

BAB II

KAJIAN TEORITIS

A. Landasan Teori

1. Tinjauan Tentang Evaluasi Hasil Belajar

a. Pengertian Evaluasi Hasil Belajar

Evaluasi secara bahasa berarti *evaluation* (Inggris), *al-taqdiir* (Arab), dan penilaian (Indonesia). Istilah “evaluasi pendidikan” memiliki arti sebagai evaluasi kegiatan pendidikan atau evaluasi pada bidang pendidikan. Nilai yang berguna untuk mengukur sejauh mana tujuan pembelajaran telah tercapai, dan dapat berguna untuk pengambilan keputusan pada pembelajaran setelahnya.

Berdasarkan Undang Undang Nomor 20 Tahun 2003 tentang Sistem Persekolahan Negeri Bagian 1 Pasal 1 ayat 21, memberi pengertian bahwa “Evaluasi pendidikan merupakan sebuah kegiatan guna mengendalikan, memastikan dan memutuskan mutu pendidikan dari berbagai komponen pada setiap jalur, jenjang dan jenis pendidikan sebagai suatu jenis tanggung jawab mengenai pelaksanaan sekolah”.²¹ Adapun hasil dari evaluasi hasil belajar dapat digunakan oleh pendidik untuk membuat kebijakan atau mengambil keputusan dalam merencanakan program pendidikan selanjutnya.

²¹) Undang Undang Republik Indonesia Nomor 20 Tahun 2003 tentang Sistem Pendidikan Nasional, diakses 20 Mei 2024, <https://peraturan.bpk.go.id/Details/43920/uu-no-20-tahun-2003>.

b. Tujuan dan Fungsi Evaluasi Hasil Belajar

Tujuan evaluasi menurut Sudijono dapat dibagi menjadi dua, antara lain²²:

1) Tujuan umum

Pengumpulan data yang akan digunakan sebagai bukti bahwa terdapat perkembangan yang ditempuh oleh peserta didik setelah mengikuti serangkaian proses pembelajaran dalam jangka waktu tertentu dan untuk mengetahui tingkat keefektifan metode pembelajaran yang digunakan pada proses pembelajaran dalam jangka waktu tertentu pula.

2) Tujuan khusus

Mendorong peserta didik untuk dapat berpartisipasi dalam program pendidikan dan menyelidiki faktor-faktor yang menentukan berhasil atau tidaknya peserta didik sehingga dapat ditemukan solusi dan cara untuk memperbaikinya.

Fungsi evaluasi adalah untuk mengukur kemajuan, menyusun rencana, memperbaiki, serta melakukan penyempurnaan. Terdapat beberapa fungsi evaluasi berdasarkan beberapa aspek dalam dunia pendidikan, antara lain²³:

²²⁾ Anas Sudijono, *Pengantar Evaluasi Pendidikan*, (Yogyakarta: PT Raja Grafindo Persada, 2015), 7-17.

²³⁾ Suharsimi Arikunto, *Dasar-Dasar Evaluasi Pendidikan*, (Jakarta: Bumi Aksara, 2013), 18-19.

1) Evaluasi Berfungsi sebagai Selektif

Melalui evaluasi, guru dapat mengadakan seleksi bagi peserta didiknya. Seleksi ini berguna untuk berbagai kondisi, seperti penerimaan peserta didik, kenaikan kelas, dan kelulusan.

2) Evaluasi Berfungsi sebagai Diagnostik

Melalui evaluasi, guru dapat mendiagnosis peserta didik mengenai kebaikan, kelemahan dan penyebab kelemahan tersebut. Setelah mengetahui penyebab kelemahan-kelemahan tersebut, maka guru dapat lebih mudah dalam mencari solusinya.

3) Evaluasi Berfungsi sebagai Penempatan

Melalui evaluasi, guru dapat menempatkan seorang peserta didik pada suatu kelompok belajar. Hal tersebut dilakukan karena setiap peserta didik memiliki kemampuan dan pembawaan yang berbeda-beda.

4) Evaluasi Berfungsi sebagai Alat Ukur Keberhasilan

Evaluasi dilakukan untuk mengetahui apakah suatu program di sekolah berhasil diterapkan. Adapun keberhasilan suatu program pembelajaran juga dipengaruhi oleh beberapa faktor, antara lain guru, strategi pembelajaran, kurikulum, fasilitas, dan sistem administrasi.

c. Prinsip-Prinsip Evaluasi Hasil Belajar

Prinsip berguna untuk memberikan gambaran menyeluruh tentang proses pendidikan atau pembelajaran, kemajuan siswa, dan prestasi

siswa. Oleh karena itu, pendidik atau guru perlu mewujudkan prinsip-prinsip evaluasi pendidikan dalam konteks yang sesungguhnya. Menurut Slameto yang dikutip oleh Sukardi, evaluasi pendidikan harus memiliki tujuh prinsip berikut: (1) terpadu, (2) menganut cara belajar, (3) kontinuitas, (4) koherensi dengan tujuan, (5) menyeluruh, (6) membedakan, dan (7) pedagogis.²⁴ Sedangkan menurut Sudijono, evaluasi hasil belajar dapat dikatakan terlaksana dengan baik apabila dalam pelaksanaannya berpegang teguh pada tiga prinsip berikut:

1) Prinsip Keseluruhan

Kegiatan evaluasi hasil belajar dapat dilaksanakan secara menyeluruh dan mencakup semua aspek, yaitu kognitif, afektif, dan psikomotorik peserta didik. Hal ini dilakukan agar guru memperoleh informasi yang lengkap dari peserta didik mengenai keadaan dan perkembangannya.

2) Prinsip Kesenambungan

Sebuah evaluasi hasil belajar harus dilaksanakan secara teratur, sehingga akan diketahui kemajuan yang dimiliki dari peserta didik. Hal ini dilakukan agar pendidik dapat menentukan apa saja langkah-langkah yang harus dilakukan selanjutnya.

3) Prinsip Objektivitas

Kegiatan evaluasi hasil belajar harus dilakukan berdasarkan keadaan nyata tanpa melalui campur tangan yang bersifat subjektif.

²⁴) Sukardi, *Evaluasi Pendidikan Prinsip dan Operasionalnya*, (Jakarta: Bumi Aksara, 2009), 4-5.

Hal ini dilakukan agar guru sebagai evaluator dapat berfikir dan bertindak secara benar.

d. Teknik-Teknik Evaluasi Hasil Belajar

Teknik evaluasi hasil belajar pada proses pembelajaran di sekolah dikenal menjadi dua macam, yaitu²⁵:

1) Teknik Tes

Pendidik dapat menggunakan tes sebagai alat ukur untuk menilai kemampuan siswa sepanjang proses pembelajaran. Teknik tes dapat dilakukan dengan cara memberikan tugas atau memberikan pertanyaan yang harus dijawab oleh peserta tes untuk mengukur perkembangan peserta didik.

2) Teknik Non-Tes

Teknik non-tes merupakan sebuah teknik penilaian yang dilakukan tanpa melakukan ujian bagi peserta didik, namun dilakukan melalui observasi, wawancara, survei, dan analisis catatan (pemeriksaan naratif).

2. Tinjauan Tentang Instrumen Tes

a. Pengertian Tes

Tes adalah instrumen yang digunakan sebagai pengukuran, khususnya alat untuk mengumpulkan data tentang kualitas suatu item.²⁶

Beberapa pendapat mengenai definisi tes. Menurut Djemari, tes

²⁵) Anas Sudijono, *Pengantar Evaluasi Pendidikan*, (Jakarta: PT Raja Grafindo Persada, 2015), 65.

²⁶) Eko Putro Widoyoko, *Penilaian Hasil Pembelajaran di Sekolah*, (Yogyakarta: Pustaka Pelajar, 2016), 117.

merupakan salah satu cara untuk menaksir besarnya kemampuan seseorang secara tidak langsung, yaitu melalui respons seseorang terhadap stimulus atau pertanyaan.²⁷ Cronbach mengartikan tes sebagai *“a methodical approach to observing and describing a person's behavior using a numerical scale or category system”*.²⁸

Berdasarkan pengertian berikut, dapat ditarik kesimpulan bahwa tes adalah suatu alat evaluasi yang digunakan untuk menguji dan mengetahui sejauh mana kemampuan peserta didik dalam menempuh proses pembelajaran.

b. Pengertian Tes Hasil Belajar

Tes hasil belajar (THB) merupakan sebuah tes penguasaan, karena tes ini berguna sebagai alat ukur tingkat kemampuan peserta didik pada materi yang diajarkan oleh pendidik atau yang dipelajari secara mandiri oleh peserta didik.²⁹ THB dilakukan untuk mengukur hasil belajar, yaitu sejauh mana tujuan pembelajaran telah dicapai oleh peserta didik. THB diujikan setelah peserta didik memperoleh sejumlah materi dan pengujian dilakukan guna mengetahui tingkat pemahaman atau penguasaan peserta didik atas materi tersebut.

²⁷⁾ *Ibid.*

²⁸⁾ Purwanto, *Evaluasi Hasil Belajar*, (Yogyakarta: Pustaka Pelajar, 2013), 64.

²⁹⁾ *Ibid.*, 66.

c. Fungsi Tes Hasil Belajar

Menurut Anas Sudijono, fungsi tes hasil belajar ada dua macam, yaitu³⁰:

- 1) Sebagai alat ukur. Tes berfungsi untuk mengukur seberapa jauh perkembangan atau kemajuan individu peserta didik setelah mengikuti kegiatan pembelajaran pada jangka waktu tertentu.
- 2) Sebagai alat ukur keberhasilan suatu program pembelajaran. Melalui tes, maka dapat diketahui tingkat keberhasilan yang dicapai pada sebuah program pembelajaran.

d. Macam-Macam Tes Hasil Belajar

Menurut peranan fungsionalnya dalam pembelajaran, THB dapat dibagi menjadi empat macam, yaitu³¹:

1) Tes Formatif

Tes formatif merupakan tes yang digunakan guna mengukur kemampuan peserta didik setelah mengikuti serangkaian proses pembelajaran. Pengujian ini dilakukan untuk mengetahui sejauh mana proses pembelajaran dapat membentuk peserta didik sesuai dengan tujuan pembelajaran yang dilakukan. Adapun tes formatif dalam praktik pembelajaran dikenal dengan ulangan harian.

³⁰) Anas Sudijono, *Op.Cit.*, 67.

³¹) Purwanto, *Op.Cit.*, 67-70

2) Tes Sumatif

Tes sumatif merupakan tes yang digunakan untuk mengetahui penguasaan peserta didik terhadap semua jumlah materi yang disampaikan pada satuan waktu tertentu seperti catur wulan atau semester. Tes sumatif diujikan setelah semua materi selesai disampaikan. Adapun dalam praktik pembelajaran, tes sumatif dikenal sebagai UAS, ujian sekolah/madrasah, atau ujian semesteran, tergantung satuan waktu yang digunakan dalam menyelesaikan materi pembelajaran.

3) Tes Diagnostik

Tes diagnostik merupakan tes yang digunakan untuk mengidentifikasi kekuatan dan kelemahan peserta didik terhadap proses pembelajaran.³² THB digunakan untuk mengidentifikasi permasalahan yang dialami peserta didik serta menelusuri jenis masalah yang dihadapi. Berdasarkan hasil permasalahan yang didapatkan melalui tes diagnostik, maka pendidik dapat mengusahakan untuk mencari pemecahan masalah yang dihadapi peserta didik.

4) Tes Penempatan

Tes penempatan merupakan tes yang digunakan untuk menempatkan peserta didik menjadi beberapa kelompok yang

³²⁾ Muhammad As'ari dan Andri Wahyu Wijayadi, "Pengembangan Tes Diagnostik Kognitif Materi Kalor dan Perpindahan Kelas VII SMP", *Jurnal Ed Humanistic* 08, no 02 (2023): 1009, <https://doi.org/10.33752/ed-humanistics.v8i02.5109>

sesuai dengan masing-masing minat dan bakatnya. Penempatan atau pengelompokan dilakukan guna memberikan pelayanan yang maksimal, yaitu sesuai dengan minat dan bakat peserta didik. Tes penempatan ini biasa dilakukan pada penempatan peserta didik baru ke dalam suatu kelompok belajar atau jurusan tertentu.

e. Bentuk Tes Hasil Belajar

Tes Hasil Belajar yang dapat digunakan pada pelaksanaan evaluasi hasil belajar ada dua, yaitu:

1) Tes Objektif

Tes obyektif, yaitu tes yang dapat dipengaruhi oleh obyek tanggapan atau reaksi yang diberikan oleh peserta didik sebagai peserta tes.³³ Adapun tes objektif dapat dibedakan menjadi lima golongan, yaitu³⁴:

- a) Tes objektif jenis benar-salah
- b) Tes objektif jenis menjodohkan
- c) Tes objektif jenis melengkapi
- d) Tes objektif jenis isian
- e) Tes objektif jenis pilihan ganda

2) Tes Subjektif

Tes yang menggunakan respon siswa sebagai peserta tes dan penilaian subjektif pendidik sebagai pemberi nilai dikenal dengan

³³) Eko Putro Widiyoko, *Evaluasi Program Pembelajaran*, (Yogyakarta: Pustaka Pelajar, 2009), 46.

³⁴) Anas Sudijono, *Op.Cit.*, 107.

tes subjektif.³⁵ Tes subjektif atau tes uraian mengharuskan peserta didik untuk dapat menjawab berbagai pertanyaan yang diajukan pendidik dalam berbagai bentuk, seperti uraian, diskusi, perbandingan, dan memberikan alasan menggunakan bahasa sendiri.³⁶

f. Langkah-Langkah dalam Penyusunan Tes Hasil Belajar

Menurut Mardapi, langkah-langkah yang dapat dilakukan dalam penyusunan tes hasil belajar, yaitu³⁷:

- 1) Menyusun kisi-kisi.
- 2) Menulis soal tes sesuai dengan kisi-kisi.
- 3) Menelaah soal tes, berupa pengecekan terhadap butir soal untuk mengetahui apakah dalam pembuatan soal masih terdapat kekurangan atau kesalahan.
- 4) Melakukan uji coba tes, guna memperoleh data empirik mengenai tingkat kualitas soal yang disusun.
- 5) Menganalisis kualitas butir soal, dilakukan pada uji tingkat daya pembeda, tingkat kesukaran soal, serta efektivitas pengecohnya.
- 6) Perbaiki soal yang belum sesuai setelah dilakukan uji coba.
- 7) Merakit tes, berupa penyusunan seluruh butir soal menjadi satu kesatuan soal yang padu dalam naskah soal.

³⁵) Eko Putro Widiyoko, *Loc.Cit.*

³⁶) Nana Sudjana, *Penilaian Hasil Proses Belajar Mengajar*, (Bandung: PT. Remaja Rosdakarya, 2011), 35.

³⁷) Tri Rahayu Setyariningsih, “Analisis Butir Soal Penilaian Akhir Tahun (PAT) Mata Pelajaran IPA Kelas VIII SMP Negeri 3 Pabelan Tahun Pelajaran 2019/2020”, (Skripsi, IAIN Salatiga, 2020), 20-21.

- 8) Melaksanakan tes, dilaksanakan dengan pantauan dan pengawasan agar tes dapat terselenggara dengan jujur dan sesuai pada ketentuan yang telah ditentukan.
- 9) Menafsirkan hasil tes, berupa penafsiran hasil data berupa skor menjadi nilai apakah termasuk rendah, menengah, atau tinggi.

g. Tes Hasil Belajar yang Baik

Tes Hasil Belajar dapat dikatakan sebagai alat pengukuran yang baik apabila telah memenuhi persyaratan tes, antara lain³⁸:

1) Validitas

Tes dikatakan valid apabila tes tersebut dapat memberikan data dengan tepat dan relevan. Sebuah data dikatakan valid sesuai dengan kenyataan yang sesungguhnya.

2) Reliabilitas

Sebuah Tes dikatakan reliable apabila tes tersebut dapat digunakan secara berulang kali dan memberikan hasil yang sama meskipun telah diteskan.

3) Objektivitas

Sebuah tes dapat memiliki objektivitas yang baik apabila dalam pelaksanaan tes tidak terdapat faktor lain yang memengaruhi.

4) Praktibilitas

Praktibilitas dalam tes diartikan bahwa tes tersebut bersifat praktis. Tes yang praktis merupakan tes yang mudah, baik dalam

³⁸⁾ Suharsimi Arikunto, *Dasar-Dasar Evaluasi Pendidikan*, (Jakarta: Bumi Aksara, 2013), 72-77.

pengerjaannya, pengkoreksiannya, dan pengadministrasiannya, serta dilengkapi dengan berbagai petunjuk yang jelas sehingga dapat digunakan kembali.

5) Ekonomis

Tes dapat dikatakan ekonomis apabila proses pelaksanaan tes tidak membutuhkan biaya yang mahal, serta tidak membutuhkan waktu dan tenaga yang banyak.

3. Tinjauan Tentang Analisis Butir Soal

a. Pengertian Analisis Butir Soal

Analisis butir soal merupakan suatu kegiatan yang harus dilakukan guru guna meningkatkan mutu soal yang telah disusun, sehingga dapat dihasilkan sebuah instrument tes yang berkualitas. Kegiatan ini mencakup tentang proses pengumpulan data, peringkasan dan penggunaan informasi dari jawaban peserta didik untuk membuat keputusan tentang tes hasil belajar.³⁹

b. Program Analisis Butir Soal

Kegiatan analisis butir soal dapat dilakukan secara manual, yaitu menggunakan kalkulator. Namun, cara tersebut memerlukan waktu yang lama, dan tingkat keakuratan perhitungan lebih rendah dibandingkan dengan pengolahan menggunakan program komputer.

³⁹⁾ Elis Ratnawulan dan A. Rusdiana, *Evaluasi Pembelajaran*, (Bandung: Pustaka Setia, 2015), 148.

Sehingga dalam menganalisis butir soal diperlukan *software* untuk membantu proses analisis butir soal tersebut.

Saat ini, terdapat banyak program computer yang telah dikembangkan dan dapat digunakan untuk melakukan analisis butir soal, antara lain: ANATES, *Microsoft Office Excell*, ITEMAN, serta SPSS. Program komputer yang akan digunakan dalam penelitian ini adalah SPSS versi 16 dan *Microsoft Office Excell*.

c. Langkah-Langkah Analisis Butir Soal

Sebagai upaya untuk menjamin kualitas dan efektivitas dalam melakukan analisis, maka akan dilakukan analisis butir soal dengan langkah-langkah yang sistematis. Proses ini meliputi tahapan persiapan data, melakukan perhitungan parameter statistik, hingga interpretasi hasil analisis untuk mengidentifikasi karakteristik setiap butir soal. Adapun langkah-langkah melakukan analisis butir soal meliputi⁴⁰:

- 1) Mengumpulkan data hasil tes belajar, dalam hal ini berupa lembar jawab asesmen.
- 2) Menyiapkan kunci jawaban yang valid.
- 3) Mendeskripsikan data yang telah dikumpulkan dan membuat tabulasi data menggunakan *Microsoft Excel* untuk memudahkan proses analisis.

⁴⁰⁾ Santosa, Sedya dan Jami Ahmad Badawi, "Analisis Butir Soal Pilihan Ganda Tema Pertumbuhan dan Perkembangan Makhluq Hidup Kelas III Madrasah Ibtidaiyah," *Jurnal Basicedu* 6, no 2 (2022): 1680, <https://jbasic.org/index.php/basicedu/article/download/2206/pdf/8949>

- 4) Menganalisis data dengan menghitung validitas, reliabilitas, derajat kesukaran, daya pembeda item, dan fungsi pengecoh/*distractor*.
- 5) Merumuskan simpulan dari hasil analisis.

4. Uji Kualitas Butir Soal

a. Validitas

Validitas secara bahasa berasal dari kata *validity*, yang berarti ketepatan, yaitu ketepatan sebuah instrument atau alat ukur mampu memenuhi fungsi ukurnya.⁴¹ Suatu tes hasil belajar dapat valid jika tes tersebut dapat memberikan hasil yang sesuai dengan perkiraan yang diharapkan. Semakin banyak butir item yang dapat dijawab dengan benar oleh peserta didik, maka nilai total hasil tes tersebut akan semakin tinggi. Sebutir item dapat dikatakan valid atau telah memiliki validitas yang tinggi apabila terdapat kesesuaian antara skor-skor pada butir soal dengan skor total.⁴²

b. Reliabilitas

Reliabilitas merujuk pada ketepatan suatu alat ukur guna menilai apa yang dinilainya. Reliabilitas instrument tes hasil belajar merupakan syarat yang dilakukan untuk pengujian validitas instrument. Meskipun

⁴¹⁾ Shodiq Abdullah, *Evaluasi Pembelajaran*, (Semarang: Pustaka Rizki Putra, 2012), 76-77.

⁴²⁾ Afrida, dkk., “Analisis Kualitas Butir Soal Ujian Semester Ganjil Pada Mata Pelajaran Pendidikan Kewarganegaraan Kelas V MI,” *Jurnal Keilmuan dan Kependidikan Dasar* 12, no 02 (2020): 117, <https://ftk.uinbanten.ac.id/journals/index.php/primary/article/download/3530/2782/11100>

instrument yang valid pada umumnya pasti reliable, namun pengujian reliabilitas instrument perlu dilakukan.⁴³

c. Derajat Kesukaran

Butir soal dapat dikatakan sebagai soal yang baik apabila butir item tersebut memiliki tingkat kesukaran yang sedang atau cukup.⁴⁴ Soal yang terlalu mudah tidak dapat mendorong peserta didik untuk berusaha mencari jawaban, dan soal yang terlalu sulit dapat menyebabkan peserta didik menjadi putus asa dan tidak ingin untuk mencoba lagi karena merasa diluar jangkauannya. Tingkat kesukaran tidak dilihat dari sudut pandang pendidik sebagai pembuat soal, namun dipandang dari tingkat kesanggupan atau kemampuan peserta didik dalam menjawabnya. Setelah mengetahui tingkat kesukaran butir soal, tindakan yang dapat dilakukan, yaitu⁴⁵:

- 1) Butir soal dengan tingkat kesukaran kategori baik dapat disimpan di bank soal agar bisa digunakan kembali dalam tes hasil belajar yang akan datang.
- 2) Butir soal dengan kategori sukar, terdapat tiga kemungkinan tindak lanjut, yaitu;
 - a) Butir soal tersebut dibuang dan tidak digunakan lagi pada tes hasil belajar yang akan datang;

⁴³) Eko Putro Widoyoko, *Evaluasi Program Pembelajaran*, (Yogyakarta: Pustaka Plajar, 2017), 144.

⁴⁴) Suharsimi Arikunto, *Op.Cit.*, 222.

⁴⁵) Anas Sudijono, *Op.Cit.*, 376-378.

- b) Diteliti, direvisi, dan ditelusuri sehingga diketahui penyebab butir soal tersebut sukar dijawab oleh peserta didik. Perbaikan dapat dilakukan dengan menyederhanakan kalimat soal, sehingga lebih mudah dipahami. Setelah dilakukan perbaikan, butir soal dapat digunakan lagi pada tes yang akan datang;
 - c) Butir soal tetap dipertahankan dan dapat digunakan lagi pada tes seleksi, dalam artian sebagian *testee* tidak akan diluluskan dalam tes seleksi tersebut.
- 3) Butir tes dengan kategori mudah, terdapat tiga kemungkinan tindak lanjut:
 - a) Butir soal tersebut harus dibuang dan tidak digunakan lagi pada tes hasil belajar yang akan datang;
 - b) Diteliti ulang, direvisi, dan ditelusuri kesulitannya. Perbaikan soal dapat dilakukan dengan menyederhanakan kalimat soal, sehingga dapat dengan mudah dipahami. Setelah dilakukan perbaikan, butir soal dapat digunakan kembali pada tes yang akan datang.
 - c) Butir soal tetap dipertahankan untuk digunakan kembali pada tes yang sifatnya seleksi, yaitu sebagian *testee* tidak akan diluluskan.

d. Daya Pembeda Item

Daya pembeda item merupakan kemampuan butir soal untuk membedakan peserta didik yang termasuk kelompok atas dengan

peserta didik kelompok bawah.⁴⁶ Sebuah tes hasil belajar dapat dikatakan memiliki daya pembeda yang baik apabila butir soal tersebut diberikan kepada peserta didik yang pandai, maka hasil belajarnya tinggi. Sebaliknya, apabila butir soal tersebut diberikan kepada peserta didik yang kurang atau lemah, maka didapatkan hasil belajar yang rendah. Adapun tindak lanjut terhadap butir soal yang telah diukur daya pembedanya adalah sebagai berikut⁴⁷:

- 1) Butir soal yang memiliki daya pembeda baik akan disimpan dalam bank soal. Butir soal tersebut dapat digunakan kembali pada tes hasil belajar yang akan datang.
- 2) Butir item dengan tingkat daya pembeda kurang baik, ada dua kemungkinan tindak lanjut, yaitu:
 - a) Ditelusuri untuk dilakukan perbaikan dan selanjutnya dapat digunakan kembali pada tes hasil belajar yang akan datang untuk mengetahui apakah daya pembedanya mengalami peningkatan.
 - b) Dibuang (didrop)
- 3) Butir item yang memiliki angka indeks diskriminasi bertanda negatif, maka soal tersebut ditolak dan dapat dibuang karena butir soalnya sangat buruk.

⁴⁶) M. Ngalim Purwanto, *Prinsip-Prinsip dan Teknik Evaluasi Pengajaran*, (Bandung: PT Remaja Rosdakarya, 2008)

⁴⁷) Anas Sudijono, *Op.Cit.*, 408-409.

e. Fungsi Pengecoh/*Distractor*

Butir soal dalam bentuk pilihan ganda memiliki alternatif jawaban (opsi) yang merupakan pengecoh. Butir soal yang baik, pengecohnya akan dipilih peserta didik dan butir soal yang kurang baik pengecohnya tidak akan dipilih secara merata.⁴⁸ Pemilihan jawaban tersebut akan memperlihatkan pola jawaban. Pola jawaban diperoleh dengan menghitung banyaknya *testee* yang memilih jawaban antara a, b, c, d atau tidak memilih jawaban apapun. Berdasarkan pola jawaban butir soal, maka dapat ditentukan apakah pengecoh/*distractor* berfungsi sebagai pengecoh yang baik atau tidak.

Soal pengecoh bertujuan untuk mengetahui dari sekian banyak peserta didik yang mengikuti tes hasil belajar, apakah ada yang terkecoh dan memilih opsi yang salah sebab dianggap jawaban yang benar. Adapun pengecoh yang tidak dipilih sama sekali oleh *testee* dapat disimpulkan bahwa pengecoh tersebut buruk. Sebaliknya, apabila sebuah pengecoh minimal dipilih 5% dari seluruh *testee*, maka pengecoh tersebut dikatakan berfungsi dengan baik. Melalui hasil tersebut, tindak lanjut yang dapat dilakukan dalam efektivitas pengecoh, antara lain⁴⁹:

- 1) Pengecoh yang berfungsi dengan baik dapat disimpan dan dipakai lagi pada tes hasil belajar yang akan datang;

⁴⁸⁾ Zainal Arifin, *Evaluasi Pembelajaran*, (Bandung: PT Remaja Rosdakarya, 2013), 279.

⁴⁹⁾ Anas Sudijono, *Op.Cit.*, 417.

- 2) Pengecoh yang belum berfungsi dengan baik dapat diganti atau diperbaiki.

B. Penelitian yang Relevan

1. Penelitian tesis yang dilakukan oleh Laela Vitrocin Maulida dengan judul “Analisis Soal Penilaian Akhir Semester (PAS) Berbasis Keterampilan Berfikir Tingkat Tinggi (Studi pada Mata Pelajaran Tematik di SDI Sabilul Khoir Kota Batu”.

Penelitian ini berfokus pada soal PAS kelas VI di SDI Sabilul Khoir. Hal tersebut dilatarbelakangi oleh guru yang belum memahami keterampilan berfikir tingkat tinggi, dan belum mengerti cara melakukan analisis butir soal yang baik, soal dibuat dengan menekankan pada silabus dan RPP tanpa memperhatikan aspek kognitif C1-C6. Analisis data dilakukan dengan analisis kuantitatif dan kualitatif deskriptif. Kegiatan analisis ini meliputi aspek bahasa, materi, konstruksi, *Taksonomi Bloom*, serta analisis data mengenai tingkat kesukaran dan daya pembeda soal.

Hasil Penelitian menunjukkan soal buatan guru SDI Sabilul Khoir pada tingkat SD kelas VI semester II tingkat kesukarannya 90% ditolak, soal uraian perlu revisi, 18 soal pilihan ganda perlu direvisi pada segi daya pembedanya. Berdasarkan aspek materi yang memenuhi berjumlah 25 soal, memenuhi aspek bahasa berjumlah 22 soal, dan yang tidak memenuhi aspek bahasa berjumlah 3 soal. Sedangkan berdasarkan aspek konstruksi yaitu berjumlah 23 soal, dan yang tidak memenuhi 2 soal. Berdasarkan aspek kognitif, dari 35 soal diperoleh rincian soal (HOTS) berjumlah 8 soal

(23%), (MOTS) berjumlah 11 soal (32%) dan 16 soal (45%) berbasis Low order thinking skills (LOTS).⁵⁰

Persamaan penelitian ini dengan penelitian yang dilakukan adalah penelitian ini sama-sama menggunakan pendekatan kuantitatif. Perbedaannya ada pada fokus penelitian, dimana penelitian ini fokus terhadap analisis pada soal PAS kelas VI di SDI Sabilul Khoir, meliputi analisis aspek bahasa, aspek materi, aspek konstruksi, analisis soal berdasarkan *Taksonomi Bloom*, serta analisis data mengenai tingkat kesukaran dan daya pembeda soal.

2. Penelitian jurnal yang dilakukan oleh Muhamad Rishan dan Sulaiman dengan judul “Analisis Butir Soal Tipe Multiple Choice Ujian Akhir Semester Ganjil Kelas X Tahun 2022/2023 Mata Pelajaran Pendidikan Agama Islam di SMA Negeri 1 Padang”.

Penelitian ini bertujuan untuk mengetahui kualitas butir soal Tipe *Multiple Choice* Ujian Akhir Semester Ganjil Kelas X Tahun 2022/2023 Mata Pelajaran Pendidikan Agama Islam di SMA Negeri 1 Padang. Hal tersebut dilatarbelakangi oleh hasil observasi peneliti bahwa terjadi penurunan nilai Ujian Akhir Semester mata pelajaran PAI secara drastis, yaitu sebanyak 61% peserta didik yang belum mencapai Kriteria Ketuntasan Minimal (KKM).

⁵⁰⁾ Laela Vitrotin Maulida, “Analisis Soal Penilaian Akhir Semester (PAS) Berbasis Keterampilan Berfikir Tingkat Tinggi (Studi pada Mata Pelajaran Tematik di SDI Sabilul Khoir Kota Batu)”, (Thesis, Universitas Islam Negeri Maulana Malik Ibrahim Malang, 2023), 159-161.

Analisis soal Ujian Akhir Semester Ganjil Mata Pelajaran Pendidikan Agama Islam Kelas X SMA N 1 Padang Tahun Pelajaran 2022/2023 dilakukan menggunakan program ANATES versi 4. Berdasarkan analisis yang dilakukan, didapatkan hasil sebagai berikut: 1) Berdasarkan analisis tingkat kesukaran, dari 50 butir soal terbagi menjadi beberapa kategori, yaitu 3 butir soal (6%) sukar, 12 butir soal (24%) sedang, dan 35 butir soal (70%) mudah; 2) Berdasarkan analisis daya pembeda, dari 50 butir soal 18 butir soal (36%) Buruk, 17 butir soal (34%) cukup, 14 butir soal (28%) baik dan 1 butir soal (2%) negatif; 3) berdasarkan analisis kualitas pengecoh soal, terdapat 15 butir soal (30%) efektif dan 35 butir soal (70%) tidak aktif.⁵¹

Persamaan penelitian di atas dengan penelitian yang akan dilakukan adalah sama-sama meneliti tentang analisis butir soal dengan metode penelitian kuantitatif. Perbedaannya ada pada *software* yang digunakan dalam melakukan analisis data, yaitu program ANATES versi 4. Analisis dilakukan pada soal Ujian Akhir Semester Ganjil Mata Pelajaran Pendidikan Agama Islam Kelas X SMA N 1 Padang Tahun Pelajaran 2022/2023, meliputi aspek tingkat kesukaran, daya pembeda, dan kualitas pengecoh soal.

⁵¹⁾ Muhamad Rishan dan Sulaiman, Analisis Butir Soal Tipe Multiple Choice Ujian Akhir Semester Ganjil Kelas X Tahun 2022/2023 Mata Pelajaran Pendidikan Agama Islam di SMA Negeri 1 Padang, *Jurnal KeIslaman dan Ilmu Pendidikan* 5, no 3 (2023): 982, <https://doi.org/10.36088/Islamika.v5i3.3517>

3. Penelitian jurnal yang dilakukan oleh Raudhatul Jannah dengan judul “Analisis Kualitas Butir Soal dan Abilitas Peserta didik Pada Asesmen Madrasah Ibtidaiyah dengan Rasch Model”.

Penelitian ini membahas mengenai kualitas alat pengukuran yang diberikan kepada peserta didik. Teknik analisis data menggunakan Rasch Model, guna menguji validitas, reliabilitas instrument, serta person dan item secara sekaligus. Analisis yang dilakukan yakni mengenai kesesuaian butir soal, tingkat kesukaran, dan abilitas peserta didik. Hasil penelitian menunjukkan bahwa, soal asesmen Ilmu Pengetahuan Sosial memiliki kesesuaian dengan karakteristik fit, karena mampu memenuhi minimal satu dari tiga kriteria yang sudah ditetapkan.⁵²

Persamaan penelitian ini dengan penelitian yang dilakukan adalah sama-sama meneliti tentang analisis kualitas butir soal Asesmen Madrasah. Penelitian dilakukan dengan metode deskriptif kuantitatif. Perbedaanya ada pada aplikasi yang digunakan untuk melakukan analisis, yaitu aplikasi Winstep 3.73. Data yang dianalisis adalah butir soal mata pelajaran Ilmu Pengetahuan Sosial yang diujikan pada Asesmen Madrasah di Madrasah Ibtidaiyah.

4. Penelitian jurnal yang dilakukan oleh Rusmayani dengan judul “Analisis Butir Soal Penilaian Akhir Semester Genap Mata Pelajaran Pendidikan Agama Islam di SMP Bintang Persada Tabanan-Bali”.

⁵²⁾ Raudhatul Jannah, “Analisis Kualitas Butir Soal dan Abilitas Peserta didik pada Asesmen Madrasah Ibtidaiyah dengan Rasch Model”, *Jurnal Didaktika Pendidikan Dasar* 8, no 1 (2024): 332, <https://doi.org/10.26811/didaktika.v8i1.1234>

Penelitian ini memiliki tujuan untuk mengetahui validitas, reliabilitas, daya pembeda, dan taraf kesukaran pada soal PAS Genap mata pelajaran PAI kelas VII di SMP Bintang Persada Tabanan-Bali Tahun Pelajaran 2018-2019. Objek penelitiannya adalah 40 butir soal bentuk pilihan ganda yang dibuat oleh guru MGMP PAI di Kabupaten Tabanan, Bali. Jumlah responden yang digunakan adalah 39 peserta didik, dengan metode pengambilan data melalui dokumentasi hasil tes dan wawancara.

Penelitian ini menggunakan analisis campuran. Berdasarkan analisis yang dilakukan, hasilnya menunjukkan, bahwa: 1) Dari 40 butir soal, terdapat 15 butir drop, dan 25 butir valid, 2) Nilai koefisien reliabilitas sebesar 0.855, 3) Terdapat 11 butir soal kategori sedang dan 14 butir soal kategori mudah, 4) Terdapat 8 butir soal kategori sedang, dan 17 butir soal kategori lemah..⁵³

Persamaan penelitian ini dengan penelitian yang dilakukan adalah sama-sama meneliti tentang Analisis Butir Soal, dengan menganalisis validitas, reliabilitas, taraf kesukaran butir soal, dan daya beda butir soal. Perbedaan penelitian yaitu penelitian ini menggunakan metode penelitian evaluative, sedangkan penelitian yang dilakukan menggunakan metode penelitian kuantitatif. Perbedaan yang lain, yaitu terdapat pada analisis data yang digunakan, yaitu analisis campuran (kualitatif dan kuantitatif).

⁵³⁾ Rusmayani, "Analisis Butir Soal Penilaian Akhir Semester Genap Mata Pelajaran Pendidikan Agama Islam di SMP Bintang Persada Tabanan-Bali," *Jurnal Widya Balina* 5, no 1 (2020):41, <https://journal.staidenpasar.ac.id/index.php/wb>

5. Penelitian jurnal yang dilakukan oleh Ina Magdalena, Septy Nurul Fauziah, Siti Nur Faizah, dan Fika Sulaehatun Nupus dengan judul “Analisis Validitas, Reliabilitas, Tingkat Kesulitan dan Daya Beda Butir Soal Ujian Akhir Semester Tema 7 Kelas III SDN Karet 1 Sepatan”.

Penelitian ini bertujuan untuk menentukan item kualitas Ujian Akhir Semester kelas IV mata pelajaran tematik tema 7 tahun pelajaran 2020/2021 dengan bentuk tes pilihan ganda. Jumlah responden yang digunakan adalah 20 peserta didik. Data terdiri dari 20 lembar jawaban peserta didik kelas social IV tahun pelajaran 2020/2021, yang kemudian dianalisis untuk mengetahui tingkat validitas, reliabilitas, tingkat kesukaran dan daya beda dengan bantuan IBM SPSS Statistik 25.

Hasil penelitian menunjukkan bahwa 16 pertanyaan (80%) valid dan 4 pertanyaan (20%) tidak valid. Nilai rtabel adalah 0,4438. Tingkat kesulitan item tidak baik dengan keseimbangan proporsinya 3-5-2 atau 3-4-3. Daya pembeda item soal katagori baik sekali 1 soal atau 5%, katagori baik ada 9 atau 45%, katagori cukup 4 atau 20%, soal yang berkatagori kurang baik ada sekitar 5 soal atau 25% dan soal yang berkatagori kurang baik ada 1 soal atau 5%.⁵⁴

⁵⁴) Ina Magdalena, dkk., “Analisis Validitas, Reliabilitas, Tingkat Kesulitan dan Daya Beda Butir Soal Ujian Akhir Semester Tema 7 Kelas III SDN Karet 1 Sepatan,” *Jurnal Pendidikan Sains* 2, no 2 (2021): 198, <https://ejournal.stitpn.ac.id/index.php/bintang>

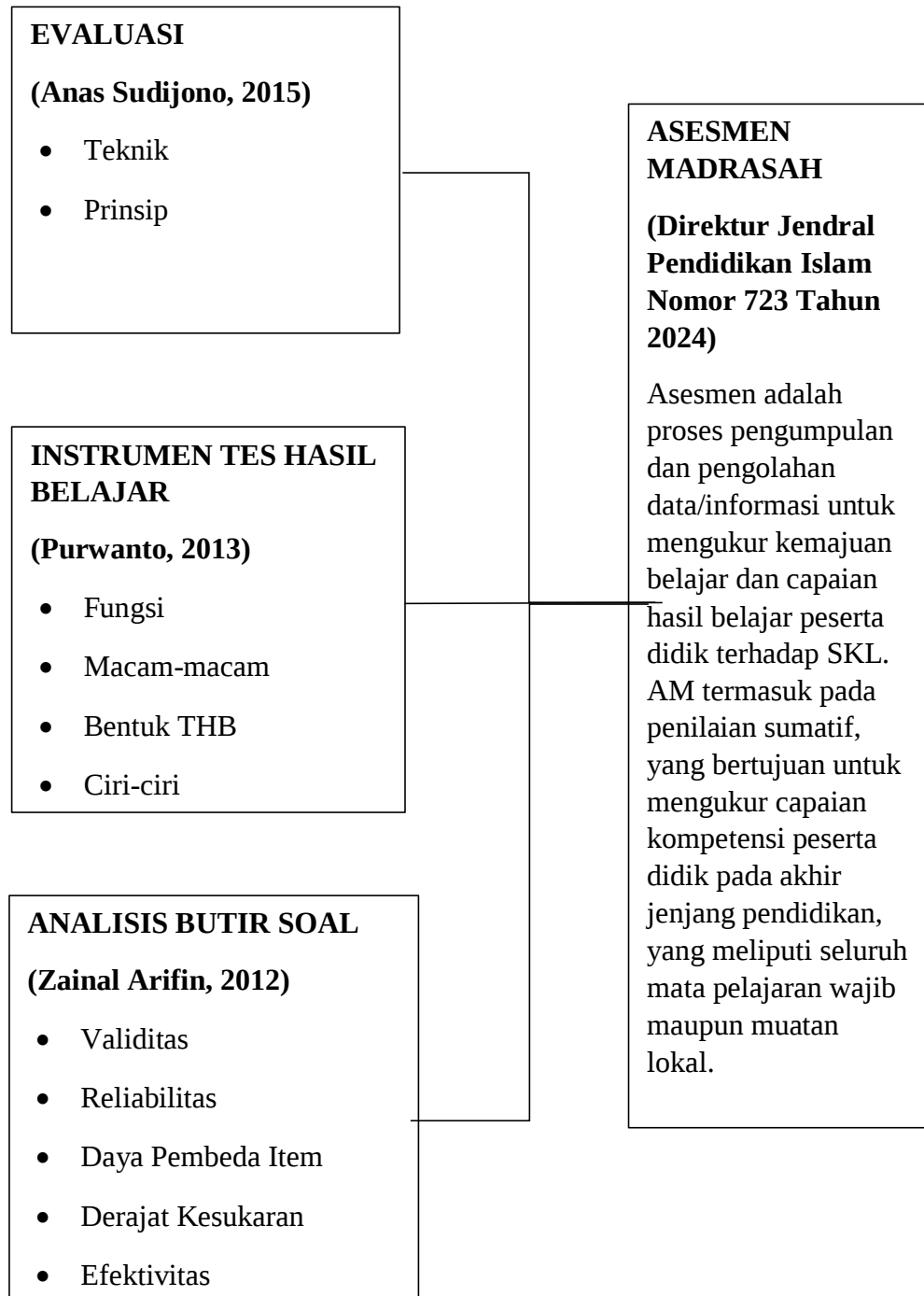
Tabel 1
Persamaan dan Perbedaan pada Penelitian yang Relevan dan
Penelitian yang Dilakukan

No	Judul Penelitian	Peneliti	Persamaan	Perbedaan
1.	Analisis Soal Penilaian Akhir Semester (PAS) Berbasis Keterampilan Berfikir Tingkat Tinggi (Studi pada Mata Pelajaran Tematik di SDI Sabilul Khoir Kota Batu	Laela Vitrotin Maulida	- Menggunakan metode kuantitatif - Membahas analisis butir soal	Fokus penelitian tersebut mengenai analisis pada soal PAS kelas VI di SDI Sabilul Khoir, meliputi analisis aspek bahasa, aspek materi, aspek konstruksi, analisis soal berdasarkan <i>Taksonomi Bloom</i> . Sedangkan penelitian yang dilakukan pada analisis butir soal AM berfokus pada analisis validitas, reliabilitas, derajat kesukaran item, daya pembeda, dan efektivitas pengecohnya.
2.	Analisis Butir Soal Tipe Multiple Choice Ujian Akhir Semester Ganjil Kelas X Tahun 2022/2023 Mata Pelajaran Pendidikan Agama Islam di SMA	- Muhamad Rishan - Sulaiman	- Menggunakan metode kuantitatif - Membahas analisis butir soal	- Menggunakan program ANATES versi 4, sedangkan penelitian dilakukan menggunakan program SPSS. - Analisis dilakukan pada soal Ujian Akhir Semester Ganjil Mata Pelajaran Pendidikan Agama Islam Kelas X SMA N 1 Padang

	Negeri 1 Padang			Tahun Pelajaran 2022/2023, meliputi aspek tingkat kesukaran, daya pembeda, dan kualitas pengecoh soal, sedangkan penelitian yang dilakukan pada analisis butir soal AM berfokus pada analisis validitas, reliabilitas, derajat kesukaran item, daya pembeda, dan efektivitas pengecohnya.
3.	Analisis Kualitas Butir Soal dan Abilitas Peserta didik Pada Asesmen Madrasah Ibtidaiyah dengan Resch Model	- Raudhatul Jannah	- Menggunakan metode kuantitatif - Membahas analisis butir soal AM	- Menggunakan aplikasi Winstep 3.73, sedangkan penelitian yang dilakukan menggunakan SPSS. - Soal AM yang dianalisis merupakan mata pelajaran Ilmu Pengetahuan Sosial, sedangkan penelitian yang dilakukan merupakan mata pelajaran Fikih.
	Analisis Butir Soal Penilaian	- Rusmayani	- Membahas tentang	- Menggunakan metode penelitian

	Akhir Semester Genap Mata Pelajaran Pendidikan Agama Islam di SMP Bintang Persada Tabanan-Bali		analisis butir soal	evaluatif dengan analisis campuran (kualitatif dan kuantitatif), sedangkan penelitian yang dilakukan menggunakan penelitian kuantitatif. - Soal yang dianalisis adalah butir soal PAS Genap Mata Pelajaran PAI, sedangkan penelitian yang dilakukan adalah soal AM mata pelajaran Fikih.
5.	Analisis Validitas, Reliabilitas, Tingkat Kesulitan dan Daya Beda Butir Soal Ujian Akhir Semester Tema 7 Kelas III SDN Karet 1 Sepatan	- Ina Magdalena - Septy Nurul Fauziah - Siti Nur Faizah - Fika Sulaehatun Nupus dengan judul	- Meneliti tentang analisis butir soal - Penelitian menggunakan metode kuantitatif, dengan menganalisis validitas, reliabilitas, tingkat kesulitan, dan daya beda - Menggunakan program SPSS	- Soal yang dianalisis adalah butir soal UAS Tema 7, sedangkan penelitian yang dilakukan adalah soal AM mata pelajaran Fikih

C. Kerangka Teori



Gambar 1
Kerangka Teori

D. Hipotesis Penelitian

Hipotesis merupakan dugaan atau jawaban sementara terhadap permasalahan yang akan diteliti. Hipotesis pada penelitian ini adalah hasil analisis kualitas butir soal Asesmen Madrasah mata pelajaran Fiqih di MI Ma'arif Pejagoan Tahun pelajaran 2023/2024 memiliki kualitas standar soal yang baik. Butir soal telah memenuhi ciri-ciri butir soal yang baik, yaitu dilihat dari segi validitas, reliabilitas, daya pembeda item, derajat kesukaran, serta fungsi pengecoh/*distractor*.