

BAB II LANDASAN TEORI

A. Deskripsi Teori

1. Mata Pelajaran Pendidikan Kewarganegaraan (PKn)

a. Pengertian Pendidikan Kewarganegaraan

Menurut Kansil, pendidikan kewarganegaraan adalah mata pelajaran yang sebagai untuk mengembangkan dan melestarikan nilai-nilai luhur, moral yang berakar pada budaya bangsa Indonesia yang diharapkan dapat mewujudkan dalam bentuk perilaku dalam kehidupan sehari-hari sebagai individu maupun sebagai anggota masyarakat dan makhluk ciptaan Tuhan Yang Maha Esa. Selain itu, pendidikan kewarganegaraan merupakan salah satu cara untuk membentuk watak bangsa Indonesia dan membentuk kepribadian manusia Indonesia yang seutuhnya sesuai dengan nilai-nilai yang terkandung dalam sila-sila Pancasila dan UUD 1945.¹ Dengan demikian, dapat dikatakan bahwa mata pelajaran PKn merupakan suatu mata pelajaran yang bertujuan untuk membentuk manusia Indonesia seutuhnya yang berlandaskan pada Pancasila, undang-undang, dan norma-norma yang berlaku di masyarakat.

Mata pelajaran PKn merupakan salah satu mata pelajaran yang diberikan sejak sekolah dasar (SD). Pendidikan kewarganegaraan di sekolah dasar atau madrasah ibtidaiyah, memberikan pelajaran kepada peserta didik untuk memahami dan membiasakan dirinya dalam kehidupan sekolah atau luar sekolah. PKn juga melatih peserta didik agar mampu melaksanakan hak-hak dan kewajibannya untuk menjadi warga negara yang cerdas, terampil, dan berkarakter sesuai Pancasila dan UUD 1945. Oleh karena itu, sebagai generasi penerus bangsa, peserta didik harus mengetahui kewajibannya dalam kehidupan masyarakat, berbangsa, dan bernegara demi keutuhan NKRI dan terciptanya masyarakat Indonesia yang berbudaya dan bermartabat.

b. Tujuan Mata Pelajaran Pendidikan Kewarganegaraan

Mata pelajaran Pendidikan Kewarganegaraan bertujuan agar peserta didik memiliki kemampuan sebagai berikut, yaitu:²

¹ Agung Suharyanto, "Peranan Pendidikan Kewarganegaraan Dalam Membina Sikap Toleransi Antar Siswa," *Jurnal Ilmu Pemerintahan dan Sosial Politik UMA* 1, no. 2, (2013): 195-196, diakses pada 2 Mei, 2019, <http://ojs.uma.ac.id/index.php/jppuma/article/download/563/403>.

² Bambang Sumantri, "Peningkatan Kualitas Pembelajaran PKn Melalui Model Pembelajaran Kooperatif Tipe Game Tournamen (TGT) Pada Siswa Kelas III SD Negeri Pelem 2 Ngawi," *Jurnal Ilmiah STKIP PGRI Ngawi* 13, no. 1 (2014): 20-

- 1) Berpikir secara kritis, rasional, dan kreatif dalam menanggapi isu kewarganegaraan.
- 2) Berpartisipasi secara aktif dan bertanggung jawab, dan bertindak secara cerdas dalam kegiatan bermasyarakat, berbangsa, dan bernegara, serta anti korupsi.
- 3) Berkembang secara positif dan demokratis untuk membentuk diri berdasarkan karakter-karakter masyarakat Indonesia, agar dapat hidup bersama dengan bangsa-bangsa lainnya.
- 4) Berinteraksi dengan bangsa-bangsa lain dalam percaturan dunia secara langsung dengan memanfaatkan teknologi informasi dan komunikasi.

c. Ruang Lingkup Mata Pelajaran Pendidikan Kewarganegaraan

Berdasarkan Permendiknas Nomor 22 tahun 2006 Bab II Pasal I ayat 2, ruang lingkup mata pelajaran Pendidikan Kewarganegaraan untuk pendidikan dasar dan menengah meliputi aspek-aspek sebagai berikut:³

- 1) Kebangsaan
- 2) Jiwa dan patriotisme bela negara
- 3) Hak asasi manusia
- 4) Kemajemukan bangsa
- 5) Pelestarian lingkungan hidup
- 6) Kesetaraan gender
- 7) Tanggung jawab sosial;
- 8) Ketaatan hukum dan pajak;
- 9) Perilaku anti korupsi, kolusi, dan nepotisme (KKN).

d. Standar Kompetensi dan Kompetensi Dasar

Khusus untuk SD/MI, lingkup isi Pendidikan Kewarganegaraan berdasarkan Standar Kompetensi (SK) dan Kompetensi Dasar (KD). Selanjutnya Standar Kompetensi (SK) dan Kompetensi Dasar (KD) mata pelajaran PKn kelas III semester 1 di MI NU Sholahiyah Pedawang seperti pada Tabel 2.1 berikut:⁴

21, diakses pada 2 Mei, 2019, http://jurnal.stkipngawi.ac.id/index.php/mp/article/viewFile/36/pdf_18.

³ Permendiknas RI, “ 22 Tahun 2006, Standar Isi untuk Satuan Pendidikan Dasar dan Menengah,” (23 Mei 2006).

⁴ Data berdasarkan dari dokumentasi yang telah disiapkan oleh guru kelas III A MI NU Sholahiyah Pedawang Bae Kudus.

Tabel 2.1.
Standar Kompetensi dan Kompetensi Dasar Semester 1

Standar Kompetensi	Kompetensi Dasar
1. Mengamalkan makna sumpah pemuda	1.1 Mengenal makna satu nusa, satu bangsa, dan satu bahasa 1.2 Mengamalkan nilai-nilai sumpah pemuda dalam kehidupan sehari-hari
2. Melaksanakan norma yang berlaku di masyarakat	2.1 Mengenal aturan-aturan yang berlaku di lingkungan masyarakat sekitar 2.2 Menyebutkan aturan-aturan yang berlaku di lingkungan masyarakat sekitar 2.3 Melaksanakan aturan-aturan yang berlaku di lingkungan masyarakat sekitar

2. Evaluasi Pembelajaran

a. Pengertian Evaluasi Pembelajaran

Evaluasi menurut bahasa, berasal dari bahasa Inggris yang dikenal dengan istilah *evaluation*, yang berarti penilaian atau penaksiran. Sedangkan menurut pengertian istilah, evaluasi merupakan kegiatan yang terencana untuk mengetahui keadaan suatu objek dengan menggunakan instrumen dan hasilnya dibandingkan dengan tolak ukur untuk memperoleh kesimpulan.⁵ Evaluasi secara umum dikenal sebagai proses sistematis untuk menentukan nilai sesuatu (tujuan, kegiatan, keputusan, unjuk kerja, proses, orang ataupun objek) berdasarkan kriteria tertentu.⁶

Menurut Suharsimi Arikunto, evaluasi adalah kegiatan untuk mengumpulkan informasi tentang bekerjanya sesuatu, selanjutnya informasi tersebut digunakan untuk menentukan alternatif yang tepat dalam mengambil keputusan. Sedangkan menurut Djemari Merdapi, evaluasi adalah proses mengumpulkan informasi untuk mengetahui pencapaian belajar kelas atau kelompok. Dalam bidang pendidikan, evaluasi sebagaimana dikatakan Gronlund merupakan proses yang sistematis tentang mengumpulkan, menganalisis, dan

⁵ Supardi, *Penilaian Autentik Pembelajaran Afektif, Kognitif, dan Psikomotor: Konsep dan Aplikasi* (Jakarta: Rajawali Pers, 2005), 12.

⁶ Elis Ratnawulan dan Rusdiana, *Evaluasi Pembelajaran* (Bandung: CV Pustaka Setia, 2015), 19.

menafsirkan informasi untuk menentukan sejauh mana tujuan pembelajaran yang telah dicapai oleh siswa.⁷

Terkait sistem pembelajaran, evaluasi merupakan salah satu komponen penting dan tahap yang harus ditempuh oleh guru untuk mengetahui keefektifan pembelajaran. Hasil yang diperoleh dapat dijadikan balikan (*feed-back*) bagi guru dalam memperbaiki dan menyempurnakan program dan kegiatan pembelajaran.⁸

Mengenai evaluasi pembelajaran, Grondlund dan Linn mendefinisikan, bahwa evaluasi pembelajaran adalah proses mengumpulkan, menganalisis, mengeinterpretasi informasi secara sistematis untuk menetapkan ketercapaian tujuan pembelajaran. Dalam rangka kegiatan pembelajaran, evaluasi dapat didefinisikan sebagai proses sistematis dalam menentukan tingkat pencapaian tujuan pembelajaran yang telah ditetapkan.

Evaluasi pembelajaran diartikan sebagai penentuan kesesuaian antara tampilan siswa dengan tujuan pembelajaran. Dalam hal ini yang dievaluasi adalah karakteristik siswa dengan menggunakan suatu tolok ukur tertentu. Karakteristik tersebut dalam ruang lingkup kegiatan belajar mengajar adalah tampilan siswa dalam bidang kognitif (pengetahuan dan intelektual), afektif (sikap, minat, dan motivasi), dan psikomotor (keterampilan, gerak, dan tindakan). Tampilan tersebut dapat dievaluasi secara lisan, tertulis, ataupun perbuatan. Dengan demikian, mengevaluasi disini adalah menentukan tampilan siswa telah sesuai dengan tujuan intruksional yang telah dirumuskan atau belum.

Pengertian evaluasi pembelajaran adalah proses untuk menentukan nilai pembelajaran yang dilaksanakan melalui kegiatan pengukuran dan penilaian pembelajaran. Pengukuran yang disini adalah proses membandingkan tingkat keberhasilan pembelajaran dengan ukuran keberhasilan pembelajaran yang telah ditentukan secara kuantitatif, sedangkan penilaian adalah proses pembuatan keputusan nilai keberhasilan secara kualitatif.⁹

b. Tujuan Evaluasi Pembelajaran

Adapun tujuan evaluasi pembelajaran yaitu:¹⁰

- 1) Untuk mengetahui keefektifan dan efisiensi sistem pembelajaran, baik tujuan, materi, metode, media, sumber belajar, lingkungan maupun sistem penilaian.

⁷ Aan Hasanah, *Pengembangan Profesi Guru* (Bandung: Pustaka Setia, 2012), 194.

⁸ Zainal Arifin, *Evaluasi Pembelajaran* (Jakarta: Direktorat Jenderal Pendidikan Islam Kementrian Agama, 2012), 6.

⁹ Elis Ratnawulan dan Rusdiana, *Evaluasi Pembelajaran*, 21.

¹⁰ Elis Ratnawulan dan Rusdiana, *Evaluasi Pembelajaran*, 26-27.

- 2) Untuk menghimpun bahan keterangan (data) yang dijadikan sebagai bukti mengenai taraf kemajuan anak didik dalam mengalami proses pendidikan selama jangka waktu tertentu.

c. Fungsi Evaluasi Pembelajaran

Adapun fungsi evaluasi pembelajaran yaitu:¹¹

- 1) Untuk mengetahui tercapai tidaknya tujuan pengajaran. Dalam hal ini adalah tujuan instruksional khusus. Dengan fungsi ini dapat diketahui tingkat penguasaan bahan pelajaran yang dikuasai oleh para siswa. Dengan perkataan lain dapat diketahui hasil belajar yang dicapai para siswa.
- 2) Untuk mengetahui keefektifan proses belajar mengajar yang telah dilakukan oleh guru. Dengan fungsi ini, guru dapat mengetahui berhasil tidaknya dalam mengajar. Rendahnya hasil belajar yang dicapai siswa tidak hanya disebabkan kurang berhasilnya guru mengajar. Melalui penilaian, menilai kemampuan guru dan hasilnya dapat dijadikan bahan dalam memperbaiki usahanya, yakni tindakan mengajar berikutnya.

d. Manfaat Evaluasi Pembelajaran

Kegunaan yang akan diperoleh dari kegiatan evaluasi pembelajaran, antara lain sebagai berikut:¹²

- 1) Adanya kemungkinan untuk dapat dihimpunnya informasi, baik yang bersifat kuantitatif maupun kualitatif tentang hasil atau kemajuan pembelajaran yang telah dicapai, dalam rangka pencapaian program pembelajaran pada khususnya, dan program pendidikan pada umumnya.
- 2) Adanya kemungkinan untuk dapat dilakukan usaha-usaha perbaikan, penyesuaian, dan penyempurnaan program pembelajaran yang dipandang perlu dan lebih berdaya guna, sehingga tujuan yang diinginkan akan dapat dicapai dengan sebaik-baiknya.

3. Pelaksanaan Penilaian Akhir Semester (PAS)

Menurut Peraturan Pemerintah No.13 Tahun 2015 tentang Perubahan Kedua atas Peraturan Pemerintah Nomor 19 Tahun 2005 tentang Standar Nasional Pendidikan Bab XII Pasal 78 Ayat butir a,

“evaluasi kinerja pendidikan yang dilakukan oleh satuan pendidikan kepada pihak-pihak yang berkepentingan”. Selanjutnya, Pasal 79 Ayat 1, dinyatakan “evaluasi sebagaimana dimaksud dalam Pasal 78 butir a dilakukan oleh satuan pendidikan pada setiap akhir semester”. Lebih lanjut, dalam Peraturan Pemerintah yang sama, Pasal 79 Ayat 2, “evaluasi sebagaimana dimaksud pada Ayat (1) sekurang-kurangnya

¹¹ Elis Ratnawulan dan Rusdiana, *Evaluasi Pembelajaran*, 28.

¹² Elis Ratnawulan dan Rusdiana, *Evaluasi Pembelajaran*, 35.

meliputi: (1) tingkat kehadiran peserta didik, pendidik, dan tenaga kependidikan; (2) pelaksanaan kurikulum tingkat satuan pendidikan dan kegiatan ekstrakurikuler; (3) hasil belajar peserta didik; dan (4) reliasasi anggaran.”¹³

Berdasarkan Peraturan Pemerintah tersebut, evaluasi hasil belajar peserta didik dilakukan pada setiap akhir pembelajaran. Hasil belajar merupakan kemampuan peserta didik setelah menempuh pengalaman belajarnya. Evaluasi hasil belajar peserta didik selama satu semester diujikan melalui Penilaian Akhir Semester (PAS). Nilai PAS ini merupakan gambaran dari penguasaan kompetensi peserta didik selama satu semester.

Kegiatan PAS dilaksanakan berdasarkan kalender pendidikan. Kegiatan PAS gasal MI NU Sholahiyah Pedawang Kecamatan Bae Kabupaten Kudus dilaksanakan pada tanggal 1-8 Desember 2018. Pelaksanaan PAS, khususnya mata pelajaran PKn pada tanggal 4 Desember 2018. Selama mengerjakan PAS, terdapat tata tertib yang harus dipatuhi peserta didik, seperti: tidak diperbolehkan menanyakan atau memberikan jawaban kepada temannya, tidak boleh mengganggu teman ketika mengerjakan UAS, dan lain-lain. Apabila ada peserta didik yang melanggar tata tertib tersebut, maka guru akan memberikan sanksi berupa teguran sehingga peserta didik yang melanggar tidak mengulangi perbuatannya lagi.

4. Karakteristik Soal Objektif (Soal Pilihan Ganda)

Tes objektif adalah bentuk tes yang mengandung kemungkinan jawaban atau respons yang harus dipilih oleh peserta tes. Peserta hanya memilih alternatif jawaban yang telah disediakan oleh penyusun butir soal.¹⁴ Akan tetapi, dikarenakan luasnya bahan pelajaran yang harus dicakup dalam tes dan mudahnya menilai jawaban yang diberikan, maka soal-soal bentuk objektif banyak digunakan dalam menilai hasil belajar.¹⁵

Secara umum ada lima tipe tes objektif, yaitu: (1) tes bentuk benar-salah (*true-false test*); (2) tes bentuk menjodohkan (*matching test*); (3) tes bentuk jawaban singkat (*short answer*) dan melengkapi

¹³ Peraturan Pemerintah RI, “13 Tahun 2015, Perubahan Kedua Atas Peraturan Pemerintah Nomor 19 Tahun 2005 Tentang Standar Nasional Pendidikan,” (6 Maret 2015)

¹⁴ Eko Putro Widoyoko, *Evaluasi Program Pembelajaran: Panduan Praktis Bagi Pendidik dan Calon Pendidik* (Yogyakarta: Pustaka Pelajar, 2017), 49.

¹⁵ Supardi, *Penilaian Autentik Pembelajaran Afektif, Kognitif, dan Psikomotor: Konsep dan Aplikasi*, 51.

(*completion*); (4) tes uraian objektif; dan (5) tes bentuk pilihan ganda (*multiple choice item test*).

Berdasarkan pembatasan masalah dalam penelitian ini, peneliti hanya akan menjelaskan tes bentuk 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 2 (dua) atau 5 (lima). Tipe tes pilihan ganda banyak digunakan dalam kelompok tes objektif, karena banyak sekali materi yang dapat dicakup.¹⁶

a. Bentuk Soal Pilihan Ganda

Dilihat dari strukturnya, bentuk soal pilihan ganda terdiri dari empat bagian, yaitu:¹⁷

- 1) *Stem* adalah pernyataan-pernyataan yang berisi permasalahan yang akan dinyatakan.
- 2) *Option* adalah sejumlah pilihan atau alternatif jawaban.
- 3) Kunci adalah jawaban yang benar atau paling tepat.
- 4) *Distractor* adalah jawaban-jawaban lain selain kunci jawaban (pengecoh).

b. Model Tes Pilihan Ganda

Menurut Eko Putro Widoyoko, tes pilihan ganda dapat dibedakan menjadi 5 (lima) model, yaitu:¹⁸

1) Pilihan Ganda Sederhana

Pada tes pilihan ganda sederhana, *stem* atau pokok soal dapat terdiri dari pertanyaan atau pernyataan yang belum lengkap dan untuk melengkapinya, harus memilih satu jawaban. Pilihan jawaban (*option*) terdiri dari beberapa alternatif pilihan jawaban yang memiliki satu jawaban yang benar.

Berikut contoh soal pilihan ganda sederhana:

Di pulau Sumatra terdapat beberapa buah danau. Salah satu danau tersebut mempunyai ciri fisik yang berbeda dari danau lain, karena di tengahnya terdapat daratan luas berupa pulau. Danau manakah yang dimaksud?

- a. Danau Ranau
- b. Danau Maninjau
- c. Danau Singkarak
- d. Danau Toba*
- e. Danau Laut Tawar

¹⁶ Eko Putro Widoyoko, *Evaluasi Program Pembelajaran: Panduan Praktis Bagi Pendidik dan Calon Pendidik*, 59.

¹⁷ Supardi, *Penilaian Autentik Pembelajaran Afektif, Kognitif, dan Psikomotor: Konsep dan Aplikasi*, 56.

¹⁸ Eko Putro Widoyoko, *Evaluasi Program Pembelajaran: Panduan Praktis Bagi Pendidik dan Calon Pendidik*, 60-64.

Kunci jawaban: d

Dari contoh di atas *stem* atau pokok soal dapat terdiri dari pernyataan dan pertanyaan. Sedangkan pilihan jawaban (*option*) terdiri dari beberapa alternatif pilihan jawaban. Salah satu dari alternatif pilihan itu adalah jawaban yang benar terhadap pertanyaan. Dalam hal ini ditandai dengan asterisk (*). Jawaban tersebut dinamakan kunci jawaban. Alternatif jawaban yang bukan kunci dinamakan pengecoh atau *distractors*. Jadi dalam pilihan (*option*) ada pilihan yang bukan kunci.

2) Pilihan Ganda Analisis Hubungan Antar Hal

Pada pilihan ganda hubungan antar hal terdiri dari dua pernyataan. Kedua pernyataan tersebut dihubungkan dengan kata “SEBAB”. Jadi ada dua kemungkinan hubungan antara kedua pernyataan tersebut, yaitu ada hubungan sebab akibat atau tidak ada hubungan sebab akibat.

Berikut contoh soal pilihan ganda analisis hubungan antar hal:

Pada ekonomi liberal persaingan merupakan suatu tatanan nilai yang baik untuk meningkatkan kualitas produksi.

SEBAB

Persaingan dapat menghasilkan produksi yang terbaik yang diterima masyarakat.

Pilihlah:

- a. Pernyataan benar, alasan benar, keduanya menunjukkan hubungan sebab akibat.
- b. Pernyataan benar, alasan benar, keduanya tidak menunjukkan hubungan sebab akibat.
- c. Pernyataan benar, tetapi alasan salah.
- d. Pernyataan salah, tetapi alasan benar.
- e. Pernyataan dan alasan salah.

Kunci jawaban: a

Pada contoh di atas pernyataan dan alasan benar. Keduanya menunjukkan hubungan sebab akibat, maka jawaban yang benar adalah “a”.

3) Pilihan Ganda Analisis Kasus

Pada bentuk pilihan ganda analisis kasus, peserta tes dihadapkan pada suatu kasus. Kasus disajikan dalam bentuk cerita, peristiwa, dan sejenisnya. Berdasarkan kasus tersebut, kepada peserta tes diajukan beberapa pertanyaan yang dibuat dalam bentuk melengkapi pilihan. Oleh karena itu, sebelum menjawab soal, peserta tes harus memahami kasus yang disajikan.

Berikut contoh soal pilihan ganda kasus:

Untuk menjawab butir soal di bawah ini, pahami kasus secara cermat, kemudian jawablah soal-soal berikutnya!

“Kadit Lantas Polda DIY Letkol Pol menjelaskan jumlah kecelakaan lalu lintas di DIY bulan Januari-November 2008 sebanyak 7.090 kasus atau meningkat 4,8 persen dibandingkan tahun 2007 periode yang sama. Meningkatnya kecelakaan lalu lintas itu antara lain karena terhentinya Operasi Zebra menjadi operasi rutin lalu lintas. Di samping itu pengguna jalan hanya disiplin jika ada petugas”. Pertanyaan:

1. Meningkatnya kecelakaan lalu lintas di DIY bukan hanya disebabkan oleh terhentinya Operasi Zebra, tetapi juga disebabkan
 - a. Pengawas lalu lintas yang tidak pernah kendor
 - b. Volume kendaraan di jalan yang semakin bertambah
 - c. Pengangkutan yang terlibat dalam pengaturan lalu lintas dikurangi jumlahnya
 - d. Potensi polisi lalu lintas belum dikerahkan secara maksimal

Kunci jawaban: d

2. Dalam periode yang sama di tahun 2008 kasus kecelakaan di DIY meningkat 4,87 persen dibandingkan dengan tahun 2007. Dari data ini:
 - a. Dapat dibuat kesimpulan bahwa kecelakaan lebih banyak terjadi pada tahun 2007 dibandingkan tahun 2008.
 - b. Belum dapat dibuat kesimpulan tentang jumlah kecelakaan bulan Desember belum dilaporkan.
 - c. Sudah dapat dipastikan bahwa kasus kecelakaan selama tahun 2008 lebih banyak terjadi daripada tahun 2007.
 - d. Tidak dapat dibuat generalisasi karena banyak kasus kecelakaan yang tidak dilaporkan yang bersangkutan kepada polisi lalu lintas.

Kunci jawaban: c

Berdasarkan contoh di atas dapat dikatakan bahwa dari satu kasus dapat dibuat beberapa pertanyaan dan aspek berpikir yang diukur banyak pada proses berpikir yang lebih tinggi daripada pemahaman.

4) Pilihan Ganda Asosiasi

Bentuk pilihan ganda asosiasi memiliki struktur soalnya sama dengan melengkapi satu pilihan. Perbedaannya adalah kalau pada melengkapi pilihan hanya ada satu jawaban yang benar atau paling benar, tetapi pada melengkapi berganda justru jawaban yang benar dapat lebih dari satu, mungkin 2, 3, 4.

Berikut contoh soal pilihan ganda asosiasi:

Petunjuk pilihan:

- a. Jika (1), (2), dan (3) betul
- b. Jika (1) dan (3) betul
- c. Jika (2) dan (4) betul
- d. Jika hanya (4) yang betul
- e. Jika semuanya betul

Berikut ini adalah unsur-unsur pengeluaran rutin anggaran belanja negara:

- (1) Belanja pegawai
- (2) Belanja barang
- (3) Pembayaran bunga utang
- (4) Pembiayaan pembangunan daerah

Kunci jawaban: a

5) Pilihan Ganda dengan Diagram, Grafik, Tabel

Bentuk pilihan ganda dengan gambar/ diagram/ grafik/ tabel mirip alias kasus, baik struktur maupun pola pertanyaannya. Bedanya dalam bentuk ini tidak disajikan kasus dalam bentuk cerita atau peristiwa tersebut berupa diagram, gambar, grafik ataupun tabel.

Berikut contoh soal pilihan ganda dengan gambar:



Kegiatan seperti gambar di samping menunjukkan contoh pelaksanaan peraturan di lingkungan

- a. Masyarakat
- b. Keluarga
- c. Sekolah
- d. Rumah

Kunci jawaban: a

Berdasarkan pembahasan tersebut, maka dapat diketahui bahwa soal pilihan ganda mempunyai lima variasi soal yang dapat digunakan sehingga penyusunan soal tidak menonton. Selain mempunyai bentuk yang bervariasi, soal bentuk pilihan ganda juga mempunyai beberapa kelebihan dan juga kekurangan.

c. Kelebihan dan Kekurangan Tes Pilihan Ganda

- 1) Kelebihan tes pilihan ganda meliputi:¹⁹
 - a) Butir soal tes pilihan ganda dapat digunakan untuk mengukur segala level tujuan pembelajaran, mulai dari yang paling sederhana sampai dengan yang paling kompleks, kecuali tujuan yang berupa kemampuan mendemonstrasikan, keterampilan menyatakan sesuatu secara ekspresif.
 - b) Karena karakteristik butir soal pilihan ganda hanya menuntut waktu mengerjakan sangat minimal, maka setiap perangkat tes yang menggunakan butir soal pilihan ganda sebagai alat ukur dapat menggunakan jumlah butir soal yang relatif banyak dan karena itu penarikan sampel pokok bahasan yang akan diujikan dapat lebih luas.
 - c) Penskoran hasil tes dapat dilakukan secara objektif.
 - d) Tiap butir soal dapat disusun sedemikian rupa sehingga menuntut kemampuan peserta tes untuk membedakan berbagai tingkatan kebenaran sekaligus.
 - e) Jumlah pilihan yang disediakan melebihi dua.
 - f) Tipe butir soal pilihan ganda memungkinkan dilakukan analisis butir soal secara baik.
 - g) Tingkat kesukaran butir soal dapat diatur, dengan hanya mengubah tingkat homogenitas alternatif jawaban.
 - h) Informasi yang diberikan lebih kaya.
- 2) Kekurangan tes pilihan ganda meliputi:²⁰
 - a) Relatif lebih sulit dalam penyusunan butir soal
 - b) Ada kecenderungan bahwa guru menyusun butir soal tipe ini dengan hanya menguji atau mengukur aspek ingatan, atau aspek yang paling rendah dalam ranah kognitif.
 - c) Adanya pengaruh kebiasaan peserta tes terhadap tes bentuk pilihan ganda (*testwise*) terhadap hasil tes peserta.

d. Pedoman Penyusunan Butir Soal Pilihan Ganda

Menurut Eko Putro Widoyoko menyatakan bahwa, ada sebelas pedoman penyusunan butir soal pilihan ganda, yaitu:²¹

- 1) Inti permasalahan harus dicantumkan dalam rumusan pokok soal, sehingga dengan membaca pokok soal siswa dapat menentukan jawaban sebelum dilanjutkan membaca pilihan jawaban.

¹⁹ Eko Putro Widoyoko, *Evaluasi Program Pembelajaran: Paduan Praktis Bagi Pendidik dan Calon Pendidik*, 68-70.

²⁰ Eko Putro Widoyoko, *Evaluasi Program Pembelajaran: Paduan Praktis Bagi Pendidik dan Calon Pendidik*, 70-71.

²¹ Eko Putro Widoyoko, *Penilaian Hasil Pembelajaran di Sekolah* (Yogyakarta: Pustaka Pelajar, 2016), 137-144.

- 2) Hindari pengulangan kata-kata yang sama dalam pilihan.
- 3) Hindari rumusan kata yang berlebihan.
- 4) Kalau pokok soal merupakan pernyataan yang belum lengkap, maka kata atau kata-kata yang melengkapi harus diletakkan pada ujung pernyataan, bukan di tengah-tengah kalimat.
- 5) Susunan alternatif jawaban dibuat teratur dan sederhana.
- 6) Semua jawaban harus homogen dan dimungkinkan sebagai jawaban yang benar.
- 7) Hindari jawaban yang benar selalu ditulis lebih panjang dari jawaban yang salah.
- 8) Hindari adanya petunjuk atau indikator pada jawaban yang benar.
- 9) Gunakan tiga atau lebih alternatif pilihan jawaban.
- 10) Pokok soal diusahakan tidak menggunakan ungkapan atau kata-kata yang bermakna tidak pasti, misalnya: kebanyakan, sering kali; kadang-kadang dan sejenisnya.
- 11) Pokok soal sedapat mungkin dalam pernyataan atau pertanyaan positif.

Berdasarkan penjelasan tersebut, dapat disimpulkan bahwa bentuk soal pilihan ganda merupakan bentuk soal yang memiliki tingkat kesulitan dan ketelitian yang tinggi dalam proses pembuatannya. Namun, hal ini dapat diantisipasi apabila penyusun soal terus berlatih dan juga memperhatikan pedoman dalam menyusun butir soal pilihan ganda, sehingga soal yang disusun dapat memiliki kualitas yang baik.

5. Analisis Butir Soal

Analisis butir soal merupakan suatu kegiatan yang harus dilakukan guru untuk meningkatkan mutu soal yang telah ditulis. Kegiatan ini merupakan proses pengumpulan, peringkasan dan penggunaan informasi dari jawaban siswa untuk membuat keputusan tentang setiap penilaian.²² Tujuan dari kegiatan analisis butir soal adalah untuk mengadakan identifikasi soal-soal yang baik, kurang baik, dan soal yang jelek. Dengan analisis soal dapat diperoleh informasi tentang kejelekan sebuah soal dan “petunjuk” untuk mengadakan perbaikan.²³ Dalam pelaksanaannya, kegiatan analisis butir soal dapat dilakukan secara kualitatif dan kuantitatif.

a. Analisis Butir Soal Secara Kualitatif

Analisis butir soal secara kualitatif dilaksanakan berdasarkan kaidah penulisan soal (tes tertulis, perbuatan, sikap). Aspek yang diperhatikan adalah penelaahan secara kualitatif adalah setiap soal

²² Elis Ratnawulan dan Rusdiana, *Evaluasi Pembelajaran*, 148.

²³ Suharsimi Arikunto, *Dasar-dasar Evaluasi Pendidikan* (Jakarta: Bumi Aksara, 2009), 206-207.

ditelaah dari segi materi, konstruksi, bahasa/ budaya, dan kunci jawaban/ pedoman penskorannya. Ada beberapa teknik yang dapat digunakan untuk menganalisis butir soal secara kualitatif, yaitu teknik moderator dan teknik panel.

Teknik moderator merupakan teknik berdiskusi yang didalamnya terdapat satu orang sebagai penengah. Berdasarkan teknik ini, setiap butir soal didiskusikan secara bersama-sama dengan beberapa ahli seperti guru yang mengajarkan materi, ahli materi, penyusun/ pengembang kurikulum, ahli penilaian, ahli bahasa, berlatar belakang psikologi. Sedangkan teknik panel merupakan suatu teknik menelaah butir soal yang setiap butir soal ditelaah berdasarkan kaidah penulisan butir soal.²⁴ Teknik panel analisis kualitatif butir instrumen penilaian dilakukan dengan cara setiap soal dianalisis dari segi materi, konstruksi, bahasa/budaya.²⁵ Setiap butir soal yang telah ditelaah dan memenuhi persyaratan dari segi materi, konstruksi, dan bahasa dapat digunakan untuk ulangan harian ataupun ulangan semesteran. Selanjutnya hasil ulangan harian ataupun ulangan semesteran dapat dianalisis kembali untuk mengetahui materi yang belum dipahami oleh peserta didik.

Lebih lanjut aspek-aspek yang ditelaah dalam menganalisis butir soal pilihan ganda secara kualitatif dapat dilihat dari segi materi, konstruksi, dan bahasa/ budaya.

1) Aspek Materi

Berkaitan dengan substansi keilmuan yang ditanyakan dalam soal. Aspek yang harus diperhatikan yaitu:

- a) Soal sesuai dengan indikator.
- b) Materi yang ditanyakan sesuai dengan kompetensi.
- c) Pilihan jawaban homogen dan logis
- d) Hanya ada satu kunci jawaban.

2) Aspek Konstruksi

Berkaitan dengan teknik penulisan soal. Aspek yang diperhatikan yaitu:

- a) Pokok soal dirumuskan dengan singkat, jelas, dan tegas.
- b) Rumusan pokok dan pilihan jawaban merupakan pernyataan yang diperlukan saja.
- c) Pokok soal tidak memberi petunjuk kunci jawaban.
- d) Pokok soal bebas dari pernyataan yang bersifat negatif ganda.
- e) Pilihan jawaban homogen dan logis ditinjau dari segi materi.
- f) Gambar, grafik, tabel, diagram, atau sejenisnya jelas dan berfungsi.

²⁴ Elis Ratnawulan dan Rusdiana, *Evaluasi Pembelajaran*, 153-154.

²⁵ Supardi, *Penilaian Autentik Pembelajaran Afektif, Kognitif, dan Psikomotor: Konsep dan Aplikasi*, 81-82.

- g) Panjang jawaban relatif sama.
 - h) Pilihan jawaban tidak menggunakan pernyataan semua jawaban di atas salah/ benar dan sejenisnya.
 - i) Pilihlah jawaban yang berbentuk angka/ waktu disusun berdasarkan urutan besar kecilnya angka atau kronologisnya.
 - j) Butir soal tidak bergantung pada jawaban soal sebelumnya.
- 3) Aspek Bahasa/Budaya

Berkaitan dengan penggunaan bahasa yang baik dan benar menurut ejaan yang sesuai. Aspek yang harus diperhatikan yaitu:

- a) Menggunakan bahasa yang sesuai dengan kaidah bahasa Indonesia.
- b) Menggunakan bahasa yang komunikatif.
- c) Tidak menggunakan bahasa yang berlaku setempat/ tabu.
- d) Pilihan jawaban tidak mengulang kata/ kelompok kata yang sama, kecuali merupakan satu kesatuan pengertian.²⁶

b. Analisis Butir Soal Secara Kuantitatif

Analisis butir soal secara kuantitatif dapat dilaksanakan berdasarkan teori tes modern dan teori tes klasik. Teori tes modern dikenal dengan sebutan teori tanggapan butir (*item response theory* atau IRT). IRT adalah suatu paradigma untuk merancang tes dan memberikan tes, skor, 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, maka digunakan teori klasik yang relatif lebih sederhana.²⁷

Teori tes klasik adalah teori mengenai analisis butir tes dimana analisis dilakukan dengan memperhitungkan kedudukan butir dalam suatu kelas atau kelompok. Karakteristik atau kualitas butir sangat bergantung pada kelompok dimana analisis butir dilakukan sehingga kualitas butir terikat pada sampel responden atau siswa yang memberikan respons (*sample bounded*).²⁸ Pada teori klasik, proses penelaahan butir soal melalui informasi dari jawaban peserta didik untuk meningkatkan mutu soal yang bersangkutan. Adapun kelebihan analisis butir soal secara klasik yaitu murah, dapat dilaksanakan sehari-hari dengan cepat menggunakan komputer, sederhana dan dapat menggunakan data dari beberapa peserta didik.²⁹ Sehingga teori klasik sering digunakan dalam menganalisis

²⁶ Elis Ratnawulan dan Rusdiana, *Evaluasi Pembelajaran*, 157-158.

²⁷ Eko Putro Widoyoko, *Penilaian Hasil Pembelajaran di Sekolah*, 174.

²⁸ Purwanto, *Evaluasi Hasil Belajar*, ed. Budi Santoso (Yogyakarta: Pustaka Pelajar, 2011), 98.

²⁹ Elis Ratnawulan dan Rusdiana, *Evaluasi Pembelajaran*, 161.

butir soal dibandingkan dengan teori modern. Oleh karena itu, peneliti hanya akan menggunakan teori klasik dalam menganalisis butir soal secara kuantitatif.

Aspek yang perlu diperhatikan analisis butir soal secara klasik yaitu: validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh. Penjelasannya adalah sebagai berikut:

1) Validitas

Validitas berasal dari kata *validity* yang mempunyai arti sejauh mana ketepatan dan kecermatan suatu instrumen pengukuran dalam melaksanakan fungsi ukurnya. Suatu alat evaluasi (tes) dikatakan mempunyai validitas yang tinggi (disebut valid) jika alat evaluasi atau tes itu dapat mengukur apa yang sebenarnya akan diukur atau memberikan hasil yang tepat dan akurat sesuai yang dimaksud dikenakannya tes tersebut. Sedangkan suatu tes yang menghasilkan data tidak relevan dengan tujuan diadakannya pengukuran dikatakan sebagai tes yang memiliki validitas rendah.³⁰

Validitas tes dibagi menjadi dua yaitu:

a) Validitas Logis

Validitas logis merupakan sebuah instrumen evaluasi yang menunjuk pada kondisi bagi sebuah instrumen yang memenuhi persyaratan valid berdasarkan hasil penalaran. Kondisi valid dipandang terpenuhi karena instrumen yang bersangkutan sudah dirancang secara baik, mengikuti teori dan ketentuan yang sudah ada. Validitas logis tidak perlu diuji kondisinya, tetapi langsung diperoleh sesudah instrumen sudah selesai disusun. Validitas logis ada dua macam, yaitu validitas isi dan validitas konstruk.³¹

(1) Validitas Isi (*Content Validity*)

Validitas isi berhubungan dengan kesanggupan tes untuk mengukur isi yang seharusnya diukur. Dengan kata lain, validitas isi menyatakan apakah tes sudah mencakup sampel yang representif dari domain perilaku yang diukur. Misalnya tes hasil belajar matematika di SMU harus mampu mengungkapkan isi pelajaran matematika di SMU.

(2) Validitas Konstruk (*Construct Validity*)

Validitas konstruk menunjuk pada sejauh mana suatu instrumen mampu mengukur pengertian-pengertian

³⁰ Mulyadi, *Evaluasi Pendidikan: Pengembangan Model Evaluasi Pendidikan Agama di Sekolah*, ed. A. Nurul Kawakip (Malang: UIN Maliki Press, 2014), 36.

³¹ Mulyadi, *Evaluasi Pendidikan: Pengembangan Model Evaluasi Pendidikan Agama di Sekolah*, 37-38.

yang terkandung dalam materi yang akan diukur. Misalnya: instrumen minat harus mampu mengukur pengertian-pengertian yang terkandung pada variabel minat. Agar lebih jelas, biasanya variabel tersebut diuraikan dulu menjadi indikator-indikator.³²

b) Validitas Empiris

Validitas empiris merupakan sebuah instrumen yang menunjuk pada kondisi bagi instrumen yang memenuhi persyaratan valid setelah diuji dari pengalaman. Contohnya: dalam kehidupan sehari-hari, seseorang dapat diakui jujur oleh masyarakat apabila dalam pengalaman dibuktikan bahwa orang tersebut jujur. Contoh lain, seseorang dapat dikatakan kreatif apabila dari pengalaman dibuktikan bahwa orang tersebut sudah banyak menghasilkan ide-ide baru yang diakui berbeda dari hal-hal yang sudah ada. Dari penjelasan dan kedua contoh tersebut diketahui bahwa validitas empiris tidak dapat diperoleh hanya dengan menyusun instrumen berdasarkan ketentuan seperti halnya validitas logis, tetapi harus dibuktikan melalui pengalaman.³³ Ada dua macam validitas empiris yaitu validitas “ada sekarang” atau kesamaan dan validitas prediksi atau ramalan.

(1) Validitas Kesamaan (*Concurrent Validity*)

Validitas kesamaan menunjuk kepada sejauh mana tes memiliki kesamaan dengan tes yang sudah ada atau yang sudah dibakukan. Kesamaan yang dimaksud meliputi: kemampuan yang diukur, objek yang diukur, dan waktu yang diperlukan.

(2) Validitas Prediksi (*Predictive Validity*)

Validitas prediksi menunjuk pada sejauh mana tes dapat menentukan atau meramalkan kriteria tertentu yang diinginkan. Misalnya apakah NEM di SLTP dapat digunakan untuk meramalkan prestasi belajar di SLTP.³⁴

2) Reliabilitas

Kata reliabilitas dalam bahasa Indonesia diambil dari kata *reliability* dalam bahasa Inggris, berasal dari kata *reliable* yang artinya dapat dipercaya. Tes dikatakan dapat dipercaya (*reliable*) jika memberikan hasil yang tetap atau ajek (*consistent*) apabila diteskan berkali-kali. Jika kepada siswa diberikan tes yang sama

³² Hamzah B.Uno dan Satria Koni, *Assesment Pembelajaran*, ed. Dewi Ispurwanti (Jakarta: Bumi Aksara, 2013), 152.

³³ Mulyadi, *Evaluasi Pendidikan: Pengembangan Model Evaluasi Pendidikan Agama di Sekolah*, 38.

³⁴ Hamzah B. Uno dan Satria Koni, *Assesment Pembelajaran*, 152.

yang pada waktu berlainan, maka setiap siswa akan tetap berada dalam urutan (rangking) yang sama atau ajek dalam kelompoknya.

Ajek atau tetap tidak selalu harus sama, tetapi mengikuti perubahan secara ajek. Jika keadaan mula-mula A berada lebih rendah dibandingkan dengan B, maka jika diadakan pengukuran ulang, A tetap berada lebih rendah dari B, itulah yang dikatakan ajek atau tetap, yaitu tetap dalam kedudukan siswa di antara anggota kelompok yang lain. Jika dihubungkan dengan validitas maka validitas berhubungan dengan ketepatan sedangkan reliabilitas berhubungan dengan ketetapan atau keajekan.³⁵

3) Tingkat Kesukaran

Soal yang baik adalah soal yang tidak terlalu mudah atau tidak terlalu sukar. Soal yang terlalu mudah tidak merangsang siswa untuk mempertinggi usaha memecahkannya. Sebaliknya soal yang terlalu sukar akan menyebabkan siswa menjadi putus asa dan tidak mempunyai semangat untuk mencoba lagi karena di luar jangkauannya.

Seorang siswa akan menjadi hafal akan kebiasaan guru-gurunya dalam hal pembuatan soal ini. Misalnya saja guru A dalam memberikan ulangan soalnya mudah-mudah, sebaliknya guru B kalau memberikan ulangan soalnya sukar-sukar. Dengan pengetahuannya tentang kebiasaan ini, maka siswa akan belajar giat jika menghadapi ulangan dari guru B dan sebaliknya jika akan mendapat ulangan dari guru A, tidak mau belajar giat atau bahkan mungkin tidak mau belajar sama sekali.³⁶

Bilangan yang menunjukkan sukar dan mudahnya sesuatu soal disebut indeks kesukaran (*difficulty index*). Besarnya indeks kesukaran antara 0,00 sampai 1,0. Indeks 0,00, indeks kesukaran ini menunjukkan bahwa soal itu terlalu sukar, sebaliknya indeks 1,0 menunjukkan bahwa soalnya terlalu mudah.

Untuk menghitung Indeks Kesukaran digunakan rumus:

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

Keterangan:

P = Indeks kesukaran

B = Banyaknya siswa yang menjawab benar untuk item soal yang dicari indeks kesukarannya

J = Jumlah seluruh siswa peserta tes

Menurut ketentuan yang sering diikuti, indeks kesukaran sering diklasifikasikan sebagai berikut:

- Soal dengan P 0,00 sampai 0,30 adalah butir instrumen sukar

³⁵ Eko Putro Widoyoko, *Penilaian Hasil Pembelajaran di Sekolah*, 184.

³⁶ Suharsimi Arikunto, *Dasar-dasar Evaluasi Pendidikan*, 207.

- Soal dengan P 0,31 sampai 0,70 adalah butir instrumen sedang
 - Soal dengan P 0,71 sampai 1,00 adalah butir instrumen mudah
- Butir-butir instrumen penilaian yang dianggap baik adalah butir instrumen dengan tingkat kesukaran sedang, yaitu dengan P antara 0,31 sampai dengan 0,70.³⁷

4) Daya Pembeda

Indeks daya pembeda merupakan indeks yang digunakan dalam membedakan antara peserta tes yang berkemampuan tinggi dengan peserta tes yang berkemampuan rendah.³⁸ Logikanya adalah siswa yang pandai (berkemampuan tinggi) lebih mampu menjawab dibandingkan dengan siswa yang kurang pandai (berkemampuan rendah).

Angka yang menunjukkan besarnya daya pembeda disebut indeks diskriminasi, disingkat D (d besar). Seperti halnya indeks kesukaran, indeks diskriminasi (daya pembeda) ini berkisar antara 0,00 sampai 1,00. Hanya bedanya, indeks kesukaran tidak mengenal tanda negatif (-), tetapi pada indeks diskriminasi ada tanda negatif. Tanda negatif pada indeks diskriminasi digunakan jika sesuatu soal “terbalik” menunjukkan kualitas testee. Yaitu anak pandai disebut bodoh dan anak bodoh disebut pandai.

Dengan demikian ada tiga titik pada daya pembeda yaitu:

$$-1,00 \leftrightarrow 0,00 \leftrightarrow 1,00$$

Keterangan: -1,00 = daya pembeda negatif

0,00 = daya pembeda rendah

1,00 = daya pembeda tinggi (positif)

Bagi suatu soal yang dapat dijawab benar oleh siswa pandai maupun siswa bodoh, maka soal itu tidak baik karena tidak mempunyai daya pembeda. Demikian pula jika semua siswa baik pandai maupun bodoh tidak dapat menjawab dengan benar. Soal tersebut tidak baik juga karena tidak mempunyai daya pembeda. Soal yang baik adalah soal yang dapat dijawab benar oleh siswa-siswa yang pandai saja.

Seluruh peserta tes dikelompokkan menjadi dua kelompok yaitu kelompok pandai atau kelompok atas (*upper group*) dan kelompok bodoh atau kelompok bawah (*lower group*).

Jika seluruh kelompok atas dapat menjawab soal tersebut dengan benar, seluruh kelompok bawah menjawab salah, maka soal tersebut mempunyai D paling besar, yaitu 1,00. Sebaliknya

³⁷ Supardi, *Penilaian Autentik Pembelajaran Afektif, Kognitif, dan Psikomotor: Konsep dan Aplikasi*, 88-89.

³⁸ Sumarna Surapranata, *Analisis, Validitas, Reliabilitas, dan Interpretasi Hasil Tes Implementasi Kurikulum 2004* (Bandung: PT Remaja Rodaskarya, 2004), 23.

jika semua kelompok atas menjawab salah, tetapi semua kelompok bawah menjawab betul, maka nilai D-nya -1,00. Tetapi jika siswa kelompok atas dan siswa kelompok bawah sama-sama menjawab benar atau sama-sama menjawab salah, maka soal tersebut mempunyai nilai D 0,00. Karena tidak mempunyai daya pembeda sama sekali.³⁹

Rumus yang digunakan untuk mengetahui daya pembeda setiap butir tes adalah:⁴⁰

$$D = \frac{Ba}{Ja} - \frac{Bb}{Jb}$$

Keterangan:

D = Daya pembeda butir

Ba = Banyaknya kelompok atas yang menjawab betul

Bb = Banyaknya kelompok bawah yang menjawab benar

Ja = Banyaknya subjek kelompok atas

Jb = Banyaknya subjek kelompok bawah

Suharsimi Arikunto mengklasifikasikan daya pembeda butir ke dalam lima kategori, yaitu.⁴¹

D : 0,00 - 0,20 = jelek (*poor*)

D : 0,21 - 0,40 = cukup (*satisfactory*)

D : 0,41 - 0,70 = baik (*good*)

D : 0,71 - 1,00 = baik sekali (*excellent*)

D : negative, semuanya tidak baik. Jadi semua butir soal yang mempunyai nilai D negatif sebaiknya dibuang saja.

5) Efektivitas Pengecoh

Pada soal bentuk pilihan ganda ada alternative jawaban (opsi) yang merupakan pengecoh. Menurut Purwanto, pengecoh (*distractor*) dikenal dengan istilah penyesat adalah pilihan jawaban yang bukan merupakan kunci jawaban. Pengecoh bukan sekedar pelengkap pilihan. Pengecoh diadakan untuk menyesatkan siswa agar tidak memilih kunci jawaban. Pengecoh menggoda siswa yang kurang begitu memahami materi pelajaran untuk memilihnya. Agar dapat melakukan fungsinya untuk mengecek maka pengecoh harus dibuat semirip mungkin dengan kunci jawaban.

Pengecoh dapat dikatakan berfungsi efektif apabila paling tidak ada siswa yang terkecoh memilih. Pengecoh yang sama sekali tidak dipilih tidak dapat melakukan fungsinya sebagai pengecoh karena terlalu menyolok dan dimengerti oleh semua siswa sebagai pengecoh soal. Pengecoh yang berdasarkan hasil

³⁹ Suharsimi Arikunto, *Dasar-dasar Evaluasi Pendidikan*, 211-212.

⁴⁰ Suharsimi Arikunto, *Dasar-dasar Evaluasi Pendidikan*, 213-214.

⁴¹ Suharsimi Arikunto, *Dasar-dasar Evaluasi Pendidikan*, 218.

uji coba tidak efektif direkomendasikan untuk diganti dengan pengecoh yang lebih menarik.⁴²

Suatu pengecoh dapat dikatakan berfungsi baik jika paling sedikit dipilih 5% peserta tes. Apabila pengecoh dipilih secara merata, maka termasuk pengecoh yang sangat baik. Apabila pengecoh lebih banyak dipilih oleh peserta tes dari kelompok atas dibandingkan dengan kelompok bawah, maka termasuk pengecoh yang menyesatkan.⁴³

Indeks pengecoh dihitung menggunakan rumus:⁴⁴

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

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 tiap soal

n = Jumlah alternatif jawaban (opsi)

1 = Bilangan tetap

Tolak ukur untuk menginterpretasikan kualitas pengecoh butir soal adalah sebagai berikut:

IP = lebih dari 200% : Sangat jelek

IP = 0% - 25% atau 176% - 200% : Jelek

IP = 26% - 50% atau 151% - 175% : Kurang baik

IP = 51% - 75% atau 126% - 150% : Baik

IP = 76% - 125% : Sangat baik

Catatan: jika semua peserta didik menjawab benar pada butir soal tertentu (sesuai kunci jawaban), maka IP = 0 yang berarti soal tersebut jelek. Dengan demikian pengecoh tidak berfungsi.

B. Penelitian Terdahulu

Penelitian tentang analisis tingkat kualitas soal pilihan ganda telah banyak dilakukan. Namun hal ini tetap menjadi suatu hal yang menarik untuk dijadikan sebuah penelitian. Berikut adalah beberapa hasil penelitian tentang analisis tingkat kualitas soal pilihan ganda yang pernah dilakukan sebelumnya, antara lain yaitu:

1. Nurul Septiana pada tahun 2016 melakukan penelitian analisis butir soal yang berjudul “*Analisis Butir Soal Ulangan Akhir Semester (UAS)*”

⁴² Purwanto, *Evaluasi Hasil Belajar*, 108.

⁴³ Sumarna Surapranata, *Analisis, Validitas, Reliabilitas dan Interpretasi Hasil Tes Implementasi Kurikulum 2004*, 43.

⁴⁴ Zainal Arifin, *Evaluasi Pembelajaran: Prinsip Teknik Prosedur* ed. Pipih Latifah (Bandung: PT Remaja Rodaskarya, 2014), 279-280.

Biologi Tahun Pelajaran 2015/2016 Kelas X dan XI pada MAN Sampit". Hasil penelitian ini menunjukkan bahwa: (1) kualitas soal UAS Biologi tahun pelajaran 2015/2016 kelas X dan XI di MAN Sampit memiliki kualitas cukup baik, karena sudah sesuai dengan soal standar, tetapi perlu perbaikan aspek materi dan konstruksi pada beberapa soal; (2) tingkat kesukaran butir soal biologi kelas X sebanyak 3 kategori sukar, 3 soal kategori sedang, dan 34 soal kategori mudah, sedangkan kelas X sebanyak 3 soal kategori sukar, 3 soal kategori sedang dan 34 soal kategori mudah, sedangkan pada kelas XI bahwa sebanyak 8 soal kategori sukar, 9 soal kategori sedang, dan 23 soal kategori mudah; (3) daya pembeda butir soal biologi kelas X soal dinyatakan kategori sangat baik tidak ada (0%), kategori baik 5%, kategori cukup sebanyak 27,5% dan kategori jelek berjumlah 67,5%, sedangkan pada kelas XI soal dinyatakan kategori sangat baik tidak ada (0%), kategori baik 5%, kategori cukup sebanyak 30%, dan kategori jelek berjumlah 65%; (4) efektivitas pengecoh soal biologi kelas X dari 40 soal terdapat 2 soal termasuk kriteria baik, 10 soal jriteria tidak baik, pada kelas XI terdapat 3 soal kriteria sangat baik, 6 soal kriteria baik, 12 soal kriteria cukup, 14 soal kriteria kurnag baik, dan 5 soal kriteria tidak baik; (5) validitas butir soal biologi kelas X dari 40 soal terdapat 21 soal (52,5%) yang dinyatakan valid sedangkan soal yang dinyatakan tidak valis sebanyak 19 soal (47,5%), pada kelas XI dari 40 soal terdapat 16 soal (40%) yang dinyatakan valid sedangkan soal yang dinyatakan tidak valid sebanyak 24 soal (60%); (6) reliabilitas butir soal biologi memiliki tingkat reliabilitas yang tinggi atau reliable yakni 0,731 pada kelas X dan 0,667 pada kelas XI.⁴⁵

Perbedaan antara penelitian yang dilakukan oleh Nurul Septiani dengan peneliti yaitu: (1) mata pelajaran soal UAS yang dianalisis berbeda, jika Nurul Septiani adalah Biologi Kelas X dan XI, maka peneliti adalah PKn kelas III MI; (2) metode yang digunakan dalam penelitian Nurul Septiani adalah Deskriptif Kualitatif, sedangkan peneliti metode penelitian yang digunakan adalah Deskriptif Kuantitatif; (3) waktu dan tempat saat penelitian berbeda; (4) subjek dan objek yang diteliti saat penelitian berbeda.

2. Halimah Wahyuningrum pada tahun 2014 telah melakukan penelitian yang berjudul "*Analisis Butir Soal Ujian Akhir Semester Mata Pelajaran Bahasa Jawa Siswa Kelas V SD Negeri Kledung Kradenan Banyuurip Purworejo Tahun Pelajaran 2013/2014*". Hasil penelitain

⁴⁵ Nurul Septiati, "Analisis Butir soal Ulangan Akhir Semester (UAS) Biologi Tahun Pelajaran 2015/2016 Kelas X dan XI pada MAN Sampit," *EduSains* 4, no. 2 (2016): 115, diakses pada 12 November, 2018, <http://e-journal.iain-palangkaraya.ac.id/index.php/edusains/article/download/514/648>.

ini menunjukkan bahwa: (1) tingkat kesukaran butir soal dari keseluruhan butir soal yang memenuhi persyaratan kelayakan sebesar 70%; (2) daya pembeda butir soal dari keluruhan butir soal yang memenuhi syarat daya pembeda sebesar 76%; (3) validitas butir soal pilihan ganda yang mempunyai validitas cukup sebesar 31,4%, validitas rendah sebesar 22,9% dan validitas sangat rendah sebesar 45,7%, sedangkan validitas isi soal esai dan uraian mempunyai kesejajaran terhadap kurikulum sebesar 86,7%, butir soal yang tidak mempunyai kesejajaran terhadap kurikulum sebesar 13,3%, dan untuk reliabilitas butir soal dengan menggunakan rumus KR-20 adalah 1,0277 dan reliabilitas butir soal dengan menggunakan KR-21 adalah 1,0271 dengan arti reliabilitas butir soal pilihan ganda memiliki korelasi yang tinggi, sedangkan reliabilitas butir soal esai menggunakan rumus Alpha sebesar -0,140 dengan arti reliabilitas soal esai kolerasinya sangat rendah, dan reliabilitas butir soal uraian sebesar 0,416 itu berarti reliabilitas soal uraian memiliki korelasi sedang. Butir soal yang memenuhi syarat tes yang baik sebesar 44% dan yang tidak memenuhi persyaratan sebesar 56%.⁴⁶

Perbedaan antara penelitian yang dilakukan oleh Halimah Wahyuningrum dengan peneliti yaitu: (1) mata pelajaran yang dianalisis berbeda, jika mata pelajaran yang dianalisis oleh Halimah Wahyuningrum adalah bahasa jawa kelas V SD maka peneliti adalah PKn kelas III MI; (2) subjek dan objek penelitian berbeda; (3) waktu dan tempat penelitian berbeda; dan (4) analisis yang digunakan berbeda jika penelitian yang dilakukan oleh Halimah Wahyuningrum adalah secara kuantitatif yang terdiri dari tingkat kesukaran soal, daya pembeda soal, validitas dan reliabilitas soal, sedangkan peneliti menganalisis secara kualitatif dan kuantitatif.

3. Muslikhah pada tahun 2014 telah melakukan penelitian yang berjudul “*Analisis Butir Soal Ujian Akhir Mata Pelajaran Akuntansi Keuangan Menggunakan Microsoft Excel 2010*”. Hasil penelitian ini menunjukkan bahwa: butir soal pilihan ganda valid 19 butir (63,33%), tidak valid 11 butir (36,67%), soal uraian valid 3 butir (75%), tidak valid 1 butir (25%); (2) soal pilihan ganda indeks reliabilitas menunjukkan angka 0,660, uraian sebesar 0,50 sehingga tidak reliable; (3) bentuk soal pilihan ganda 4 butir (13,33%) kategori sukar, 9 butir soal (30%) sedang, dan 16 butir (56,67%) mudah, bentuk soal uraian 4

⁴⁶ Halimah Wahyuningrum, “Analisis Butir Soal Ujian Akhir Semester Mata Pelajaran Bahasa Jawa Siswa Kelas V SD Negeri Kledung KRadenan Banyuurip Purworejo Tahun Pelajaran 2013/2014,” *Jurnal Program Studi Pendidikan Bahasa dan Sastra Jawa Universitas Muhammadiyah Purworejo* 5, no. 4 (2014): 1, diakses pada 2 Mei, 2019 <http://ejournal.umpwr.ac.id/index.php/aditya/article/download/1657/1576>.

butir (100%) kategori soal tingkat kesulitan sedang; (4) butir soal pilihan ganda dengan daya pembeda jelek 7 butir (23,33%), cukup 7 butir (23,33%), baik 10 butir (33,33%), baik sekali 6 butir (20%), bentuk soal uraian dengan daya pembeda jelek 1 butir (25%), cukup 1 (25%), dan baik sekali (50%); (5) bentuk soal pilihan ganda yang termasuk soal dengan pengecoh atau *distractor* yang berkualitas 3 butir (10%) memiliki pengecoh sangat baik, 10 butir (33,33%) baik, 11 butir (36,67%) kurang baik, dan 2 butir (6,67%) tidak baik.⁴⁷

Perbedaan antara penelitian yang dilakukan oleh Muslikah Purwanti dengan peneliti yaitu: (1) objek dan subjek yang berbeda; (2) tempat dan waktu yang berbeda; (3) analisis yang digunakan oleh Muslikah Purwanti adalah secara kuantitatif saja, jika peneliti secara kualitatif dan kuantitatif.

4. Heri Susanto, Achi Rinaldi, dan Novalia pada tahun 2015 telah melakukan penelitian yang berjudul “*Analisis Validitas Reliabilitas Tingkat Kesukaran dan Daya Beda pada Butir Soal Ujian Akhir Semester Ganjil Mata Pelajaran Matematika*”. Hasil penelitian ini menunjukkan bahwa: (1) terdapat 22 butir soal (55%) dinyatakan valid dan 18 butir soal (45%) dinyatakan tidak valid dari 40 butir soal tes bentuk obyektif; (2) reliabilitas soal tergolong baik berdasarkan perhitungan menggunakan beberapa rumus dan software yang menunjukkan hasil koefisien $r_{11} > r_{tabel}$; (3) tingkat kesukaran butir soal kurang baik karena keseimbangan soal yang mudah, sedang dan sukar belum proposional, keseimbangan yang proposional yaitu 3-5-2 atau 3-4-3; (4) terdapat 1 butir soal (2,50%) baik sekali, 9 butir soal (22,50%) tergolong baik, 10 butir soal (25%) tergolong sedang, 14 butir soal (35%) tergolong kurang baik dan 6 butir soal (15%) tergolong jelek sekali berdasarkan analisis daya beda.⁴⁸

Perbedaan penelitian yang dilakukan oleh Heri Susanto, Achi Rinaldi, dan Novalia dengan peneliti yaitu: (1) objek dan subjek yang diteliti berbeda; (2) tempat dan waktu berbeda; dan (3) analisis yang digunakan dalam penelitian ini adalah kuantitatif saja sedangkan analisis yang digunakan oleh peneliti adalah secara kualitatif dan kuantitatif.

⁴⁷ Muslikah Purwanti, “Analisis Butir Soal Ujian Akhir Mata Pelajaran Akuntansi Keuangan Menggunakan *Microsoft Office Excel 2010*,” *Jurnal Pendidikan Akuntansi Indonesia* 12, no. 1 (2014): 81, diakses pada 2 Mei, 2019, <https://journal.uny.ac.id/index.php/jpakun/article/download/2710/2258>.

⁴⁸ Hery Susanto, Achi Rinaldi, dan Novalia, “Analisis Validitas Reliabilitas Tingkat Kesukaran dan Daya Beda pada Butir Soal Ujian Akhir Semester Ganjil Mata Pelajaran Matematika,” *Al-Jabar: Jurnal Pendidikan Matematika* 6, no. 2 (2015): 215-216, diakses pada 2 Mei, 2019, <http://ejournal.radenintan.ac.id/index.php/al-jabar/article/download/50/44>.

5. Mahmud Alpusari pada tahun 2014 telah melakukan penelitian yang berjudul “*Analisis Butir Soal Konsep Dasar IPA 1 Melalui Penggunaan Program Komputer Anates Versi 4,0 For Windows*”. Hasil penelitian ini menunjukkan bahwa: (1) hasil analisis validitas butir soal pada taraf koefisien 1% soal yang valid berjumlah 16 butir dan soal yang tidak valid 24 butir soal, sedangkan pada taraf koefisien 5% soal yang valid 26 soal dan soal yang tidak valid 14 butir soal; (2) hasil analisis daya pembeda soal yang mendapat kategori sangat jelek 1 butir pada soal nomor 20, butir soal yang berkategori jelek 15 butir, butir soal kategori cukup 15 soal dan butir soal yang berkategori baik 9 soal; (3) hasil analisis tingkat kesukaran soal yang mendapatkan kategori sangat mudah 17 butir soal, butir soal yang mendapatkan kategori mudah 9 soal, butir soal yang mendapatkan kategori sukar 1 soal dan berkategori sangat sukar dua butir soal; (4) hasil analisis rekap butir soal yang bisa digunakan berjumlah 21 butir soal langsung saja digunakan, 5 butir soal bisa digunakan tetapi harus diperbaiki dan 14 butir soal tidak bisa dipakai.⁴⁹

Perbedaan penelitian yang dilakukan oleh Mahmud Alpusari dengan peneliti yaitu: (1) objek dan subjek berbeda; (2) tempat dan waktu penelitian berbeda; (3) analisis yang digunakan oleh Mahmud Alpusari adalah secara kuantitatif sedangkan peneliti menggunakan analisis kualitatif dan kuantitatif.

C. Kerangka Berpikir

Kemampuan peserta didik dapat diukur menggunakan alat ukur berupa tes. Pengetahuan peserta didik pada mata pelajaran PKn kelas III di MI NU Sholahiyah saat penilaian akhir semester gasal diukur menggunakan alat ukur berupa tes tertulis. Soal yang digunakan untuk PAS gasal disusun oleh lembaga Ma’arif NU Kudus dan belum dianalisis oleh guru PKn kelas III di madrasah tersebut. Untuk itu, agar soal yang digunakan untuk PAS dapat mengukur kemampuan peserta didik dengan tepat, perlu dilakukan analisis tingkat kualitas butir soal.

Kegiatan analisis tingkat kualitas butir soal merupakan suatu kegiatan yang harus dilakukan guru untuk meningkatkan kualitas soal yang telah dibuat. Analisis tingkat kualitas butir soal dalam penelitian ini ditinjau dari aspek materi, konstruksi, bahasa, validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh. Teknik analisis yang

⁴⁹ Mahmud Alpusari, “Analisis Butir Soal Konsep Dasar IPA 1 Melalui Penggunaan Program Komputer Anates Versi 4,0 For Window,” *Jurnal Primary Program Studi Pendidikan Guru Sekolah Dasar Fakultas Keguruan dan Ilmu Pendidikan Universitas Riau* 3, no. 2 (2014): 115, diakses pada 2 Mei, 2019, <https://ejournal.unri.ac.id/index.php/JPFKIP/article/download/2501/2458>.

digunakan yaitu analisis secara kualitatif dan analisis secara kuantitatif. Analisis kualitatif dilakukan untuk menganalisis data yang bersifat kualitatif yaitu pada aspek materi, konstruksi, dan bahasa. Sedangkan analisis kuantitatif dilakukan untuk menganalisis data yang bersifat kuantitatif yaitu pada aspek validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh.

Pada penelitian ini analisis dengan dengan cara kualitatif digunakan untuk menguji validitas isi dari soal bahwa soal tersebut sesuai atau tidak dengan isi atau materi yang telah diberikan. Validitas soal merupakan ketepatan mengukur yang dimiliki oleh setiap butir soal dalam mengukur apa yang seharusnya diukur berdasarkan kriteria yang telah ditentukan. Soal dinyatakan valid atau signifikan jika soal tersebut menunjukkan ketepatan dan kecermatan dalam mengukur hasil belajar.

Reliabilitas soal digunakan untuk mengukur tingkat konsistensi dari suatu soal. Soal dikatakan reliabel jika ia mampu memberikan hasil yang sama jika diberikan pada peserta didik yang sama namun pada waktu yang berbeda. Tingkat kesukaran merupakan pengukuran yang digunakan untuk mengukur besar derajat kesukaran suatu soal. Daya pembeda merupakan kemampuan soal untuk membedakan peserta didik yang telah menguasai materi dengan yang belum menguasai. Inti dari daya pembeda adalah kemampuan untuk mengukur kemampuan peserta didik. Efektivitas pengecoh adalah kemampuan soal untuk mengecoh peserta didik yang belum menguasai materi secara keseluruhan.

Dari berbagai aspek yang dianalisis tersebut terhadap Tingkat Kualitas Soal Pilihan Ganda Penilaian Akhir Semester Gasal Mata Pelajaran Pendidikan Kewarganegaraan Kelas III MI NU Sholahiyah Pedawang Kecamatan Bae Kabupaten Kudus Tahun Pelajaran 2018/2019, akan menunjukkan hasil analisis apakah soal tersebut merupakan soal yang baik atau tidak. Hasil analisis tersebut, yang menunjukkan mana soal yang baik, kurang baik dan tidak baik. tindak lanjut untuk soal yang baik akan dimasukkan ke dalam bank soal, sedangkan untuk soal yang kurang baik dapat direvisi dan soal yang tidak baik akan dibuang.

Berikut gambar kerangka berpikir yang menggambarkan kegiatan analisis tingkat kualitas soal pilihan ganda mata pelajaran PKn kelas III MI NU Sholahiyah Pedawang Kecamatan Bae Kabupaten Kudus, sebagai berikut:

Gambar 2.1
Kerangka Berpikir Penelitian

