

BAB II

KAJIAN TEORI DAN KERANGKA PEMIKIRAN

A. Kajian Teori

1. Konsep Dasar Penilaian

Ada tiga istilah yang sering digunakan dan berkaitan dengan penilaian, yaitu tes, pengukuran, dan evaluasi. (*test, measurement, and evaluation*). Dalam kehidupan sehari-hari orang sering menyamakan pengertian ke empat istilah tersebut (*test, measurement, assessment and evaluation*), padahal ke empat istilah tersebut memiliki makna yang berbeda. Beberapa orang juga sering rancu menggunakan istilah-istilah tersebut karena ke empat istilah digunakan untuk merujuk kegiatan yang sama.

Tes (*test*) merupakan suatu cara untuk memprediksi tingkat pengetahuan seseorang secara tidak langsung, yaitu melalui respons seseorang terhadap stimulus atau pertanyaan. Respons peserta tes terhadap sejumlah pertanyaan menggambarkan tingkat pengetahuan peserta tes dalam bidang tertentu. Tes merupakan alat ukur untuk memperoleh informasi hasil belajar siswa yang memerlukan jawaban atau respon benar atau salah. Tes merupakan bagian tersempit dari evaluasi. (Eko Putro Widoyoko, 2009, hlm 150).

“Pengukuran dinyatakan sebagai proses penetapan angka terhadap individu atau karakteristiknya menurut aturan tertentu.” (Ebel & Frisbiedalam).

Esensi dari pengukuran adalah kuantifikasi atau penetapan angka tentang karakteristik atau keadaan individu menurut aturan-aturan tertentu. Hasil pengukuran berupa skor atau angka. Pengukuran memiliki konsep yang lebih luas dari pada tes. Kita dapat mengukur karakteristik suatu objek tanpa menggunakan tes, misalnya dengan pengamatan, wawancara, atau cara lain untuk memperoleh informasi dalam bentuk kuantitatif (Eko Putro Widoyoko, 2016, hlm. 4).

Penafsiran harga koefisien korelasi dilakukan dengan membandingkan harga r_{xy} dengan harga kritik. Adapun harga kritik besar atau sama dengan 0,3 ($r_{xy} \geq 0,3$), nomor butir tersebut dapat dikatakan valid. Sebaliknya apabila r_{xy} lebih kecil dari 0,3 ($r_{xy} < 0,3$),

nomor butir tersebut dikatakan tidak valid. Penentuan batas minimal suatu butir instrumen dianggap valid apabila memiliki korelasi 0,3 terhadap skor total dengan asumsi bahwa besarnya pengaruh atau determinan butir terhadap total instrumen = $(r)^2 = (0,3)^2 = 0,09$, dibulatkan menjadi 0,1 atau 10%. Butir instrumen yang memiliki sumbangan terhadap total butir instrumen kurang dari 10% dianggap butir tersebut kurang bermakna terhadap keberadaan instrumen secara keseluruhan.

“Penilaian dalam konteks hasil belajar diartikan sebagai kegiatan menafsirkan atau memaknai data hasil pengukuran tentang kompetensi yang dimiliki siswa setelah mengikuti kegiatan pembelajaran. Data hasil pengukuran dapat diperoleh melalui tes, pengamatan, wawancara, portofolio, jurnal, maupun instrumen lainnya” (Eko Putro Widoyoko, 2015, hlm. 10).

Berdasarkan pendapat di atas dapat disimpulkan bahwa evaluasi merupakan proses yang sistematis dan berkelanjutan untuk mengumpulkan, mendeskripsikan, mengintepretasikan dan menyajikan informasi tentang suatu program untuk dapat digunakan sebagai dasar membuat keputusan, menyusun kebijakan maupun menyusun program selanjutnya.

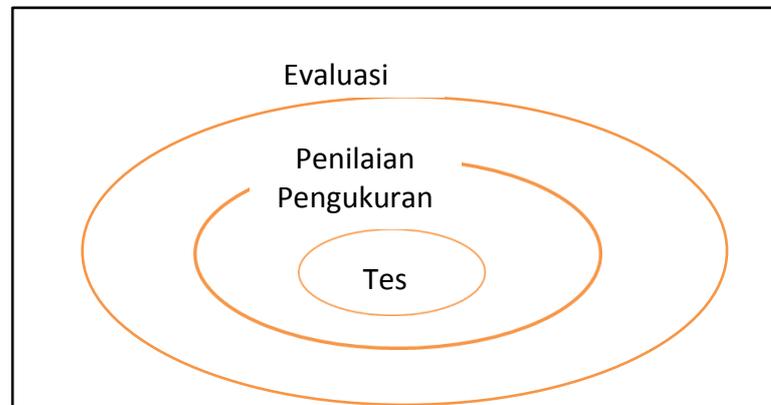
Adapun tujuan evaluasi adalah untuk memperoleh informasi yang akurat dan objektif tentang suatu program. Informasi tersebut dapat berupa proses pelaksanaan program, dampak/hasil yang dicapai, efisiensi serta pemanfaatan hasil evaluasi yang difokuskan untuk program itu sendiri, yaitu untuk mengambil keputusan apakah dilanjutkan, diperbaiki atau dihentikan. Selain itu, juga dipergunakan untuk kepentingan penyusunan program berikutnya. (Eko Putro Widoyoko, 2016, hlm. 8)

Dalam konteks pembelajaran lingkup atau cakupan penilaian hanya pada individu siswa dalam kelas, sedangkan lingkup evaluasi adalah seluruh komponen dalam program pembelajaran, mulai dari input, proses, sampai pada hasil pembelajaran. Dalam proses pembelajaran, cakupan evaluasi meliputi siswa, guru, kurikulum, sarana dan prasana atau media pembelajaran, iklim kelas, sikap siswa dalam pembelajaran dan sebagainya. Dengan demikian perbedaan prinsip antara penilaian

dengan evaluasi adalah pada cakupan. Penilaian mencakup satu aspek, sedangkan evaluasi mencakup beberapa aspek dalam program. Kegiatan evaluasi selalu terkait dengan program. Cakupan evaluasi lebih luas dibandingkan dengan cakupan penilaian. Adapun persamaannya yaitu sama-sama proses atau kegiatan menafsirkan, memaknai dan mendeskripsikan atau menetapkan kualitas hasil pengukuran. Dengan adanya persamaan tersebut tidak mengherankan apabila banyak orang yang tidak bisa membedakan bahkan menyamakan antara penilaian dengan evaluasi, walaupun secara esensial berbeda.

Tes, pengukuran, penilaian dan evaluasi bersifat hierarkhi. Evaluasi didahului dengan penilaian (*assessment*), sedangkan penilaian didahului dengan pengukuran. Salah satu alat ukurnya adalah tes. Tes diartikan sebagai alat ukur untuk memperoleh informasi hasil belajar siswa yang memerlukan jawaban atau respon benar atau salah. Pengukuran diartikan sebagai kuantifikasi atau penetapan angka (skor) tentang karakteristik atau keadaan individu menurut aturan, kriteria atau standar tertentu. Penilaian (*assessment*) merupakan kegiatan menafsirkan, memaknai dan mendeskripsikan hasil pengukuran, sedangkan evaluasi merupakan penetapan kualitas suatu program beserta tindak lanjutnya berdasarkan penilaian aspek-aspek program.

Berdasarkan uraian di atas hubungan antara tes, pengukuran, penilaian, dan evaluasi dapat dibuat dalam bentuk gambar sebagai berikut:



Gambar 2.1 Hubungan antara Tes, Pengukuran, Penilaian, dan Evaluasi

Menurut Suharsimi Arikunto dalam Eko putro Widoyoko (2016, hlm. 10) “Guru maupun pendidik lainnya perlu mengadakan penilaian terhadap hasil belajar siswa karena dalam dunia pendidikan, khususnya dunia persekolahan penilaian hasil belajar mempunyai makna yang penting, baik bagi siswa, guru maupun sekolah.” Adapun makna penilaian bagi ketiga pihak tersebut adalah:

a. Makna Bagi Siswa

Dengan diadakannya penilaian hasil belajar, maka siswa dapat mengetahui sejauh mana telah berhasil mengikuti pelajaran yang disajikan oleh guru. Hasil yang diperoleh siswa dari penilaian hasil belajar ini ada dua kemungkinan:

- 1) Memuaskan Jika siswa memperoleh hasil yang memuaskan dan hasil itu menyenangkan, tentu kepuasan itu ingin diperolehnya lagi pada kesempatan lain waktu. Akibatnya, siswa akan mempunyai motivasi yang cukup besar untuk belajar lebih giat, agar lain kali mendapat hasil yang lebih memuaskan. Keadaan sebaliknya dapat juga terjadi, yakni siswa sudah merasa puas dengan hasil yang diperoleh dan usahanya menjadi kurang gigih untuk lain kali.
- 2) Tidak memuaskan Jika siswa tidak puas dengan hasil yang diperoleh, ia akan berusaha agar lain kali keadaan itu tidak terulang

lagi. Maka ia selalu belajar giat. Namun demikian dapat juga sebaliknya, bagi siswa yang lemah kemauannya, akan menjadi putus asa dengan hasil kurang memuaskan yang telah diterimanya.

b. Makna Bagi Guru

- 1) Berdasarkan hasil penilaian yang diperoleh, guru akan dapat mengetahui siswa-siswa mana yang sudah sudah berhak melanjutkan pelajarannya karena sudah berhasil mencapai kegiatan belajar mengajar (KBM) kompetensi yang diharapkan, maupun mengetahui siswa-siswa yang belum berhasil mencapai KBM kompetensi yang diharapkan. Dengan petunjuk ini guru dapat lebih memusatkan perhatiannya kepada siswa-siswa yang belum berhasil mencapai KBM kompetensi yang diharapkan.
- 2) Berdasarkan hasil penilaian yang diperoleh, guru akan dapat mengetahui apakah pengalaman belajar (materi pelajaran) yang disajikan sudah tepat bagi siswa sehingga untuk kegiatan pembelajaran di waktu yang akan datang tidak perlu diadakan perubahan.
- 3) Berdasarkan hasil penilaian yang diperoleh, guru akan dapat mengetahui apakah strategi pembelajaran yang digunakan sudah tepat atau belum. Jika sebagian besar dari siswa memperoleh hasil penilaian yang kurang baik maupun jelek pada penilaian yang diadakan, mungkin hal ini disebabkan oleh strategi atau metode pembelajaran yang kurang tepat. Apabila demikian halnya, maka guru harus introspeksi diri dan mencoba mencari strategi lain dalam kegiatan pembelajaran yang dilaksanakan.

c. Makna Bagi Sekolah

- 1) Apabila guru-guru mengadakan penilaian dan diketahui bagaimana hasil belajar siswa-siswanya, maka akan dapat diketahui pula apakah kondisi belajar maupun kultur akademik yang diciptakan oleh sekolah sudah sesuai dengan harapan atau belum. Hasil belajar siswa merupakan cermin kualitas suatu sekolah.

- 2) Informasi hasil penilaian yang diperoleh dari tahun ke tahun dapat digunakan sebagai pedoman bagi sekolah untuk mengetahui apakah yang dilakukan oleh sekolah sudah memenuhi standar pendidikan sebagaimana dituntut standar nasional pendidikan (SNP) atau belum. Pemenuhan berbagai standar akan terlihat dari bagusnya hasil penilaian belajar siswa.
- 3) Informasi hasil penilaian yang diperoleh dapat dijadikan sebagai pertimbangan bagi sekolah untuk menyusun berbagai program pendidikan di sekolah untuk masa-masa yang akan datang.

3. Objek Penilaian Hasil Belajar

Penilaian hasil belajar siswa di sekolah mencakup aspek atau ranah kompetensi pengetahuan, sikap dan keterampilan (kognitif, afektif, dan psikomotor) yang dilakukan secara berimbang sehingga dapat digunakan untuk menentukan posisi relative setiap siswa terhadap standar yang telah ditetapkan.

a. Pengetahuan (Kognitif)

Berdasarkan temuan-temuan baru dalam riset tentang belajar dan perbedaan-perbedaan dalam taksonomi pembelajaran, Anderson dan Krathwohl membedakan pengetahuan (*knowledge*) menjadi dua dimensi, yaitu komponen kata kerja dengan istilah “dimensi pengetahuan” dan komponen kata kerja dengan istilah “dimensi proses kognitif”. Pembagian tersebut merupakan revisi terhadap Taksonomi Bloom yang telah digunakan dalam dunia pendidikan sejak tahun 1956. (Anderson & Krathwohl. 2001).

1) Dimensi Pengetahuan

Terdapat banyak jenis pengetahuan dan lebih banyak lagi istilah yang digunakan untuk mendeskripsikan pengetahuan-pengetahuan tersebut. Sebagian istilah menggambarkan pengetahuan-pengetahuan yang sangat berbeda, sedangkan sebagian istilah lainnya sekadar label-label yang berbeda untuk kategori pengetahuan yang sama. Menurut Anderson dan Krathwohl (2001) membedakan dimensi pengetahuan menjadi

empat jenis yaitu: pengetahuan faktual (*factual knowledge*), pengetahuan konseptual (*conceptual knowledge*), pengetahuan prosedural (*procedural knowledge*), dan pengetahuan metakognitif (*metacognitive knowledge*).

2) Dimensi Proses Kognitif

Siswa melakukan proses kognitif secara aktif, yakni memperhatikan informasi yang relevan yang datang, menata informasi menjadi gambaran yang koheren, dan memadukan informasi tersebut dengan pengetahuan yang telah dimilikinya. Proses kognitif merupakan cara yang dipakai siswa secara aktif dalam proses mengkonstruksi makna. Menurut Anderson dan Krathwohl (2001) “Proses kognitif dalam pembelajaran dibedakan menjadi enam jenjang dari jenjang yang rendah ke jenjang yang tinggi, yaitu: mengingat, memahami, mengaplikasikan, menganalisis, mengevaluasi, dan mencipta.”

3) Sikap (Afektif)

Stiggins dalam Eko Putro Widoyoko (2016, hlm. 48) menyatakan “Bahwa siswa yang memiliki sikap positif dan motivasi memiliki peluang yang lebih baik untuk mencapai prestasi belajar yang lebih baik daripada siswa yang memiliki sikap yang negatif.”

Menurut Krathwohl, Bloom dan Maisa dalam Eko Putro Widoyoko (2016, hlm. 52). Jenjang afektif dibedakan menjadi lima jenjang dari jenjang yang sederhana sampai jenjang yang kompleks, yaitu : *receiving/attending* (menerima / memperhatikan), *responding* (menanggapi), *valuing* (menilai/menghargai), *organization* (mengatur mengorganisasikan), dan *characterization by value or value complex* (karakterisasi dengan suatu nilai atau kompleks nilai).

4) Keterampilan (Psikomotorik)

Aspek keterampilan atau psikomotor merupakan hasil belajar yang pencapaiannya melibatkan otot dan kekuatan fisik. Dengan

kata lain ranah psikomotor adalah ranah yang berhubungan dengan aktivitas fisik, misalnya lari, melompat, menari, memukul, megetik dan sebagainya.

Ranah keterampilan atau psikomotor menurut Bloom dibedakan menjadi tujuh, yaitu: *perception* (persepsi), *set* (kesiapan), *guided response* (respons terpimpin), *mechanism* (mekanisme), *complex overt response* (respons tampak yang kompleks), *adaptation* (penyesuaian), dan *origination* (penciptaan). (Eko Putro Widoyoko, 2016, hlm. 10-12)

4. Fungsi Penilaian Hasil Belajar

Penilaian hasil belajar oleh pendidik memiliki fungsi untuk memantau kemajuan belajar, memantau hasil belajar, dan mendeteksi kebutuhan perbaikan hasil belajar peserta didik secara berkesinambungan. Berdasarkan fungsinya penilaian hasil belajar oleh pendidik meliputi:

- a. Formatif yaitu memperbaiki kekurangan hasil belajar peserta didik dalam sikap, pengetahuan, dan keterampilan pada setiap kegiatan penilaian selama proses pembelajaran dalam satu semester, sesuai dengan prinsip Kurikulum 2013 agar peserta didik tahu, mampu dan mau. Hasil dari kajian terhadap kekurangan peserta didik digunakan untuk memberikan pembelajaran remedial dan perbaikan RPP serta proses pembelajaran yang dikembangkan guru untuk pertemuan berikutnya;
- b. Sumatif yaitu menentukan keberhasilan belajar peserta didik pada akhir suatu semester, satu tahun pembelajaran, atau masa pendidikan di satuan pendidikan. Hasil dari penentuan keberhasilan ini digunakan untuk menentukan nilai rapor, kenaikan kelas dan keberhasilan belajar satuan pendidikan seorang peserta didik. (Eko Putro Widoyoko, 2016, hlm. 18)
- c. Teknik Penilaian Penilaian hasil belajar siswa mencakup aspek kompetensi sikap, pengetahuan, dan aspek keterampilan yang dilakukan secara berimbang sehingga dapat digunakan untuk menentukan posisi relatif setiap siswa terhadap standar yang telah ditetapkan. Tiap-tiap

aspek penilaian (sikap, pengetahuan, dan keterampilan) memiliki karakteristik yang berbeda antara yang satu dengan yang lainnya sehingga memerlukan teknik penilaian yang berbeda. Tidak ada suatu teknik penilaian yang terbaik yang dapat digunakan untuk menilai semua aspek kompetensi, karena masing-masing teknik penilaian memiliki kelebihan yang kekurangan. Memerlukan kejelian dan kecerdasan guru untuk memilih teknik penilaian yang paling sesuai dengan aspek yang akan dinilai. Secara garis besar ada sembilan teknik penilaian yang dapat dipilih guru untuk menilai hasil pembelajaran siswa, yaitu: tes, observasi, penilaian diri (*self assessment*), penilaian antar peserta (*peer assessment*), penilaian kinerja (*performance assessment*), penilaian portofolio (*portofolio assessment*), penilaian proyek (*project assessment*), penilaian produk (*product assessment*), dan penilaian jurnal (*journal assessment*). Tiap-tiap teknik penilaian memiliki penggunaan yang berbeda-beda. Tes lebih cocok digunakan untuk menilai aspek pengetahuan. Observasi, penilaian diri, penilaian antar teman, dan penilaian jurnal lebih cocok digunakan untuk menilai aspek sikap siswa. Teknik penilaian portofolio dan penilaian produk lebih cocok digunakan untuk menilai aspek keterampilan, sedangkan penilaian kinerja dan penilaian proyek dapat digunakan untuk menilai aspek pengetahuan dan keterampilan. (Eko Putro Widoyoko, 2016, hlm. 63-64) Hubungan antara teknik penilaian dengan aspek penilaian dapat disajikan dalam bentuk tabel 1 berikut ini :

Tabel 2. 1 Hubungan antara Teknik Penilaian dengan Aspek Penilaian

Teknik Penilaian	Aspek Penilaian		
	Pengetahuan	Keterampilan	Sikap
Tes	✓		
Observasi			✓
Penilaian diri			✓
Penilaian antar teman			✓
Penilaian kinerja	✓	✓	
Penilaian portofolio		✓	
Penilaian proyek	✓	✓	

Penilaian produk		✓	
Penilaian Jurnal			✓

Sumber: (Eko Putro Widoyoko, 2016, hlm. 115-116)

d. Tes

Tes merupakan salah satu alat untuk melakukan pengukuran, yaitu alat untuk mengumpulkan informasi karakteristik suatu objek. Di antara objek tes adalah kemampuan siswa. Respons peserta tes terhadap sejumlah pertanyaan atau pernyataan menggambarkan kemampuan peserta tes dalam bidang tertentu. Tes merupakan alat ukur untuk memperoleh informasi hasil belajar siswa yang memerlukan jawaban atau respon benar atau salah. Tes lebih cocok digunakan untuk mengetahui kemampuan siswa dalam aspek pengetahuan (kognitif), tidak cocok digunakan untuk mengukur sikap, karena sikap tidak dapat diinterpretasi ke dalam kategori benar atau salah, namun untuk mendapatkan deskripsi tentang profil sikap siswa. Tes merupakan salah satu teknik penilaian hasil belajar siswa. (Eko Putro Widoyoko, 2016, hlm. 117-118).

Berdasarkan sistem penskoran, tes dapat dikategorikan menjadi dua, yaitu tes objektif dan tes subjektif.

1) Tes Objektif

Tes objektif memiliki arti siapa saja yang memeriksa lembar jawaban tes akan menghasilkan skor yang sama. Skor tes ditentukan oleh jawaban yang diberikan oleh peserta tes. Dengan kata lain dapat dikatakan bahwa tes objektif adalah tes yang penskorannya bersifat objektif, yaitu hanya dipengaruhi oleh objek jawaban atau respon yang diberikan oleh peserta tes. Hal ini bisa terjadi karena kemungkinan jawaban atau respon telah disediakan oleh penyusun butir soal. Peserta hanya memilih alternatif jawaban yang telah disediakan. Dengan demikian pemeriksaan atau penskoran jawaban/respon peserta tes sepenuhnya dapat dilakukan secara objektif oleh pemeriksa. Karena sifatnya yang objektif ini maka tidak perlu harus dilakukan oleh manusia. Pekerjaan tersebut dapat dilakukan oleh mesin, misalnya mesin *scanner*. Dengan demikian skor hasil tes dapat dilakukan secara objektif.

(Eko Putro Widoyoko, 2016, hlm. 71)

Secara umum ada empat tipe tes objektif, yaitu: benar salah (*true false*), menjodohkan (*matching*), pilihan ganda (*multiple choice*) dan uraian objektif. (Eko Putro Widoyoko, 2016, hlm. 118)

e. Tes pilihan ganda

Tes Pilihan Ganda adalah tes di mana setiap butir soalnya memiliki jumlah alternatif jawaban lebih dari dua. Pada umumnya jumlah alternatif jawaban berkisar antara 3 (tiga) atau 5 (lima). Tentu saja jumlah alternatif tersebut tidak boleh terlalu banyak. Bila alternatif lebih dari lima maka akan sangat membingungkan peserta tes, dan juga akan sangat menyulitkan penyusunan butir soal. Tipe tes ini dalam bahasa Inggris dikenal dengan nama *multiple choice item* (butir tes pilihan majemuk atau ganda). (Eko Putro Widoyoko, 2016, hlm. 126).

Tes pilihan ganda dapat dibedakan menjadi 5 model, yaitu: pilihan ganda sederhana, pilihan ganda analisis hubungan antar hal, pilihan ganda analisis kasus, pilihan ganda asosiasi, dan pilihan ganda dengan diagram, grafik, tabel dan sebagainya. Kelima ragam tes objektif pilihan ganda tersebut sama struktur (formatnya), yaitu ada pokok soal (*stem*) yang diikuti oleh sejumlah pilihan (*option*). Di antara pilihan ini ada satu jawaban yang benar atau paling benar sebagai kunci (*key*). Pilihan di luar yang benar atau yang paling tepat berfungsi sebagai pengecoh (*distractors*).

(Eko Putro Widoyoko, 2016, hlm. 126-127)

Tes uraian objektif sering digunakan pada bidang sains dan teknologi atau bidang sosial yang jawaban soalnya sudah pasti, dan hanya satu jawaban yang benar. Bentuk tes uraian objektif sering digunakan pada mata pelajaran yang batasnya jelas, misalnya mata pelajaran fisika, matematika, kimia, biologi, dan sebagainya. Soal pada tes ini jawabannya hanya satu, mulai dari memilih rumus yang tepat, memasukkan angka dalam rumus, menghitung hasil, dan menafsirkan hasilnya.

(Eko Putro Widoyoko, 2016, hlm. 144)

2) Tes Subjektif

Tes subjektif adalah tes yang penskorannya dipengaruhi oleh jawaban peserta tes dan pemberi skor. Jawaban yang sama dapat memiliki skor yang berbeda oleh pemberi skor yang berlainan. Di antara subjektivitas yang dapat mempengaruhi hasil penskoran hasil tes di antaranya adalah: ketidak konsistenan penilai (*rater unreliability*), halo effect, pengaruh urutan pemeriksaan (*order effect*), dan pengaruh bentuk tulisan dan bahasa atau *mechanic and language effect*. (Eko Putro Widoyoko. 2016, hlm. 147)

“Tes subjektif, pada umumnya berbentuk uraian (esai). Tes bentuk uraian adalah butir soal yang mengandung pertanyaan atau tugas yang jawaban atau pengerjaan soal tersebut harus dilakukan dengan

cara mengekspresikan pikiran peserta tes.” (Asmawi Zaenul dan Noehi Nasution dalam Eko Putro Widoyoko 2016, hlm. 75).

Ciri khas tes uraian adalah jawaban terhadap soal tersebut tidak disediakan oleh penyusun soal, tetapi harus disusun oleh peserta tes. Butir soal tipe uraian (*essay test*) hanya terdiri dari pertanyaan atau tugas dan jawaban sepenuhnya harus dipikirkan oleh peserta tes. Ciri-ciri pertanyaannya didahului dengan kata-kata seperti: uraikan, jelaskan, bandingkan, mengapa, bagaimana, simpulkan dan sebagainya. (Suharsimi Arikunto dalam Eko Putro Widoyoko 2016, hlm. 75).

Jumlah butir soal dalam tes uraian biasanya tidak banyak, hanya sekitar 5 – 10 butir soal dalam waktu kira-kira 90 s.d. 120 menit. Soal-soal bentuk uraian ini menuntut kemampuan peserta tes untuk dapat mengorganisir, menginterpretasi, menghubungkan pengertian-pengertian yang dimiliki. Secara singkat dapat dikatakan bahwa tes uraian menuntut peserta tes untuk dapat mengingat-ingat dan mengenal kembali, dan terutama harus mempunyai daya kreativitas yang tinggi. (Eko Putro Widoyoko, 2016, hlm. 147)

f. Analisis Butir Tes

Analisis butir tes adalah pengkajian tentang isi butir-butir soal dalam instrumen tes agar diperoleh seperangkat butir tes yang memiliki kualitas yang memadai. Analisis butir tes bertujuan untuk mengidentifikasi butir-butir soal dalam instrumen tes yang baik, kurang baik dan tidak baik. Butir tes buatan guru pada umumnya tidak diujicobakan sebelum digunakan. Akibatnya banyak butir tes yang digunakan dalam ujian tidak dapat menghasilkan data yang benar atau akurat tentang hasil belajar siswa. Hal ini dapat berakibat jauh, karena hasil ujian seringkali digunakan untuk mengambil keputusan tentang masa depan siswa. Bila keputusan yang diambil didasarkan pada data yang tidak benar atau tidak akurat, yang disebabkan oleh instrumen yang digunakan untuk mengumbulkan data tidak disusun secara baik, maka tentu saja keputusan demikian merupakan keputusan yang tidak dapat dipertanggung jawabkan.

Ada beberapa alasan mengapa diperlukan analisis butir tes (Asmawi Nainul dan Noehi Nasution, dalam Eko Putro Widoyoko, 2016, hlm. 173). Alasan tersebut antara lain:

- 1) Untuk dapat mengetahui kekuatan dan kelemahan butir tes, sehingga dapat ditentukan butir yang baik atau yang harus direvisi.
- 2) Untuk menyediakan informasi tentang spesifikasi butir tes secara lengkap, sehingga akan lebih memudahkan bagi guru dalam menyusun perangkat soal yang akan memenuhi kebutuhan ujian dalam bidang dan tingkat tertentu.
- 3) Untuk dapat segera dapat diketahui masalah yang terkandung dalam butir soal, seperti: kesalahan meletakkan kunci jawaban, soal yang terlalu sulit atau terlalu mudah, atau soal yang tidak dapat membedakan siswa yang mempersiapkan diri dengan baik atau tidak dalam menghadapi ujian. Masalah ini, bila dapat diketahui dengan segera, akan memungkinkan guru mengambil keputusan apakah butir soal tersebut yang bermasalah itu akan digugurkan atau tidak dalam menentukan nilai siswa.
- 4) Untuk dijadikan alat guna menilai butir tes yang akan disimpan dalam bank soal. Bila seorang guru telah memiliki sejumlah butir tes (bank soal) yang baik, maka ia akan dengan mudah dapat menyusun suatu perangkat soal yang baik untuk digunakan sesuai dengan tujuan. Untuk memperoleh informasi tentang butir tes sehingga memungkinkan untuk menyusun beberapa perangkat soal yang paralel. Penyusunan perangkat seperti ini sangat bermanfaat bila akan melakukan ujian ulang atau mengukur hasil belajar beberapa kelompok siswa dalam waktu berbeda. (Eko Putro Widoyoko, 2016, hlm. 173-174)

Analisis butir tes dapat dilaksanakan berdasarkan teori tes modern dan teori tes klasik (*classical test theory and true score theory*). Teori tes modern dikenal dengan sebutan teori tanggapan butir (*item response theory atau IRT*). IRT adalah suatu paradigma untuk merancang tes dan memberikan skor tes, angket, serta perangkat tes yang mirip lainnya untuk mengukur kecakapan, sikap, maupun variabel lainnya. Teori modern terlalu kental unsur matematisnya, oleh karena itu untuk memudahkan guru, dalam uraian berikut ini difokuskan pada teori klasik yang relatif lebih sederhana. Analisis butir tes meliputi dua hal, yaitu karakteristik butir tes dan spesifikasi butir soal. Karakteristik butir tes merupakan parameter kuantitatif butir soal. Sedangkan spesifikasi butir tes merupakan parameter kualitatif butir tes yang ditentukan atas dasar penilaian ahli (*expert judgement*).

Biasanya hal-hal yang dianalisis dalam spesifikasi butir tes adalah hal-hal yang berkaitan dengan materi tes, konstruksi soal dan kaitannya dengan bahan serta budaya di masyarakat tempat butir tes itu disusun. Dalam bidang pengukuran, dikenal beberapa karakteristik butir soal. Untuk tes hasil belajar umumnya dipertimbangkan 3 (tiga) karakteristik butir soal, yaitu: tingkat kesulitan (*difficulty index*), daya beda (*discriminating power*), dan efektivitas pengecoh (*distractor effectivity*).

a) Tingkat Kesulitan

Tingkat kesulitan (*difficulty index, difficulty level*) butir tes adalah proporsi peserta tes menjawab dengan benar terhadap suatu butir tes. Sedangkan angka yang menunjukkan sulit atau mudahnya suatu butir tes dinamakan dengan indeks kesulitan yang dilambangkan dengan p (*proportion correct*). Makin besar nilai p berarti makin besar proporsi peserta tes yang menjawab benar terhadap suatu butir tes, makin rendah tingkat kesulitan butir tes itu, yang berarti butir tes itu makin mudah. Sebaliknya semakin kecil nilai p berarti semakin kecil proporsi peserta tes menjawab dengan benar suatu butir tes, makin tinggi tingkat kesulitan butir tes itu, yang berarti butir tes itu makin sulit. Tingkat kesulitan butir tes berkisar antara 0,0 sampai dengan 1,0. Bila butir tes mempunyai tingkat kesulitan 0,0 berarti tidak ada seorangpun peserta tes yang dapat menjawab dengan benar butir tes tersebut. Tingkat kesulitan 1,0 berarti semua peserta tes dapat menjawab dengan benar butir tes itu.

Nilai ideal tingkat kesulitan butir adalah lebih tinggi antara titik tengah peluang ($1,0$ dibagi dengan jumlah pilihan jawaban) dengan nilai sempurna ($1,0$) bagi setiap butir soal. Misalnya untuk soal pilihan ganda dengan 4 alternatif pilihan, peluang menjawab secara benar adalah $\frac{1}{4} = 0,25$, dengan demikian tingkat kesulitan optimalnya $0,25 + (1,0 - 0,25)/2 = 0,62$. Dalam pilihan benar salah, peluang menjawab benar adalah $\frac{1}{2} = 0,5$, sehingga tingkat kesulitan optimal adalah $0,5 + (1,0 - 0,5)/2 = 0,75$.

Rumus untuk menghitung tingkat kesulitan adalah:

$$p = \frac{\Sigma b}{N}$$

Keterangan:

p = tingkat kesulitan butir

Σb = jumlah peserta yang menjawab benar

N = jumlah peserta tes

Berdasarkan rumus di atas dapat diketahui bahwa tingkat kesulitan butir tes sangat dipengaruhi oleh tingkat kemampuan anggota kelompok peserta tes. Bila kelompok peserta tes mempunyai tingkat kemampuan yang berbeda, maka hasil indeks tingkat kesulitan juga akan berbeda pula. Dengan demikian tingkat kesulitan butir tes tidak sepenuhnya merupakan ukuran karakteristik butir tes saja, tetapi lebih merupakan kemampuan rata-rata kelompok peserta tes. Oleh karena itu apabila suatu butir test mempunyai tingkat kesulitan 0,60, maka interpretasinya adalah bahwa butir tes itu mempunyai tingkat kesulitan 0,60 untuk kelompok peserta tes tersebut, belum tentu berlaku untuk kelompok peserta tes lain. Kriteria yang digunakan untuk menentukan tingkat kesulitan dan kualitas butir tes di sajikan dalam tabel sebagai berikut: (lihat tabel 2)

Tabel 2. 2 Hubungan antara Tingkat Kesulitan dengan Kualitas Butir Tes

Tingkat Kesulitan	Kualitas Butir Soal
0,91 - 1,00	Sangat mudah, butir tes tidak baik, tidak digunakan
0,71 - 0,90	Mudah, butir tes kurang baik, direvisi
0,31 - 0,70	Sedang, butir tes cukup baik, digunakan
0,21 - 0,30	Sulit, butir tes kurang baik, direvisi
0,00 - 0,20	Sangat sulit, butir tes tidak baik, tidak digunakan

Sumber: (Eko Putro Widoyoko, 2016, hlm. 177)

Untuk tes hasil belajar, tingkat kesulitan yang dianggap baik adalah berkisar sekitar 0,50. Dengan kata lain, makin dekat tingkat kesulitan butir butir tes hasil belajar ke 0,50, makin baik butir tes tersebut bagi kelompok tertentu. Sebaliknya semakin jauh tingkat

kesulitan dari 0,50 maka makin kurang informasi yang diperoleh tentang butir tes dan kelompok peserta tes.

Untuk mengetahui tingkat kesulitan butir tes keseluruhan dalam suatu naskah soal tes maka dapat dilakukan dengan menjumlahkan tingkat kesulitan semua butir tes, kemudian dibagi dengan jumlah butir tes. Secara singkat tingkat kesulitan naskah soal dapat dirumuskan sebagai berikut:

$$P_s = \frac{\sum P_b}{N}$$

Keterangan:

P_s = tingkat kesulitan naskah tes

$\sum P_b$ = Jumlah tingkat kesulitan butir tes

N = jumlah butir tes

Untuk menyusun suatu naskah butir tes sebaiknya digunakan butir tes yang tingkat kesulitannya berimbang, yaitu sulit = 25%, sedang = 50%, dan mudah = 25%. Dengan komposisi seperti itu maka dapat diterapkan penilaian berdasarkan acuan norma maupun acuan patokan. Bila komposisi butir tes dalam naskah soal tidak berimbang, maka penggunaan penilaian acuan norma tidak tepat, karena informasi kemampuan yang dihasilkan tidak akan terdistribusi dalam suatu kurva normal. (Eko Putro Widoyoko, 2016:175-179)

Daya Beda Menurut Asmawi Zainul dan Noehi Nasution dalam Eko Putro Wdoyoko (2016, hlm. 180) “Daya beda (*discriminating power*) butir tes adalah indeks yang menunjukkan tingkat kemampuan butir tes membedakan antara peserta tes yang pandai (kelompok atas) dengan peserta tes yang kurang pandai (kelompok bawah) di antara peserta tes.”

Tujuan mencari daya beda adalah untuk menentukan apakah butir tes tersebut memiliki kemampuan membedakan kelompok dari aspek yang diukur, sesuai perbedaan yang ada pada kelompok tersebut. Karena daya beda dihitung dari hasil tes kelompok peserta ujian tertentu, maka penafsiran daya bedapun haruslah selalu dikaitkan dengan kelompok peserta tertentu tersebut. Daya beda suatu butir tes yang didasarkan pada hasil tes suatu kelompok belum tentu akan berlaku pada kelompok lain, apabila tingkat kemampuan masing-masing kelompok peserta tes itu berbeda.

Adapun rumus untuk mencari indeks daya beda adalah sebagai berikut:

$$D = \frac{Ba - Bb}{\frac{1}{2}N}$$

Keterangan:

D = daya beda

Ba = jumlah jawaban benar kelompok atas

Bb = jumlah jawaban benar kelompok bawah

N = jumlah peserta tes dalam kelompok atas dan bawah

Sumber: (Eko Putro Widoyoko, 2016 :180)

Indeks daya beda berkisar antara +1,0 sampai -1,0. Daya beda +1,0 berarti semua anggota kelompok atas menjawab benar terhadap butir tes tersebut, sedangkan kelompok bawah menjawab salah terhadap butir tes tersebut. Sebaliknya daya beda -1,0 berarti bahwa semua anggota kelompok atas menjawab salah butir tes tersebut, sedangkan kelompok bawah seluruhnya menjawab benar terhadap butir tes tersebut. Bila daya beda negatif maka butir tes sama sekali tidak dapat dipakai sebagai alat ukur prestasi belajar siswa. Karena itu butir tes tersebut harus dikeluarkan dari naskah soal. Makin tinggi daya beda suatu butir soal, maka semakin baik butir tes tersebut, dan sebaliknya makin rendah daya bedanya, maka butir tes tersebut dianggap makin tidak baik. (Eko Putro Widoyoko, 2016, hlm. 180)

Kriteria yang digunakan untuk menentukan indeks daya beda dan kualitas butir tes di sajikan dalam tabel berikut: (lihat table 3)

Tabel 2. 3 Hubungan antara Daya Beda dengan Kualitas Butir Soal

Daya Beda	Kualitas Butir Soal
0,51 - 1,00	Amat baik, dapat digunakan tanpa revisi
0,41 - 0,50	Baik, dapat digunakan tanpa revisi
0,31 - 0,40	Cukup baik, dapat digunakan dengan revisi
0,21 - 0,30	Kurang baik, perlu pembahasan dan revisi
0,00 - 0,20	Tidak baik, dibuang atau diganti

Sumber: (Eko Putro Widoyoko, 2016 :181)

Langkah-langkah menghitung daya beda:

- a. Susunlah urutan peserta tes berdasarkan skor yang diperolehnya, mulai dari skor yang tertinggi sampai ke skor terendah.
- b. Bagilah peserta tes menjadi dua kelompok yang sama jumlahnya. Bila jumlah peserta tes ganjil, maka peserta yang ditengah-tengah tidak perlu dimasukkan dalam salah

- satu kelompok (tidak dihitung). Kelompok pertama disebut kelompok prestasi tinggi atau kelompok atas (Ba) dan kelompok kedua disebut kelompok prestasi rendah atau kelompok bawah (Bb). Bila jumlah peserta cukup banyak (lebih dari 50), maka kelompok atas dan bawah masing-masing diambil 27%.
- c. Hitunglah jumlah kelompok atas yang menjawab benar terhadap butir tes yang akan dihitung indeks daya bedanya. Demikian pula untuk kelompok bawah.
 - d. Hitung dengan menggunakan rumus di atas
- Sumber: (Eko Putro Widoyoko, 2016 :181-82)

Berbeda dengan tingkat kesulitan, daya beda butir tes secara langsung menentukan kualitas butir tes dalam arti kualitas konstruksi butir tes. Bila suatu butir tes rendah daya bedanya, maka konstruksi butir tes tersebut dinilai tidak baik. Oleh karena itu maka bila akan merevisi butir tes, patokan utama yang akan digunakan adalah daya beda butir tes.

Efektivitas Pengecoh (*distractor effectivity*) berlaku pada tes pilihan ganda. Di antara pilihan jawaban yang ada, hanya satu yang benar. Selain jawaban yang benar tersebut, ada jawaban yang salah. Jawaban yang salah itulah yang dikenal dengan *distractor* (pengecoh). Butir tes yang baik, pengecohnya akan dipilih secara merata oleh peserta didik yang menjawab salah. Sebaliknya, butir tes yang kurang baik, pengecohnya akan dipilih secara tidak merata oleh peserta didik. Tujuan utama dari pemasangan *distractor* pada setiap butir tes adalah agar dari sekian banyak peserta tes yang mengikuti tes hasil belajar ada yang tertarik untuk memilihnya.

Distractor akan mengecoh peserta didik yang kurang mampu untuk dapat dibedakan dengan yang mampu. *Distractor* yang baik adalah yang dapat dihindari oleh peserta didik yang pandai dan akan dipilih oleh peserta didik yang kurang pandai. Dengan demikian *distractor* baru dapat dikatakan telah berfungsi dengan baik apabila distraktor tersebut telah memiliki daya rangsang atau daya tarik yang baik bagi peserta didik. Analisis efektivitas pengecoh menurut Hamzah B. Uno dan Satria Koni (2013) Analisis efektivitas pengecoh (*distractor*) atau analisis pola jawaban dilakukan dengan menghiyung peserta tes yang memilih tiap alternative jawaban pada masing-masing item.

Kriteria pengecoh yang baik adalah apabila pengecoh tersebut dipilih oleh paling sedikit 5% dari peserta tes. Efektivitas pengecoh dapat diukur menggunakan rumus:

$$IP = \frac{P}{N - B} \times 100\%$$

Keterangan :

IP = indeks pengecoh

P = jumlah peserta tes yang memilih pengecoh

N = jumlah peserta tes

B = jumlah peserta tes yang menjawab benar

Sumber: (Zainal Arifin, 2009, hlm. 279) Aspek yang menentukan kualitas butir tes selain tingkat kesulitan, daya beda dan efektivitas pengecoh adalah validitas butir.

5. Kemampuan Siswa

Kemampuan siswa dalam belajar adalah kecakapan seorang peserta didik, yang dimiliki dari hasil apa yang telah dipelajari yang dapat ditunjukkan atau dilihat melalui hasil belajarnya (Syah, 2010, hlm. 150). Ada tiga ranah (aspek) yang terkait dengan kemampuan siswa dalam belajar, yaitu ranah kognitif (pengetahuan), ranah afektif (sikap), dan ranah psikomotorik (keterampilan). Contoh ranah kognitif adalah kemampuan siswa dalam menganalisis suatu masalah berdasarkan pemahaman yang dimilikinya. Contoh ranah afektif adalah siswa mampu menentukan sikap untuk menerima atau menolak suatu objek. Contoh ranah psikomotorik adalah siswa mampu berekspresi dengan baik.

Setiap siswa dikatakan berhasil dalam belajar apabila memiliki kemampuan dalam belajar sebagaimana dikemukakan di atas. Akan tetapi yang menjadi masalah adalah tidak semua siswa memiliki kemampuan yang sama. Banyak faktor yang mempengaruhi kemampuan siswa dalam belajar, antara lain faktor internal, faktor eksternal, dan faktor pendekatan belajar. Contoh faktor internal yang mempengaruhi kemampuan siswa dalam belajar adalah kesehatan siswa dan intelegensinya. Siswa yang sehat dan mempunyai intelegensi yang baik akan mempunyai kesiapan yang lebih baik dalam belajar sehingga kemampuan belajarnya dapat optimal.

Sebaliknya siswa yang kurang sehat (sedang sakit) akan sulit menerima pelajaran sehingga kurang optimal kemampuan belajarnya. Contoh faktor eksternal yang mempengaruhi kemampuan siswa dalam belajar adalah lingkungan keluarga.

Lingkungan keluarga yang mendukung akan membuat siswa mudah untuk menerima pelajaran, sebaliknya lingkungan keluarga yang tidak mendukung, akan membuat siswa tidak tenang dalam belajar sehingga kemampuan siswa menjadi tidak optimal. Faktor pendekatan belajar yang berbeda juga akan memberikan kemampuan belajar yang berbeda. Siswa yang belajar secara mendalam akan memiliki kemampuan belajar yang lebih baik daripada siswa yang hanya belajar sambil lalu saja (tidak mendalam).

6. Ujian Nasional

Ujian Nasional adalah kegiatan penilaian hasil belajar peserta didik yang telah menyelesaikan jenjang pendidikan pada jalur sekolah/madrasah yang diselenggarakan secara nasional. Berdasarkan Permendikbud No. 66 tahun 2013 tentang Standar Penilaian Pendidikan, Ujian Nasional yang selanjutnya disebut UN merupakan kegiatan pengukuran kompetensi tertentu yang dicapai peserta didik dalam rangka menilai pencapaian Standar Nasional Pendidikan, yang dilaksanakan secara nasional. Eko (2010, hlm. 45)

Ujian Nasional adalah salah satu jenis dari kegiatan tes, dan kegunaan yang utama adalah untuk mengambil keputusan tentang orang yang diuji, misalnya untuk keperluan sertifikasi/kelulusan, seleksi, penjurusan, dan sebagainya. Dari beberapa pengertian di atas, dapat disimpulkan bahwa secara sederhana UN merupakan alat untuk menilai ketercapaian standar nasional pendidikan dalam rangka memberikan informasi dalam pengambilan keputusan bagi pemegang kebijakan pendidikan di Indonesia. Selanjutnya bertujuan akhir dalam rangka meningkatkan mutu pendidikan Indonesia. Sesuai dengan amanat Peraturan Pemerintah nomor 19 tahun 2005 tentang standar nasional pendidikan, tujuan penyelenggaraan UN adalah untuk menilai pencapaian kompetensi lulusan secara nasional pada mata pelajaran tertentu dalam kelompok mata pelajaran ilmu pengetahuan dan teknologi.

Fungsi ujian nasional sebagai alat pengendali mutu pendidikan secara nasional, pendorong peningkatan mutu pendidikan, bahan dalam menentukan kelulusan peserta didik, dan bahan pertimbangan dalam seleksi penerimaan peserta didik baru pada jenjang yang lebih tinggi. Sesuai dengan pernyataan tersebut dapat dikatakan jika hasil ujian nasional yang telah dilaksanakan dapat digunakan untuk mengetahui lulusan memiliki kompetensi yang memadai sehingga dapat digunakan untuk menentukan langkah ke jenjang berikutnya sesuai dengan kemampuannya berdasarkan hasil ujian.

Beberapa alasan yang mendasari UN menjadi sangat diperlukan untuk dilaksanakan antara lain (Puspendik, 2009:29-30).

- a) Ujian Nasional mendorong peningkatan mutu pendidikan di sekolah yaitu dengan mendorong siswa tekun belajar, guru mengajar lebih baik, mendorong pihak sekolah untuk memberikan perhatian yang ekstra, dan mendorong orangtua murid untuk memberikan perhatian dan motivasi terhadap pembelajaran anak.
- b) Ujian nasional merupakan entry point untuk meningkatkan mutu SDM dan daya saing bangsa. Ujian nasional dapat mendorong peningkatan mutu pendidikan selain berfungsi untuk mengukur dan menilai pencapaian kompetensi lulusan dalam mata pelajaran tertentu, serta pemetaan mutu pendidikan pada tingkat pendidikan dasar dan menengah, juga berfungsi sebagai motivator bagi pihak-pihak terkait untuk bekerja lebih baik guna mencapai hasil ujian yang baik.
- c) Penyelenggaraan UN merupakan salah satu upaya pemerintah dalam rangka peningkatan dan pemetaan mutu pendidikan di seluruh Indonesia. Meskipun UN sudah tidak lagi menjadi penentu kelulusan, tetapi hingga saat ini masih tetap dilaksanakan. Pelaksanaan UN tersebut tidak serta merta diselenggarakan pemerintah hanya karena peraturan yang telah ada sebelumnya, akan tetapi ada beberapa hal yang mendasari tetap dilaksanakannya UN. Sesuai dengan tulisan Suyanto (Kedaulatan Rakyat, 2016, hlm. 7) yang menyatakan

“bahwa beberapa hal yang mendasari tetap diselenggarakan UN yaitu pertama, untuk pemetaan mutu program dan/atau satuan pendidikan. Kedua, untuk dasar seleksi masuk jenjang

pendidikan berikutnya dan ketiga untuk pembinaan dan pemberian bantuan kepada satuan pendidikan dalam upayanya untuk meningkatkan mutu pendidikan”.

Data nilai UN secara nasional dapat digunakan untuk melihat mutu sekolah di seluruh tanah air, dari yang tinggi, sedang dan rendah. Dari hasil tersebut dapat digunakan untuk mempermudah melakukan pembinaan berbagai variabel penentu mutu dari satuan pendidikan, pendidik dan peserta didik. Selain itu, data tentang nilai jawaban soal UN dapat digunakan untuk menilai capaian kompetensi lulusan. Dari data tersebut sekaligus dapat diketahui pokok bahasan yang mana dalam suatu mata pelajaran yang para peserta mayoritas menjawab salah. Ketika sebagian besar siswa menjawab salah terhadap pertanyaan untuk materi tertentu, dapat disimpulkan kemampuan guru untuk menyampaikan materi ajar masih lemah. Sehingga dapat dilaksanakan pembinaan guru untuk meningkatkan kompetensinya dalam mengimplementasikan kurikulum dalam proses belajar-mengajarnya.

B. Hasil Penelitian Terdahulu

Tabel 2. 4 Penelitian Terdahulu

No	Nama peneliti/tahun	Judul	Tempat penelitian	Hasil Penelitian	Persamaan	Perbedaan
1	Aditya Melia Nugrahanti/tahun 2013	Analisis Butir Soal Ujian Akhir Semester Gasal Mata Pelajaran Akuntansi Keuangan Kelas XI Kompetensi keahlian Akuntansi SMK Negeri 1 Yogyakarta Tahun Ajaran 2012/2013	SMK N 1 Yogyakarta	<ol style="list-style-type: none"> 1) dari 30 butir soal pilihan ganda dan 4 butir soal uraian yang dianalisis, berdasarkan tingkat validitas item pada bentuk soal pilihan ganda butir soal yang valid sebesar 70% dan soal yang tidak valid sebesar 30%, sedangkan bentuk soal uraian semua soal dinyatakan valid; 2) berdasarkan reliabilitas termasuk soal yang reliabilitasnya rendah, pada bentuk soal pilihan ganda 0,610 sedangkan bentuk soal uraian 0,49; 3) berdasarkan daya pembeda butir soal yang daya pembedanya jelek sebesar 20%, cukup sebesar 10%, baik sebesar 10%, baik sekali sebesar 60%, sedangkan pada bentuk soal uraian yang memiliki daya pembeda jelek sebesar 75% dan cukup sebesar 25%; 	deskriptif kuantitatif	Pada subjek dan objek penelitian, tempat, dan waktu penelitian.

No	Nama peneliti/tahun	Judul	Tempat penelitian	Hasil Penelitian	Persamaan	Perbedaan
				<p>4) berdasarkan tingkat kesukaran butir soal pilihan ganda yang sukar sebesar 10%, sedang sebesar 53,33%, dan mudah sebesar 36,67%, sedangkan 37 bentuk soal uraian yang termasuk sukar sebesar 50%, sedang 25%, dan mudah 25%;</p> <p>5) berdasarkan efektivitas pengecoh butir soal yang memiliki pengecoh yang berfungsi sangat baik sebesar 33,33%, baik sebesar 23,33%, cukup sebesar 20%, kurang baik sebesar 16,67%, dan tidak baik sebesar 6,67%</p>		
2	Tri Setya Ernawati pada tahun 2013	Analisis Butir Soal Ujian Akhir Semester Genap Buatan Guru Akuntansi Program Keahlian Akuntansi Kelas X di SMK Negeri 1 Bantul Tahun Ajaran 2013/2014	SMK Negeri 1 Bantul	<p>1) dari 80 butir yang dianalisis, berdasarkan tingkat validitas item butir soal yang valid berjumlah 61 butir (76,25%), dan soal yang tidak valid berjumlah 19 butir (23,75%)</p> <p>2) berdasarkan reliabilitas termasuk soal yang memiliki reliabilitas sangat tinggi yaitu 0,820</p> <p>3) butir soal yang termasuk</p>	deskriptif kuantitatif	subjek dan objek penelitian, tempat, dan waktu penelitian.

No	Nama peneliti/tahun	Judul	Tempat penelitian	Hasil Penelitian	Persamaan	Perbedaan
				<p>kriteria mudah sejumlah 57 butir (71,25%), berkriteria sedang sebanyak 19 butir (23,75%) dan berkriteria sukar sebanyak 4 butir (5%)</p> <p>4) butir soal dengan daya pembeda jelek berjumlah 38 butir (47,5%), cukup berjumlah 28 butir (35%), baik berjumlah 12 butir (15%), baik sekali berjumlah 0 butir (0%), dan tidak baik berjumlah 2 butir (2,5%); (5) butir soal yang memiliki penyebaran jawaban soal yang berkategori sangat baik berjumlah 6 butir 38 (7,5%), berkategori baik berjumlah 9 butir (11,25%), kategori cukup sebanyak 22 butir (27,5%), kategori kurang baik sebanyak 21 butir (26,25%), dan yang tidak baik berjumlah 22 butir (27,5%)</p>		
3	Amelia Dina pada tahun 2017/2018	Analisis Kemampuan siswa dalam menjawab soal ujian Nasional	SMA Pasundan 1 Bandung		deskriptif kuantitatif	subjek dan objek penelitian, tempat, dan waktu penelitian.

No	Nama peneliti/tahun	Judul	Tempat penelitian	Hasil Penelitian	Persamaan	Perbedaan
		materi pelajaran ekonomi di SMA Pasundan 1 Bandung dan SMA Pasundan 2 Bandung	dan SMA Pasundan 2 Bandung			

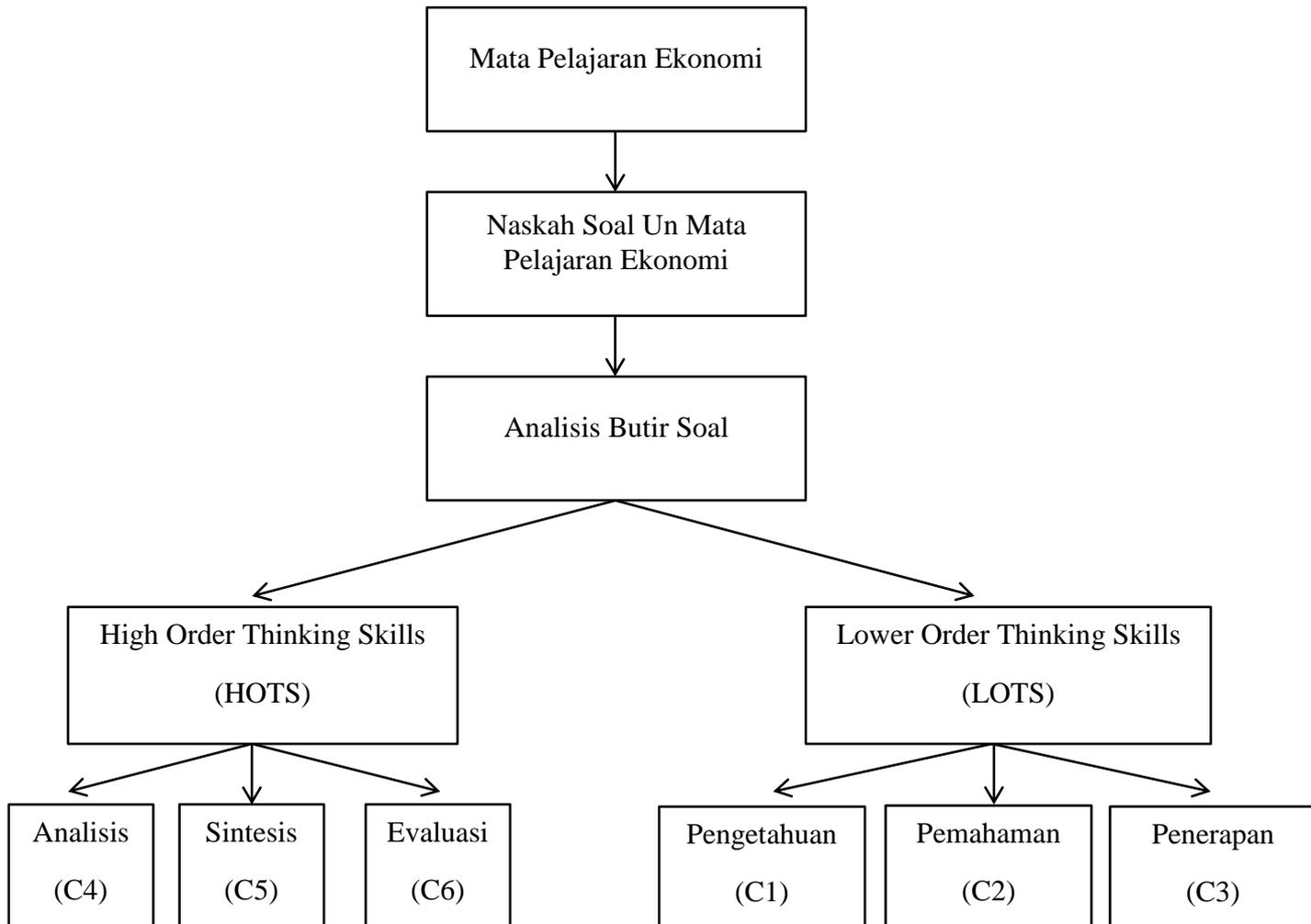
C. Kerangka Pemikiran

Guru seharusnya memiliki pengetahuan dan kemampuan untuk melaksanakan evaluasi terhadap proses maupun hasil belajar siswa. Evaluasi merupakan salah satu tahap yang harus ditempuh oleh guru untuk mengetahui keefektifan pembelajaran. Hasil evaluasi dapat digunakan guru sebagai umpan balik untuk memperbaiki program dan kegiatan pembelajaran. Evaluasi pembelajaran salah satu yang dapat dilakukan oleh guru adalah melalui teknik tes. Tes merupakan alat ukur untuk mengetahui seberapa besar proses belajar di dalam kelas tersebut berhasil. Seorang guru perlu memiliki kemampuan untuk membuat sebuah tes. Tes yang dapat mengukur seberapa besar pemahaman siswa mengenai materi yang disampaikan guru.

Kualitas tes dapat diketahui guru dengan melakukan analisis kualitas tes. Analisis kualitas tes tersebut dapat diukur menurut indikator kualitas soal antara lain validitas, reliabilitas, tingkat kesukaran, daya pembeda butir, dan efektivitas pengecoh. Analisis butir soal bermanfaat untuk mengetahui tingkat kualitas soal. Dari hasil analisis butir soal dapat diketahui soal yang perlu diperbaiki atau direvisi, dibuang, dan disimpan di bank soal. Dengan dilakukan analisis butir soal maka diharapkan soal yang dijadikan tes mendatang dapat berkualitas baik sehingga dapat mengukur hasil belajar peserta didik secara tepat.

Analisis butir soal terhadap soal Ujian Nasional Mata Pelajaran Ekonomi Kelas XI IPS dilaksanakan pada tahun ajaran 2016/2017. Kegiatan analisis butir soal tersebut dilakukan dengan lembar jawaban siswa kelas XI IPS di SMA Pasundan 1 dan SMA Pasundan 2 Bandung. Lembar jawab tersebut dibandingkan dengan kunci jawaban yang telah dibuat oleh guru.

Bagan berikut merupakan skema dari kerangka berpikir:



Tabel 2. 5 Gambar Kerangka Pemikiran

D. Asumsi dan Hipotesis Penelitian

1. Asumsi

Asumsi menurut Arikunto (2010, hlm. 106) adalah suatu hal yang diyakini kebenarannya oleh peneliti harus dirumuskan secara jelas. Peneliti harus merumuskan asumsi

- a) Siswa menjawab soal ujian nasional dengan tepat karena soal yang sesuai dengan kemampuan siswa baik dari tingkat kemudahan dan tingkat kesukaran yang sebelumnya sudah di analisis terlebih dahulu.
- b) Guru ekonomi di anggap memiliki kemampuan dan keterampilan yang memadai dalam menganalisis soal ujian nasional mata pelajaran ekonomi.

2. Hipotesis

Menurut Sugiyono (2016, hlm. 63) mendefinisikan bahwa hipotesis adalah jawaban sementara terhadap rumusan masalah penelitian telah dinyatakan dalam bentuk kalimat pertanyaan. Dikatakan sementara, karena jawaban yang diberikan baru didasarkan pada teori yang relevan, belum didasarkan pada fakta-fakta empiris yang diperoleh melalui pengumpulan data.

Berdasarkan kajian pustaka serta kerangka pemikiran yang terlebih dahulu di kemukakan, maka hipotesis yang diajukan dalam penelitian ini adalah “Analisis Kualitas Butir Soal Ujian Nasional dan Kemampuan Siswa dalam Menjawab Soal Ujian Nasional Mata Pelajaran Ekonomi”.

