penilaian bagi ahli yang memuat sejumlah komponen yang dinilai keterampilan menjelaskan fenomena secara ilmiah dan tanggapan oleh ahlinya.

Komponen yang dinilai adalah materi yang disajikan sesuai dengan tuntutan KD 3.4 (Menganalisis struktur, replikasi dan peran virus dalam kehidupan) kurikulum 2013 kelas X, Materi yang disajikan sesuai dengan konten pandemik COVID-19 yang dipilih, Indikator yang dikembangkan sesuai dengan uraian keterampilan indikator–indikator *framework* keterampilan menjelaskan fenomena secara ilmiah, dan Indikator yang dikembangkan merepresentasikan/ memberi gambaran mengenai soal yang akan dikembangkan.

4. Lembar Validasi Logis Soal

Lembar validasi logis soal berisi tentang indikator dari keterampilan menjelaskan fenomena secara ilmiah. Logis soal ini digunakan untuk menilai indikator logis soal pada keterampilan siswa pada keterampilan menjelaskan fenomena secara ilmiah yang dilakukan secara individu. Memiliki tiga indikator utama yaitu konstruksi, isi butir soal, dan kejelasan bahasa. Instrumen yang dinilai ini disusun menggunakan skala tanggapan sangat kurang, kurang, baik, dan sangat baik. Adapun lembar validasi dapat dilihat pada lampiran 7 yang digunakan dalam penelitian. Komponen yang dinilai pada kisi-kisi logis soal keterampilan menjelaskan fenomena secara ilmiah adalah konstruksi, isi butir soal, dan kejelasan bahasa.

5. Lembar Soal

Lembar validasi soal ini dibuat dan disesuaikan dengan indikator dari keterampilan menjelaskan fenomena secara ilmiah. Soal ini digunakan untuk menilai keterampilan siswa pada keterampilan menjelaskan fenomena secara ilmiah yang dilakukan secara individu. Tes tertulis yang diberikan berupa soal pilihan ganda, dan jawaban singkat yang terdiri dari tiga paket. Hasil penilaian digunakan sebagai bahan perbaikan soal dari menjelaskan fenomena secara ilmiah pada konten pandemik COVID-19. Lembar validasi ini akan digunakan untuk memperoleh data kualitas dari draf soal yang dinilai dengan beberapa komponen oleh ahli yang memiliki lima indikator.

Kisi-kisi menjelaskan fenomena secara ilmiah siswa pada lampiran 8 yang digunakan dalam penelitian terdiri dari indikator mengingat dan menerapkan pengetahuan ilmiah yang sesuai dengan situasi tertentu, mengidentifikasi, menggunakan, dan membuat model dan gambaran sederhana untuk menjelaskan fenomena secara ilmiah yang dijumpai dalam kehidupan sehari-hari, membuat dan memberikan prediksi yang sesuai, menawarkan hipotesis yang bersifat menjelaskan, dan menjelaskan keterlibatan potensial pengetahuan ilmiah bagi masyarakat.

3.5 Prosedur Penelitian

Prosedur penelitian yang akan dilaksanakan terdiri dari tiga tahapan yaitu tahap persiapan, tahap pelaksanaan, dan tahap pengolahan data dan penyusunan laporan. Adapun masing-masing tahapan dipaparkan sebagai berikut:

1. Tahap persiapan penelitian

Adapun yang dilakukan pada tahapan persiapan penelitian yakni meliputi kegiatan seperti:

- a. Melakukan studi kepustakaan atau literatur terkait pengembangan instrumen tes kompetensi abad ke-21 dan keterampilan menjelaskan fenomena secara ilmiah.
- b. Melakukan analisis terhadap potensi, masalah dan kebutuhan.
- c. Menyusun proposal penelitian dengan merumuskan tujuan dari pengembangan instrumen tes kompetensi abad ke-21 dengan menganalisis *framework*

keterampilan menjelaskan fenomena secara ilmiah.

- d. Melakukan studi kurikulum.
 - e. Melakukan seminar proposal penelitian.
 - f. Melakukan perbaikan (*revisi*) proposal penelitian berdasarkan kritik dan saran dari penguji seminar.
2. Tahap pelaksanaan penelitian

Setelah tahap persiapan selesai, kemudian dilanjutkan ke tahap pelaksanaan serangkaian kegiatan dilakukan meliputi:

- a. Menentukan subjek yang akan dijadikan subyek penelitian.
- b. Mengumpulkan soal-soal LKS, UH, PTS, dan PAS dari SMA Kota Bandung dengan kategori sekolah sesuai dengan subyek penelitian.
- c. Menganalisis keberadaan aspek keterampilan menjelaskan fenomena secara ilmiah pada mata pelajaran Biologi konten virus di SMA yang digunakan guru/sekolah pada soal LKPD, UH, PTS, dan PAS.
- d. Menganalisis persentase soal LKS, UH, PTS, dan PAS tentang keterampilan menjelaskan fenomena secara ilmiah pada pokok bahasan Biologi konten virus.
- e. Menganalisis kriteria soal pada LKPD, UH, PTS, dan PAS.
- f. Mendesain *draft framework* PISA keterampilan menjelaskan fenomena secara ilmiah dengan melakukan pemeriksaan validasi oleh ahli lalu melakukan perbaikan desain *framework* PISA keterampilan menjelaskan fenomena secara ilmiah sampai dengan membuat *draft final framework*.
- g. Menyusun *test blueprint* keterampilan menjelaskan fenomena secara ilmiah dengan tahapan mengidentifikasi kompetensi dasar, menentukan lingkup materi, menentukan level kognitif, menyusun kisi-kisi soal, menyusun instrumen soal. Kemudian melakukan pemeriksaan validasi oleh ahli (3 ahli) lalu melakukan perbaikan desain *test blueprint framework* PISA keterampilan menjelaskan fenomena secara ilmiah sampai dengan membuat *draft final test blueprint*.
- h. Menyusun soal keterampilan menjelaskan fenomena secara ilmiah dengan

melakukan pemeriksaan validasi oleh ahli lalu melakukan perbaikan desain soal *framework* PISA keterampilan menjelaskan fenomena secara ilmiah sampai dengan membuat *draft final* soal.

- i. Menentukan soal *linking* 20% dari jumlah soal yang diperoleh dari hasil *jugjement* oleh ahli.
 - j. Menentukan subyek penelitian.
 - k. Membuat surat izin penelitian untuk melakukan studi pendahuluan.
 - l. Melakukan observasi ke sekolah yang dijadikan lokasi penelitian dan melakukan komunikasi dengan pihak sekolah yang dijadikan lokasi penelitian terkait teknis pengujian atau uji coba butir soal ke lapangan.
 - m. Menyampaikan informasi kepada siswa terkait teknis pengujian atau uji coba butir soal yang akan dilakukan pada siswa SMA dengan konten virus pada konten pandemik COVID-19 yang sudah diajarkan pada waktu kelas X.
 - n. Melakukan tes uji coba butir soal kepada subyek penelitian dengan menggunakan *google form* untuk menganalisis keterampilan menjelaskan fenomena secara ilmiah
3. Tahap pengolahan data dan penyusunan laporan penelitian

Pada tahap akhir dari prosedur penelitian kegiatan yang dilakukan meliputi hal-hal berikut seperti:

- a. Melakukan pengolahan data hasil uji keterampilan menjelaskan fenomena secara ilmiah siswa dalam bentuk tabel, grafik dan deskripsi secara statistik *deskriptif*.
- b. Melakukan analisis data hasil penelitian menggunakan model RASCH dan dibantu oleh *software Winstep versi 3.73*.
- c. Melakukan interpretasi terhadap data hasil penelitian dan menyusun pembahasan penelitian.
- d. Membuat penarikan kesimpulan penelitian berdasarkan data temuan dan pembahasan penelitian yang mengacu pada pertanyaan penelitian.
- e. Menyusun laporan penelitian dalam bentuk tesis.

3.6 Teknik Pengumpulan data.

Penelitian ini menggunakan data yang diperoleh oleh instrumen yang telah disiapkan. Instrumen penelitian yang digunakan sebagai alat pengumpul data berbentuk tes. teknik pengumpulan data untuk mengetahui keterampilan menjelaskan fenomena secara ilmiah yang diuji pada siswa kelas XI SMA dengan materi virus yang sudah diajarkan di kelas X SMA. Rincian teknik pengumpulan data terdapat pada tabel 3.6 berikut ini.

Tabel 3.8 Teknik Pengumpulan Data.

No	Instrumen	Jenis Data
1	Tabel analisis soal keterampilan menjelaskan fenomena secara ilmiah	Data hasil analisis oleh peneliti dengan menganalisis keadaan/penggunaan tes/soa-soal tentang keterampilan menjelaskan fenomena secara ilmiah.
2.	Tabel analisis kriteria soal keterampilan menjelaskan fenomena secara ilmiah	Data hasil analisis oleh peneliti dengan menganalisis persentase dari pokok uji tes kaitannya dengan keterampilan menjelaskan fenomena secara ilmiah pada mata pelajaran Biologi materi virus
3.	Draf <i>framework</i>	Data indikator –indikator keterampilan menjelaskan fenomena secara ilmiah konten pandemik COVID-19 yang akan menjadi <i>test blueprint</i> .
4.	Lembar validasi <i>test blueprint</i>	Data hasil penilaian oleh ahli dari penyusunan <i>test blueprint</i> keterampilan menjelaskan fenomena secara ilmiah .
5.	Lembar validasi logis soal	Data hasil penelitian oleh ahli dari desain soal keterampilan menjelaskan fenomena secara ilmiah.
6.	Soal tes tertulis	Nilai kemampuan menjawab soal atau skor dalam soal keterampilan menjelaskan fenomena secara ilmiah.

Lembar validasi *blueprint* yang merupakan data hasil penilaian oleh ahli dari penyusunan *test blueprint* keterampilan menjelaskan fenomena secara ilmiah sehubungan dengan karakteristik prasyaratnya yang divalidasi oleh ahli, lembar validasi logis soal yang merupakan nilai kemampuan menjawab soal keterampilan menjelaskan fenomena secara ilmiah.

Asesmen berupa *test blueprint* yang menjembatani keterampilan menjelaskan fenomena secara ilmiah dengan kurikulum Biologi di SMA (lampiran 1) yang disempurnakan berdasarkan masukan/pendapat oleh ahli, selain itu lembar pemeriksaan perangkat asesmen berupa tiga perangkat soal tes tertulis paralel tes uraian dan tes objektif (Pilihan Ganda) tentang keterampilan menjelaskan fenomena secara ilmiah yang disempurnakan berdasarkan masukan/pendapat oleh ahli yang merupakan nilai kemampuan menjawab soal keterampilan menjelaskan fenomena

secara ilmiah. Dan juga angket siswa yang merupakan Deskripsi dan presentase respon siswa terhadap tanggapan pemberian soal, tanggapan keterampilan menjelaskan fenomena secara ilmiah, dan tanggapan terhadap kelebihan dan kekurangan soal.

3.7 Analisis Data

Analisis data yang dilakukan berdasarkan sifat data. Seluruh data yang diperoleh kemudian akan dianalisis secara menyeluruh. Analisis data menggunakan analisis kuantitatif dan kualitatif, data diolah sebagai persentase. Analisis data secara kuantitatif dilakukan melalui statistik deskriptif. data diproses dalam bentuk persentase dan digunakan untuk keterampilan menjelaskan fenomena secara ilmiah pada konten pandemik COVID-19 pada dokumen yang digunakan disekolah. Melakukan analisis data kualitatif dengan mengorganisir dan mencari pola standar yang terkait dengan masalah pada LKPD, UH, PTS, dan PAS. analisis validitas *test blueprint*, dan validasi soal keterampilan menjelaskan fenomena secara ilmiah berdasarkan validasi para ahli. Analisis butir soal dengan menggunakan Model *RASCH* dengan bantuan program aplikasi *Winstep* versi 3.73 untuk mengetahui *pilot profile* siswa dan kemampuan siswa dalam *proportion correct* dari *item* yang dihasilkan. Adapun rincian pengolahan data penilaian ini adalah sebagai berikut

Analisis data dilakukan berdasarkan sifat data yang ditemukan. Data dianalisis secara kuantitatif presentase dengan statistik kualitatif. Diperoleh data presentase untuk data analisis asesmen tes tertulis pada soal tentang keterampilan menjelaskan fenomena secara ilmiah mata pelajaran Biologi di SMA yang merujuk pada *framework* PISA 2018. Data kualitatif diperoleh melalui karakteristik asesmen tes tertulis tentang keterampilan menjelaskan fenomena secara ilmiah dari soal sehubungan dengan karakteristik prasyaratnya yang dikelompokkan dengan kriteria soal “tidak sesuai”, “mendekati sesuai” dan “sesuai”.

Soal “tidak sesuai” yaitu soal yang tidak memenuhi/tidak memuat karakteristik prasyarat soal keterampilan menjelaskan fenomena secara ilmiah, soal “mendekati sesuai” adalah soal yang memuat sebagian karakteristik prasyarat

keterampilan menjelaskan fenomena secara ilmiah, sedangkan kriteria soal “sesuai” merupakan soal yang memenuhi karakteristik prasyarat soal menjelaskan fenomena secara ilmiah menjelaskan fenomena secara ilmiah. Selain itu penyempurnaan perangkat asesmen berupa perangkat asesmen yang terdiri dari model *test blueprint* dan tiga perangkat soal tes tertulis paralel, didasarkan terhadap masukan dan perbaikan oleh pakar/ahli.

Data hasil *judgement* ahli dianalisis secara kualitatif dan kuantitatif. pengujian empirik butir soal dilakukan dengan analisis menggunakan aplikasi pemodelan RASCH yang baik digunakan pada asesmen pendidikan. Pemodelan RASCH dapat mentransformasi data, dari data skor mentah menjadi data interval yang sama. Proses transformasi data ini mampu menghasilkan skala pengukuran yang linear, presisi dan mempunyai satuan (Sumintono, 2016). selain itu, pemodelan RASCH juga bisa digunakan menganalisis kualitas instrumen tes, menentukan tingkat abilitas siswa dan menentukan tingkat kesulitan tes sampai pada mendeteksi miskonsepsi, atau bias pada instrumen tes. Analisis pemodelan RASCH ini menggunakan *software Winstep* dalam pelaksanaannya.

Analisis data dilakukan berdasarkan sifat data. Seluruh data yang diperoleh kemudian akan dianalisis secara menyeluruh. Kemudian analisis butir soal dilakukan dengan menggunakan Model *RASCH* yang didukung oleh aplikasi *Winstep* versi 3.73 untuk melihat kualitas soal, profil awal siswa, dan *proportion correct* dari item yang dihasilkan. Adapun rincian pengolahan data penelitian ini adalah sebagai berikut.

1. Analisis kriteria soal keterampilan menjelaskan fenomena secara ilmiah.

Mengolah hasil analisis data pada soal LKPD, UH, PTS dan PAS maka digunakan deskriptif kualitatif dengan menghitung persentase soal yang sesuai indikator tentang keterampilan mengevaluasi dan merancang penyelidikan ilmiah pada pokok bahasan Biologi di SMA. Pada perhitungan persentase yang dihasilkan dapat melihat perbandingan besar kecilnya frekuensi pada setiap soal LKPD, UH, PTS dan PAS. Persentase yang diperoleh dengan membandingkan soal sesuai dengan jumlah soal yang terkait dengan angka 100%, rumusnya adalah sebagai berikut:

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

(Arikunto, 2006)

Keterangan :

P = Presentase hasil analisis.

F = Frekuensi soal yang sesuai

N = Jumlah soal

100% = Bilangan tetap

Kriteria penilaian untuk memudahkan dalam penentuan kriteria penilaian, maka dibuatlah pedoman penilaian yang yang diinterpretasi dan dikategorikan dengan menggunakan kriteria yang dikemukakan oleh koentjaraningrat (1990) yang disajikan sebagai berikut:

Tabel 3.7 Pedoman penilaian

Persentase	Tafsiran
0%	Tidak ada
1% - 25%	Sebagian kecil
26% - 49%	Hampir separuhnya
50%	Separuhnya
51% - 75%	Sebagian besar
76% - 99%	Hampir seluruhnya
100%	Seluruhnya

Sumber : koentjaraningrat, 1990

2. Analisis butir soal Keterampilan menjelaskan fenomena secara ilmiah.

Teknik analisis data menggunakan analisis model *RASCH*. Analisis model *RASCH* merupakan teori analisis modern yang dikembangkan oleh B. Sumintono dan W. Widhiarso baru-baru ini. Analisis model *RASCH* memiliki beberapa keunggulan, salah satunya adalah mampu mengidentifikasi kemampuan instrumen untuk mengukur suatu variabel secara akurat. (Hartanti *et al.*, 2020)

Model *RASCH* digunakan dalam penelitian ini untuk menganalisis data yang dikumpulkan dari keterampilan menjelaskan fenomena secara ilmiah. Data yang

terkumpul dimasukkan ke dalam *microsoft excel* dan kemudian diimpor ke perangkat lunak *Winsteps* versi 3.73. Untuk memvalidasi, validitas konten dan reliabilitas konsistensi internal dari keterampilan menjelaskan fenomena secara ilmiah ditentukan. Keandalan item, pemisahan item, keandalan personel, dan pemisahan person juga diperiksa. Kesesuaian kualitas item diperiksa menggunakan beberapa kriteria: *Outfit mean-square* (MNSQ), *Outfit z-standardized* (ZSTD), dan *Point-measure korelasi* (Pt-Measure Corr) untuk setiap *item*.

Peta Wright ditampilkan untuk mendemonstrasikan secara komprehensif tentang kesulitan item dan kemampuan siswa. Selanjutnya, analisis *logit value person* (LVP) dilakukan untuk mengidentifikasi kemampuan keterampilan menjelaskan fenomena secara ilmiah siswa SMA. Analisis *person-fit* juga diuji dengan menggunakan tiga kriteria, yaitu *Outfit mean-square* (MNSQ), *Outfit z-standardized* (ZSTD), dan *Point-measure korelasi* (Pt-Measure Corr). Untuk mengeksplorasi keberadaan item tes yang berfungsi berbeda untuk jenis kelamin dan tingkat kelas siswa, analisis *Differential Item Functioning* (DIF) juga dilakukan dengan menggunakan *Winsteps*. Data kemampuan menjawab soal keterampilan menjelaskan fenomena secara ilmiah menggunakan *Item Response Theory* (IRT) model RASCH dengan bantuan *software* dari *Winstep*.

- a. Tingkat kesesuaian butir soal (*Item Fit*) dan tingkat kesesuaian individu (*Person Fit*).

Penetapan *fit* tiap butir soal menggunakan pengujian berdasarkan *fit* menurut model RASCH bila memiliki nilai *INFIT* MNSQ < 0,77 dan > 1,30. Untuk melihat validitas butir soal, digunakan model RASCH. Langkah selanjutnya adalah mengukur validitas item dengan uji *item fit* untuk memastikan bahwa semua item sudah sesuai dengan model RASCH. Proses tersebut bertujuan untuk mengetahui apakah item tes dapat mengukur aspek yang ingin diukur, atau menguji validitas (Linacre, 2012; Sumintono & Widhiarso, 2014). Dan Menurut Boone *et al.*, (2014) serta Bond dan Fox (2015) dalam Sumintono dan Widhiarso (2013), nilai *outfit means-square* (MNSQ), *outfitz standard* (ZSTD), dan *point measure correlation* (Pt Mean Corr) adalah kriteria yang digunakan untuk melihat tingkat kesesuaian butir (*item fit*). Seperti halnya dalam pemeriksaan butir soal dalam hal kesesuaian

dengan model, pada pemeriksaan *person* pun digunakan kriteria yang sama (Boone *et al.*, 2014 dalam Sumintono dan Widhiarso, 2013). Pada Model RASCH, untuk melihat kualitas butir soal dari aspek validitas adalah jika memenuhi beberapa kriteria berikut ini (Sumintono & Widhiarso, 2015).

- 1) Nilai *Outfit Mean Square* (MNSQ) yang diterima adalah : $0,5 < \text{MNSQ} < 1,5$
- 2) Nilai *Outfit Z-Standard* (ZSTD) yang diterima adalah : $-2,0 < \text{ZSTD} < +2,0$
- 3) Nilai *Point Measure Correlation* (Pt Mean Corr) yang diterima adalah: $0,4 < \text{Pt Measure Corr} < 0,85$, pada prinsipnya sama dengan korelasi *point biserial* pada teori *test* klasik mengklasifikasikan nilai tersebut menjadi

Tabel 3.9 Klasifikasi Nilai

Sangat bagus	(>0,40)
Bagus	(0,30-0,39)
Cukup	(0,20-0,29)
Tidak mampu mendiskriminasi	(0,00-0,19)
Memerlukan pemeriksaan terhadap butir soal	(<0,00)

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	S.E.	MNSQ	ZSTD	Pt-MCORR	RE	TE	FACT	MATCH	Item
3	3	20	3.46	1.1211.39	1.97	51	34	31	95.0	94.91	53	
2	4	20	1.52	631.86	1.83	91	50	43	85.0	83.21	52	
1	6	20	.64	551.82	1.69	51	56	45	80.0	74.91	53	
10	6	20	.84	551.76	1.57	71	61	45	80.0	74.91	53	
7	7	20	.55	531.87	1.71	51	56	46	75.0	71.21	57	
6	10	20	-.25	511.09	1.96	81	43	47	65.0	70.91	56	
5	13	20	-1.06	541.11	1.21	61	39	47	70.0	74.81	55	
8	14	20	-1.35	561.55	1.42	1.01	74	46	85.0	76.91	58	
9	16	20	-2.05	631.12	1.28	1.41	72	42	85.0	82.41	59	
4	17	20	-2.49	7011.60	1.13	90	40	39	80.0	86.31	54	
MEAN	9.4	20.0	.00	6311.02	11.05	11		30.0	79.01			
S.D.	5.1	0	1.71	421.29				8.1	7.21			

Gambar 3. 2 *Item fit* Sumber: Sumintono & Widhiarso, (2015)

b. Tingkat Kesukaran (*Measure*)

Tingkat kesukaran soal adalah besaran yang digunakan untuk menyatakan apakah suatu soal termasuk kedalam kategori mudah, sedang atau sukar. Soal yang baik adalah soal yang tidak terlalu mudah atau tidak terlalu sukar. Pada penelitian ini, uji tingkat kesukaran dilakukan menggunakan RASCH model dengan bantuan program *winsteps*. Pada program tersebut, disajikan urutan butir soal dari yang tersulit sampai pada butir soal yang termudah. Pada model RASCH tingkat kesulitan butir soal ditinjau menurut (Sumintono dan Widhiarso, 2015) yang dapat dikelompokkan menjadi empat kategori berdasarkan nilai *measure* yang diperoleh

pada analisis model RASCH seperti berikut.

- 1) Nilai *measure* < -1 = item sangat mudah
- 2) Nilai *measure* -1 s.d. 0 = item mudah
- 3) Nilai *measure* 0 s.d. 1 = item sulit
- 4) Nilai *measure* > 1 = item sangat sulit

Item STATISTICS: MEASURE ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT	OUTFIT	[PT-MEASURE]	[E-GACT]	[MATCH]	Item			
3	1	20	1.46	1.12	1.39	.71	.97	.51	.14	.33	95.0	94.9	53
2	4	20	1.52	.63	.86	-.31	.61	.01	.50	-.43	85.0	83.2	52
1	6	20	.84	.55	.82	-.01	.69	-.51	.56	-.65	80.0	74.9	51
10	6	20	.84	.55	.76	-1.01	.57	-.71	.63	-.45	80.0	74.9	51B
7	7	20	.55	.53	.87	-.61	.71	-.51	.56	-.46	75.0	73.2	57
6	10	20	-.25	.53	.99	.51	.96	.01	.43	-.47	65.0	70.9	56
5	13	20	-1.06	.54	1.33	.51	1.21	.61	.39	-.47	70.0	74.8	59
8	14	20	-1.35	.56	.55	-1.01	.42	-1.01	.74	-.46	85.0	76.9	58
9	16	20	-2.05	.63	1.12	-.61	.20	1.41	.22	-.42	85.0	82.4	59
4	17	20	-2.49	.70	1.60	1.31	1.90	1.01	.00	-.39	80.0	86.3	54
MEAN	9.4	20.0	.00							1	30.0	79.0	1
S.D.	5.1	.0	1.73							1	8.1	7.2	1

Annotations in the image:

- Nomor urut soal (points to Entry Number)
- Nilai logit untuk masing-masing butir diurutkan dari tertinggi ke terendah (points to Measure column)
- Total skor menyatakan jumlah jawaban benar untuk tiap soal (points to Total Score)
- Nilai rata-rata item selalu 0,00 logit. Nilai deviasi standar logit item berguna untuk identifikasi kelompok item (separation) (points to MEAN and S.D. rows)
- Label butir soal (points to Item column)

Gambar 3. 3 Item measure Sumber: Sumintono & Widhiarso, (2015)

c. Analisis instrumen

Dalam pemodelan RASCH, selain dapat melakukan analisis sampai ke tingkat butir soal dan tingkat individu, analisis di tingkat instrumen pun bisa dilakukan.

- 1) *Person measure*
 - 2) Nilai *alpha Cronbach* (mengukur reliabilitas, yaitu interaksi antara person dan butir-butir soal secara keseluruhan)
 - 3) Nilai *person reliability* dan *item reliability*
 - 4) Daya pembeda untuk *item* dan *person (separation)*
 - 5) Peta *person-item* (peta wright)
- 1) *Person measure*.

Person Measure menunjukkan rata-rata nilai seluruh siswa dalam mengerjakan butir-butir soal yang diberikan. Nilai rata-rata yang lebih kecil dari nilai *logit* 0.0 menunjukkan kecenderungan nilai abilitas siswa yang lebih kecil dari

pada kesulitan soal. Dapat dilihat pada tabel *Summary statistik*, hasil output yaitu *summary of measured* tersebut memberikan info secara keseluruhan tentang kualitas pola respon siswa secara keseluruhan, kualitas instrumen yang digunakan, maupun interaksi antara *person* dengan butir.

2) Nilai *alpha Cronbach*

Kualitas butir soal dari aspek reliabilitas butir soal pada pendekatan teori tes klasik menggunakan rumus *Alpha Cronbach* (KR-20). Butir soal dianggap reliabel jika memenuhi kriteria koefisien korelasi reliabilitas instrumen. Pengujian reliabilitas instrumen menggunakan model *RASCH*, selanjutnya guna mempermudah penafsiran hasil uji reliabilitas, peneliti menggunakan patokan nilai reliabilitas item untuk melihat reliabilitas tes sudah *fit* dengan model *RASCH*. Mengukur *reliabilitas*, yaitu interaksi antara *person* dan butir-butir soal secara keseluruhan. Sumintono dan Widhiarso (2013) untuk melihat nilai reliabilitas item menurut model *RASCH*, menggunakan kategori dengan kriteria skor sebagai berikut:

- a) $0,00 \leq r < 0,50$ reliabilitas buruk
- b) $0,50 \leq r < 0,60$ reliabilitas jelek
- c) $0,60 \leq r < 0,70$ reliabilitas cukup
- d) $0,70 \leq r < 0,80$ reliabilitas bagus
- e) $0,80 \leq r < 1,00$ reliabilitas bagus sekali

d. Daya pembeda untuk *item* dan *person* (*Separation*)

Menurut Nasyidah *et al* (2020), menyatakan Perhitungan daya pembeda digunakan untuk mengukur sejauh mana suatu butir soal mampu membedakan peserta didik yang sudah menguasai kompetensi dengan peserta didik yang belum/kurang menguasai kompetensi berdasarkan kriteria tertentu. Semakin tinggi indeks diskriminasi daya pembeda maka termasuk soal yang sangat baik, ini berarti bahwa soal tersebut dapat membedakan antara peserta didik yang sudah menguasai kompetensi dengan peserta didik yang belum/kurang menguasai kompetensi. Untuk menghitung daya pembeda setiap butir soal dapat menggunakan analisis *RASCH* Model dengan aplikasi WINSTEP 3.7.3 dengan

output tabel 10 *Item Fit Order* yang dianalisis dengan analisis *RASCH*.

Pada analisis melalui model *RASCH*, untuk dapat membedakan kemampuan siswa yang tinggi dan rendah digunakan analisis pada tingkat abilitas individu. Selain itu dapat pula dilihat dengan cara mengidentifikasi kelompok responden berdasarkan indeks separasi responden. Menurut (Sumintono dan Widhiarso, 2015). Daya pembeda butir soal merupakan kemampuan soal untuk membedakan peserta didik yang mampu menjawab soal atau peserta didik yang memiliki tingkat kemampuan yang tinggi dengan peserta didik yang memiliki kemampuan menjawab soal yang rendah. (Erfan *et al.*, 2020). semakin besar nilai separasi maka kualitas instrumen dalam hal keseluruhan responden dan butir makin bagus, karena dapat mengidentifikasi kelompok responden dan butir. Persamaan lain untuk melihat pengelompokkan secara lebih teliti digunakan persamaan pemisahan strata (H):

$$H = \frac{(4 \times SEPARATION) + 1}{3}$$

3

Keterangan:

H : Nilai strata (Pengelompokan)

Separation : Nilai *separation* untuk responden yang dihasilkan

1. Peta *Wright*

Peta *person-item* (*peta wright*) Peta *Wright* menyajikan distribusi kesulitan soal dan kemampuan siswa pada skala logit yang sama (Bond dan Fox, 2015). Kesulitan item ditempatkan di sisi kiri *peta Wright*, sedangkan kemampuan siswa terletak di sisi kanan *peta Wright*. Semakin tinggi logit berarti semakin sulit butir soal dan siswa dengan kemampuan lebih tinggi. Item yang lebih mudah dan siswa berkemampuan rendah diwakili oleh logit yang lebih rendah (Boone *et al.* 2014). tingkat kesulitan dalam pembagiannya, konsistensi responden dalam menjawab, reliabilitas butir soal, dan reliabilitas responden dalam menjawab soal.

Selain itu, analisis ini dapat digunakan untuk mengetahui kemampuan responden dalam menjawab pertanyaan yang telah disajikan. (Hartanti *et al.*, 2020)

Peta Wright adalah mengukur konsistensi tingkat kesukaran soal dan tes kemampuan siswa yang dibangun. Semakin tinggi tingkat kesukaran soal, semakin tinggi pula tingkat kemampuan yang dihasilkan. Informasi *Peta Wright: Item-Peta* Orang ditampilkan pada *Peta Wright* sebelumnya menunjukkan bahwa semua item instrumen mencakup hampir semua kemampuan siswa. Peta tersebut menghasilkan varians dari siswa dengan kemampuan sangat tinggi (> 3.0 logit), hingga siswa dengan kemampuan sangat rendah (< -2.0 logit) juga. Selain itu, disparitas (tidak ada item yang sesuai dengan kemampuan siswa) diamati pada interval -3.0 logit hingga -0.5 logit dan pada interval $+1.0$ logit hingga $+3.7$ logit. Ini menandakan bahwa informasi yang dihasilkan dalam rentang interval agak terbatas dan membutuhkan elaborasi lebih lanjut. Sedangkan tingkat kesulitan item sebagian besar berada pada interval -1.0 logit hingga $+1.0$ logit; Selain itu, item cenderung terjadi pada tingkat kesulitan yang sama. Item Q14 / PS / LG merupakan item tersulit dengan logit $+0.97$, sedangkan item Q17 / Pmo / LG merupakan item termudah dengan logit -0.71 . (Laliyo *et al.*, 2020)

Peta ini biasa disebut *Wright Map* yang tidak lain adalah peta *person-item* secara komprehensif. Peta *Wright* sebelah kiri menggambarkan abilitas siswa. Peta *Wright* sebelah kanan menjelaskan sebaran nilai *logit* butir. Peta *person-item* yang dihasilkan dari Ministep untuk data politomi juga berbentuk text (*.txt), yang bisa dilakukan modifikasi sesuai keperluan (Sumintono dan Widhiarso, 2013).

e. Nilai *Person Reliability* dan *Item Reliability*.

Nilai *Person Reliability* dan *Item Reliability* digunakan untuk mencari Reliabilitas. *Person Reliability* digunakan untuk mengetahui konsistensi jawaban siswa dan *Item Reliability* digunakan untuk mengetahui konsistensi butir soal secara keseluruhan (Sumintono dan Widhiarso, 2013), dengan kriteria sebagai berikut :

< 0.67 = Lemah

$0.67 - 0.80$ = Cukup

$0.80 - 0.90$ = Bagus

0.91 – 0.94 = Bagus Sekali

>0.94 = Istimewa

f. Analisis pengecoh (*distractor analysis*)

Analisis distraktor adalah pendekatan untuk memeriksa pola respons pada item *Multiple Choice* di mana pilihan jawaban siswa diperiksa bersama dengan ukuran pencapaian. Ketika *item Multiple Choice* dibangun sedemikian rupa sehingga pengganggu terkait dengan aspek penting dari pemahaman siswa tentang konsep tertentu, pendekatan analitik ini memberikan wawasan yang berharga tentang jenis konsepsi yang benar dan alternatif yang ada dalam subyek siswa, serta wawasan ke dalam hierarki perkembangan kesalahpahaman yang termasuk dalam pengalih item. Informasi ini dapat digunakan untuk menginformasikan keputusan instruksional untuk menargetkan area pemahaman siswa tertentu tanpa menggunakan prosedur penilaian kinerja yang lebih memakan waktu. (Wind *et al.*, 2019). Pengecoh dianggap tidak bekerja dengan baik jika peserta tes dengan kemampuan rendah (logit rendah) bisa mendapat jawaban benar.

Analisis distraktor sebagai alat diagnostik. Terlepas dari keterbatasan yang umumnya dikaitkan dengan penilaian berdasarkan item pilihan ganda (MC), banyak praktisi dan peneliti menggunakan penilaian dengan item MC dalam pengaturan pendidikan. Salah satu alasannya adalah bahwa item MC mudah diterapkan dan dinilai secara objektif, sehingga memungkinkan pengguna pengujian untuk menilai berbagai konten dalam satu administrasi penilaian. Selain itu, dimungkinkan untuk membangun item MC sedemikian rupa sehingga memberikan pengguna tes informasi diagnostik yang berguna tentang pemahaman siswa (Gierl *et al.* 2017; Thissen *et al.* 1989). Gierl dkk. (2017) melakukan tinjauan literatur yang komprehensif tentang penggunaan distraktor item MC dalam pendidikan, dan menyimpulkan bahwa, ketika dibangun dengan tujuan diagnostik dalam pikiran dan dianalisis dengan tepat, pengalih item MC memberikan informasi diagnostik yang berguna (Wind *et al.*, 2019).