

BAB II

KAJIAN PUSTAKA

Pada bab kajian pustaka diuraikan mengenai teori-teori yang melandasi permasalahan pada penelitian, diantaranya yaitu landasan teori, penelitian terdahulu, paradigma penelitian, dan hipotesis penelitian yang dipaparkan sebagai berikut.

A. Landasan Teori

1. Pengetian Evaluasi Hasil Belajar

Evaluasi hasil belajar adalah suatu kegiatan untuk dapat mengetahui dan menentukan tercapai tidaknya tujuan pembelajaran. Menurut Sudjana (1991: 3), definisi evaluasi secara etimologi berasal dari bahasa Inggris yaitu *Evaluation* yang artinya penilaian, maksudnya memberikan nilai atau harga terhadap sesuatu dengan menggunakan kriteria tertentu. Sedangkan menurut istilah, evaluasi adalah proses yang sengaja dilakukan untuk memperoleh informasi atau data yang diperlukan yang kemudian digunakan sebagai alternatif-alternatif untuk membuat keputusan.

Dalam dunia pendidikan, guru bisa dikatakan memiliki kompetensi pedagogik apabila bisa melaksanakan kegiatan evaluasi dan penilaian program pembelajaran. Menurut Indriani (dalam Susetyo, 2020: 189), mengemukakan bahwa salah satu tugas seorang guru yaitu menilai dan mengevaluasi peserta didik. Evaluasi merupakan suatu proses dan tindakan yang sengaja direncanakan dengan tujuan untuk mengumpulkan informasi-informasi mengenai pertumbuhan, perkembangan, dan kemajuan peserta didik, sehingga dapat disusun penilaiannya yang bisa dijadikan acuan

atau dasar dalam membuat keputusan akhir. Evaluasi merupakan sebuah mekanisme yang sangat penting untuk bisa melihat dan menilai tingkat kemajuan proses pembelajaran yang telah dilakukan. Adanya kegiatan evaluasi adalah sebagai bahan yang sangat signifikan untuk bisa melakukan perbaikan-perbaikan dalam kegiatan pembelajaran di masa yang akan datang (Haryanto, 2020: 5).

Evaluasi adalah proses yang sengaja dilakukan untuk merencanakan, memperoleh, dan menyediakan informasi yang diperlukan untuk membuat jalan alternatif dalam mengambil keputusan (Yunita, 2018: 2). Tanpa adanya evaluasi, proses pembelajaran tidak dapat dinilai keberhasilannya. Dengan adanya kegiatan evaluasi dalam pendidikan, hal-hal yang baik akan dipertahankan dan ditingkatkan, sedangkan hal-hal yang menjadi kendala dalam kegiatan pembelajaran akan dicari penyebabnya dan solusinya sehingga meningkatkan mutu program pembelajaran selanjutnya.

Kegiatan evaluasi memiliki dua fungsi yaitu fungsi formatif dan fungsi sumatif. Evaluasi yang dilakukan sebagai perbaikan dan pengembangan kegiatan yang sedang berjalan adalah fungsi formatif. Sedangkan evaluasi yang dilakukan untuk pertanggungjawaban, keterangan, seleksi atau lanjutan disebut dengan fungsi sumatif. Jadi adanya kegiatan evaluasi bisa membantu pengembangan, implementasi, kebutuhan dan perbaikan program, pertanggungjawaban, seleksi, menambah pengetahuan dari pihak-pihak yang terlibat dalam kegiatan evaluasi.

Sedangkan menurut Sudaryono (dalam Berkah 2019: 12), fungsi evaluasi dibagi menjadi dua yaitu secara umum dan secara khusus. Fungsi evaluasi secara umum adalah sebagai suatu tindakan untuk mengetahui kemajuan, menunjang penyusunan rencana, dan memperbaiki atau melakukan penyempurnaan kembali. Adapun fungsi evaluasi secara khusus dibagi lagi menjadi enam, antara lain:

- a. Evaluasi berfungsi sebagai selektif, guru mengadakan selektif untuk penilaian terhadap peserta didiknya.
- b. Evaluasi sebagai diagnosis terhadap peserta didik tentang kebaikan dan kelemahan yang dimiliki.
- c. Evaluasi sebagai penempatan, menentukan tempat peserta didik sesuai dengan kemampuannya, sehingga peserta didik dengan hasil penilaian yang sama akan berada dalam lingkup kelompok belajar yang sama.
- d. Evaluasi sebagai pengukur keberhasilan program pendidikan yang dilaksanakan.
- e. Evaluasi sebagai formatif, yaitu memantau perkembangan belajar peserta didik dengan tujuan untuk memberikan umpan balik, baik kepada peserta didik maupun guru.
- f. Evaluasi sebagai sumatif, untuk mengetahui sejauh mana program pembelajaran berhasil diterapkan.

Berdasarkan penjelasan di atas dapat disimpulkan bahwa evaluasi merupakan kegiatan yang sengaja dilakukan untuk mengetahui kemampuan peserta didik dalam menguasai serta memahami hal-hal yang

sudah disampaikan oleh guru, keberhasilan guru dalam menyampaikan materi, serta seberapa banyak tujuan pembelajaran yang bisa dicapai.

2. Teknik-Teknik Evaluasi

Pada umumnya, teknik evaluasi yang digunakan dalam dunia pendidikan ada dua macam yaitu teknik tes dan teknik non tes (Rizqiyah, 2018: 13). Adapun penjelasannya adalah sebagai berikut.

a. Teknik Tes

Tes merupakan pemberian suatu tugas atau perintah dalam kegiatan evaluasi yang harus dikerjakan oleh seseorang agar diperoleh suatu informasi dari orang tersebut. Thoha (1996: 43), dalam bukunya yang berjudul Teknik Evaluasi Pendidikan mengemukakan bahwa kata “tes” berasal dari bahasa latin yang artinya alat untuk mengukur tanah. Sedangkan dalam bahasa Prancis kuno yaitu *Testum*, yang memiliki arti ukuran yang digunakan untuk membedakan emas dengan perak dan logam lainnya. Tes adalah alat yang digunakan untuk mengumpulkan informasi bersifat resmi. Berbentuk tugas atau suruhan yang harus dikerjakan dan bisa juga berbentuk pertanyaan-pertanyaan ataupun soal-soal yang harus diselesaikan/ dijawab. Pelaksanaan tes bisa dilakukan secara tertulis maupun lisan. Keberadaan tes digunakan untuk mengukur kemampuan, keahlian, dan pengetahuan seseorang (Rizqiyah, 2018: 6).

Sedangkan menurut Arikunto (2010: 53), tes berarti instrumen/ alat yang digunakan untuk mengetahui atau mengukur sesuatu dengan menggunakan cara dan aturan-aturan yang telah ditentukan. Sejalan

dengan pendapat Arikunto, Chaplin (2005), mengemukakan bahwa tes adalah seperangkat pertanyaan yang telah dibakukan, yang ditujukan kepada seseorang dengan maksud untuk mengukur perolehan, kemampuan, atau bakat yang dimiliki pada suatu bidang tertentu.

Dalam pendidikan, tes adalah cara atau prosedur yang dilakukan untuk mengukur dan menilai pencapaian dibidang pendidikan dalam bentuk pemberian tugas atau serangkaian tugas baik berupa soal-soal yang harus dijawab atau perintah-perintah yang harus dikerjakan oleh *testee* yaitu pihak yang dikenai tes/ peserta didik (Sudijono, 2015: 67). Selanjutnya Depdiknas (Departemen Pendidikan Nasional), mendefinisikan tes adalah himpunan pertanyaan yang wajib dijawab, atau tugas-tugas yang harus dilakukan oleh orang yang dites untuk mengetahui dan mengukur suatu aspek (perilaku) tertentu dari orang tersebut.

Penggunaan tes lebih cocok diaplikasikan untuk mengetahui kemampuan peserta didik dibidang aspek pengetahuan (kognitif), tes tidak cocok digunakan untuk mengukur sikap peserta didik karena tidak bisa diinterpretasikan ke dalam kategori salah atau benar. Tapi jika untuk mendeskripsikan profil tentang siswa, tes bisa digunakan sebagai salah satu alat dalam teknik penilaian hasil belajar peserta didik.

Dari beberapa definisi-definisi yang telah dijelaskan di atas, maka dapat disimpulkan bahwa tes yang ada di dalam dunia pendidikan adalah suatu alat atau prosedur yang dilakukan untuk memperoleh

informasi sekaligus mengukur kemampuan, pemahaman, pencapaian dari peserta didik. Tes yang disajikan berupa serangkaian pertanyaan-pertanyaan yang harus diselesaikan oleh peserta didik, bisa juga berupa perintah-perintah yang harus dilakukan. Sehingga dari hasil pengukuran dapat diperoleh data yang menggambarkan tentang perilaku ataupun prestasi dari peserta didik.

b. Teknik Non-Tes

Teknik non tes adalah cara penilaian hasil belajar dari peserta didik yang cara melakukannya tanpa menguji peserta didik melainkan dengan melakukan pengamatan secara sistematis (Mulyadi, 2010: 61). Teknik non tes ini umumnya digunakan untuk menilai sikap, tingkah laku, sifat, dan segala hal yang berkaitan dengan kepribadian anak. Dalam pendidikan, teknik ini untuk menilai hal-hal yang berhubungan dengan kegiatan belajar peserta didik, baik secara individu maupun secara berkelompok.

Teknik non-tes dalam pendidikan adalah teknik evaluasi hasil belajar yang dilakukan tanpa menguji peserta didik. Tetapi melakukan pengamatan secara sistematis, melakukan wawancara, menyebarkan angket, dan memeriksa atau meneliti dokumen-dokumen. Teknik non-tes dilakukan untuk menilai kepribadian peserta didik secara menyeluruh yang meliputi sikap, tingkah laku, sifat, sikap sosial dan lain-lain. Yang ada hubungannya dengan kegiatan belajar baik secara individu (Rizqiyah, 2018: 11).

Dengan menggunakan teknik non tes, melaksanakan penilaian atau evaluasi pada peserta didik tidak perlu menggunakan tes, tetapi dengan pengamatan secara sistematis, wawancara, atau bisa dengan menyebar angket. Adapun jenis-jenis teknik non tes adalah sebagai berikut.

1) Pengamatan (*Observation*)

Menurut Sudijono (dalam Indrawati, 2014: 6), observasi atau pengamatan adalah cara menghimpun bahan-bahan (data) yang dilakukan dengan cara melakukan pengamatan dan mencatat secara sistematis terhadap fenomena-fenomena yang dijadikan sebagai sasaran pengamatan. Dalam kegiatan pembelajaran, observasi dilakukan untuk menilai proses dan hasil belajar peserta didik ketika belajar di kelas, berdiskusi, mengerjakan tugas, dan lain-lain. Selain menilai peserta didik, observasi juga bisa dilakukan untuk menilai penampilan guru dalam kegiatan pembelajaran, suasana di kelas, hubungan sosial sesama peserta didik, hubungan sosial guru dengan peserta didik, atau perilaku sosial lainnya.

2) Wawancara (*Interview*)

Wawancara merupakan cara mengumpulkan data dengan cara melakukan tanya jawab lisan antara pewawancara dengan informan, berhadapan, baik secara langsung atau tidak langsung (menggunakan alat komunikasi), dengan arah tujuan yang sudah ditentukan. Dengan wawancara, maka seseorang dapat mengajukan

pertanyaan-pertanyaan terkait permasalahan yang dihadapi secara luwes, mengetahui perilaku nonverbal informan, dan memperoleh data yang benar-benar diperoleh dari sumbernya, bukan dibuat oleh orang lain.

3) Kuisisioner

Kuisisioner atau angket adalah sebuah rangkaian pertanyaan yang harus diisi oleh responden (Indrawati, 2014: 130). Dalam pembelajaran, keberadaan kuisisioner bertujuan untuk memperoleh data mengenai latar belakang dari peserta didik sebagai salah satu bahan dalam menganalisis hal-hal yang terjadi dalam proses belajar peserta didik. Misalnya, kesulitan-kesulitan yang dialami peserta didik dalam mengikuti pembelajaran. Kuisisioner atau angket umumnya digunakan untuk menilai hasil belajar peserta didik pada ranah afektif, yang disajikan dalam bentuk pilihan ganda atau skala sikap.

3. Analisis Butir Soal

Analisis butir soal adalah kegiatan pengkajian tentang isi butir soal-soal di dalam instrumen tes agar diperoleh seperangkat soal-soal yang memiliki kualitas yang baik (Isnaeni, 2017: 23). Selama ini, soal-soal yang disajikan siswa umumnya belum diujicobakan sebelum digunakan. Akibatnya banyak butir soal yang sudah digunakan dalam penilaian namun tidak dapat menghasilkan data yang diharapkan tentang hasil belajar siswa.

Menurut Arikunto (2013: 220), analisis butir soal merupakan suatu kegiatan mengkaji setiap item atau butir soal secara mendetail sesuai dengan prosedur yang sistematis, bertujuan untuk mengetahui kualitas dari butir soal tersebut. Sejalan dengan pendapat Daryanto (2008: 179), yang menyatakan bahwa menganalisis butir soal adalah kegiatan mengidentifikasi setiap soal-soal, baik yang memiliki kualitas baik, kurang baik, dan soal dengan kualitas buruk dengan harapan mampu memperoleh petunjuk untuk melakukan perbaikan pada setiap butir soal.

Berdasarkan beberapa pendapat di atas mengenai definisi analisis butir soal, dapat diambil kesimpulan bahwa analisis butir soal merupakan suatu kegiatan menelaah, mengidentifikasi, mengkaji suatu butir soal untuk memperoleh beberapa informasi terkait baik buruknya soal, layak tidaknya soal untuk dijadikan alat untuk mengukur pemahaman dan kemampuan peserta tes/ siswa setelah mengikuti proses pembelajaran serta memperoleh petunjuk untuk melakukan perbaikan pada butir soal.

Teknik analisis butir soal dapat dilakukan melalui dua cara yaitu dengan analisis butir soal secara kualitatif dan analisis secara kuantitatif. Menganalisis butir soal secara kualitatif berkaitan dengan isi dan bentuk soal. Aspek yang perlu diperhatikan dalam mengkaji butir soal secara kualitatif adalah dilihat dari segi materi, konstruksi, dan bahasa yang digunakan. Soal yang baik harus memenuhi ketiga kaidah penulisan soal. Sedangkan aspek-aspek yang terdapat dalam mengkaji butir soal secara kuantitatif Analisis butir soal secara kuantitatif mencakup validitas,

reliabilitas, tingkat kesukaran, daya beda dan efektivitas pengecoh/fungsi distraktor.

Pengkajian butir soal secara kuantitatif didasarkan pada data empirik yang ada di dalam butir soal. Terdapat dua pendekatan dalam menganalisis butir soal secara kuantitatif, yaitu pendekatan secara klasik dan pendekatan secara modern. Menurut Millman dkk (dalam Destiniar, Octaria, & Mulbasari, 2018), analisis butir soal secara kuantitatif menggunakan pendekatan klasik adalah menelaah soal melalui informasi dari jawaban siswa untuk meningkatkan kualitas butir soal tersebut. Analisis butir soal dengan pendekatan klasik memiliki kelebihan diantaranya murah, dapat dilaksanakan dengan cepat menggunakan komputer, sederhana, dapat menggunakan data dari beberapa siswa (sampel kecil).

Analisis soal dengan pendekatan modern dilakukan dengan menggunakan *Item Response Theory* (IRT), yaitu teori yang menggunakan fungsi matematika untuk menghubungkan antara peluang menjawab benar suatu soal dengan kemampuan siswa. Aspek yang perlu diperhatikan dalam menganalisis butir soal secara klasik meliputi tingkat kesukaran butir soal, daya pembeda, serta penyebaran pilihan jawaban atau frekuensi jawaban pada setiap pilihan jawaban. Pada penelitian ini, peneliti menganalisis butir soal Penilaian Tengah Semester genap secara kuantitatif ditinjau dari aspek tingkat kesukaran, daya pembeda, dan efektivitas pengecoh pada setiap butir soal.

a. Tingkat kesukaran

Menurut Arifin (dalam Rahmaini & Taufiq, 2018: 6), tingkat kesukaran soal adalah ukuran seberapa besar derajat kesukaran dalam setiap butir. Soal yang baik harus memiliki keseimbangan dari tingkat kesulitan soal. Maksudnya antara soal yang mudah, sedang, dan sulitimbang secara proporsional. Kriteria soal yang berkualitas baik adalah tidak terlalu mudah dan juga tidak terlalu sulit. Tingkat kesukaran soal bisa dilihat dari mampu tidaknya peserta didik dalam menjawab soal, bukan dilihat dari guru sebagai pembuat soal.

Soal yang terlalu mudah, tidak bisa membantu meningkatkan perkembangan daya berfikir siswa dengan baik, siswa tidak memiliki usaha untuk memecahkan permasalahan di dalam soal. Sebaliknya, jika soal terlalu sulit, bisa menjadikan siswa mudah putus asa dan tidak lagi memiliki keinginan untuk mencoba memecahkan soal tersebut karena di luar kemampuannya.

Menurut Sudjana (2017: 225), butir soal yang baik adalah butir soal yang memiliki tingkat kesukaran sesuai dengan tujuan dilaksanakannya tes. Misalnya, jika tes dilakukan untuk keperluan ujian semester, maka menggunakan butir soal dengan tingkat kesukaran sedang, untuk keperluan diagnosis maka menggunakan butir soal dengan tingkat kesukaran berkategori mudah, kemudian menggunakan butir soal dengan tingkat kesukaran tinggi untuk keperluan seleksi.

Menurut Arikunto (2013: 223), cara melakukan analisis untuk menentukan tingkat kesukaran soal adalah dengan menggunakan rumus sebagai berikut.

$$P = \frac{B}{Js}$$

P = Indeks kesukaran butir soal

B = Jumlah siswa yang menjawab butir soal dengan benar

Js = banyaknya siswa yang menjawab butir soal

Kriteria yang digunakan adalah semakin kecil jumlah indeks yang diperoleh, maka semakin sulit soal tersebut. Sebaliknya, semakin besar indeks yang diperoleh, maka makin mudah soal tersebut. Menurut Arikunto (dalam Halik, Mania, & Nur, 2019: 12), kriteria indeks kesulitan soal itu adalah sebagai berikut.

Tabel 2.1
Interpretasi Indeks Tingkat Kesukaran

P	Interpretasi
P = 0,00	Sangat Sukar
$0,00 < P \leq 0,30$	Sukar
$0,30 < P \leq 0,70$	Sedang
$0,70 < P < 1,00$	Mudah
P = 1,00	Sangat Mudah

Besarnya tingkat kesukaran berkisar nol sampai satu. Butir soal yang memiliki nilai p mendekati nol, maka dikategorikan sukar sedangkan butir soal yang bernilai p mendekati satu, maka dikategorikan mudah. Fernandes (dalam Kartowagiran, 2012: 15), mengemukakan bahwa butir soal yang menghasilkan rerata skor

sekitar 50% dari skor maksimum, maka dapat dikatakan mempunyai tingkat kesukaran yang tepat.

Untuk menyusun naskah soal, sebaiknya digunakan butir soal yang memiliki tingkat kesukaran yang seimbang, dengan proporsi soal sukar sebanyak 30%, soal yang sedang 40%, dan soal yang mudah 30%. Dengan komposisi soal yang seimbang, maka dapat diterapkan penilaian berdasarkan Penilaian Acuan Patokan (PAP) atau Penilaian Acuan Norma (PAN). Jika komposisi butir soal tidak seimbang, maka penggunaan penilaian berdasarkan acuan norma atau acuan patokan tidaklah tepat. Nurgiyantoro (2012: 195) menegaskan bahwa butir soal terlalu mudah atau sulit maka soal itu perlu direvisi atau diganti.

b. Daya pembeda

Daya pembeda butir soal adalah kemampuan suatu butir soal untuk dapat membedakan antara kelompok atas dan kelompok bawah dalam aspek yang diukur sesuai dengan perbedaan yang ada dalam kelompok itu. Menurut Arifin (2013: 273), mengemukakan salah satu tujuan analisis daya pembeda butir soal adalah untuk menentukan mampu tidaknya suatu butir soal membedakan antara peserta pelatihan berkemampuan tinggi dengan peserta pelatihan yang berkemampuan rendah. Daya pembeda pada dasarnya dihitung atas dasar pembagian peserta pelatihan ke dalam dua kelompok, yaitu kelompok atas merupakan kelompok yang tergolong pandai, dan kelompok bawah yaitu kelompok siswa yang tergolong tidak pandai.

Dalam hubungan ini, jika sebuah butir soal memiliki angka indeks diskriminasi butir soal dengan tanda positif, hal ini membuktikan bahwa butir soal tersebut memiliki daya pembeda, dalam arti bahwa peserta yang termasuk kategori pandai lebih banyak yang dapat menjawab dengan benar terhadap butir soal yang bersangkutan, sedangkan peserta yang termasuk kategori tidak pandai lebih banyak menjawab salah.

Klasifikasi daya pembeda ditentukan berdasarkan angka indeks diskriminasi butir soal. Dengan kata lain, apabila suatu butir soal memiliki angka pembeda yang baik maka dapat diartikan bahwa butir soal itu mampu membedakan antara peserta pelatihan yang berkemampuan tinggi dengan peserta pelatihan yang berkemampuan rendah. Menurut Sundayana (2016: 77), cara menentukan jumlah kelompok atas dan kelompok bawah yaitu dengan mengurutkan perolehan skor yang didapat oleh peserta tes, lalu jika jumlah peserta tes lebih dari 30 maka 27% menjadi kelompok atas dan 27% menjadi kelompok bawah. Sedangkan jika jumlah peserta tes kurang dari 30 maka diambil 50% tiap kelompoknya.

Menurut Arikunto (2013: 228), nilai angka indeks diskriminasi dapat ditentukan menggunakan persamaan berikut.

$$D = \frac{A_B}{A} - \frac{B_B}{B}$$

$$D = P_A - P_B$$

Keterangan:

D = Indeks diskriminasi

A = Jumlah peserta kelompok bawah

- A_B = Peserta kelompok atas yang menjawab benar
 B = Jumlah peserta kelompok bawah
 B_B = Peserta kelompok bawah yang menjawab benar
 P_A = Tingkat kesukaran kelompok atas
 P_B = Tingkat kesukaran kelompok bawah

Tabel 2.2
Interpretasi Indeks Daya Pembeda

D	Interpretasi
$D = 0,00 - 0,20$	Jelek Sekali
$0,20 < D \leq 0,30$	Jelek
$0,30 < D \leq 0,40$	Cukup
$0,40 < D < 50$	Baik
$0,50 < D \leq 1,00$	Baik Sekali

Seperti halnya indeks kesukaran (P), nilai indeks diskriminasi (D) berkisar 0,00 sampai 1,00. Perbedaannya, dalam indeks kesukaran tidak mengenal tanda negatif (-), sedangkan di dalam indeks diskriminasi ada kemungkinan menghasilkan tanda negatif. Jika suatu butir soal mempunyai $D = 0$, maka hal tersebut menunjukkan bahwa butir soal tersebut tidak memiliki daya pembeda sama sekali. Artinya jumlah peserta kelompok atas yang menjawab benar sama dengan jumlah peserta kelompok bawah yang sama-sama menjawab benar. Jadi butir soal yang memiliki nilai $D = 0$ tidak dapat membedakan antara kelompok atas dengan kelompok bawah.

Selanjutnya, apabila nilai D suatu butir soal menghasilkan tanda negatif (-), maka dapat diartikan bahwa butir soal tersebut lebih banyak dijawab benar oleh peserta tes kelompok bawah dibandingkan

peserta tes kelompok atas. Itu artinya butir soal tersebut sama sekali tidak dapat digunakan sebagai alat ukur prestasi belajar (Arikunto, 2013: 232). Daya pembeda pada butir soal dianggap masih memadai jika nilainya sama atau lebih besar dari 0,30. Jika lebih kecil dari 0,30, maka butir soal tersebut dianggap kurang mampu membedakan antara kelompok atas dengan kelompok bawah.

c. Efektivitas pengecoh/ Distraktor

Pengecoh/ distraktor adalah pilihan jawaban yang bukan merupakan kunci jawaban. Pengecoh diadakan untuk menyesatkan peserta didik agar tidak dengan mudahnya memilih kunci jawaban. Menurut Arikunto (2008: 220) sebuah distraktor yang baik adalah distraktor yang memiliki daya tarik yang kuat bagi peserta tes yang kurang memahami atau menguasai bahan.

Pada tes objektif bentuk *multiple choice* (pilihan ganda), untuk setiap butir soalnya dilengkapi beberapa kemungkinan jawaban atau sering disebut dengan *option* atau alternatif jawaban. Opsi atau alternative jawaban tersebut biasanya berjumlah tiga sampai lima buah. Dari kemungkinan jawaban-jawaban yang ada pada setiap butir soal, salah satunya adalah jawaban yang benar, sedangkan yang lainnya adalah jawaban yang salah. Pilihan jawaban yang seperti itulah yang sering disebut dengan pengecoh/ distraktor.

Menurut Thoha (dalam Fatimah, 2019: 59), keberadaan distraktor atau pengecoh pada soal digunakan untuk mengecoh peserta tes yang kurang mampu untuk bisa dibedakan dengan yang mampu. Pengecoh

yang baik adalah pengecoh yang mampu dihindari oleh peserta tes yang pandai dan dipilih oleh peserta tes yang kurang pandai. Menurut Sumarna (2006: 43), distraktor dikatakan berfungsi dengan baik apabila paling sedikit dipilih oleh 5% dari banyaknya peserta tes.

Sudijono (2010: 409), mengemukakan bahwa tujuan utama dipasangnya pengecoh pada setiap butir soal agar dari sekian banyak peserta tes yang mengikuti tes ada yang tertarik untuk memilih salah satu dari pilihan jawaban, karena mereka pasti mengira bahwa distraktor yang mereka pilih merupakan jawaban yang benar. Semakin banyak peserta tes yang terkecoh, maka distraktor yang dipasang dapat diartikan mampu menjalankan fungsinya dengan baik.

Menurut Arifin (2013: 279), indeks daya pengecoh dapat dihitung menggunakan rumus sebagai berikut.

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

Keterangan:

IP = Indeks Pengecoh

P = Jumlah peserta didik yang memilih pengecoh

N = Jumlah peserta didik yang ikut tes

B = Jumlah peserta didik yang menjawab benar pada setiap soal

n = Jumlah alternatif jawaban (opsi)

1 = Bilangan tetap

Tabel 2.3
Interpretasi Indeks Pengecoh/ Distraktor

IP	Interpretasi
Lebih dari 200%	Sangat Jelek
0% - 25% atau 176% - 200%	Jelek
26% - 50% atau 151% - 175%	Kurang Baik
51% - 75% atau 126% - 150%	Baik
76% - 125%	Sangat Baik

B. Penelitian Terdahulu

Penelitian terkait dengan analisis butir soal pernah dilakukan oleh Farihatin Nihayah dengan judul “Analisis Butir Soal Ujian Madrasah Mata Pelajaran Matematika di MTs I’anatut Tholibin Cebolek Kidul Kecamatan Margoyoso Kabupaten Pati Tahun Ajaran 2019/ 2020”. Tujuan dilakukannya penelitian ini untuk mengetahui validitas, reliabilitas, tingkat kesukaran, daya pembeda, efektivitas pengecoh, dan tingkat ranah kognitif soal ujian madrasah mata pelajaran matematika di MTs I’anatut Tholibin Cebolek Kidul Kecamatan Margoyoso Kabupaten Pati Tahun Ajaran 2019/ 2020.

Dalam penelitiannya Nihayah menyatakan bahwa (1) ditinjau dari validitas, butir soal yang valid berjumlah 26 butir soal (86,67%) dan yang tidak valid ada 4 butir (13,33%). (2) ditinjau dari reliabilitas, butir soal memiliki tingkat reliabilitas yang tinggi dengan koefisien reliabilitas sebesar 0,87. (3) ditinjau dari tingkat kesukaran, ada 5 (16,67%) butir soal yang sukar, 25 (83,33%) butir soal dengan kategori sedang, dan tidak ada kategori mudah. (4) ditinjau dari daya pembeda, terdapat 1 (3,33%) soal dengan daya pembeda tidak baik, 3 butir soal (10%) dengan daya pembeda jelek, tidak ada soal dengan daya pembeda cukup, 19 butir soal (63,34%) dengan daya

pembeda berkategori baik dan 7 butir soal (23,33%) dengan daya pembeda baik sekali.

Hasil yang selanjutnya (5) ditinjau dari efektivitas pengecoh, ada 29 (96,67%) butir soal dengan efektivitas pengecoh tinggi, kemudian hanya ada 1 (3,33%) butir soal dengan pengecoh berkategori baik. Pengecoh dengan kategori kurang baik, jelek, dan sangat jelek tidak ditemukan. (6) terakhir ditinjau dari tingkat ranah kognitif, tidak ditemukan butir soal yang termasuk tingkat Mengingat (C1) atau (0%), soal yang termasuk tingkat Memahami (C2) berjumlah 6 butir (20%), kemudian soal yang termasuk ke dalam tingkat Mengaplikasikan (C3) berjumlah 19 butir (63,34%), soal yang termasuk tingkat Menganalisis (C4) berjumlah 4 butir (13,33%), soal yang termasuk tingkat Mengevaluasi (C5) 1 butir (3,33%), dan butir soal yang termasuk tingkat Membuat (C6) tidak ada (0%).

Penelitian kedua dilakukan oleh Sukma Sacita Dewi, dkk (2019) dengan judul “Analisis Tingkat Kesukaran dan Daya Pembeda Soal Olimpiade Matematika (OMI) Tingkat SMP Tahun 2018”. Tujuan dari penelitian ini adalah untuk mendeskripsikan tingkat kesukaran dan daya pembeda pada butir soal OMI (Olimpiade Matematika Integral) tingkat SMP khususnya pada butir soal berbentuk pilihan ganda. Hasil penelitian ini adalah tingkat kesukaran pada 20 butir soal OMI tingkat SMP terdapat 11 soal dengan kategori sukar, 8 kategori sedang, dan 1 soal dengan kategori mudah. Sedangkan daya pembeda pada butir soal, ditemukan 10 soal dengan kategori baik, 5 soal berkategori cukup baik, 2 butir soal kategori sedang, dan 3 soal dengan kategori buruk. Hal ini menunjukkan bahwa tingkat kesukaran pada

soal OMI masih perlu diperbaiki, karena daya pembedanya masuk kategori buruk dan sedang, bahkan beberapa ada yang harus dibuang dan digantikan dengan pilihan jawaban yang baru.

Penelitian ketiga ditulis oleh Elviana (2020) dalam jurnalnya berjudul “Analisis Butir Soal Evaluasi Pembelajaran Pendidikan Agama Islam Menggunakan Program Anates”. Tujuan dilakukannya penelitian ini adalah untuk mendeskripsikan kualitas butir soal evaluasi pembelajaran Pendidikan Agama Islam dengan menggunakan aplikasi Anates. Hasil penelitian menunjukkan bahwa ditinjau dari daya beda soal diperoleh 5 soal dengan kategori sangat baik, 2 soal dengan baik, 8 soal dengan kategori cukup, 6 soal dengan kategori buruk, dan 4 soal dengan kategori sangat buruk dan harus dibuang.

Kemudian ditinjau dari tingkat kesukaran, soal yang sangat mudah ada 2 soal, soal mudah ada 1 soal, soal dengan kategori sedang ada 15 soal, soal sukar berjumlah 4 butir soal, dan soal sangat sukar ada 3 soal. Validitas yang diperoleh dengan menggunakan aplikasi anates sebesar 0,24 (sangat rendah) terdapat pada 6 butir soal dari 25 soal. Reliabilitas butir soal diperoleh sebesar 0,39 dengan menggunakan teknik belah yaitu ganjil dan genap. Butir soal yang memiliki daya pengecoh baik sejumlah 8 soal.

Penelitian keempat ditulis oleh A. Andriyani Asra (2017) dengan judul “Analisis Butir Soal Try Out Bahasa Indonesia yang Disusun Oleh Tim MGMP Madrasah Tsanawiyah Kabupaten Bulukumba Tahun 2017”. Asra melakukan penelitian ini dengan tujuan untuk mengetahui tingkat kesukaran dan daya beda soal try out bahasa Indonesia yang disusun oleh TIM MGMP

MTs Kabupaten Bulukumba tahun 2017. Hasil penelitian ini menunjukkan bahwa indeks kesukaran dalam soal ditemukan sebanyak 28 butir soal dari 50 soal yang memiliki kategori baik, sedangkan ada 22 butir soal dengan kategori tidak baik. Kemudian indeks daya beda, ditemukan 30 butir soal dengan kategori baik dan 20 butir soal dengan kategori tidak baik. Kesimpulannya, ada sebanyak 23 butir soal yang layak, 13 butir soal harus direvisi, dan 14 butir soal tidak layak, artinya harus diganti/ dibuang.

Penelitian kelima ditulis oleh Nurul Hasanah (2018) dengan judul “Kualitas Soal Ujian Sekolah atau Madrasah (US/M) Mata Pelajaran Matematika Tahun Ajaran 2016-2017 Kota Surabaya Berdasarkan Teori Respon Butir”. Penelitian ini dilakukan untuk mengetahui kualitas butir soal Ujian Madrasah tingkat SMP kelas VII secara kuantitatif dengan teori respon butir soal, dengan indikator yang dianalisis meliputi reliabilitas, daya beda, tingka kesukaran, dan peluang menebak atau *guessing*. Hasil dari penelitian menunjukkan bahwa reliabilitas soal secara keseluruhan 0,8213 termasuk kategori reliabel. 40 butir soal dengan tingkan kesukaran secara keseluruhan -2,089 berkategori kurang baik. Daya pembeda 1,063 masuk kategori baik. Peluang menebak dari 40 butir soal secara keseluruhan adalah 0,250 dengan kategori baik.

Tabel 2.4
Penelitian Terdahulu

No	Nama Peneliti dan Judul Penelitian	Persamaan Penelitian	Perbedaan Penelitian
1.	Farihatin Nihayah, Analisis Butir Soal	Penelitian ini sama-sama menganalisis	Aspek yang diteliti meliputi validitas,

	<p>Ujian Madrasah Mata Pelajaran Matematika di Mts I'anatut Tholibin Cebolek Kidul Kecamatan Margoyoso Kabupaten Pati Tahun Ajaran 2019/2020.</p>	<p>tentang kualitas butir soal jenjang SMP/MTs. Menggunakan metode penelitian deskriptif kuantitatif.</p>	<p>reliabilitas, tingkat kesukaran, daya pembeda, efek pengecoh, dan ranah kognitif. Lokasi penelitian di Mts I'anatut Tholibin Cebolek Kidul Kecamatan Margoyoso Kabupaten Pati. Analisis data penelitian sebelumnya menggunakan program <i>ANATES 4.09</i>. Soal yang dianalisis soal mata pelajaran matematika</p>
2.	<p>Sukma Sacita Dewi, dkk. Analisis Tingkat Kesukaran dan Daya Pembeda Soal Olimpiade Matematika (OMI) Tingkat SMP Tahun</p>	<p>Meneliti tentang kualitas butir soal. Aspek yang diteliti meliputi tingkat kesukaran dan daya pembeda butir soal.</p>	<p>Penelitian ini merupakan jenis penelitian kuantitatif-kualitatif (campuran). Soal yang dianalisis soal mata pelajaran matematika</p>

	2018.		
3.	Elviana dengan judul Analisis Butir Soal Evaluasi Pembelajaran Pendidikan Agama Islam Menggunakan Program Anates.	Menganalisis butir soal. Penelitian ini menggunakan pendekatan deskriptif kuantitatif.	Aspek yang diteliti meliputi validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efek pengecoh. Soal yang dianalisis soal mata pelajaran Pendidikan Agama Islam. Penelitian sebelumnya cara analisis data menggunakan program <i>ANATES 4.09</i> .
4.	A. Andriyani Asra dengan judul Analisis Butir Soal Try Out Bahasa Indonesia yang disusun Oleh Tim MGMP Madrasah Tsanawiyah Kabupaten Bulukumba Tahun 2017	Metode yang digunakan adalah deskriptif kuantitatif. Aspek yang diteliti meliputi tingkat kesukaran dan daya pembeda butir soal. Soal yang dianalisis soal mata pelajaran Bahasa Indonesia	Penelitian sebelumnya Lokasi penelitian di MTs Babul Khaer Kabupaten Bulukumba. Analisis data menggunakan program <i>ANATES 4.09</i> .

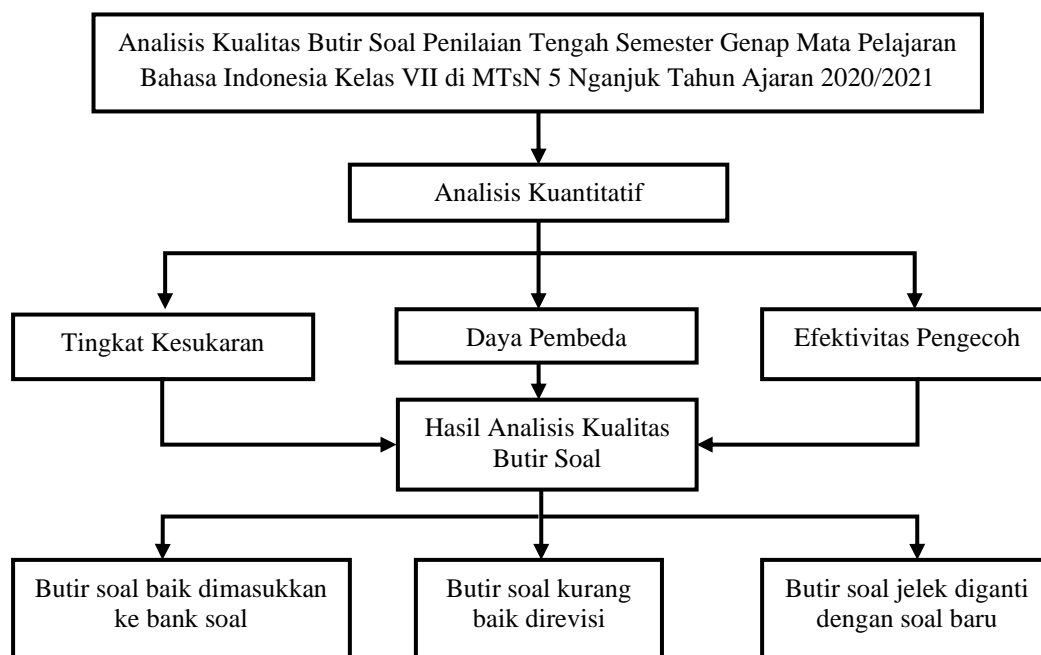
5.	Nurul Hasanah dengan judul Kualitas Soal Ujian Sekolah atau Madrasah (US/M) Mata Pelajaran Matematika Tahun Ajaran 2016-2017 Kota Surabaya Berdasarkan Teori Respon Butir.	Penelitian ini sama-sama menganalisis tentang kualitas butir soal jenjang SMP/MTs. Jenis penelitian ini adalah penelitian deskriptif kuantitatif. Aspek yang diteliti meliputi tingkat kesukaran, daya pembeda, pengecoh	Kualitas soal dideskripsikan berdasarkan teori respon butir dengan menggunakan bantuan program BILOG MG.
----	---	--	--

C. Paradigma Penelitian

Kegiatan analisis butir soal berperan penting dalam dunia pendidikan untuk mengetahui kelayakan dari butir soal yang akan disajikan kepada peserta didik sebagai bahan evaluasi dan alat tolak ukur dari kemampuan peserta didik sejauh mana materi yang telah mereka terima ketika proses pembelajaran berlangsung. Selain itu, dengan melaksanakan analisis soal, guru bisa memperbaiki dan meningkatkan kualitas soal yang telah dibuat. Soal yang memiliki kualitas baik adalah soal yang mampu memberikan informasi akurat sesuai dengan tujuannya yaitu mengukur kemampuan peserta didik. Penelitian ini hanya menganalisis kualitas butir soal pilihan ganda yang ditinjau dari tingkat kesukaran, daya pembeda, dan keefektifan pengecoh.

Bagan 2.1

Kerangka Berpikir



D. Hipotesis Penelitian

1. Soal penilaian tengah semester genap mata pelajaran bahasa Indonesia pada tahun ajaran 2020/2021 di MTsN 5 Nganjuk memiliki kategori baik.
2. Daya pembeda soal penilaian akhir semester genap mata pelajaran bahasa Indonesia pada tahun ajaran 2020/2021 MTsN 5 Nganjuk dapat dinyatakan baik.
3. Keefektifan pengecoh soal akhir semester genap mata pelajaran bahasa Indonesia pada tahun ajaran 2020/2021 MTsN 5 Nganjuk dapat dinyatakan efektif.