

BAB II KAJIAN TEORI

A. Deskripsi Teori

1. Tinjauan tentang Penilaian, Pengukuran dan Evaluasi

a. Pengertian Pengukuran dan Penilaian

Pengukuran, penilaian dan evaluasi memiliki kaitan yang erat antara satu dengan yang lainnya. Pengukuran adalah suatu proses kegiatan dalam rangka untuk menentukan kuantitas dari sesuatu (Zainal Arifin, 2012:7). Sesuatu yang dimaksudkan dapat berarti peserta didik, guru, sarana-prasarana sekolah, dan sebagainya. Selain itu pengukuran juga dapat berarti pemberian angka atau usaha dalam memperoleh deskripsi numerik dari suatu tingkatan dimana seorang peserta didik telah mencapai karakteristik tertentu (Siahaan:2010).

Menurut Ahmann dan Glock dalam Zainal (2012:7) pengukuran menyediakan informasi dimana evaluasi dapat didasarkan. Selanjutnya penilaian pendidikan adalah proses yang dilakukan untuk mendapatkan representasi secara kuantitatif tentang sejauh mana suatu ciri yang dimiliki oleh peserta didik. Pendapat yang sama juga disampaikan oleh Kizlik (2012:3), karena selain definisi umum tersebut, pengukuran mengacu pada serangkaian prosedur dan prinsip-prinsip dalam tes dan penilaian pendidikan.

Berdasarkan beberapa pemaparan para ahli tersebut maka dapat disimpulkan apabila pengukuran merupakan suatu proses dalam mengumpulkan informasi yang nantinya akan menghasilkan suatu angka secara kuantitatif.

Pengukuran itu sendiri juga tidak terlepas dari penilaian. Menurut Kizlik (2012:3) penilaian adalah suatu proses pengumpulan informasi untuk memantau kemajuan dan membuat keputusan bila dirasa perlu. Suatu penilaian tidak hanya mencakup tes saja, tapi juga bisa mencakup beberapa metode seperti observasi, wawancara, pengamatan perilaku dan sebagainya. Pendapat yang sama juga disampaikan oleh Zainal Arifin (2012:8), dilihat dari hubungannya dengan proses dan hasil belajar, penilaian adalah suatu proses yang secara sistematis dan berkesinambungan dalam mengumpulkan informasi tentang proses dan hasil belajar peserta didik dalam rangka membuat keputusan-keputusan berdasarkan kriteria dan pertimbangan tertentu.

Penilaian juga turut melayani kebutuhan di semua tingkatan hirarki pendidikan, misalnya, penilaian membantu guru untuk menetapkan standar, memotivasi kinerja, memberikan umpan balik, menilai atau mengevaluasi kemajuan dan mengkomunikasikan kemajuan terhadap pihak terkait (Joan L. Herman, 1992:2). Hasil dari penilaian nantinya dapat berupa angka (setelah melewati proses pengukuran) dan dapat digunakan untuk mengetahui kelemahan dari peserta didik yang memerlukan bantuan lebih lanjut.

b. Pengertian Evaluasi

Secara garis besar, evaluasi merupakan suatu proses yang disengaja dan terencana untuk memperoleh informasi atau data. Kemudian berdasarkan data yang sudah diperoleh tersebut dibuatlah sebuah keputusan (Purwanto, 2017: 3). Pengertian lain dari evaluasi juga dikemukakan oleh Cronbach dan Stufflebeam dalam Suharsimi (2016: 3) bukan hanya digunakan untuk mengukur sejauh mana

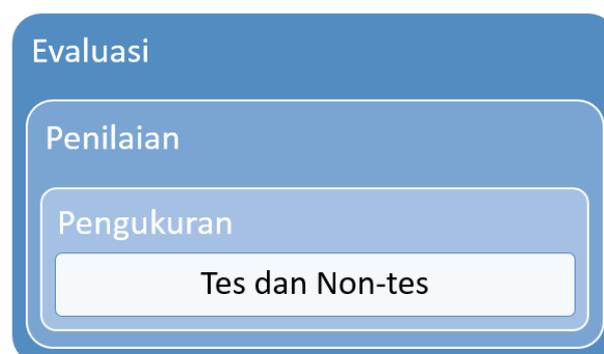
tujuan tercapai tetapi juga sebagai pengambilan keputusan merupakan tujuan dari proses evaluasi. Pengertian evaluasi menurut pendapat para ahli tersebut bukan hanya mengenai proses dan tujuan yang tercapai, akan tetapi juga mengenai kesimpulan serta pengambilan keputusan untuk proses berikutnya.

Hubungan antara evaluasi dengan pembelajaran, menurut Bob Kizlik (2012:3) evaluasi merupakan prosedur yang digunakan untuk menentukan apakah peserta didik sudah memenuhi kriteria yang ditentukan karena evaluasi sendiri bertujuan untuk membuat penentuan kualifikasi sesuai dengan kriteria. Norman E Grounland dalam Purwanto (2017: 3) merumuskan jika evaluasi merupakan suatu proses yang sistematis dan bertujuan untuk menentukan atau memutuskan sejauh mana tujuan-tujuan dari proses pengajaran telah dicapai oleh peserta didik. Martubi (2004: 12) menjelaskan hal yang serupa bahwa evaluasi pembelajaran merupakan suatu proses untuk memperoleh informasi secara menyeluruh dan kontinyu tentang suatu proses dan hasil belajar yang telah dilakukan oleh siswa sehingga dapat dijadikan acuan untuk penentuan perlakuan lanjut. Menurut UU Nomor 20 tahun 2003 pasal 1 ayat 21, evaluasi pendidikan merupakan suatu kegiatan pengendalian, penjaminan dan penetapan mutu pendidikan di setiap jenjang, jalur dan jenis pendidikan pada setiap komponennya sebagai bentuk pertanggungjawaban dalam pelaksanaan pendidikan.

Berdasarkan pemaparan para ahli dan juga UU Nomor 20 tahun 2003 maka didapatkan kesimpulan jika evaluasi merupakan suatu rangkaian proses yang dilaksanakan dengan cara mengumpulkan informasi atau data untuk mengetahui seberapa jauh nilai-nilai dan tujuan sudah tercapai, dalam hal ini terutama pada

proses pembelajaran. Informasi data yang diambil sebagai bahan evaluasi adalah hasil belajar siswa sejak awal hingga akhir pembelajaran. Data yang telah terkumpul kemudian menjadi dasar untuk membuat kesimpulan serta penentu langkah selanjutnya yang akan digunakan oleh satuan pendidikan untuk proses atau sistem pembelajaran berikutnya.

Dalam pelaksanaannya, antara penilaian, pengukuran dan evaluasi memiliki keterkaitan. Penilaian dan pengukuran dalam prosesnya akan selalu berkesinambungan. Sementara itu antara penilaian dan evaluasi keduanya memiliki pengertian yang sama yaitu menilai atau menentukan nilai sesuatu, perbedaannya terletak pada ruang lingkup dan pelaksanaannya (Zainal Arifin, 2012:11). Evaluasi dan penilaian lebih bersifat komprehensif yang meliputi pengukuran, sedangkan tes adalah salah satu alat pengukuran. Hubungan antara evaluasi, penilaian, pengukuran dan tes dapat dilihat pada gambar berikut ini.



Gambar 1. Hubungan Evaluasi-Penilaian-Pengukuran dan Tes

Dari gambar berikut maka dapat diartikan apabila evaluasi pembelajaran adalah suatu proses atau kegiatan yang sistematis, berkelanjutan dan menyeluruh dalam rangka pengendalian, penjaminan dan penetapan kualitas pembelajaran terhadap berbagai komponen pembelajaran. Sedangkan penilaian hasil belajar

adalah suatu proses atau kegiatan yang sistematis, berkelanjutan dan menyeluruh dalam rangka pengumpulan dan pengolahan informasi untuk menilai pencapaian proses dan hasil belajar peserta didik (Zainal Arifin, 2012:8).

c. Tujuan dan Fungsi Evaluasi Hasil Belajar

Tujuan dan fungsi dari evaluasi adalah dua hal yang selalu berkaitan. Menurut Rusydi dan Tien (2017: 7) tujuan evaluasi memiliki dua fungsi, yaitu fungsi formatif dan fungsi sumatif. Fungsi sumatif yaitu evaluasi digunakan untuk keterangan, pertanggungjawaban, seleksi atau lanjutan sedangkan fungsi formatif dari evaluasi yaitu evaluasi digunakan untuk pengembangan dan perbaikan kegiatan yang sedang berlangsung. Dalam batasan mengenai evaluasi hasil belajar, menurut Purwanto (2017: 5) mendapatkan bukti data yang menunjukkan hasil tingkat keberhasilan dan kemampuan peserta didik dalam mencapai tujuan belajar merupakan tujuan dari kegiatan evaluasi itu sendiri. Menurut Ngalim Purwanto (2017:5) evaluasi dikelompokkan menjadi empat fungsi. Fungsinya yang pertama adalah untuk mengetahui perkembangan dan kemajuan serta keberhasilan peserta didik setelah melaksanakan kegiatan pembelajaran selama jangka waktu tertentu. Setelah itu fungsi ke dua evaluasi adalah sebagai tolak ukur untuk mengetahui seberapa besar tingkat keberhasilan dari program pengajaran yang dilaksanakan. Kemudian sebagai bahan keperluan bimbingan dan konseling merupakan fungsi evaluasi yang ke tiga. Selanjutnya fungsi evaluasi yang terakhir adalah untuk keperluan perbaikan dan pengembangan kurikulum yang digunakan oleh sekolah yang bersangkutan.

Martubi (2004: 12) menjelaskan bahwa evaluasi pembelajaran memiliki fungsi yang strategis dan terbagi menjadi tiga, yaitu:

1) Fungsi evaluasi pembelajaran bagi Siswa

Bagi siswa, evaluasi berfungsi untuk mengetahui kemajuan belajar yang selama ini sudah dilaksanakan. Selain itu juga sebagai sarana untuk memberikan dorongan dan pengalaman belajar siswa.

2) Fungsi evaluasi bagi Guru

Bagi guru, evaluasi berfungsi untuk menyeleksi siswa dan memperkirakan keberhasilan dari siswa tersebut. Mengetahui penyebab kesulitan belajar siswa juga dibersamai dengan memberi bimbingan supaya siswa tersebut mampu mengatasi kesulitan belajar yang dialami. Dalam pemberian pedoman dalam mengajar, ketepatan metode mengajar dan juga mengelompokkan peserta didik dalam kelas yang sesuai dengan tingkat kepandaianya juga menjadi fungsi evaluasi bagi guru.

3) Fungsi evaluasi bagi Lembaga/ Organisasi Pendidikan

Dengan adanya fungsi evaluasi, lembaga atau organisasi pendidikan juga mendapatkan fungsinya yaitu untuk mempertahankan standar mutu pendidikan yang sudah ada, menilai ketepatan kurikulum dan juga untuk menilai kemajuan yang sudah dicapai oleh sekolah.

Menurut Peraturan Menteri Pendidikan dan Kebudayaan Nomor 23 Tahun 2016, Standar Penilaian Pendidikan pada pasal 4 menjelaskan mengenai tujuan dari penilaian, yaitu:

- 1) Penilaian hasil belajar oleh pendidik bertujuan untuk memantau dan mengevaluasi proses, kemajuan belajar, dan perbaikan hasil belajar peserta didik secara berkesinambungan.
 - 2) Penilaian hasil belajar oleh satuan pendidikan bertujuan untuk menilai pencapaian Standar Kompetensi Lulusan untuk semua mata pelajaran.
 - 3) Penilaian hasil belajar oleh Pemerintah bertujuan untuk menilai pencapaian kompetensi lulusan secara nasional pada mata pelajaran tertentu.
- Suharsimi (2016: 18-19) menjelaskan tentang tujuan dan fungsi evaluasi

hasil belajar, yaitu:

- 1) Penilaian berfungsi selektif

Dengan dilakukannya penilaian, guru dapat menyeleksi atau menilai peserta didiknya. Penilaian ini mempunyai tujuan untuk memilih dan memilah penerimaan peserta didik, kenaikan kelas peserta didik, pemberian beasiswa kepada peserta didik dan juga memilih peserta didik yang dinyatakan lulus.

- 2) Penilaian berfungsi diagnostik

Dengan melihat hasil dari penilaian siswa maka guru dapat mengetahui kelemahan dan kelebihan siswa tersebut. diketahuinya kelemahan dan penyebabnya, maka secara tidak langsung dalam pelaksanaan penilaian guru melakukan diagnosis untuk kemudian mencari cara dalam mengatasi kelemahan tersebut.

- 3) Penilaian berfungsi sebagai penempatan

Fungsi penempatan dari penilaian digunakan untuk dengan benar menentukan kelompok mana peserta didik tersebut ditempatkan. Penempatan pengelompokan peserta didik dilakukan dengan menempatkan siswa yang memiliki hasil penilaian belajar yang sama pada kategorinya.

4) Penilaian berfungsi sebagai pengukur keberhasilan

Sebagai pengukur keberhasilan artinya adalah penilaian dapat digunakan untuk mengetahui sejauh mana tujuan dari suatu program tersebut berhasil dilakukan. Dalam hal ini pendidikan atau pembelajaran harus dilakukan evaluasi, supaya tujuan dari proses pembelajaran tersebut dapat diketahui, apakah sudah sesuai atau masih gagal dalam mencapai target sehingga dapat diketahui penyebabnya dan kemudian dilakukan perbaikan.

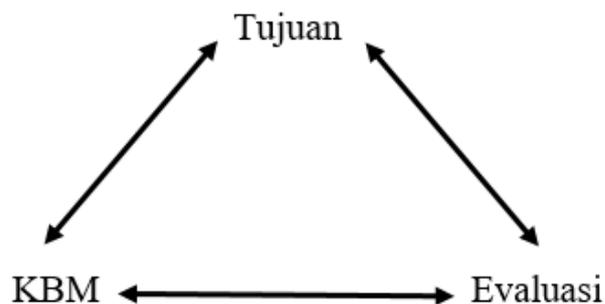
Dapat disimpulkan apabila tujuan dari evaluasi hasil belajar adalah untuk mengetahui seberapa jauh keberhasilan dari tujuan pembelajaran tersebut dapat dicapai oleh para peserta didik dalam jangka waktu satu periode tertentu. Sehingga dengan diketahuinya keberhasilan tersebut, maka guru mampu untuk mengambil suatu keputusan. Keputusan yang diambil oleh guru didasarkan pada hasil evaluasi yang menampilkan kelemahan dan kelebihan dari peserta didik yang juga berhubungan erat dengan materi dan metode pembelajaran yang digunakan. Keputusan yang diambil oleh guru juga akan menentukan keberhasilan pembelajaran berikutnya.

d. Prinsip-prinsip Evaluasi Hasil Belajar

Menurut Permendikbud Nomor 23 Tahun 2016 yang tercantum pada pasal 5 terdapat sembilan prinsip penilaian atau evaluasi hasil belajar. Prinsip-prinsip yang harus dimiliki pada penilaian hasil belajar antara lain sah yang berarti berdasarkan pada kemampuan yang akan diukur, objektif dan adil yang berarti tidak terpengaruh oleh subjektivitas penilai serta tidak merugikan atau pun menguntungkan peserta didik karena beberapa alasan. Prinsip lainnya yang harus

dimiliki adalah terpadu, terbuka, menyeluruh dan berkesinambungan yang berarti penilaian merupakan kesatuan komponen yang tidak terpisahkan dari pembelajaran, proses penilaiannya juga dapat diketahui oleh beberapa pihak yang berkepentingan. Penilaian ini juga mencakup keseluruhan aspek kompetensi dan menggunakan teknik penilaian yang sesuai untuk memantau dan mengetahui perkembangan kemampuan peserta didik. Kemudian prinsip yang harus dimiliki adalah akuntabel, sistematis dan beracuan pada kriteria.

Martubi (2004: 13) menyampaikannya supaya evaluasi pembelajaran dapat digunakan dan berfungsi dengan sebagai mana mestinya, maka evaluasi tersebut harus memiliki prinsip-prinsip, antara lain adalah menyeluruh/komprehensif, berkesinambungan, objektif, valid, reliable dan edukatif. Menurut Suharsimi (2016: 38) terdapat satu prinsip umum dan penting dalam evaluasi, yaitu triangulasi atau hubungan erat tiga komponen. Komponen yang termasuk dalam triangulasi yaitu tujuan pembelajaran, kegiatan pembelajaran atau KBM dan evaluasi. Triangulasi tersebut dapat digambarkan seperti berikut:



Gambar 2. Triangulasi Komponen Evaluasi

Penjelasan lebih lanjut dari gambar diatas pada Suharsimi (2016: 39) adalah:

1) Hubungan antara tujuan dengan KBM

Sebelum melaksanakan kegiatan atau proses belajar mengajar penting untuk adanya perencanaan mengajar yang disusun oleh guru. Dalam bagan Triangulasi ini, anak panah yang menunjuk hubungan antara KBM dengan tujuan belajar yang bermakna jika KBM mengacu pada tujuan. Akan tetapi pada bagan, panah juga mengarah dari tujuan ke KBM yang artinya dalam penyusunan tujuan mengacu pada tujuan proses belajar.

2) Hubungan antara tujuan dengan evaluasi

Evaluasi merupakan suatu proses pengumpulan data yang bertujuan untuk mengetahui sejauh mana tujuan belajar sudah tercapai. Oleh karena itu dalam penyusunan alat dan teknik evaluasi harus didasarkan pada tujuan yang telah direncanakan.

3) Hubungan antara KBM dengan evaluasi

Evaluasi yang dilakukan seharusnya tidak hanya berorientasi pada tujuannya saja, melainkan juga disesuaikan dengan proses KBM. Misalnya dalam proses pembelajaran yang dilaksanakan guru lebih berorientasi pada aspek keterampilan, maka evaluasi yang dibuat juga harus mengukur aspek keterampilan dari peserta didik.

Didapat dari diktat perkuliahan Penilaian Pembelajaran Kejuruan yang diampu oleh Edy Supriyadi, prinsip-prinsip dasar evaluasi yang harus dimiliki antara lain:

- 1) Memiliki tujuan yang jelas
- 2) Sesuai dengan tujuan penggunaannya
- 3) Dapat digunakan untuk meningkatkan belajar peserta didik

- 4) Menunjukkan hubungan yang erat antara prestasi atau hasil pembelajaran dengan substansi yang dinilai
- 5) Memiliki keandalan setinggi mungkin dan diinterpretasikan dengan cermat
- 6) Memiliki validitas yang memadai
- 7) Mencakup sampel tugas-tugas pembelajaran secara representatif

Berdasarkan pemaparan beberapa para ahli dan Permendikbud diatas,

secara garis besar prinsip evaluasi merupakan acuan atau nilai yang seharusnya diterapkan dan dilaksanakan oleh guru dalam melaksanakan kegiatan evaluasi yang bertujuan untuk mengetahui seberapa jauh pemahaman peserta didik terhadap proses pembelajaran yang telah dilakukan sudah tercapai.

e. Langkah-langkah Evaluasi Hasil Belajar

Menurut Sudijono (2012: 59) dalam melaksanakan kegiatan evaluasi hasil belajar dapat dilaksanakan dengan enam langkah pokok, yaitu:

- 1) Menyusun rencana evaluasi hasil belajar

Menyusun rencana evaluasi dilakukan untuk merumuskan tujuan dilaksanakannya evaluasi. Sebelum melaksanakan evaluasi, penting untuk menetapkan aspek apa saja yang akan dinilai, memilih dan menentukan teknik evaluasi apa yang akan digunakan, menyusun alat pengukur evaluasi, menetapkan tolak ukur dalam menginterpretasikan data hasil evaluasi, serta menentukan frekuensi dari evaluasi.

- 2) Menghimpun data evaluasi hasil belajar

Langkah selanjutnya setelah penyusunan rencana evaluasi hasil belajar adalah dengan mengumpulkan atau menghimpun data dan melaksanakan pengukuran.

3) Melakukan verifikasi data evaluasi hasil belajar

Setelah data terhimpun kemudian melakukan verifikasi dari data yang sudah terkumpul dengan tujuan untuk mengetahui keabsahan data yang akan dievaluasi.

4) Mengolah dan menganalisis data evaluasi hasil belajar

Data yang telah terverifikasi selanjutnya siap untuk diolah dan dianalisis agar memperoleh hasil yang baik.

5) Memberikan interpretasi dan menarik kesimpulan

Data yang telah selesai diolah dan dianalisis kemudian diinterpretasikan dan selanjutnya dapat ditarik sebuah kesimpulan.

6) Tindak lanjut hasil evaluasi

Berdasarkan pada data evaluasi yang telah disusun, direncanakan, dihimpun, diolah, dianalisis dan disimpulkan maka pada akhirnya evaluator akan mampu untuk mengambil sebuah keputusan yang selanjutnya merumuskan kebijakan atau langkah yang dirasa perlu sebagai suatu bentuk tindak lanjut dari kegiatan evaluasi yang telah dilaksanakan.

Permendikbud Nomor 23 tahun 2016 mengenai prosedur penilaian yang terdapat dalam pasal 12, penilaian aspek pengetahuan dilakukan melewati beberapa tahapan, yaitu:

- 1) Menyusun perencanaan penilaian;
- 2) Mengembangkan instrumen penilaian;
- 3) Melaksanakan penilaian;
- 4) Memanfaatkan hasil penilaian; dan
- 5) Melaporkan hasil penilaian dalam bentuk angka dengan skala 0-100 dan deskripsi.

f. Teknik Evaluasi Hasil Belajar

Terdapat beberapa teknik evaluasi hasil belajar yang bisa digunakan oleh guru dalam mengevaluasi peserta didiknya. Ruhimat, dkk (2013: 56-57) menyampaikan jika teknik evaluasi dapat dikelompokkan kedalam dua jenis, yaitu:

1) Teknik tes

Tes biasanya difungsikan guna mengukur tingkat kemampuan peserta didik pada aspek kognitif (pengetahuan) atau tingkat penguasaan materi yang hasil dari teknik tes ini biasanya diolah secara kuantitatif. Pelaksanaan proses tes hasil belajar dilakukan ketika telah menyelesaikan satu pokok bahasan.

2) Teknik non-tes

Teknik penilaian non-tes biasanya digunakan untuk mengukur tingkat kemampuan peserta didik pada aspek tingkah laku termasuk minat, sikap dan motivasi.

Anas Sudijono (2012: 65) juga menyampaikan hal yang sama mengenai teknik dalam mengevaluasi hasil belajar, yaitu teknik tes dan non-tes. Teknik tes merupakan cara atau prosedur yang perlu dilalui dalam rangka penilaian dan pengukuran bidang pendidikan. Tes dapat berupa pemberian tugas atau serangkaian tugas yang menghasilkan nilai untuk melambangkan prestasi atau tingkah laku *testee*. Kemudian teknik non-tes melakukan penilaian atau evaluasi hasil belajar peserta didik tanpa menguji pesesrta didik. Cara yang dilakukan pada teknik ini adalah dengan melakukan pengamatan secara sistematis (*observation*), melakukan wawancara (*interview*), angket (*questionnaire*), dan memeriksa atau menilai dokumen (*documentary analysis*).

2. Tinjauan tentang Tes Sebagai Alat Evaluasi Belajar

a. Pengertian Tes

Menurut Suharsimi (2016: 67) tes merupakan alat atau prosedur yang dipakai guna mengetahui atau mengukur sesuatu dalam suasana, cara dan beberapa aturan sebelumnya sudah ditentukan. Bachman (1990) mengamati bahwa tes adalah salah satu jenis instrumen pengukuran oleh karena itu tes mengukur karakteristik individu sesuai dengan prosedur. Sedangkan F.L Goodenough dalam Anas Sudijono (2012: 67) menyampaikan jika tes merupakan sesuatu atau serangkaian tugas yang diberikan kepada individu atau sekelompok individu yang memiliki tujuan untuk membandingkan kompetensi mereka antara satu dengan yang lainnya.

Ngalim Purwanto (2017: 33) menyampaikan apabila tes hasil belajar atau *achievement test* digunakan untuk menilai hasil-hasil pelajaran yang telah diberikan oleh pengajar atau guru terhadap para peserta didiknya, atau dosen terhadap mahasiswanya pada jangka waktu tertentu.

Anas Sudijono (2012: 67) menyebutkan tes merupakan pemberian tugas kepada *testee* yang harus dikerjakan dalam suatu cara atau prosedur pengukuran dan penilaian untuk memperoleh data yang menghasilkan nilai untuk menunjukkan prestasi atau tingkah laku *testee*. Selanjutnya nilai yang telah didapatkan dapat dibandingkan dengan standar nilai tertentu.

Berdasarkan pemaparan para ahli tersebut maka dapat disimpulkan jika tes merupakan suatu alat yang digunakan dalam evaluasi guna mengetahui kemampuan atau kompetensi peserta didik. Pelaksanaannya dilakukan dengan

memberikan tugas kepada peserta didik yang kemudian hasilnya menjadi acuan untuk memberikan nilai dalam periode waktu tertentu.

b. Fungsi Tes Hasil Belajar

Menurut Anas Sudijono (2012: 67) sebagai alat pengukur serta alat untuk mengukur keberhasilan peserta didik pada program pembelajaran yang telah terlaksana merupakan fungsi dari tes hasil belajar. Selain itu menurut Suharsimi (2016: 165-166) fungsi dari tes dapat ditinjau dari tiga hal, yaitu:

- 1) Bagi kelas, berguna untuk mendiagnosis kesulitan belajar, peningkatan prestasi dan pengelompokan peserta didik yang selanjutnya dilakukan bimbingan.
- 2) Untuk bimbingan berarti tes berfungsi mengarahkan peserta didik maupun orang tua peserta didik dalam hal mengatasi masalah kesulitan belajar yang dialami agar mencapai tujuan dan mampu menentukan pilihan peserta didik selanjutnya.
- 3) Untuk administrasi artinya tes berfungsi guna keperluan seleksi peserta didik baru, perbaikan pembelajaran dan kurikulum yang berlaku serta memberikan dan menyediakan informasi kepada para pihak yang terkait.

c. Bentuk Tes Hasil Belajar

Bentuk dari tes hasil belajar mempengaruhi hasil akhir dari evaluasi yang diberikan kepada peserta didik. Secara umum, bentuk tes hasil belajar dapat terbagi menjadi dua, yaitu:

1) Ditinjau dari segi kegunaan

a) Tes formatif

Anas Sudijono (2012: 71) menyampaikan tes formatif adalah tes hasil belajar yang bertujuan guna mengetahui seberapa jauh peserta didik telah terbentuk sesuai dengan tujuan pengajaran yang ditentukan setelah mengikuti proses pembelajaran dalam jangka waktu tertentu. Sementara itu Ngalim Purwanto (2017: 26) menjelaskan penilaian formatif adalah penilaian yang bertujuan mencari umpan balik (*feedback*). Kemudian hasilnya dapat digunakan dalam memperbaiki proses belajar- mengajar yang sedang atau telah dilaksanakan.

Tes formatif dapat dilakukan pada saat proses pembelajaran sedang berlangsung dan berakhirnya satu pokok bahasan tertentu. Hasil dari tes formatif ini dapat menjadi acuan tindak lanjut, apabila hasil tes menunjukkan penguasaan materi sudah baik maka dapat melanjutkan ke materi berikutnya, namun jika hasil yang diperoleh sebaliknya maka ada bagian-bagian yang dapat diulang atau dijelaskan kembali kepada peserta didik.

b) Tes sumatif

Anas Sudijono (2012: 72) mengemukakan tes sumatif merupakan tes hasil belajar yang dilaksanakan setelah serangkaian atau sekumpulan satuan program pengajaran selesai diberikan. Menurut Ngalim Purwanto (2017: 26-27) penilaian sumatif adalah penilaian yang dilaksanakan untuk memperoleh data atau informasi mengenai penguasaan atau pencapaian belajar siswa terhadap bahan pelajaran yang telah dipelajari selama jangka waktu tertentu.

Tujuan dan fungsi utama dari tes sumatif adalah untuk menentukan pencapaian hasil akhir dari pembelajaran dan penentuan nilai akhir pada periode waktu tertentu, sehingga dapat untuk menentukan kedudukan peserta didik di dalam kelompok dan laporan kemajuan peserta didik kepada orang tua atau pihak terkait lainnya.

c) Tes diagnostik

Anas Sudijono (2012: 70) menuturkan jika tes diagnostik merupakan tes yang dilakukan guna menentukan dengan akurat dan tepat jenis kesulitan apa yang dialami oleh peserta didik dalam proses pembelajaran yang berlangsung pada suatu mata pelajaran tertentu. Sejalan dengan hal tersebut, didapatkan dari diktat perkuliahan Penilaian Pembelajaran Kejuruan oleh Edy Supriyadi, tes diagnostik merupakan sebuah tes yang dilakukan untuk dapat menentukan penyebab-penyebab baik pada aspek intelektual, fisik, emosi, maupun lingkungan kesulitan belajar yang seringkali ditemui oleh peserta didik.

Pertanyaan pada tes ini biasanya termasuk dalam kategori sukar untuk dikerjakan atau dipahami, tujuannya adalah supaya guru dapat mengetahui kelemahan peserta didik. Apabila hasil tes menunjukkan termasuk dalam kategori rendah, maka guru dapat memutuskan untuk melakukan bimbingan khusus terhadap siswa tersebut untuk memperbaiki penguasaan materi belajar oleh peserta didik.

2) Ditinjau dari segi sistem penskoran

a) Tes subjektif

Suharsimi (2016: 177) menjelaskan yang dimaksud dengan tes subjektif adalah tes yang pada umumnya berbentuk esai atau uraian yang membutuhkan jawaban yang sifatnya berupa pembahasan atau uraian kata-kata. Tes uraian atau esai (*essay test*) atau yang seringkali juga dikenal dengan istilah tes subjektif (*subjective test*) merupakan salah satu jenis tes yang memiliki karakteristik menurut Anas Sudijono (2012: 99) berupa pertanyaan atau perintah yang jawabannya berupa uraian atau penjelasan kalimat yang pada umumnya cukup panjang. Karakteristik lainnya adalah bentuk pertanyaan atau perintah terhadap *testee* menuntut jawaban yang berupa pemberian penjelasan, penafsiran, komentar, membandingkan, membedakan dan lain sebagainya. Pada umumnya butir soal pertanyaan tes uraian diawali dengan kata-kata “Uraikan..”, “Jelaskan...”, “Mengapa...”, atau kata lain yang serupa dengan itu.

Menurut Suharsimi (2016: 178) tes subjektif juga memiliki beberapa kelebihan antara lain adalah mudah disiapkan dan disusun. Kelebihan lainnya memberi dorongan kepada peserta didik untuk mengemukakan pendapatnya dengan menggunakan kalimat yang bagus dan lugas dengan gaya bahasa mereka sendiri. Kemudian kelebihan soal tes subjektif berikutnya adalah tidak memberi banyak ruang untuk berspekulasi terhadap peserta didik sehingga dapat mengetahui seberapa jauh peserta didik tersebut mendalami masalah yang sedang diujikan.

Beberapa kekurangan yang dimiliki tes subjektif antara lain kadar reliabilitas dan validitas yang rendah sehingga kurang merepresentasikan pada seluruh bagian pelajaran yang diujikan. Kekurangan lainnya adalah banyaknya pengaruh subjektif pada saat pemeriksaan jawaban dan membutuhkan banyak pertimbangan individual dari penilai yang tidak dapat diwakilkan oleh orang lain sehingga membutuhkan waktu pemeriksaan yang lama.

Dapat disimpulkan apabila tes subjektif adalah tes yang mampu memberikan peserta didik kebebasan saat menjawab soal dalam bentuk uraian sesuai dengan perintah yang terdapat pada pertanyaan. Pemberi skor pada tes subjektif sangat mempengaruhi hasil penilaian.

b) Tes objektif

Menurut pendapat Suharsimi (2016: 179) untuk mengurangi kelemahan-kelemahan yang terdapat pada tes subjektif dalam pemeriksaannya maka dapat dengan menggunakan tes objektif sehingga penilaian dan pemeriksaan itu dilakukan secara objektif. Adapun menurut Anas Sudijono (2012: 106) tes objektif merupakan salah satu jenis tes yang menggunakan atau terdiri dari butir-butir soal (*items*) yang dapat dipilih oleh peserta tes. Terdapat empat macam tes objektif menurut Suharsimi (2016: 181-190) yaitu tes benar-salah (*true-false*), tes pilihan ganda (*multiple choice test*), menjodohkan (*matching test*), dan tes isian (*completion test*).

Kelebihan tes objektif yang dikemukakan oleh Suharsimi (2016: 180) adalah lebih mudah dan cepat pemeriksaannya karena menggunakan kunci jawaban tes bahkan pemeriksaan juga dapat dilakukan menggunakan alat hasil

kemajuan teknologi. Selain itu tes objektif juga lebih mampu dalam merepresentasikan isi dan dapat mengurangi unsur-unsur subjektif baik dari peserta didik maupun guru sebagai pemeriksa.

Kekurangan yang dimiliki oleh tes objektif antara lain banyak soal yang harus diteliti untuk mengurangi dan menghindari kelemahan-kelemahan sehingga dalam persiapan menyusunnya jauh lebih sulit daripada tes subjektif. Pada soal tes objektif sukar untuk mengukur proses mental yang tinggi karena lebih mengandalkan daya ingat dan pengenalan saja. Selain itu juga banyaknya kesempatan peserta tes untuk memperoleh jawaban secara untung-untungan dan bekerjasama dengan peserta tes lainnya.

d. Langkah-langkah Penyusunan Tes

Dalam merumuskan perencanaan ataupun penyusunan tes maka diperlukan langkah-langkah yang harus dilakukan. Suharsimi (2016: 167) menampakan terdapat beberapa langkah dalam penyusunan sebuah tes secara umum yaitu sebagai berikut:

- 1) Menentukan tujuan diadakannya sebuah tes.
- 2) Membuat batasan terhadap bahan atau materi yang akan dijadikan tes.
- 3) Merumuskan tujuan instruksional secara khusus pada tiap bagian.
- 4) Menyusun tabel persiapan yang didalamnya memuat pokok materi dan penulisan soal yang juga didasarkan pada aspek perilaku.

Sementara itu tidak jauh berbeda, Ngalim Purwanto (2017: 31) mengemukakan langkah-langkah penyusunan soal, yaitu sebagai berikut:

- 1) Menentukan dan atau merumuskan tujuan tes,

- 2) Mengidentifikasi hasil belajar (*learning outcomes*) apa saja yang akan diukur dari tes tersebut
- 3) Menentukan dan atau menandai hasil-hasil belajar secara spesifik.
- 4) Merinci mata pelajaran/ bahan pelajaran yang akan diukur.
- 5) Menyiapkan tabel spesifikasi.
- 6) Menggunakan tabel spesifikasi tersebut sebagai dasar penyusunan tes.

e. Ciri-ciri Tes yang Baik

Sebuah tes dapat dikatakan berfungsi dengan baik sebagai alat pengukur apabila mampu memenuhi beberapa persyaratan. Menurut Suharsimi (2016: 72) ciri tes yang baik yaitu bisa memenuhi persyaratan yang berupa validitas, reliabilitas, objektivitas, praktibilitas dan ekonomis. Adapun penjelasannya adalah sebagai berikut:

1) Validitas

Validitas artinya tes yang diberikan sudah sesuai atau tepat dengan fungsinya sebagai alat pengukur prestasi hasil belajar dari peserta didik.

2) Reliabilitas

Reliabilitas artinya tes yang digunakan dapat dipertanggungjawabkan keajegannya. Maksudnya adalah apabila tes diberikan atau diujikan beberapa kali akan tetap memberikan hasil yang relatif sama, oleh karena itu tes dapat dikatakan reliabel.

3) Objektivitas

Suatu tes dikatakan mempunyai objektivitas apabila tidak adanya unsur-unsur subjektif yang mempengaruhi soal pada saat dilaksanakan.

4) Praktibilitas

Praktibilitas pada suatu tes adalah apabila tes tersebut memiliki sifat yang praktis dan mudah pada proses pelaksanaan maupun pengadministrasiannya sehingga tes tersebut tidak memerlukan proses yang rumit. Tes yang praktis merupakan tes yang dalam pelaksanaannya dan pemeriksaannya mudah serta dilengkapi petunjuk yang jelas dan mudah untuk dipahami.

5) Ekonomis

Suatu tes dapat dikatakan memiliki nilai ekonomis apabila dalam pelaksanaannya tidak memerlukan banyak biaya, tenaga dan waktu yang lama.

3. Tinjauan tentang Analisis Butir Soal

a. Pengertian Analisis Butir Soal

Mengevaluasi tes hasil belajar peserta didik merupakan salah satu upaya efektif yang bermanfaat guna memperbaiki proses belajar mengajar. Menurut Suharsimi (2016: 220) analisis soal (*item analysis*) ialah suatu prosedur yang sistematis dan akan memberikan informasi yang spesifik dan khusus terhadap butir soal yang disusun.

Menurut Arifin (2016: 246) analisis kualitas tes adalah suatu tahapan yang harus dilakukan untuk mengetahui derajat kualitas tes, baik secara keseluruhan ataupun pada tiap butir soal yang menjadi bagian dari tes tersebut.

Berdasarkan pendapat dari kedua sumber maka dapat diambil kesimpulan bahwa analisis butir soal merupakan suatu proses yang tersusun secara sistematis,

bertujuan untuk mencari dan memastikan apabila soal-soal yang digunakan untuk tes merupakan soal yang berkualitas baik.

b. Manfaat Analisis Butir Soal

Manfaat dari analisis butir soal yang dilakukan adalah dapat meningkatkan kualitas butir soal selain validitas dan reliabilitas juga meliputi tiga aspek lainnya yaitu tingkat kesukaraan, daya beda dan efektivitas pengecoh. Dengan diadakannya analisis butir soal ini juga dapat mengetahui kualitas dari setiap butir soal yang diujikan apakah sudah berfungsi dengan baik atau belum sehingga nantinya dapat sebagai bahan acuan untuk melakukan perbaikan atau revisi. Perbaikan atau revisi dilakukan kepada soal-soal yang masih belum relevan dengan materi pengejaran, ditandai oleh banyaknya peserta didik yang tidak mampu untuk menjawab butir soal tersebut secara benar. Selain itu manfaat yang berikutnya adalah dapat memilih dan memilah kemudian menyimpan butir soal mana yang sudah memiliki kualitas baik ke bank soal untuk kemudian dapat digunakan kembali pada tes berikutnya dengan beberapa pengembangan.

c. Teknik Analisis Butir Soal

Tes diharapkan mampu menampilkan hasil yang objektif dan akurat dalam peranannya sebagai alat evaluasi hasil belajar. Analisis kualitas butir tes merupakan langkah yang harus dilaksanakan yang bertujuan untuk mengetahui derajat kualitas soal tersebut. Adapun pembuat soal pada analisis butir soal dapat melakukan dua jenis analisis, yaitu analisis secara teoritik (kualitatif) dan empirik (kuantitatif), yang penjelasannya adalah sebagai berikut:

1) Analisis teoretik

Analisis soal teoretik atau kualitatif menurut Mardapi (2007: 140) dilakukan sebelum pengujian dilaksanakan, yaitu dengan mencermati atau menelaah butir-butir soal yang telah disusun kemudian melihat kesesuaian butir soal dengan kemampuan dasar dan indikator soal yang akan diukur. Pada analisis secara teoretik juga melihat apakah soal tersebut sudah memenuhi prasyarat yang dilihat dari aspek materi, bahasa dan konstruksi.

Dalam analisis butir soal secara teoretik pada soal pilihan ganda terdapat kaidah penulisan soal yang merupakan petunjuk atau pedoman yang harus digunakan. Kaidah penulisan ini digunakan oleh penulis agar soal yang digunakan mempunyai kualitas atau mutu yang baik. Soal dapat dikatakan memiliki mutu yang baik apabila mampu untuk menjaring informasi yang dibutuhkan dalam soal secara optimal. Kaidah penulisan soal menurut BSNP (2010: 4-6) yang meliputi aspek isi atau materi, konstruksi soal dan bahasa adalah sebagai berikut:

a) Aspek materi

Validitas dari segi materi atau isi dilihat melalui hubungan antara isi dari tes dengan konstruk yang akan diukur. Isi tes mengacu terhadap tema, kata-kata, format butir, tugas atau pertanyaan pada tes (Mardapi, 2007:17). Soal yang baik adalah antara soal dengan indikator atau konstruk memiliki kesesuaian sehingga pada soal harus menanyakan materi dan perilaku sesuai dengan tuntutan yang terdapat dalam indikator. Alternatif jawaban harus bersifat homogen dan logis dilihat dari aspek materi. Alternatif jawaban juga seharusnya berasal dari materi yang sama seperti pada pokok soal. Penulisan alternatif jawaban yang setara juga harus dilakukan supaya seluruh alternatif

atau pilihan jawaban dapat berfungsi secara baik. Pada soal yang baik dalam segi materi, harus memiliki satu jawaban soal yang paling benar pada setiap soal sehingga akan hanya ada satu kunci jawaban di masing-masing soal.

b) Aspek konstruksi

Soal dapat dikatakan baik pada aspek konstruksi apabila dapat memenuhi prasyarat seperti berikut ini:

- Pokok soal dirumuskan secara jelas dan tegas, artinya materi yang ada pada soal harus ditanyakan secara jelas dan tidak mengandung pengertian ganda sehingga hanya mempunyai satu permasalahan pada tiap soal
- Rumusan soal dan pilihan atau alternatif jawaban harus berisi pernyataan yang hanya dibutuhkan saja
- Soal yang ditanyakan tidak memberi petunjuk apapun ke arah pilihan jawaban benar
- Pokok soal tidak boleh mengandung dua pernyataan negatif ganda, artinya tidak boleh sampai ada dua atau lebih kata yang bersifat negatif untuk mencegah terjadinya salah penafsiran soal oleh peserta tes
- Alternatif jawaban yang ada harus logis dan homogen dilihat dari segi materi, artinya semua alternatif jawaban harus sesuai dengan materi yang ditanyakan pada pokok soal, penulisannya harus setara dan harus berfungsi semuanya
- Alternatif jawaban tidak mengandung sebuah pernyataan. Seperti contohnya alternatif jawaban yang memiliki pernyataan “semua jawaban di atas benar” atau “semua jawaban di atas salah”

- Panjang alternatif jawaban harus relatif sama, artinya ada kecenderungan bagi peserta tes untuk memilih alternatif jawaban yang panjang karena seringkali dinilai lebih panjang, lebih lengkap ataupun sebaliknya dan dianggap merupakan kunci jawaban
- Alternatif jawaban yang berupa angka atau waktu harus disusun atau diurutkan berdasarkan urutan besar kecilnya nilai angka atau urutan kronologi waktu. Pengurutan angka dilakukan dari nilai yang terkecil ke nilai terbesar ataupun sebaliknya, dan pengurutan waktu berdasarkan kronologi waktu. Hal ini dilakukan guna mempermudah peserta tes dalam melihat alternatif jawaban
- Apabila terdapat gambar, tabel diagram atau grafik dan sejenisnya pada pokok soal haruslah jelas dan berfungsi dengan baik. Hal tersebut berarti apa saja yang terdapat pada soal harus ditanyakan dengan jelas, terbaca dan dapat dimengerti oleh peserta tes. Jika soal dapat terjawab tanpa harus melihat gambar, grafik atau tabel dan sejenisnya maka itu berarti gambar, tabel diagram atau grafik dan sejenisnya tidak berfungsi dalam soal tersebut
- Butir soal tidak tergantung dengan alternatif jawaban butir soal sebelumnya karena sangat dikhawatirkan apabila peserta tes yang tidak dapat menjawab soal sebelumnya dengan benar tidak akan menjawab soal berikutnya dengan benar

c) Aspek bahasa

Soal dapat dikatakan baik dilihat dari segi aspek bahasa apabila:

- Pada setiap butir soal harus menggunakan bahasa Indonesia yang sesuai dengan kaidah penulisan
- Bahasa yang digunakan harus komunikatif dan mudah dimengerti atau dipahami oleh peserta tes
- Tidak menggunakan bahasa daerah setempat yang berlaku apabila akan digunakan untuk beberapa daerah atau secara nasional

2) Analisis Empirik

Analisis soal secara empirik atau kuantitatif lebih menekankan pada analisis karakteristik tes melalui data yang telah didapatkan secara empiris dari soal yang sudah diujikan terhadap peserta didik. Data yang telah diperoleh kemudian ditampilkan dalam bentuk angka. Dua pendekatan yang digunakan pada analisis butir secara kuantitatif atau empirik antara lain melalui pendekatan secara teori tes klasik dan respon butir.

a) Teori tes klasik

Penelaahan butir soal melalui informasi yang didapatkan dari jawaban peserta didik yang bertujuan untuk meningkatkan kualitas butir soal tersebut merupakan pengertian analisis butir soal secara klasik. Dalam teori tes klasik menurut Smith (2016:6) skor total pada tes memberikan perkiraan kemampuan dari seseorang. Biasanya skor pada pengukuran dalam teori tes klasik biasanya adalah jumlah item dari peserta tes yang menjawab benar. Pada pengukuran tes klasik, Dali S Naga (1992: 4-6) menyebutkan jika kelompok butir yang sama akan diisi oleh kelompok peserta yang sama. Artinya jumlah skor akhir nantinya bergantung pada soal dan peserta tes.

Menurut Hambleton & Swaminathan (1985:310) prosedur analisis tes klasik sangat kuat, dan mudah dipahami. Selain itu menurut BSNP (2010:9) kelebihan dari analisis butir soal secara klasik adalah dapat dilakukan dengan cepat menggunakan komputer, mudah, sederhana, familiar dan dapat menggunakan data dari beberapa peserta didik atau sampel kecil. Pada analisis butir soal secara klasik terdapat beberapa aspek yang diperhatikan. Setiap butir soal akan dianalisis baik dari aspek reliabilitas, daya beda, tingkat kesukaran dan efektivitas pengecoh.

b) Teori respon butir

Teori respon butir pada analisis butir soal atau lebih dikenal dengan istilah *Item Response Theory* (IRT) pada dasarnya dipakai dalam memperbaiki kelemahan yang ada pada teori tes klasik. Kelemahannya yaitu ketergantungan ciri peserta pada kelompok butir tes (Dali, 1992: 160). Teori respon butir merupakan bentuk pengukuran modern yang biasanya digunakan pada analisis butir soal dengan menggunakan model matematis untuk menghubungkan karakteristik butir soal terhadap kemampuan responden yang digambarkan melalui kurva karakteristik butir soal.

Analisis butir dengan IRT dapat menggunakan model logistik satu, dua dan tiga parameter yang diterapkan pada tes kognitif. Menurut Hambleton dalam Firmanto (2016:65) nama model disesuaikan dengan banyaknya jumlah parameter yang digunakan. Parameter yang dimaksud adalah taraf kesukaran *item*, daya diskriminasi *item* dan tebakan semu. Hubungan antara perfomansi

responden terhadap *item* membentuk suatu fungsi yang disebut *Item Characteristic Curve* (ICC).

Persamaan ICC model logistik satu parameter dapat menggunakan rumus:

$$P_i(\theta) = \frac{e^{(\theta-b_i)}}{1 + e^{(\theta-b_i)}}$$

$P_i(\theta)$ merupakan probabilitas responden terpilih secara random, θ adalah responden yang menjawab *item i* dengan benar. Kemudian b_i merupakan tingkat kesukaran *item i*, dan e adalah nilai 2,718. Semakin besar nilai parameter b_i maka semakin besar abilitas yang diperlukan bagi responden menjawab *item* benar dengan peluang sebesar 50%. Parameter b_i merupakan titik pada kontinum abilitas (Hambleton:1991).

Selanjutnya persamaan ICC model logistik dua parameter dapat menggunakan rumus:

$$P_i(\theta) = \frac{e^{Dai(\theta-b_i)}}{1 + e^{Dai(\theta-b_i)}}$$

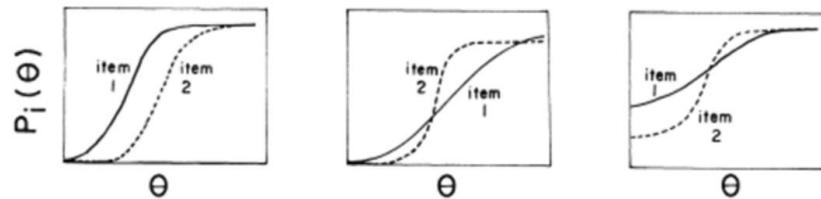
Persamaan tersebut menggunakan dua parameter, dimana $D = 1,7$ dan a_i merupakan parameter diskriminasi *item* sedangkan parameter lain yang digunakan adalah parameter b_i . Parameter diskriminasi adalah kemampuan *item* untuk membedakan responden dengan tingkat abilitas tinggi dan rendah. Grafik yang dihasilkan akan lebih curam apabila harga a_i menunjukkan angka yang tinggi.

Kemudian perumusan matematis untuk model logistik tiga parameter adalah sebagai berikut:

$$P_i(\theta) = c_i + (1 - c_i) \frac{e^{Dai(\theta - bi)}}{1 + e^{Dai(\theta - bi)}}$$

Terdapat parameter baru yaitu c_i atau *pseudochance level* yang akan menunjukkan probabilitas responden dengan kemampuan rendah menjawab *item* dengan benar (Hambleton:1991).

Berikut ini merupakan grafik *Item Characteristic Curve* (ICC) dua *item* pada model logistik Satu, Dua dan Tiga Parameter menurut Hambleton & Swaminathan (1985:28):



Gambar 3. ICC Dua *Item* Pada Model Logistik Satu, Dua dan Tiga Parameter

d. Analisis Kualitas Butir Soal

Untuk memperoleh soal dengan kualitas yang baik maka diperlukan beberapa tahapan atau proses yang harus dilalui. Tahapan yang harus dilalui antara lain adalah proses pembuatan soal yang harus sesuai dengan prosedur penyusunan soal diawali dari proses perencanaan, pembuatan, pelaksanaan sampai dengan proses evaluasi. Soal yang telah dibuat hendaknya dilakukan analisis karena dirasa perlu supaya memastikan apakah soal yang dibuat dan diujikan termasuk ke dalam

soal yang memiliki kualitas baik dan layak untuk dijadikan alat ukur penilaian bagi peserta didik. Analisis soal yang dapat dilakukan meliputi:

1) Validitas

Menurut Ngalim Purwanto (2017: 137) validitas atau kesahihan dapat diartikan sebagai kualitas yang menunjukkan hubungan antara tujuan kriteria belajar dengan pengukuran (diagnosis). Sementara itu Anderson dalam Suharsimi (2016: 80) menyebutkan apabila sebuah tes dapat dikatakan valid jika mampu mengukur apa yang hendak diukur.

Validitas suatu tes mengacu pada tingkat kebenaran penafsiran skor tes tersebut (Mardapi, 2007:17). Perlu diketahui, penafsiran skor ini berdasarkan pada tujuan tes tersebut dibuat. Menurut Cronbach dalam Mardapi (2007:17) sebenarnya bukan validasi tes yang menjadi tujuan, tetapi melakukan validasi terhadap interpretasi data yang terkumpul dari suatu prosedur tertentu. Ketepatan interpretasi hasil suatu tes berdasarkan pada bukti-bukti yang mendukung sesuai dengan tujuan dilakukannya tes. Bukti validitas suatu tes terbagi menjadi empat kelompok, yaitu bukti berdasarkan isi tes, proses respons, struktur internal dan hubungan terhadap variabel lain.

Berdasarkan beberapa pendapat tersebut maka dapat disimpulkan apabila validitas merupakan salah satu aspek atau syarat yang harus terpenuhi oleh sebuah tes dalam ketepatannya mengukur hal yang seharusnya diukur oleh tes tersebut.

Anas Sudijono (2012: 163) membagi validitas menjadi 2 macam, yaitu:

a) Validitas tes

Validitas tes terdiri dari 2 macam, yaitu validitas rasional dan validitas empirik. Validitas tes digunakan dalam mengukur tes secara keseluruhan. Menurut Anas Sudijono (2012: 164) validitas rasional dapat dimiliki oleh sebuah tes apabila tes tersebut secara rasional mampu untuk mengukur apa yang seharusnya diukur dengan tepat. Validitas rasional diperoleh melalui hasil pemikiran atau penalaran, atau dapat dikatakan diperoleh dengan berpikir logis.

Kemudian Anas Sudijono (2012: 167) mengemukakan validitas empirik adalah ketepatan mengukur yang didasarkan pada hasil analisis yang bersifat empirik. Apabila sebuah tes sudah diuji dari pengalaman maka tes tersebut sudah dapat dikatakan memiliki validitas empirik.

b) Validitas item

Anas Sudijono (2012: 182) mengemukakan bahwa validitas item merupakan ketepatan dalam mengukur apa yang seharusnya diukur pada sebutir soal atau item. Selain itu menurut Suharsimi (2016: 90) validitas item digunakan untuk mengetahui manakah dari keseluruhan soal tersebut yang menyebabkan sebuah tes memiliki validitas yang rendah. Untuk menghitung validitas item dapat menggunakan rumus sebagai berikut:

$$\gamma_{bpi} = \frac{Mp - Mt}{St} \sqrt{\frac{p}{q}}$$

Keterangan:

γ_{bpi} : koefisien korelasi *point biserial*

- Mp : rerata skor dari subjek yang menjawab betul bagi item yang dicari validitasnya
- Mt : rerata skor total
- St : standar deviasi dari skor total
- p : proporsi peserta didik yang menjawab benar

$$\left(p = \frac{\text{banyak peserta didik yang benar}}{\text{jumlah seluruh peserta didik}} \right)$$
- q : proporsi peserta didik menjawab salah ($q = 1 - p$)
 (Suharsimi, 2016: 93)

Indeks korelasi *point biserial* (y_{bpi}) yang diperoleh dari hasil perhitungan kemudian dibandingkan dengan harga r tabel pada taraf signifikansi 5% sesuai dengan jumlah peserta tes. Jika nilai $y_{bpi} > r$ tabel maka butir soal tersebut termasuk valid, begitu pula sebaliknya. Pada penelitian ini menggunakan bukti validitas berdasarkan proses respons, dimana bukti ini merupakan kesesuaian antara konstruk dengan respons peserta tes yang dalam hal ini merupakan pilihan jawaban dari peserta tes.

2) Reliabilitas

Menurut Ngalim Purwanto (2017: 139) reliabilitas atau keandalan merupakan ketetapan atau ketelitian dari suatu alat evaluasi yang dapat dipertanggungjawabkan sejauh mana tes tersebut dapat dipercaya. Kemudian Suharsimi (2016: 74) menyebutkan sebuah tes dapat memiliki reliabilitas apabila jika dilakukan selama berulang kali pada waktu yang berbeda, pengujian tersebut menunjukkan hasil yang relatif sama atau tidak adanya perubahan signifikan. Dapat disimpulkan apabila reliabilitas merupakan salah satu aspek yang harus dimiliki oleh tes dalam hal kepercayaan dan ketetapan soal untuk bisa diujikan

kepada peserta tes. Terdapat 3 macam metode yang dapat digunakan dalam menghitung realibilitas dapat, yaitu:

a) Metode bentuk paralel (*equivalent*)

Dengan metode ini perhitungan reliabilitas berasal dari dua buah tes yang paralel dimana tujuan, tingkat kesukaran, dan susunan tes memiliki keasamaan antara satu dengan yang lainnya namun butir soalnya berbeda. Metode ini diujikan terhadap kelompok peserta didik yang sama yang hasilnya kemudian dikorelasikan dengan nilai koefisien. Apabila memperoleh nilai tinggi maka tes paralel tersebut termasuk reliabel.

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

Pada metode tes ulang, satu seri soal diujikan sebanyak dua kali dalam waktu yang berbeda kepada kelompok peserta tes yang sama. Hasil dari dua kali pengujian tersebut kemudian dikorelasikan agar mendapat nilai reliabilitas. Pada umumnya pengujian yang kedua mendapatkan hasil yang lebih baik daripada pengujian pertama.

c) Metode belah dua (*split-half method*)

Metode ini hanya melakukan satu kali pengujian terhadap satu kelompok peserta didik yang dibagi menjadi dua pada waktu tertentu. Pada metode ini dapat menggunakan rumus K-R 20 untuk mencari nilai reliabilitas dalam tes bentuk objektif. Rumus K-R 20 adalah sebagai berikut:

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

Keterangan:

r_{11}	: reliabilitas tes secara keseluruhan
p	: proporsi subjek yang menjawab item dengan benar
q ($q=1.p$)	: proporsi subjek yang menjawab item dengan salah
Σpq	: jumlah hasil perkalian p dan q
n	: banyak item
S	: standar deviasi dari tes

(Suharsmi, 2016: 115)

3) Tingkat kesukaran

Kriteria soal yang baik adalah soal yang tidak terlalu mudah atau terlalu sukar. Soal yang terlalu sukar akan membuat peserta didik merasa mudah putus asa dan kurang termotivasi untuk belajar karena soal tersebut melebihi jangkauan kemampuannya. Begitupun apabila soal terlalu mudah maka peserta didik tidak akan termotivasi meningkatkan usahanya untuk belajar lebih. Perbandingan tingkat kesukaran yang ideal antara soal mudah : sedang : sukar adalah 30%:50%:20%. Tingkat kesukaran dapat ditentukan sesuai indeks dengan menggunakan rumus sebagai berikut:

$$TK = \frac{(WL + WH)}{(nL + nH)10}$$

Keterangan:

TK	: indeks kesukaran/tingkat kesukaran
WL	: banyaknya kelompok bawah menjawab soal salah
WH	: banyaknya kelompok unggul yang menjawab soal salah
nL	: jumlah kelompok bawah
nH	: jumlah kelompok unggul

(Zainal Arifin, 2012: 342)

4) Daya pembeda

Menurut Suharsimi (2016: 226) daya pembeda soal adalah kemampuan suatu butir soal tes hasil belajar dalam membedakan antara peserta didik yang

berkemampuan rendah dengan yang berkemampuan tinggi. Daya pembeda dapat dihitung menggunakan rumus:

$$DP = \frac{(WL - WH)}{n}$$

Keterangan:

DP : indeks daya pembeda
WL : banyaknya peserta yang gagal dari kelompok bawah
WH : banyaknya peserta yang gagal dari kelompok atas
n : jumlah peserta tes kelompok unggul
(Zainal Arifin, 2012:350)

Setelah memperoleh hasil dari perhitungan daya beda, langkah berikutnya adalah mengklasifikasikan berdasarkan kualitas soalnya. Pengklasifikasian ini bertujuan guna memudahkan dalam menentukan kualitas suatu butir soal.

Klasifikasi Daya Pembeda (DP) adalah sebagai berikut:

DP : 0,00 – 0,19 : Jelek
DP : 0,20 – 0,39 : Cukup
DP : 0,40 – 0,69 : Baik
DP : 0,70 – 1,00 : Baik Sekali

(Martubi, 2004:44)

5) Efektivitas pengecoh

Menurut Suharsimi (2016: 233-234) pola sebaran jawaban soal dapat menentukan besarnya nilai efektivitas pengecoh/ *distractor*. Pengecoh dapat dikatakan berfungsi dengan baik apabila memiliki daya tarik yang besar bagi *testee* yang kurang memahami materi untuk memilih. Pengecoh yang sama sekali tidak dipilih berarti tidak berfungsi dengan baik.

Martubi (2004: 45) mengemukakan apabila pada soal pilihan ganda kriteria pengecoh yang baik adalah tersebar secara merata, mirip dengan kunci jawaban, dipilih oleh minimal 5% *testee* dan jumlah kelompok bawah yang memilihnya lebih banyak daripada kelompok atas. Sependapat dengan itu, Suharsimi (2016: 234) juga menyampaikan pengecoh dapat dikatakan memiliki fungsi yang baik apabila pengecoh tersebut dipilih oleh sebanyak 5% dari *testee* atau peserta tes.

Berdasarkan beberapa pendapat tersebut dapat diambil kesimpulan apabila efektivitas pengecoh adalah seberapa baik alternatif jawaban salah yang akan dipilih oleh peserta yang kurang memahami materi. Semakin banyak *testee* yang memilih pengecoh tersebut maka efektivitas pengecoh juga akan semakin baik. Peserta tes yang mengabaikan semua option atau tidak memilih satupun alternatif jawaban yang tersedia disebut *omit*. Sebuah tes yang baik apabila jumlah *omit*nya tidak lebih dari 10% peserta tes.

4. Tinjauan tentang Mata Pelajaran Dasar Listrik dan Elektronika

Dalam Undang-undang Nomor 20 tahun 2003 disebutkan apabila pendidikan nasional diselenggarakan melalui tiga jalur pendidikan yaitu formal, non-formal dan informal. Pada jalur pendidikan formal terbagi lagi menjadi tiga tingkatan yang kesemuanya saling berkelanjutan antara satu dengan yang lainnya. Ketiga tingkatan tersebut terdiri atas pendidikan dasar, pendidikan menengah dan pendidikan tinggi. Menurut Pasal 18 UU Nomor 20 tahun 2003 pada pendidikan menengah terdiri atas pendidikan menengah kejuruan (*vocational education*) dan pendidikan menengah umum (*general education*). Pada jenjang pendidikan tinggi juga berhak untuk menyelenggarakan program akademik, profesi dan atau vokasi.

Clarke & Winch dalam Herminarto (2018: 24) mendefinisikan jika pendidikan kejuruan merupakan pendidikan yang diselenggarakan dalam rangka menyiapkan anak-anak muda dan remaja untuk memasuki lapangan kerja. Kemudian proses pembelajaran pendidikan kejuruan juga erat kaitannya dengan masalah teknik dan praktik.

Karakteristik yang dimiliki oleh pendidikan kejuruan berbeda dengan pendidikan umum. Menurut Wardiman dalam Herminarto (2018: 17) terdapat beberapa karakteristik yang dimiliki oleh pendidikan kejuruan, antara lain pendidikan kejuruan dipersiapkan dan diarahkan untuk memasuki dunia kerja, didasarkan atas *demand driver*, memiliki hubungan erat terhadap DUDI (Dunia Usaha Dunia Industri), lebih ditekankan kepada pengetahuan, keterampilan, sikap, *learning by doing* dan berpengalaman (*hands on experience*). Selain itu juga pendidikan kejuruan sangat *responsif* dan *antisipatif* terhadap kemajuan teknologi saat ini. Fasilitas mutakhir yang digunakan untuk menunjang keperluan kegiatan praktik pada pendidikan kejuruan membuat biaya investasi dan operasional lebih tinggi dibandingkan pendidikan umum.

Berdasarkan pada beberapa pendapat tersebut tentang pendidikan kejuruan, pada dasarnya terdapat perbedaan karakteristik antara pendidikan kejuruan dengan pendidikan yang lainnya. Dapat disimpulkan juga apabila pendidikan kejuruan bertujuan untuk mempersiapkan peserta didiknya dalam pengetahuan, kepribadian serta keterampilan peserta didik supaya lebih siap dan mandiri dalam pendidikan yang lebih lanjut sesuai dengan program kompetensi atau keahlian yang dipilih. Selain itu, mempersiapkan peserta didik menjadi tenaga kerja yang

profesional dan ahli sesuai dengan bidang atau kompetensinya juga merupakan tujuan dari pendidikan kejuruan.

Sejalan dengan hal tersebut, menurut keputusan Direktur Jenderal Pendidikan Dasar dan Menengah nomor 330 tahun 2017 mengenai kompetensi inti dan kompetensi dasar mata pelajaran pada SMK/MAK, mata pelajaran Dasar Listrik dan Elektronika masuk ke dalam dasar program keahlian (C2) pada program keahlian Teknik Ketenagalistrikan bidang keahlian Teknologi dan Rekayasa. Mata pelajaran DLDE merupakan salah satu mata pelajaran dasar yang harus dikuasai oleh peserta didik. Terdapat 14 kompetensi dasar yang harus dimiliki oleh peserta didik selama melaksanakan proses pembelajaran.

Menurut Silabus Mata Pelajaran DLDE di SMK Negeri 2 Purwokerto (dapat dilihat di lampiran 2) pada UAS Gasal kelas X tahun ajaran 2018/2019 hanya terdapat 5 kompetensi dasar yang diujikan. Kompetensi yang diujikan antara lain menerapkan konsep listrik dan elektronika dengan materi pada kompetensi ini adalah arus listrik dan elektron. Kompetensi kedua yang diujikan adalah menganalisis bahan-bahan komponen listrik dan elektronika. Pada kompetensi ini menyajikan beberapa bahan-bahan listrik yang terdiri dari konduktor, isolator dan bahan semikonduktor. Kemudian kompetensi ketiga yang diujikan adalah sifat elemen pasif pada rangkaian listrik arus searah dan rangkaian peralihan. Dimana materi yang diujikan berupa resistor-resistansi, induktor-induktansi dan kapasitor-kapasitansi. Kompetensi keempat yang diujikan adalah teorema rangkaian listrik searah mengenai rangkaian seri, paralel, seri-paralel,

hukum Ohm dan hukum Kirchoff. Kemudian kompetensi terakhir yang diujikan adalah sifat elemen aktif pada sumber arus dan sumber tegangan.

B. Penelitian yang Relevan

Untuk lebih memperkuat dasar penelitian, maka dibutuhkan beberapa penelitian yang relevan dengan penelitian ini. Beberapa penelitian yang relevan dengan Analisis Butir Soal pada penelitian ini adalah:

1. Penelitian yang dilakukan oleh Muhammad Abdul Rochim pada tahun 2018 berjudul “Analisis Butir Soal Ujian Akhir Semester Gasal Mata Pelajaran *Chasis* Kelas XII Teknik Kendaraan Ringan di SMK Muhammadiyah 2 Tempel Tahun Ajaran 2017/2018” dilakukan menggunakan program *Anates Version 4.09* memperoleh hasil penelitian:
 - a. Dilihat dari segi validitas, sebanyak 21 butir soal (43%) termasuk dalam kategori valid dan 28 butir soal (57%) kategori tidak valid.
 - b. Reliabilitas soal pada hasil analisis mendapatkan hasil sebesar 0,50 yang berarti memiliki inteprestasi yang cukup.
 - c. Berdasarkan hasil analisis tingkat kesukaran soal, 5 butir soal (10%) kategori mudah, 27 butir soal (55%) kategori sedang dan 17 butir soal (35%) kategori sukar.
 - d. Hasil analisis daya pembeda pada penelitian ini menunjukkan 3 butir soal (6%) kategori buruk sekali, 15 butir soal (31%) kategori buruk, 20 butir soal (41%) kategori cukup dan 11 butir soal (22%) kategori baik.

- e. Efektivitas pengecoh pada penelitian ini memperoleh hasil analisis sebanyak 2 butir soal (4%) kategori tidak baik, 9 butir soal (18%) kategori kurang baik, 16 butir soal (33%) kategori cukup baik, 15 butir soal (31%) kategori baik dan 7 butir soal (14%) masuk dalam kategori sangat baik.
 - f. Dari hasil analisis Validitas, Raliabilitas, Daya Pembeda, Tingkat Kesukaran, dan Efektivitas Pengecoh dapat disimpulkan jika dari 49 butir soal pilihan gand terdapat 3 butir soal (6%) kategori sangat tidak baik, 11 butir soal (22,5%) kategori tidak baik, 12 butir soal (24,5%) kategori cukup, 11 butir soal (22,5%) kategori baik dan 12 butir soal (24,5%) masuk kategori sangat baik.
2. Penelitian yang dilakukan oleh Shodiq pada tahun 2018 berjudul “Analsis Butir Soal Ujian Akhir Semester Gasal Mata Pelajaran Pemeliharaan Listrik Kendaraan Ringan Kelas XII di SMK Negeri 2 Klaten Tahun Ajaran 2017/2018” dilakukan secara kualitatif dan kuantitatif. Analisis kuantitatif menggunakan program *Anates*. Dalam penelitian ini didapatkan hasil:
- a. Ditinjau dari beberapa aspek didapatkan jika 23,33% soal tidak memenuhi aspek materi, 53,33% soal tidak memenuhi aspek konstruksi dan 73,33% soal tidak memenuhi aspek bahasa.
 - b. Dilihat dari aspek validitas, hasil analisis menunjukkan sebanyak 40% soal termasuk valid dan sebanyak 60% soal termasuk tidak valid.
 - c. Aspek reliabilitas soal didapatkan koefisien reliabilitas 0,67 yang berarti memiliki reliabilitas tinggi.

- d. Daya pembeda soal yang dianalisis menunjukkan hasil 3,33% soal termasuk kategori baik sekali, 23,33% soal kategori baik, 13,33% soal kategori cukup, 46,67% soal kategori jelek dan sisanya 13,33% termasuk soal yang harus dibuang.
- e. Hasil analisis tingkat kesukaran soal menunjukkan sebanyak 50% masuk kategori rendah, 23,33% kategori sedang dan 26,67% kategori sukar.
- f. Efektivitas pengecoh yang dianalisis menunjukkan hanya 13,33% soal yang pengecohnya berfungsi secara efektif.

Persamaan yang terdapat pada penelitian yang dilakukan oleh Shodiq dan Muhammad Abdul Rochim adalah sama-sama berupa penelitian deskriptif kuantitatif dan meneliti mengenai analisis butir soal dengan menggunakan program *Anates*. Perbedaan dari kedua penelitian diatas adalah mata pelajaran dan juga lokasi penelitian. Berdasarkan penelitian yang relevan diatas, didapatkan hasil butir soal yang tidak valid lebih besar daripada butir soal yang valid. Secara keseluruhan hasil analisis ini berarti kualitas soal tersebut masih perlu untuk adanya perbaikan supaya soal tersebut menjadi layak dan baik untuk dijadikan alat ukur evaluasi.

C. Kerangka Berfikir

Pengetahuan dan kemampuan dalam melaksanakan evaluasi harusnya dikuasai oleh guru supaya proses pelaksanaan kegiatan evaluasi dapat berjalan secara sistematis dan terstruktur. Sebelum melaksanakan kegiatan evaluasi, diperlukan langkah-langkah yang dilakukan yaitu meliputi perencanaan,

pelaksanaan, penilaian, mengolah data dan menganalisis hasil penilaian. Kegiatan evaluasi yang dilakukan dapat memberikan guru informasi mengenai perkembangan peserta didiknya. Selain itu juga sebagai tolok ukur keberhasilan proses pembelajaran yang telah dilakukan dalam periode waktu tertentu.

Soal ulangan akhir semester merupakan salah satu alat atau instrumen evaluasi yang digunakan untuk mengetahui tingkat pemahaman peserta didik terhadap materi yang telah diberikan selama satu semester. Langkah-langkah yang dapat dilakukan dalam pengembangan tes hasil belajar adalah dengan menyusun spesifikasi tes, menulis soal, menelaah soal, melakukan uji coba tes, menganalisis butir soal, memperbaiki tes, melaksanakan tes dan menafsirkan hasil dari tes tersebut. Langkah-langkah tersebut seharusnya diterapkan oleh guru saat akan, sedang dan setelah melaksanakan kegiatan evaluasi supaya instrumen yang digunakan berkualitas. Oleh karena itu, penting bagi guru untuk melakukan uji coba terhadap instrumen evaluasi untuk memperoleh informasi mengenai validitas, reliabilitas, tingkat kesukaran, daya beda, efektivitas pengecoh, materi, konstruksi dan bahasa pada soal.

Informasi mengenai aspek materi, konstruksi dan bahasa pada soal dapat dilakukan sebelum tes tersebut diujikan. Hal ini dikarenakan analisis dilakukan secara teoritik atau kualitatif dilihat dari kaidah penulisan soal tersebut. Kemudian pada aspek validitas, reliabilitas, tingkat kesukaran, daya beda dan efektivitas pengecoh dapat dilakukan setelah instrumen tersebut diujikan kepada peserta didik. Tujuannya adalah untuk mengetahui apakah soal tersebut sudah berkualitas

atau belum. Selanjutnya soal yang telah di analisis atau ditelaah di intepretasikan menjadi beberapa kategori soal.

Berdasarkan hasil analisis tersebut maka guru dapat melakukan perbaikan-perbaikan yang dirasa perlu terutama untuk soal yang masih dalam kategori jelek. Bagi soal yang sudah berkualitas baik maka dapat dimasukkan ke dalam bank soal untuk selanjutnya digunakan kembali pada kegiatan evaluasi berikutnya. Dengan dilakukannya analisis butir soal ini, maka guru akan memiliki alat atau instrumen evaluasi yang berkualitas sehingga mampu mencerminkan prestasi belajar peserta didik dengan tepat.

D. Pertanyaan Peneliti

1. Dilihat dari analisis secara teoritik:
 - a. Bagaimanakah kualitas Soal Ujian Akhir Semester Gasal Mata Pelajaran DLDE tahun pelajaran 2018/2019 SMK Negeri 2 Purwokerto kelas X TITL ditinjau dari aspek materi?
 - b. Bagaimanakah kualitas Soal Ujian Akhir Semester Gasal Mata Pelajaran DLDE tahun pelajaran 2018/2019 SMK Negeri 2 Purwokerto kelas X TITL ditinjau dari aspek konstruksi?
 - c. Bagaimanakah kualitas Soal Ujian Akhir Semester Gasal Mata Pelajaran DLDE tahun pelajaran 2018/2019 SMK Negeri 2 Purwokerto kelas X TITL ditinjau dari aspek bahasa?

2. Dilihat dari analisis empirik:
 - a. Bagaimanakah kualitas Soal Ujian Akhir Semester Gasal Mata Pelajaran DLDE tahun pelajaran 2018/2019 SMK Negeri 2 Purwokerto kelas X TITL ditinjau dari segi validitas?
 - b. Bagaimanakah kualitas Soal Ujian Akhir Semester Gasal Mata Pelajaran DLDE tahun pelajaran 2018/2019 SMK Negeri 2 Purwokerto kelas X TITL ditinjau dari segi reliabilitas?
 - c. Bagaimanakah kualitas Soal Ujian Akhir Semester Gasal Mata Pelajaran DLDE tahun pelajaran 2018/2019 SMK Negeri 2 Purwokerto kelas X TITL ditinjau dari segi tingkat kesukaran?
 - d. Bagaimanakah kualitas Soal Ujian Akhir Semester Gasal Mata Pelajaran DLDE tahun pelajaran 2018/2019 SMK Negeri 2 Purwokerto kelas X TITL ditinjau dari segi daya pembeda?
 - e. Bagaimanakah kualitas Soal Ujian Akhir Semester Gasal Mata Pelajaran DLDE tahun pelajaran 2018/2019 SMK Negeri 2 Purwokerto kelas X TITL ditinjau dari segi efektivitas pengecoh/*distractor*?