

## BAB II

### KAJIAN PUSTAKA

#### A. Penelitian Terdahulu

Pada kajian riset terdahulu, peneliti mengadakan telaah terhadap riset yang relevan dengan penelitian ini, sehingga nampaklah perbedaan yang jelas dan nyata. Oleh sebab itu, pada kajian ini akan dipaparkan sejumlah riset yang senada dengan penelitian ini, antara lain:

Penelitian yang dilakukan Margrethe H. Møller tahun 2013 yang berjudul “*Usability Testing of User Manuals*” menjelaskan bahwa ketika kita hendak menggunakan *software* atau *web* pasti kita memerlukan buku manual atau buku panduan. Namun, muncul masalah ketika menggunakan *manual book* tersebut dimana produk tersebut sulit dipahami oleh orang dengan usia lanjut. Berangkat dari hal ini peneliti mengembangkan produk dengan maksud dan tujuan agar produknya nanti bisa digunakan bukan hanya oleh kalangan muda melainkan juga usia lanjut. Produk yang dihasilkan adalah *User Manuals for older adults* (Møller, 2013). Pengembangan produk awal menjadi produk baru yang lebih adaptif terhadap kondisi tertentu, jenis dan bahan yang dikembangkan merupakan hal yang menjadi persamaan antara penelitian yang dilakukan oleh Moller dengan penelitian ini. Sedangkan perbedaanya terletak pada teknik dan cara pengembangan.

Penelitian yang dilakukan oleh Fatikhatun Najikhah dan Christina Ismaniati tahun 2019 yang berjudul “Pengembangan Buku Manual sebagai

Sumber Belajar Mandiri untuk Pekerja Migran Indonesia dengan Destinasi Malaysia” menjelaskan bahwa berawal dari *training* yang diselenggarakan Balai Latihan Kerja Luar Negeri (BLKLN) lebih fokus pada keterampilan dan bahasa tanpa dibekali dengan pengetahuan umum mengenai Negara tujuan. Berangkat dari sini peneliti membuat sebuah produk Buku Panduan Tenaga Kerja Indonesia (TKI) di Malaysia (Najikhah & Ismaniati, 2019). Persamaan dengan penelitian ini adalah sama-sama menghasilkan sebuah produk cetak. Sedangkan perbedaan dengan penelitian ini adalah pendekatan pengembangan model yang digunakan adalah *4D (Devine, Design, Develop, and Disseminate)*, sedangkan penelitian ini menggunakan model *ADDIE*.

Penelitian yang dilakukan oleh N.H. Mohd Amin et.al tahun 2019 yang berjudul “*Development of E-Help Manual Using Graphical User Interface (GUI) for Battery Management System (BMS)*” menjelaskan bahwa *Battery Management System (BMS)* pada kendaraan listrik sering terjadi *overcharge* sehingga menyebabkan baterai pada kendaraan listrik menjadi tidak terawat dengan baik. Solusi yang ditawarkan dengan menambahkan *GUI (Graphical User Interface)* pada *BMS*. Akan tetapi, diperlukan modul *E-Help Manual GUI* agar kesalahan di awal tidak terulang kembali. Penelitian ini menghasilkan *E-Help Manual Using Graphical User Interface (GUI) for Battery Management System (BMS)* (Amin et al., 2019). Persamaan dengan penelitian ini adalah jenis dan bahan yang

dikembangkan. Teknik, cara, dan jenis produk yang dihasilkan merupakan perbedaan antara penelitian Amin dengan penelitian ini.

Penelitian yang dilakukan oleh Ade Kurniawan dan Masjudin tahun 2019 yang berjudul “Pengembangan Buku Ajar *Microteaching* Berbasis Praktik untuk Meningkatkan Keterampilan Mengajar Calon Guru” menjelaskan bahwa tidak tersedianya buku ajar dalam perkuliahan *microteaching* di IKIP Mataram menjadi kendala tersendiri dalam pelaksanaan perkuliahan *microteaching*. Akibatnya, kegiatan perkuliahan kurang berjalan dengan baik karena tidak tersedianya panduan perkuliahan baik bagi dosen maupun mahasiswa (A. Kurniawan & Masjudin, 2017). Hal ini menjadi dasar pemikiran peneliti untuk membuat produk bahan ajar. Penelitian ini menghasilkan Buku ajar *microteaching* berbasis praktik. Teknik dan metode dalam pengembangan produk merupakan hal yang menjadikan persamaan antara penelitian ini dengan penelitian Kurniawan. Sedangkan hal yang berbeda dengan penelitian ini adalah jenis bahan yang dikembangkan, bidang pengembangan, dan produk hasil pengembangan.

Penelitian yang dilakukan oleh Cut Misni Mulasiwi, Siswandari dan Djoko Santosa tahun 2016 yang berjudul “Pengembangan Buku Panduan Praktik Laboratorium Bank Mini Dalam Rangka Meningkatkan Keterampilan Pencatatan Transaksi Keuangan Pada Program Keahlian Akuntansi” menjelaskan bahwa peserta didik program jurusan keahlian akuntansi belum memiliki buku panduan khusus dalam melakukan kegiatan praktik laboratorium bank mini (Mulasiwi et al., 2016). Produk yang

dihasilkan adalah Buku Panduan Praktik Laboratorium Bank Mini. produk pengembangan dan metode yang digunakan menjadi persamaan antara penelitian ini dengan penelitian Mulasiwi. Sedangkan perbedaannya terletak pada jenis bahan serta tahap pengembangan.

Penelitian yang dilakukan oleh Waluyo & Parmin tahun 2014 yang berjudul “Pengembangan Panduan Praktikum IPA Terpadu Berbasis Inkuiri Terbimbing Tema Fotosintesis Untuk Menumbuhkan Keterampilan Kerja Ilmiah Siswa SMP” menjelaskan bahwa untuk setiap tema belum ada buku panduan praktikum, dengan harapan apabila buku panduan praktikum untuk setiap tema sudah ada (Waluyo & Parmin, 2014). Panduan Praktikum IPA Terpadu Berbasis Inkuiri Terbimbing Tema Fotosintesis merupakan produk hasil pengembangan. Sedangkan perbedaannya terletak pada jenis bahan dan produk yang dihasilkan.

Penelitian yang dilakukan oleh Zulaiha, Hartono, A. Rachman Ibrahim tahun 2014 yang berjudul “Pengembangan Buku Panduan Praktikum Kimia Hidrokarbon Berbasis Keterampilan Proses Sains di SMA” kegiatan praktikum yang ada di SMA Negeri 11 Palembang menggunakan buku LKS yang ada di pasaran. Akan tetapi LKS tersebut masih banyak terdapat kekurangan diantaranya tidak ada indikator, judul praktikum tidak ada, tampilan LK jelek dan kurang menggugah semangat pemakai, ruang pengamatan sempit, LK belum mencerminkan buku berbasis keterampilan ilmu pengetahuan (Zulaiha et al., 2014). Penelitian ini menghasilkan produk berupa buku panduan praktikum Kimia. Model

pengembangan *ADDIE* merupakan persamaan antara penelitian ini dengan penelitian Zulaiha. Sedangkan yang menjadi pokok pengembangan adalah produk serta hasil pengembangan.

Penelitian yang dilakukan oleh Purwaningsih dan Harjito tahun 2019 yang berjudul “Pengembangan Panduan Praktikum Virtual Larutan Penyangga Untuk Menunjang Pelaksanaan Praktikum Nyata” menjelaskan bahwa praktikum merupakan hal yang sangat penting dalam pembelajaran karena dalam kegiatan ini peserta didik melakukan pembelajaran secara langsung yang kesemuanya dapat melatih keterampilan peserta didik. Namun, ada kendala yang dihadapi dalam kegiatan praktikum diantaranya terbatasnya sarana prasarana, petunjuk praktikum kurang jelas, dan kurang melatih peserta didik dalam menggali potensi diri (Purwaningsih & Harjito, 2019). Penelitian ini menghasilkan Panduan Praktikum Virtual Larutan Penyangga untuk Menunjang Pelaksanaan Praktikum Nyata. Jenis penelitian merupakan hal penting yang menjadi persamaan penelitian. Sedangkan model pengembangan dalam penelitian ini menjadi perbedaan yaitu model *4D* dengan *ADDIE*.

Penelitian yang dilakukan oleh Roihatul Miskiyah tahun 2013 dengan judul “Pengembangan Buku Panduan Praktikum Berbasis Inkuiri Terbimbing Pada Materi Benda dan Sifatnya untuk Meningkatkan Motivasi dan Prestasi Belajar Siswa Kelas II MI Bahrul Ulum Ngoro Mojokerto” menjelaskan bahwa buku panduan praktikum IPA untuk kelas 2 yang digunakan oleh guru dan siswa belum tersedia. Adanya buku panduan

praktikum IPA diharapkan akan mempermudah dalam proses pembelajaran IPA terutama pada materi benda dan sifatnya (Miskiyah, 2013). Penelitian ini menghasilkan Buku Panduan Berbasis Inkuiri Terbimbing Pada Materi Benda dan Sifatnya. Persamaan dengan penelitian ini adalah pada jenis penelitian dan pengembangan. Sedangkan perbedaan dengan penelitian ini adalah pada model pengembangan menggunakan model prosedural *Dick and Carey* dan produk yang dihasilkan.

Penelitian yang dilakukan oleh Syaiful Arif tahun 2020 yang berjudul “Pengembangan Panduan Praktikum Kimia Dasar 1 dengan Model *Learning Cycle*” menjelaskan bahwa Kegiatan percobaan adalah kegiatan yang menunjang proses belajar dalam rangka menemukan atau menjelaskan prinsip-prinsip yang dikembangkan. Kegiatan praktikum di laboratorium belum memiliki petunjuk atau pedoman yang terstruktur dan mencerminkan kegiatan ilmiah, sehingga perlu dilakukan pengembangan buku panduan praktikum Kimia (Arif, 2020). Penelitian ini menghasilkan Panduan Praktikum Kimia Dasar 1 dengan Model *Learning Cycle*. Persamaan dengan penelitian ini adalah model pengembangannya yaitu *ADDIE* sedangkan perbedaannya dengan penelitian ini adalah bidang dan produk hasil pengembangan.

## B. Kerangka Teori

### 1. Analisis Tes Hasil Belajar

#### a. Hasil Belajar

Belajar adalah sebuah tindakan dengan akibat terjadinya potensi perilaku yang bersumber pada pengalaman dan latihan. Kegiatan belajar lazimnya dilakukan pada kondisi, tempat, waktu, dan tujuan yang telah ditentukan. Kegiatan belajar demikian sering kita artikan sebagai kegiatan belajar formal. Kegiatan belajar formal dilakukan dalam sebuah institusi pendidikan dari jenjang pendidikan usia dini sampai pada perguruan tinggi.

Kegiatan belajar di institusi pendidikan adalah sebuah proses *transfer of knowledge* maupun *transfer of skill* yang biasa kita kenal dengan kegiatan belajar mengajar (KBM). Kegiatan pembelajaran diawali dengan perencanaan, aksi, sampai pada analisis kegiatan pembelajaran atau evaluasi. Evaluasi pembelajaran dirancang dan dibuat oleh pendidik yang nantinya akan diujikan kepada peserta didik. Hasil dari evaluasi tersebut sering kita sebut dengan hasil belajar.

Menurut Dimiyati & Mudjiono dalam Andriani & Rasto (2019), hasil belajar adalah sebuah pola tindakan, *rule*, definisi, afeksi, penghargaan, dan keterampilan yang merupakan hasil dari sebuah interaksi pembelajaran. Pencapaian prestasi belajar atau

hasil belajar peserta didik mengacu pada pencapaian semua ranah yang bersifat kognitif, afektif, dan psikomotorik (Wahyuningtyas & Sulasmono, 2020).

Hasil belajar pada dasarnya merupakan berbagai kompetensi yang mencakup aspek pengetahuan, keterampilan, sikap dan nilai-nilai yang terwujud dalam kebiasaan berpikir maupun bertindak. Sebuah tujuan dapat dilihat ketercapaiannya melalui hasil belajar yang sudah terkumpul dan terukur serta diamati indikatornya (Widarta, 2020).

Prestasi belajar akan tampak manakala dilakukan sebuah kegiatan *testing* atau uji kompetensi dengan menggunakan instrumen evaluasi atau eksamen.

#### b. Tes Hasil Belajar

Tes atau eksamen adalah salah satu instrumen yang digunakan dalam evaluasi pembelajaran. Kegiatan evaluasi dimaksudkan untuk mengukur kelebihan teknik pembelajaran, kurikulum, kegiatan belajar, membantu peserta didik, mengetahui potensi peserta didik, dan mempersiapkan data untuk menilai dan mengembangkan kurikulum yang kemudian dijadikan dasar dalam mengambil keputusan (Asrul et al., 2014).

Menurut Iskandar & Rizal (2018), Kata *testum* yang berasal dari bahasa Latin memiliki arti alat pengukur tanah, yang merupakan akar kata dari tes. Tes dapat didefinisikan sebagai seperangkat pertanyaan yang harus direspon oleh *testee* sebagai pengukur kualitas seseorang atau bagian tertentu yang ingin dilihat.

Hal tersebut senada dengan pendapat Cahyaningrum (2019), bahwa tes adalah kumpulan pertanyaan yang berlaku bagi *testee* itu soal objektif, atau seperangkat tugas yang dikerjakan dengan sukses atau

berhasil, akhirnya kemampuan *testee* dapat dilihat dengan skor atau nilai. Eksamen adalah instrumen evaluasi yang paling dominan digunakan guru untuk mengukur kemampuan peserta didik dalam kegiatan pembelajaran. Sebagaimana pendapat dari (Asrul et al., 2014), bahwa teknik pengukuran yang banyak digunakan dalam kegiatan penilaian atau evaluasi pembelajaran adalah jenis tes. Walaupun teknik yang ini bukanlah yang terbaik dan sempurna untuk beberapa tujuan.

Menurut Sudijono dalam Sarifuddin (2020), tes adalah instrumen yang digunakan dalam bentuk tugas atau instruksi berupa kumpulan pertanyaan atau soal yang harus dijawab. Pelaksanaan dapat dilaksanakan secara lisan maupun tulis untuk mengukur kompetensi, *skill*, maupun pengetahuan. Berdasarkan deskripsi tersebut, tes adalah instrumen yang mesti dirancang terlebih dahulu agar dapat berfungsi dengan baik terutama dalam mengukur kualitas dan kompetensi yang dimiliki anak. Berbeda dengan pendapat Lee J. Cronbach dalam Suharman (2018), eksamen merupakan prosedur yang sistematis dan dapat digunakan untuk mengamati tingkah laku seseorang dan mengajarkannya menggunakan sebuah alat berskala angka-angka atau sistem kategori.

Menurut Azwar (2015), ada beberapa macam batasan mengenai tes dapat diperoleh sejumlah definisi dari tes atau eksamen, yaitu:

- 1) Eksamen merupakan sebuah instruksi sistematis. Artinya: 1) penyusunan tes mengacu pada aturan tertentu, 2) kejelasan dalam

administrasi dan penskoran hasilnya, dan 3) jenis dan kondisi item-item yang diberikan responden sama.

- 2) Tes merupakan sampel perilaku. maksudnya 1) tes tidak mampu mendeskripsikan secara keseluruhan isi materi, 2) item-item dalam tes secara representatif mewakili kawasan domain sehingga tes tersebut bisa dikatakan layak.
- 3) eksamen mengidentifikasi perilaku. Maksudnya item-item dalam eksamen dapat memberikan informasi yang telah diperoleh mengenai *testee* melalui respon dan jawaban *testee*.

Berdasarkan beberapa argumen yang sudah diutarakan, kita dapat menyimpulkan bahwa tes atau eksamen adalah seperangkat alat yang terdiri dari beberapa pertanyaan, perintah, dan prosedur tertentu dan berlaku bagi *testee* serta dengan maksud dan tujuan tertentu. Jadi, tes hasil belajar adalah seperangkat instrumen yang berisi berbagai pertanyaan, instruksi, dan prosedur tertentu yang harus dilakukan oleh peserta didik sebagai bentuk uji kompetensi setelah kegiatan belajar berakhir.

### c. Jenis Tes Hasil Belajar

Eksamen sebagai alat dalam evaluasi pembelajaran dibedakan dalam dua bentuk secara umum, antara lain. 1) tes bentuk obyektif yaitu tes yang jawabannya diberi nilai sesuai dengan pedoman yang telah dibuat. Adapun jenisnya seperti pilihan jawaban, pengisian titik-titik dan menjodohkan. 2) tes bentuk subjektif yaitu tes/alat evaluasi yang jawabannya berupa uraian dan tidak diberi skor atau nilai pasti karena beragam gaya jawaban dari responden (Wisman et al., 2021)

Senada dengan pendapat Suharsimi (2013), bahwa terdapat dua jenis instrumen evaluasi berbentuk tes. Adapun eksamen tersebut antara lain.

#### 1) Tes Subjektif

Tes subjektif adalah tes berbentuk esai. Eksamen berbentuk esai bersifat homogen untuk mengukur peningkatan dalam belajar peserta didik melalui jawaban berbentuk uraian kata-kata. Karakteristik pertanyaan berbentuk esai menggunakan istilah; mengapa, bagaimana, uraikan, jelaskan, simpulkan, dan lain sebagainya.

#### 2) Tes Objektif

Tes objektif merupakan bentuk eksamen dimana dalam praktiknya dapat dilaksanakan secara objektif. Tujuannya adalah untuk mengurangi atau bahkan menghilangkan kekurangan yang terdapat dalam tes dengan bentuk esai. Bentuk tes baik objektif atau subjektif, lebih menitik beratkan pada tujuan penyelenggaraan tes tersebut. Dengan kata lain bahwa kecenderungan tes tidak dapat disamaratakan antara tes yang satu dengan yang lainnya karena setiap tes mempunyai tujuan dan tanggapan tersendiri.

Pada kondisi sebenarnya di lapangan, antara tes subjektif dan tes objektif diujikan dalam saat dan momen yang sama. Sebagai contoh evaluasi dalam pembelajaran yaitu sumatif, tengah semester, maupun akhir semester, dan biasanya guru membuat soal dengan perpaduan komposisi antara tes subjektif dan tes objektif.

#### d. Syarat Tes Hasil Belajar

Hasil pengukuran akan menunjukkan kondisi sebenarnya di lapangan apabila dibuat sebuah instrumen evaluasi yang baik yaitu eksamen yang berisi sejumlah item dengan kualitas baik. Menurut Fitriawanawati dalam Rizalia et al (2021), tes hasil belajar masuk dalam kategori baik apabila secara teori dan praktik dapat menunjukkan hasil terbaik dalam mengukur kemampuan kognitif peserta didik.

Eksamen atau tes yang baik harus memenuhi hal-hal berikut, antara lain. (1) valid, artinya hanya mengukur satu aspek saja, (2) reliabel, artinya konsisten, (3) bersifat standar, maksudnya tes mendapat perlakuan yang sama dalam hal pelaksanaan, bahan yang digunakan, skoring, dan penafsiran hasil tes, (4) objektif, artinya perlakuan yang sama pada peserta tes, (5) eksamen harus bersifat diskriminatif, maksudnya ia harus dapat membedakan tiap individu peserta tes dalam hal kemampuan mengerjakan tes (Suharman, 2018).

Tes yang baik atau berkualitas tinggi dapat diperoleh dengan langkah-langkah tertentu. Adapun langkah pertama dalam penyusunan dan pembuatan eksamen menggunakan penekanan yang jelas, pembuatan item soal didasarkan pada dalil yang sudah pasti, aturan penulisan item sudah pasti, dan menggunakan latihan yang produktif, dan keahlian yang bermanfaat.

Eksamen atau tes yang telah disusun menggunakan prosedur demikian secara teori masuk dalam kategori layak (Azwar, 2015). Artinya, untuk mendapatkan instrumen tes dengan kualitas yang bagus tidak bisa dilakukan secara instan, perlu dilakukan kaidah-kaidah tertentu saat dibuat atau sebelum diujikan maupun analisis setelah instrumen tes tersebut sudah diujikan.

Langkah kedua untuk memperoleh item-item tes (eksamen) dengan mutu baik yaitu melakukan telaah atau analisis terhadap item-item tes tersebut.

Kegiatan analisis tes menjadi hal yang harus dilakukan oleh pembuat tes setelah tes tersebut diujikan kepada peserta atau responden. Kegiatan pembuatan tes merupakan tugas seorang *evaluator*. Peran sebagai *evaluator* dalam kegiatan pembelajaran adalah guru. Menurut Kunandar dalam Nur & Palobo (Nur & Palobo, 2018), kegiatan analisis item soal merupakan sebuah cara sistematis sebagai pemberi beberapa informasi spesifik terhadap item soal yang telah dibuat atau disusun.

Menurut Nitko dalam Nasir (2015), bahwa kegiatan telaah item soal sebagai rutinitas dari guru dengan maksud mutu soal kedepannya akan menjadi lebih baik. Ada beberapa tahap dalam kegiatan ini antara lain mengumpulkan, mereview, dan menggunakan data berupa respon *testee* sebagai dasar dalam sebuah penelitian. Setiap guru (pembuat soal) dituntut memiliki kompetensi bukan hanya sebagai pembuat soal, tetapi juga harus memiliki kemampuan dalam menganalisis soal tersebut. Adapun kemampuan tersebut sebagai bukti bahwa guru mempunyai peran sebagai *evaluator*. Hal ini dikuatkan oleh Darmadi (2015), bahwa seorang guru dalam menjalankan fungsinya sebagai *evaluator* adalah ia harus dapat memberikan penilaian yang bagus dan jujur, serta memberikan nilai dalam ruang yang luas. Adapun maksudnya adalah setelah guru membuat tes, mengujikannya, kemudian menganalisis kualitas tersebut sebagai langkah selanjutnya. Kegiatan analisis hasil tes oleh guru dimaksudkan agar diperoleh instrumen tes dengan kualitas tinggi sehingga instrumen tersebut dapat disimpan dan dapat dipergunakan lagi pada waktu yang akan datang.

Alasan dilakukannya analisis terhadap soal antara lain. (1) respon *testee* terhadap soal merupakan sumber informasi diagnostik untuk meneliti kegiatan pembelajaran dan kegagalan-kegagalan dalam belajar, dan selanjutnya untuk memandu dan membentuk peserta didik untuk menemukan cara belajar yang lebih efektif, (2) hasil respon peserta tes atau *testee* menjadi dasar dalam pembuatan soal atau bank soal untuk tahun selanjutnya. Pemanfaatan hasil analisis soal oleh guru dapat menunjang peningkatan mutu pembelajaran demi tercapainya tujuan (Srika Ningsih & Yusrizal, 2018)

#### e. Fungsi dan Tujuan Analisis Tes Hasil Belajar

Telaah terhadap item soal merupakan kegiatan yang menjadi rutinitas penting dalam sebuah pembelajaran. Dikatakan demikian karena dengan kegiatan tersebut guru dapat membuat atau menyusun instrumen evaluasi yang memiliki kualitas baik. Kegiatan telaah item soal memiliki beberapa fungsi, yaitu.

- 1) Berperan dalam kegiatan evaluasi terhadap mutu tes yang telah digunakan oleh *tester*.
- 2) dapat digunakan dalam menyusun eksamen informal, seperti tes yang dipersiapkan pendidik untuk peserta didiknya.
- 3) membantu pendidik dalam pembuatan atau penyusunan eksamen yang berkualitas dan efektif.
- 4) validitas dan reliabilitas soal meningkat sehingga akan tercipta soal yang berkualitas baik.

Kegiatan utama dari telaah terhadap soal adalah untuk menemukan dan menentukan kualitas soal yang kita buat atau susun apakah soal tersebut baik, kurang baik, dan soal itu buruk (Suharsimi, 2013).

Kegiatan telaah soal memiliki manfaat praktis terhadap kegiatan pembelajaran di sekolah, antara lain: (1) Membantu tester dalam mengevaluasi eksamen, (2) membantu pendidik dalam mempersiapkan eksamen atau tes untuk kelasnya, (3) acuan dalam membuat soal yang efektif dan efisien, (4) meningkatkan kualitas eksamen, (5) meningkatkan kevalidan dan konsistensi pertanyaan, (6) meningkatkan kualitas eksamen melewati tiga komponen yaitu daya beda, tingkat kesulitan, dan distraktor (Yuniawatika et al., 2020).

Jadi, manfaat dari analisis item soal antara lain sebagai instrumen evaluasi eksamen (tes), mempersiapkan tes untuk kelas, membuat soal yang efektif, dan meningkatkan kualitas tes yang telah dibuat atau disusun.

#### **f. Prosedur Analisis Tes Hasil Belajar**

Analisis kualitatif dan kuantitatif merupakan dua teknik yang digunakan dalam kegiatan analisis item soal. Senada dengan pendapat Thoha dalam Umi Fatimah & Alfath (2019), telaah pada item soal bisa dilakukan secara kualitatif maupun kuantitatif. Setiap cara penilaian tersebut memiliki kelebihan dan kekurangan, sehingga langkah bijak yang terbaik adalah teknik kombinasi dari keduanya.

##### **1) Teknik Analisis Secara Kualitatif**

Teknik moderator dan panel merupakan teknik yang digunakan dalam kegiatan analisis secara kualitatif. Teknik moderator adalah kegiatan telaah soal dengan jalan diskusi antara beberapa orang dengan satu orang sebagai moderator. Pada teknik ini, setiap item soal dikaji secara mendalam oleh para ahli *assessment* seperti guru sebagai pemateri, validator materi, maupun tim pengembang kurikulum. Berbeda dengan teknik yang pertama, teknik panel adalah teknik analisis item soal dengan fokus kajian pada konten, konstruk, bahasa, dan pedoman penilaian.

## 2) Teknik Analisis Secara Kuantitatif

Analisis secara kuantitatif merupakan analisis item soal yang disandarkan pada data lapangan. Data tersebut didapatkan dari soal yang telah diujikan. Telaah secara kuantitatif menggunakan dua pendekatan yaitu *classical theory* dan item respon butir.

Kegiatan telaah item soal melalui informasi jawaban testee dengan maksud untuk meningkatkan kualitas item soal yang terkait dengan teori tes klasik disebut dengan analisis soal secara klasik. Bagian yang paling pokok dalam pendekatan klasik meliputi tingkat kesukaran, fungsi distraktor, keandalan, dan keshahihan soal (Umi Fatimah & Alfath, 2019).

Seperti teknik pada umumnya, baik teknik kualitatif maupun kuantitatif masing-masing memiliki kelebihan maupun kekurangan, oleh sebab itu kita dapat menyiasatinya dengan menggabungkan kedua Teknik ini dengan pertimbangan kelebihan maupun kekurangan masing-masing (Nurdiana et al., 2020). Kombinasi antara teknis kualitatif dan kuantitatif diharapkan akan menghasilkan produk analisis soal yang lebih lengkap dan terinci.

Menurut Anastasi Anne dan Urbina dalam Adawiah & Ruchliyad (2019), kegiatan analisis eksamen (tes) yang dilakukan oleh *tester* dapat dikerjakan secara kualitatif terkait isi maupun bentuknya, sedangkan yang terkait dengan ciri statistik dengan mekanisme peningkatan secara empiris *tester* bisa melakukan analisis secara kuantitatif. Analisis secara kualitatif meliputi pengukuran kesahihan konten maupun konstruk, sedangkan analisis secara kuantitatif meliputi pengukuran validitas, reliabilitas, daya beda, dan uji tingkat kesukaran. Pada penelitian ini fokus pada analisis secara kuantitatif

karena kegiatan analisis yang dilakukan oleh guru umumnya adalah analisis kuantitatif.

Analisis secara kuantitatif adalah kegiatan telaah terhadap instrumen tes (eksamen) yang dilakukan ketika kegiatan ujian atau penilaian telah dilakukan. Respon atau jawaban *testee* menjadi dasar dalam kegiatan analisis ini. Adapun dalam analisis ini meliputi tingkat kesukaran, daya beda, keefektifan distraktor, dan konsistensi (Prabowo et al., 2017). Jadi, analisis secara kuantitatif dilakukan setelah soal diujikan, artinya analisis kuantitatif dapat dilakukan berdasarkan respon atau jawaban dari peserta tes.

Menurut Fitriyanawati dalam Rizalia et al (2021), suatu eksamen dikatakan berkualitas baik apabila memenuhi beberapa ketentuan sebagai berikut yaitu eksamen harus sah artinya mengukur satu aspek saja, handal artinya dapat mengukur dengan tepat, cermat, dan konsisten, serta adil artinya bersifat objektif dan tidak memihak. Jadi, dapat disimpulkan bahwa syarat sebuah soal dikatakan berkualitas apabila soal tersebut, valid, reliabel, dan objektif.

Hal ini dikuatkan oleh A Son (2019), bahwa untuk mengungkap kualitas butir soal maka perlu adanya analisis terhadap validitas, kejelasan, tingkat kesulitan, dan daya beda. Jadi, untuk mendapatkan informasi terkait kualitas butir soal yang sudah kita buat perlu dilakukan analisis terhadap validitas, reliabilitas, tingkat kesukaran, keefektifan distraktor atau pengecoh, dan daya beda soal. Berikut kita uraikan masing-masing bagian yang bisa dilihat saat kita melakukan analisis terhadap soal yang dibuat.

## 1) Validitas

Validitas adalah keandalan dalam sebuah pengukuran. Suatu instrumen apabila memiliki tingkat kesalahan pengukuran kecil bisa diambil kesimpulan bahwa instrumen tersebut valid. Maksudnya adalah nilai skor dengan menggunakan instrumen pengukuran tidak jauh berbeda dengan nilai skor sebenarnya (Azwar, 2018). Ada sebuah keterkaitan kuat antara validitas suatu eksamen dengan tujuan dari pembuatan eksamen tersebut. Artinya apabila eksamen tersebut dapat memberikan informasi sesuai dengan maksud dan tujuan penggunaannya, maka bisa dikatakan bahwa tersebut memiliki validitas tinggi atau valid (Magdalena et al., 2021).

Syarat penting yang harus dipenuhi dalam kegiatan evaluasi yaitu instrumen evaluasi tersebut harus memiliki nilai validitas yang tinggi atau valid. Maksudnya adalah alat atau instrumen evaluasi itu harus dapat menjalankan tugasnya sesuai dengan tujuan pembuatannya. Suatu eksamen masuk dalam kategori valid jika eksamen tersebut mampu mengukur apa yang menjadi tujuan pembuatannya (Sudirman, 2022).

Menurut Fraenkel, Wallen, & Hyun dalam Yusup (2018), nilai kevalidan sebuah eksamen dapat dilihat pada nilai korelasinya. Nilai korelasi akan nampak pada koefisien korelasi dengan simbol  $r$  kecil ( $r$ ). Koefisien korelasi ( $r$ ) digunakan untuk menyatakan validasi suatu eksamen. Rentang nilai koefisien korelasi antara +1 sampai dengan -1. Nilai +1 mengindikasikan bahwa hasil yang diperoleh individu sama baik dalam uji instrumen maupun uji kriteria. Sedangkan jika nilai korelasi 0

menunjukkan bahwa antara instrumen dan kriteria tidak berhubungan (tidak ada korelasi).

Menurut Koyan dalam Styawarno (2017), untuk menghitung nilai korelasi *Point Biserial* dapat menggunakan formula sebagai berikut.

$$r_{pbi} = \frac{M_p - M_t}{S_t} \sqrt{\frac{p}{q}}$$

$r_{pbi}$  = koefisien korelasi *point biserial*

$M_p$  = rata-rata skor dari subjek yang menjawab benar bagi butir yang dicari nilai validitasnya

$M_t$  = rerata skor total

$S_t$  = standar deviasi dari skor total

$p$  = rasio peserta didik yang menjawab benar (banyaknya peserta didik yang menjawab benar dibagi dengan jumlah seluruh peserta didik)

$q$  = rasio peserta didik yang menjawab salah ( $q = 1 - p$ ).

Sebuah kondisi dimana  $r$  hitung lebih besar daripada  $r$  tabel dengan taraf signifikansi sebesar 5% ( $r$  hitung  $>$   $r$  tabel dengan taraf sig. 5%), maka bisa dikatakan bahwa item ujian tersebut valid. Adapun untuk mengetahui derajat korelasi dapat diperhatikan keterangan di bawah ini.

- 1) 0,00 – 0,20 korelasi sangat rendah (hampir tidak ada korelasi)
- 2) 0,20 – 0,40 korelasi rendah
- 3) 0,40 – 0,70 korelasi cukup
- 4) 0,70 – 0,90 korelasi tinggi

5) 0,90 – 1,00 korelasi sempurna

Nilai koefisien validitas suatu alat berbanding lurus dengan kualitas alat tersebut. Apabila nilai koefisien validitas tinggi, bisa dipastikan bahwa kualitas instrumen tersebut baik.

## 2) Reliabilitas

*Reliability* merupakan akar kata dari reliabilitas yang artinya dapat dipercaya. Suatu kegiatan pengukuran dikatakan memiliki nilai reliabilitas yang tinggi apabila dalam pengukuran tersebut dihasilkan data yang sah. Andal, ajeg, konsisten, stabil, kepercayaan dan sebagainya. Namun, sejauh mana hasil suatu penilaian dapat dipercaya merupakan ide pokok dalam konsep reliabilitas (Azwar, 2018).

Ada beberapa interpretasi yang dapat digunakan untuk memahami konsep reliabilitas dalam teori skor murni. Sekali lagi reliabilitas adalah keajegan atau konsistensi suatu eksamen ketika diujikan pada subjek yang berbeda. Korelasi *product moment* merupakan teknik analisis data guna mengukur validitas, kesejajaran hasil dalam reliabilitas eksamen (Suharsimi, 2013). Soal dikatakan reliabel apabila soal tersebut mampu mempertahankan konsistensinya dalam hal penggunaannya. Maksudnya adalah soal tersebut memberikan *output* yang sama saat diujikan pada subjek yang berbeda yang berdampak pada perolehan nilai kepercayaan yang tinggi.

Tujuan dari penentuan koefisien reliabilitas sebuah tes adalah untuk dapat memahami dan mengukur ketepatan dan konsistensi sebuah eksamen. Indeks reliabilitas mempunyai rentang nilai 0 - 1. Apabila dalam sebuah kegiatan analisis item tes mendekati 1, semakin tinggi juga ketepatannya.

Menurut Suparwoto dalam Styawarno (2017), bahwa formula yang dipergunakan untuk mencari koefisien Alpha-Cronbach dapat menggunakan rumus berikut..

$$r_{1.1} = \left( \frac{k}{k-1} \right) \left( \frac{SDt^2 - \sum SDi^2}{SDt^2} \right)$$

Keterangan:

$r_{1.1}$  = koefisien reliabilitas eksamen

$k$  = jumlah item eksamen

$SDi^2$  = varians skor tiap item

$SDt^2$  = varian skor total.

Menurut Huda & Wahyuni (2020), bahwa kajian terhadap reliabilitas sebuah tes terdiri dari beberapa rumus atau formula. Adapun formula-formula tersebut antara lain formula belah dua, *Kuder Richardson*, dan *crombach alpha*. Sedangkan hasil perhitungan dari formula-formula tersebut dinamakan koefisien reliabilitas. Hasil dari analisis item soal dengan menggunakan iteman 4.3 ada tiga formula antara lain *alpha (KR-20)*, *split-half*, dan *Spearman-Brown*. Untuk dapat menafsirkan koefisien reliabilitas dapat dilihat pada tabel 1 berikut.

**Tabel 1. Kriteria Koefisien Reliabilitas**

Koefisien Reliabilitas	Makna
$r \leq 0,20$	Sangat Rendah
$0,20 < r \leq 0,40$	Rendah
$0,40 < r \leq 0,60$	Sedang
$0,60 < r \leq 0,80$	Tinggi
$0,80 < r \leq 1,00$	Sangat Tinggi

Sumber: (Styawarno, 2017).

### 3) Tingkat Kesukaran

Menurut Arikunto dalam Syarifah et al (2020), soal masuk dalam kategori baik jika soal tersebut apabila dikerjakan oleh *testee* tidak terlalu mudah dan tidak terlalu susah. Hal yang perlu dipertimbangkan dalam menentukan proposionalitas jumlah soal mudah, sedang, dan susah. Menurut Arikunto dalam Sudarsono (2021), kajian terhadap tingkat kesukaran soal mempunyai tujuan untuk mengetahui soal yang telah dibuat atau disusun masuk dalam kategori mudah, sedang, ataukah susah. Jadi, yang dimaksud dengan tingkat kesukaran adalah angka yang menunjukkan suksr tidaknya sebuah soal.

Menurut Zaenal Arifin dalam Al Fathiyah (2019), Perhitungan tingkat kesukaran soal adalah pengukuran derajat kesukaran suatu soal. Apabila soal mempunyai derajat kesukaran yang seimbang, maka soal tersebut masuk dalam kategori baik. Tingkat kesukaran suatu item soal dapat dinyatakan suatu indeks yang disebut indeks kesukaran item dengan huruf *P* sebagai simbolnya. Indeks Kesukaran item adalah perbandingan antara *testee* yang menjawab benar dan jumlah total *testee* penjawab item. Secara teori dapat dikatakan bahwa *P* merupakan peluang empiris untuk lulus item tertentu bagi kelompok peserta didik tertentu. Menurut Azwar (2015), untuk menghitung nilai dari *P* diperlukan formulasi indeks kesukaran item sebagai berikut.

$$P = ni/N$$

*ni* = Jumlah *testee* yang menjawab benar

*N* = Jumlah *testee* yang menjawab salah

Kajian terhadap tingkat kesukaran item soal dengan menggunakan bantuan program Iteman 4.3 dapat dilihat pada bagian kolom *P* (Huda & Sri wahyuni, 2020). Adapun untuk menafsirkan nilai dari *P* dapat digunakan pada Tabel 2 sebagai berikut.

**Tabel 2. Kriteria Tingkat Kesukaran**

Nilai P	Keterangan
$P > 0,70$	Mudah
$0,30 \leq P \leq 0,70$	Sedang
$P < 0,30$	Susah

Sumber: (Prabayanti et al., 2018).

Sebuah soal tergolong baik jika saat dikerjakan oleh *testee* tidak terlalu mudah dan tidak terlalu sukar. Soal terlalu mudah tidak memotivasi peserta didik dalam menyelesaikannya. Begitu juga sebaliknya, soal yang terlalu susah dapat menyebabkan peserta didik putus asa dan tidak bersemangat untuk mencoba kembali (Arikunto, 2012). Menurut ketentuan dalam menafsirkan nilai dari *P* dapat digolongkan sebagai berikut.:

- Nilai *P* 0,00 - 0,30 soal tergolong sukar
- Nilai *P* 0,31 - 0,70 soal tergolong sedang
- Nilai *P* 0,71 - 1,00 soal tergolong mudah

Meskipun demikian, ada pendapat lain bahwa soal dalam kategori baik apabila memiliki nilai *P* (indeks kesukaran) sampai 0,70 (Suharsimi, 2013).

#### 4) Daya Pembeda

Menurut Z. Arifin (2012), bahwa penilaian daya pembeda adalah pengukuran yang dilakukan pada sebuah soal tentang kemampuannya dalam membedakan peserta didik yang telah, kurang, atau belum menguasai kompetensi tertentu. Analisis terhadap daya pembeda adalah telaah sejumlah item soal untuk mengetahui dan membedakan kemampuan peserta didik dalam kegiatan testing (Uno & Koni, 2013). Sedangkan menurut Umi Fatimah (2019), bahwa item soal mampu mengklasifikasikan peserta didik dengan tingkat kemampuan penguasaan materi, potensi personal, dan kelompok. Kemampuan semacam ini dinamakan daya beda. Maksudnya, daya pembeda dimaksudkan untuk mengelompokkan peserta tes antara yang *high class* dan *low class* dilihat dari sisi kemampuan untuk menjawab butir soal dalam tes. Daya pembeda pun baik secara langsung atau tidak sangat membantu guru terutama dalam hal *placement*/penempatan peserta didik misalnya. Dengan demikian daya beda merupakan angka yang dapat menunjukkan sejauh mana soal tersebut dapat membedakan peserta didik yang pintar dan kurang pintar. Dengan kata lain, daya beda eksamen adalah koefisien angka yang dapat menunjukkan sejauh mana soal dapat membedakan peserta didik yang pintar dan yang kurang. Daya pembeda item soal memiliki manfaat diantaranya. (1) melalui data empirik kualitas item soal dapat ditingkatkan. Keputusan terhadap soal setelah ditelaah dengan status baik, revisi, bahkan ditolak didasarkan pada nilai daya pembeda, (2) Mengukur sejauh mana efektifitas tiap item dalam membedakan potensi

peserta didik, yaitu antara peserta didik yang sudah tuntas dan belum. Mengacu pada paparan yang telah disajikan, ada titik temu tentang pengertian dari daya beda/pembeda adalah kemampuan atau daya sebuah instrumen eksamen dalam menggolongkan peserta didik khususnya dalam kemampuan menyelesaikan soal, belajar atau tidak, dan Peserta didik yang masuk dalam kelas atas maupun bawah.

Menurut Lababa (2018), bahwa untuk memperoleh koefisien atau nilai dari daya beda, dapat menggunakan formula di bawah ini.

$$Da = \frac{nB_A}{n_A} - \frac{nB_B}{n_B}$$

$nB_A$  = jumlah peserta yang menjawab benar pada kelompok atas

$nB_B$  = jumlah peserta yang menjawab benar pada kelompok bawah

$n_A$  = jumlah peserta yang masuk kelompok atas

$n_B$  = jumlah peserta yang masuk kelompok bawah.

*Rpbis* merupakan kolom yang menunjukkan nilai daya pembeda soal hasil analisis item soal dengan menggunakan program Iteman 4.3. Adapun kriteria dari daya pembeda bisa dilihat pada tabel 3 berikut.

**Tabel 3. Kriteria Daya Pembeda**

<b>Rentang Nilai</b>	<b>Keterangan</b>
0,40 – 1,00	Baik
0,30 – 0,39	Sedang
0,20 – 0,29	Cukup Baik
Negatif – 0,19	Buruk

Sumber: (Iskandar & Rizal, 2018)

Dengan demikian, sebuah eksamen dikatakan memiliki daya beda atau daya pembeda yang bagus jika ia mampu menggolongkan peserta didik (*testee*) berdasarkan yang ada pada diri mereka yaitu *testee* dengan motivasi belajar tinggi dengan *testee* yang memiliki motivasi belajar rendah.

### 5) Distraktor (Pengecoh)

Analisis efektifitas pengecoh atau analisis pola distribusi jawaban item soal pada soal berbentuk *multiple choice*. Pola tersebut didapatkan dari jumlah peserta eksamen yang memilih opsi jawaban pada item soal atau yang tidak memilih sama sekali. Melalui pola ini, kita dapat menentukan apakah pengecoh efektif atau tidak. Pengecoh atau distraktor dapat dikategorikan berfungsi dengan baik apabila pengecoh tersebut dipilih oleh minimal 5% dari seluruh *testee*.

Menurut Arifin dalam Sukendro et al (2018), untuk mengukur efektivitas pengecoh dapat dihitung dengan menggunakan rumus :

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

Keterangan:

IP = indeks distraktor

P = jumlah *testee* yang memilih pengecoh.

N = jumlah *testee* yang ikut tes.

B = jumlah *testee* yang menjawab betul.

n = jumlah alternatif jawaban.

1 = bilangan tetap

Efektifitas distraktor dapat diperoleh dari hasil telaah menggunakan program Iteman 4.3. Nilai efektifitas distraktor atau pengecoh dapat dilihat pada *Prop*. Menurut Mardapi dalam Huda & Wahyuni (2020), Apabila nilai *Prop* pada tiap opsi pilihan ganda lebih dari 0,05 ( $Prop > 0,05$ ) bisa dikatakan distraktor tersebut berfungsi dengan baik. Namun, apabila nilai *Prop* pada opsi setiap item soal kurang dari nilai 0,05 maka distraktor tersebut harus di perbaiki.

Telaah terhadap tingkat kesalahan, daya beda, serta fungsi pengecoh lebih utama dan sebaiknya dilakukan sebelum soal digunakan sebagai instrumen penilaian yang dilakukan secara berulang gara diperoleh eksamen dengan kualitas yang baik (Sudarsono, 2021).

#### 6) SEM (*Standard Error Measurement*)

SEM sangat berguna bagi *tester* untuk memahami kesalahan secara random dan berpengaruh secara langsung terhadap perolehan nilai skor peserta dalam sebuah kegiatan tes. Kesalahan pengukuran dapat dicari dengan rumus.

$$\sigma_E = \sigma_x \sqrt{1 - p_x}$$

Dimana  $\sigma_x$  merupakan deviasi standar dari skor total dan  $p_x$  adalah koefisien reliabilitas. Kesalahan baku pengukuran nilainya harus sekecil mungkin. Menurut Crocker dan Algina dalam D. D. Kurniawan (2020), *the standard error of measurement (SEM)* dapat didefinisikan dengan  $\sigma E = \sigma X \sqrt{1 - \rho_{XX'}}$  dapat dikategorikan juga sebagai rerata standar deviasi dari

distribusi kesalahan individu responden untuk sebuah penggetesan ulang. Sedangkan menurut Miller (2008:93), penilaian dari *standard error of measurement (SEM)* merupakan pengukuran besarnya kesalahan (*error*) dalam sebuah dalam penskoran tes berdasarkan reliabilitas tes.

Menurut Sukardi dalam Huda & Wahyuni (2020), Seorang evaluator harus dapat memperkirakan kemungkinan kesalahan pengukuran terhadap eksamin atau tes yang dibuatnya. Perkiraan nilai kesalahan pengukuran disebut dengan *SEM (Standard error of measurement)*. Nilai *SEM* dan reliabilitas eksamin berbanding lurus dengan reliabilitas. Jika nilai *SEM* rendah maka rendah pula nilai reliabilitasnya. Demikian pula sebaliknya, jika nilai *SEM* tinggi bisa dipastikan nilai reliabilitas sebuah tes juga tinggi. Kesalahan pengukuran harus memiliki nilai minimal dan kesalahan pengukuran berbanding terbalik dengan reliabilitas. Semakin kecil nilai *SEM* semakin tinggi reliabilitas sebuah tes, dan semakin tinggi nilai *SEM* semakin rendah reliabilitas sebuah tes.

Kegiatan analisis item soal dalam jumlah sedikit dapat dilakukan secara manual maupun dengan bantuan kalkulator saja. Akan tetapi, lain halnya jika analisis item soal dengan jumlah relatif banyak, kita memerlukan sebuah alat atau piranti khusus untuk melakukan analisis eksamin (Azwar, 2015). Adapun paket program khusus untuk analisis butir soal diantaranya Program Bilog, Iteman, Anates, dan lain sebagainya. *Item and Test Analysis* (Iteman) adalah program yang digunakan dalam penelitian ini untuk melakukan analisis terhadap item soal. Analisis menggunakan Iteman belum

familiar di kalangan guru sehingga memerlukan panduan untuk dapat mengoperasikannya.

## 2. Item and Test Analysis (Iteman)

### a. Pengertian Iteman

Analisis item soal merupakan bagian dari kegiatan guru untuk mengukur kualitas eksamen dan menafsirkan hasil belajar. Era digital menuntut semua pekerjaan dapat dikerjakan dengan tepat dan cepat, tanpa terkecuali dalam kegiatan analisis item soal. Hal demikian dapat dilakukan dengan bantuan aplikasi pada piranti komputer (Nur & Palobo, 2018).

Menurut Abidin dalam Suzana (2018), Iteman (*Item and Test Analysis*) adalah sebuah program komputer yang digunakan untuk menganalisis item soal secara kuantitatif. Hasil analisis dengan menggunakan program Iteman sangat membantu untuk melihat kualitas eksamen.

Program ini termasuk dalam jenis MicroCAT<sup>o</sup>n yang dikembangkan oleh *Assessment System Corporation* selaku pemilik. Iteman mulai dikembangkan pada tahun 1984, 1986, 1988, dan 1993; mulai dari versi 2.00 sampai dengan versi 3.50 dan akan terus berkembang sampai nanti (Yuniawatika et al., 2020).

Iteman merupakan salah satu program *software* yang dapat dioperasikan pada piranti komputer. Iteman merupakan salah satu *software* yang dapat digunakan untuk menganalisis item soal dengan pendekatan

*classical theory* untuk mengukur kualitas item soal. Hasil analisis item soal antara lain tingkat kesukaran, daya pembeda, reliabilitas, *SEM (Standard Error of Measurement)* serta distribusi jawaban peserta tes.

Iteman ini dapat digunakan untuk beberapa hal dibawah ini, yaitu.

- 1) Analisis data jawaban item soal melalui *manual entry* atau dengan menggunakan mesin *scanner*.
- 2) Skoring dan analisis data soal pilihan ganda
- 3) Analisis tes 10 skala dan dapat menginformasikan validitas setiap butir soal (Styawarno, 2017).

Menurut Huda & Wahyuni (2020), program Iteman 4.3 merupakan tool yang digunakan analisis item soal dengan pendekatan teori klasik. Adapun hasil dari analisis item soal antara lain validitas, reliabilitas, tingkat kesukaran, daya pembeda, fungsi distraktor maupun SEM (kesalahan pengukuran).

Menurut Styawarno (2017), program Iteman 4.3 memiliki keunggulan tersendiri dibandingkan dengan Iteman 3.0. Adapun keunggulan Iteman 4.3 antara lain.

- 1) Penambahan grafis menghasilkan plot yang teratur dan memudahkan pengguna dalam mengatur jumlah titik pada plot.
- 2) Iteman versi 4.3 dapat digunakan untuk *pretest* butir yang tidak masuk dalam nilai akhir.

- 3) Produk statistik banyak dihasilkan seperti alpha (KR-20), reliabilitas dengan konfirmasi butir soal yang harus dihapus, reliabilitas (Speramen-Brown), Standar Deviasi, dan sub kelompok P (proporsi benar) statistik.
- 4) Hasil dari olah data file RTF sebagai laporan formal, dan bentuk CSV
- 5) Output dapat ditambah dengan skor skala dan sub scores.
- 6) Ada dua kelompok dalam penskoran dan dapat dilabeli oleh pengguna
- 7) Jumlah maksimum butir yang dapat dianalisis maksimal 10.000 item.
- 8) Menggunakan Multiple Runs File yang memungkinkan pengguna grafis setiap run data.

#### **b. Kelebihan dan Kekurangan Iteman 4.3**

- 1) Kelebihan Iteman
  - Input data data dapat dibantu dengan *Ms. Excel* (Styawarno, 2017).
  - Hasil analisis disajikan dalam bentuk tabel untuk memudahkan dalam pemahaman
  - *Layout* Iteman 4.3 lebih menarik daripada program lain
  - Versi selalu *uptodate*.
- 2) Kekurangan Iteman
  - Kemungkinan terjadi kesalahan dalam penulisan data respon peserta didik (Huda, 2021).
  - Pengoperasian sering terjadi kendala karena masih menggunakan bahasa Inggris
  - Input format menjadi masalah tersendiri bagi pengguna.

### c. Manfaat Iteman 4.3

Iteman 4.3 sebagai program analisis memiliki manfaat antara lain sebagai berikut.

- 1) Memperkenalkan *software* aplikasi Iteman 4.3 kepada pengguna (Huda & Sri wahyuni, 2020).
- 2) Memberi keterampilan pengguna dalam menjalankan *software* aplikasi Iteman 4.3
- 3) Memberi keterampilan pengguna untuk dapat menginterpretasikan output *software* aplikasi Iteman 4.3

Panduan analisis program Iteman 4.3 yang merupakan produk dari ASC, s3 masih sulit digunakan oleh *user* dalam hal ini guru dikarenakan antara lain.

- a. Masih menggunakan Bahasa Inggris.
- b. Masih menggunakan Bahasa pemrograman.
- c. Tidak dapat digunakan secara mandiri
- d. Tidak bersifat aplikatif praktis

Kondisi panduan analisis Iteman versi 4.3 yang demikian, menjadi dasar perlunya dilakukan pengembangan panduan analisis program Iteman versi 4.3 yang berupa *user's manual for Iteman version 4.3* dengan jargon panduan analisis program Iteman 4.3 yang *user friendly* terutama bagi guru selaku pengguna (*user*).

### 3. Panduan Program Analisis Soal

Panduan merupakan dokumen yang berisi komunikasi secara teknis dengan maksud memberikan solusi terhadap penggunaan sistem tertentu, piranti tertentu, baik *software* maupun *hardware*. Panduan analisis program Iteman versi 4.3 merupakan

sebuah buku pedoman praktik dalam menjalankan sebuah program. Menurut KBBI (2021), panduan adalah kompilasi yang berisi aturan dasar dengan maksud memberi petunjuk tentang apa yang harus dilakukan. Sedangkan praktik adalah prosedur nyata dari sebuah teori. Jadi, pedoman praktik adalah ketentuan dasar yang memberikan arah bagaimana pelaksanaan secara nyata dari sebuah teori.

Panduan merupakan segala ketentuan yang berisi informasi, petunjuk/prosedur penggunaan dan cara kerja sebuah produk. Sebagaimana pendapat dari Karen A. Schriver dalam Møller (2013), bahwa prosedur yang bisa digunakan berdasarkan petunjuk pengguna: Pertama, evaluasi petunjuk pengguna berdasarkan seperangkat petunjuk (review dari ahli) dan kemudian tes dilaksanakan oleh pengguna yang diminta untuk menyelesaikan beberapa tugas dan petunjuk serta produk.

Panduan merupakan satu paket lengkap dengan produk yang digunakan. Artinya, lahirnya sebuah produk harus disertai dengan panduan di dalamnya. Panduan praktis tersebut dimaksudkan untuk mengenalkan dan mempermudah para pengguna dalam mengoperasikan produk tersebut. Panduan dibuat dalam bentuk sebuah buku dan lazim dikenal dengan istilah buku panduan.

Buku Panduan atau biasa dikenal dengan istilah buku pintar, hal ini dikarenakan apabila orang membaca buku ini menambah informasi sehingga dianggap menjadi pintar pada bidang tertentu yang sedang dikerjakan dan memperjelas terhadap informasi yang masih samar. Ada sebagian orang yang menyebut buku panduan sebagai *handbook* dan *manual book*. Menurut Sudarnoto dalam Sinambela et.al (2022), bahwa *handbook* adalah perpaduan dari beberapa

informasi yang tersusun secara sistematis dan siap pakai. Sedangkan buku pedoman (*manual book*) merupakan buku yang berisi aturan, prosedur, maupun cara untuk melakukan suatu pekerjaan tertentu secara bertahap.

Buku panduan umumnya digunakan sebagai piranti dalam menguji data. Terdapat dua macam buku panduan yaitu panduan umum dan khusus. Menurut Undang-undang tentang sistem pembukuan, menjelaskan bahwa ada dua jenis buku yaitu buku pendidikan dan buku umum (Republik Indonesia, 2017). Berdasarkan klasifikasi di atas maka jenis panduan atau buku panduan yang dikembangkan adalah buku umum.

Menurut Yusuf dalam Permata & Marlina (2017), buku pedoman dibedakan menjadi dua antara lain (1) pedoman umum, adalah salah satu jenis referensi yang berbentuk kamus, ensiklopedia, dan berbagai pembahasan berbagai disiplin ilmu; (2) pedoman khusus, adalah jenis referensi yang terbatas pada disiplin ilmu tertentu dan bersifat terbatas. Jadi, panduan program analisis soal adalah buku panduan (*manual book*) yang berisi teknis, prosedur, maupun langkah-langkah dalam menjalankan program analisis soal.

#### **a. Fungsi Panduan Program Analisis Soal**

Menurut Wiji Suwarno dalam Muhammad Jabbar (2019), handbook merupakan buku referensi yang berisi intisari pokok bahasan tertentu berkenaan dengan ilmu pengetahuan tertentu berupa pengajaran atau praktik. Handbook juga merupakan perpaduan dari berbagai berita yang disusun tepat dalam bidang tertentu. Buku pegangan dan manual sering kali diturunkan dan mengkodifikasi teks dan tulisan lain, misal coretan awal seperti protokol atau

resep, daftar dan tabel, komentar, sejumlah artikel jurnal singkat, buku catatan lab, bahkan mungkin buku harian (Creager et al., 2020).

Mengacu pada pengertian dari *handbook* di atas, maka buku panduan atau panduan program analisis soal memiliki fungsi memberikan informasi, instruksi, memberi solusi, memberi harapan pengguna buku panduan, terarah, dan sesuai petunjuk sehingga tercapai tujuan yang diinginkan

#### **b. Syarat Panduan Program Analisis Soal**

Menurut Rahayuningsih dalam Lestari & Rukiyah (2016) bahwa ada beberapa jenis buku panduan, antara lain: (1) *Handbook*, adalah buku panduan yang berisi deskripsi singkat pokok bahasan tertentu dan digunakan sebagai petunjuk teknis cara kerja alat tertentu; (2) buku petunjuk, berisi petunjuk teknis terhadap cara kerja suatu alat tertentu.

Bersumber dari argumen di atas dapat diambil kesimpulan bahwa buku panduan adalah buku teks dimana terdapat unsur-unsur tertentu di dalamnya, yaitu.

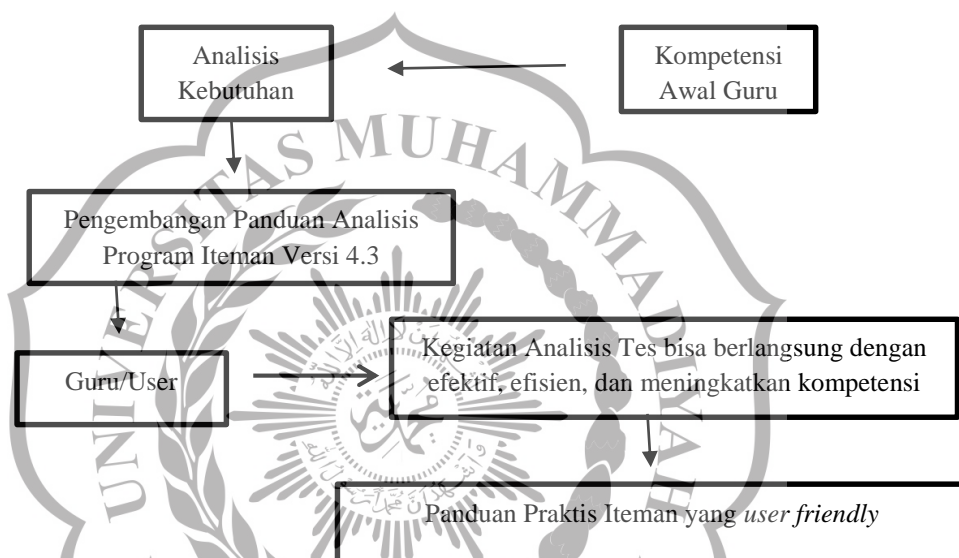
- a. Terdapat petunjuk
- b. Terdapat instruksi
- c. Masuk dalam kategori buku pintar
- d. Menjadi referensi
- e. Mengandung prosedur atau aturan pada materi
- f. Terdapat trik dan tips yang harus dilakukan
- g. adanya material dan cara kerja
- h. menjadi solusi pada bidang tertentu.

### C. Kerangka Pikir

Kegiatan analisis merupakan hal rutin yang dikerjakan guru. Kegiatan tersebut dilakukan dengan tujuan untuk mengukur kualitas soal yang telah dibuat atau disusun. Adapun maksud lainnya dari kegiatan analisis item soal adalah untuk mengukur tingkat kesuksesan kegiatan belajar di kelas yang telah dilakukan melalui eksamen yang diujikan. Kegiatan analisis soal yang dilakukan oleh guru cenderung masih menggunakan rumus manual serta hanya terbatas pada satu media tertentu, sehingga waktu pengerjaan analisis soal menjadi lebih lama. Hal inilah yang menjadi penyebab kegiatan analisis menjadi jarang dilakukan oleh guru. Keterbatasan media yang digunakan dalam menganalisis soal juga menjadi kendala tersendiri bagi guru, media yang digunakan hanya mampu menghasilkan angka/skor hasil analisis sehingga guru masih harus menafsirkan sendiri maksud dari hasil analisis yang masih berupa angka/skor. Perihal penafsiran soal bisa terselesaikan dengan menggunakan program IteMan versi 4.3 yang merupakan sebuah aplikasi yang digunakan dalam kegiatan analisis soal secara kuantitatif, kemampuan guru Pendidikan Agama Islam di Kabupaten banyumas dalam mengoperasikan komputer sudah dalam kategori lancar sehingga peneliti mengambill asumsi bahwa program IteMan versi 4.3 akan mudah mereka digunakan. IteMan versi 4.3 sebagai program analisis soal yang belum familiar bagi guru sehingga memerlukan panduan cara menjalankan aplikasi tersebut. Panduan analisis program IteMan versi 4.3 yang tersedia masih menggunakan bahasa Inggris dan bahasa pemrograman sehingga menjadi kendala tersendiri dalam menggunakannya. Hal demikian yang menjadikan alasan bagi peneliti

untuk mengembangkan produk panduan analisis program Iteman Versi 4.3 yang *User friendly* bagi guru.

Berdasarkan eksplorasi pemikiran penulis, hal yang menjadi fokus dalam riset ini adalah menganalisis prosedur pengembangan panduan analisis program Iteman versi 4.3. Berikut kami sajikan diagram kerangka pikir untuk mempermudah dalam menjelaskan prosedur pengembangan dalam penelitian ini.



**Gambar 1. Kerangka Pikir**

#### D. Pertanyaan Penelitian

1. Bagaimana proses pengembangan Panduan analisis program Iteman versi 4.3 menginterpretasikan kemampuan guru dalam menganalisis tes menggunakan Iteman.
2. Bagaimana kelayakan panduan analisis program Iteman versi 4.3 menginterpretasikan kapasitas guru dalam kegiatan analisis item tes dengan menggunakan program Iteman versi 4.3.
3. Bagaimana keefektifan panduan analisis program Iteman versi 4.3 yang dikembangkan untuk mendukung guru dalam menganalisis tes.