

BAB II

TINJAUAN PUSTAKAN DAN KERANGKA TEORI

A. Tinjauan Pustaka

Beberapa penelitian terdahulu tentang analisis butir soal diantaranya adalah :

Penelitian ditulis oleh Hikmatul Fitroh, mahasiswa jurusan Pendidikan Agama Islam Fakultas Tarbiyah dan Keguruan Universitas Islam Negeri Sunan Kalijaga Yogyakarta berjudul “*Analisis Kualitas Soal Ujian Semester I Mata Pelajaran Fiqih Kelas X Madrasah Aliyah Nurul Ummah Kota Gede Yogyakarta*” Pembahasan skripsi ini tentang analisis butir soal pilihan ganda secara kuantitatif meliputi validitas, reliabilitas, tingkat kesukaran butir soal, daya pembeda, dan fungsi pengecoh. Pembahasan lain tentang analisis secara kualitatif meliputi aspek materi, konstruksi, dan bahasa.

Penelitian ditulis oleh Moh Mustaghfiri, mahasiswa jurusan Pendidikan Agama Islam Fakultas Tarbiyah Keguruan Universitas Negeri Sunan Kalijaga Yogyakarta berjudul “*Analisis Kualitas Soal UAMBN Mata Pelajaran Al-Qur'an Hadist*”. Dalam penelitian ini yang dibahas hanya analisis kualitatif yang meliputi aspek materi, konstruksi, dan bahasa.

Penelitian ditulis oleh Solihati, mahasiswa jurusan Pendidikan Agama Islam Fakultas Tarbiyah dan Keguruan Universitas Islam Negeri Sunan Kalijaga Yogyakarta berjudul “*Analisis Kualitas Butir soal Ulangan Akhir Semester 1 Mata Pelajaran Akidah Akhlak kelas VII MTsN Maguwoharjo*”. Dalam skripsi ini yang dibahas analisis secara kuantitatif meliputi validitas, reliabilitas, tingkat kesukaran soal, daya pembeda, dan fungsi pengecoh.

Persamaan penelitian ini dengan yang sebelumnya adalah sama-sama meneliti butir soal sedangkan perbedaan dalam penelitian ini dengan penelitian-penelitian yang sebelumnya adalah subyek penelitian, objek penelitian, dan tempat penelitiannya.

B. Landasan Teori

1. Analisis Butir Soal

a. Pengertian Analisis Butir Soal

Analisis butir soal bisa membantu para pengguna tes dalam mengevaluasi tes-tes yang digunakan. Disamping itu analisis butir soal amat relevan bagi penyusunan tes-tes informal dan lokal, seperti ujian yang disiapkan guru untuk digunakan dikelas. Dalam melaksanakan analisis butir soal, para penulis soal dapat menganalisis secara kualitatif, dalam kaitan dengan isi dan bentuknya, dan

kuantitatif dalam kaitan dengan ciri-ciri statistiknya. Analisis kualitatif mencakup pertimbangan validitas isi dan konstruk, sedangkan analisis kuantitatif mencakup pengukuran kesulitan butir soal dan diskriminasi soal yang termasuk validitas soal dan reliabilitasnya, validitas dan reliabilitas yang tinggi pun tergantung dulu dalam tes melalui analisis butir soal. Tes-tes dapat diperbaiki melalui seleksi, substitusi, ataupun revisi butir-butir soal (Anastasi dan Urbina, 2007: 190).

b. Manfaat Analisis Butir Soal

Kegiatan analisis butir soal memiliki banyak manfaat, di antaranya adalah: (1) dapat membantu para pengguna tes dalam evaluasi atas tes yang digunakan, (2) sangat relevan bagi penyusunan tes informal dan lokal seperti tes yang disiapkan guru untuk siswa di kelas, (3) mendukung penulisan butir soal yang efektif, (4) secara materi dapat memperbaiki tes di kelas, (5) meningkatkan validitas soal dan reliabilitas (Anastasi and Urbina, 2007:190).

2. Teknik Analisis Butir Soal

Analisis kualitas tes merupakan suatu tahapan yang harus ditempuh untuk mengetahui derajat kualitas suatu tes. Dalam penilaian

hasil belajar diharapkan tes dapat menggambarkan hasil yang objektif dan akurat. Dalam melaksanakan analisis butir soal, pembuat soal dapat melakukan analisis secara kualitatif, dalam kaitannya dengan isi dan bentuk, dan analisis secara kuantitatif dalam kaitannya dengan ciri-ciri statistiknya atau prosedur peningkatan secara judgment dan prosedur peningkatan secara empirik. Analisis kualitatif mencakup pertimbangan validitas isi dan konstruk, sedangkan analisis kuantitatif mencakup pengukuran kesulitan butir soal dan diskriminasi soal yang termasuk validitas dan reliabilitas soal (Ata Nayla Amalia and Ani Widayati, 2012:5)

a. Validitas

Menurut Anas Sudijono (2012: 163), validitas adalah salah satu ciri yang menandai tes hasil belajar yang baik. Untuk dapat menentukan apakah tes hasil belajar telah memiliki validitas atau daya ketetapan mengukur, dapat dilakukan dari dua segi, yaitu: validitas tes dan validitas item

1) Validitas Tes

Menurut Anas Sudijono (2012: 163), penganalisisan terhadap tes hasil belajar sebagai suatu totalitas dapat dilakukan

dengan dua cara yaitu penganalisisan dengan jalan berpikir secara rasional (*logical analysis*) dan penganalisisan yang dilakukan dengan mendasarkan diri pada kenyataan empiris (*empirical analysis*).

2) Validitas Item

Menurut Anas Sudijono (2012: 182), validitas item dari suatu tes adalah ketepatan mengukur yang dimiliki oleh sebutir item (yang merupakan bagian tak terpisahkan dari tes sebagai suatu totalitas), dalam mengukur apa yang seharusnya diukur lewat butir item tersebut. Hubungan antara butir item dengan tes hasil belajar sebagai suatu totalitas adalah bahwa semakin banyak butir-butir item yang dapat dijawab oleh peserta didik, maka skor total hasil tes tersebut akan semakin tinggi. Untuk sampai pada kesimpulan bahwa item-item yang ingin diketahui validitasnya, dapat digunakan teknik korelasi sebagai teknik analisisnya. Sebutir item dapat dinyatakan valid apabila skor item yang bersangkutan terbukti memiliki kesejajaran dengan skor total.

b. Reliabilitas

Reliabilitas merujuk pada konsistensi skor yang dicapai oleh orang yang sama ketika mereka diuji ulang atas dengan tes yang sama pada kesempatan berbeda, atau dengan seperangkat butir-butir ekuivalen (*equivalent item*) yang berbeda. Dalam pengertian luas, reliabilitas tes menunjukkan sejauh mana perbedaan-perbedaan individu dalam skor tes yang dianggap sebagai sebab oleh perbedaan-perbedaan yang sesungguhnya dalam karakteristik yang dipertimbangkan sejauh mana dapat dianggap disebabkan kesalahan peluang (Anastasi dan Urbina 2007: 94)

Pendapat lain reliabilitas adalah tingkat atau derajat konsistensi dari suatu instrumen. Reliabilitas tes berkenaan dengan pertanyaan apakah suatu tes teliti dan dapat dipercaya sesuai dengan kriteria yang telah ditetapkan. Suatu tes dikatakan reliabel jika selalu memberikan hasil yang sama bila diteskan pada kelompok yang sama pada waktu atau kesempatan yang berbeda (Zainal Arifin, 2011: 258) dalam (Ata Nayla Amalia and Ani Widayati, 2012:6)

Menurut Anastasi dan Urbina (2007:102-112). Pengukuran reliabilitas dapat dilakukan dengan lima cara:

1) Reliabilitas Tes-Retes

Metode paling jelas untuk menemukan reliabilitas skor tes ialah mengulang tes yang sama pada kesempatan kedua. Koefisien reliabilitas dalam kasus ini hanyalah korelasi antara skor-skor yang didapatkan oleh orang yang sama pada penyelenggaraan tes.

2) Reliabilitas Bentuk Alternatif

Satu cara untuk menghindari kesulitan yang ditemukan dalam reliabilitas tes dan tes ulang adalah melalui penggunaan bentuk-bentuk tes lainnya. Dengan demikian, orang yang sama bisa dites dengan satu bentuk pada kesempatan pertama dan bentuk lain yang ekuivalen pada kesempatan kedua. Korelasi anatar skor-skor yang didapatkan pada dua bentuk itu merupakan koefisien reliabilitas res.

3) Reliabilitas Belah Separuh (*Split Half Reliability*)

Dua skor didapatkan untuk setiap orang dengan mambagi tes menjadi paruhan-paruhan yang ekuivalen. Tampak bahwa reliabilitas belah separuh merupakan ukuran konsisten dalam kaitan dengan sampling isi. Stabilitas sementara skor-skor tidak masuk dalam reliabilitas semacam itu karena hanya

ada satu kesempatan tes. Jenis reliabilitas ini kadang kala disebut koefisien konsistensi internal, karena hanya dibutuhkan penyelenggaraan tunggal atas bentuk tes.

4) Reliabilitas Kuder Richardson dan Koefisien Alpha

Didasarkan pada konsistensi respons terhadap semua butir soal dalam tes. Konsistensi antarpersonal ini dipengaruhi oleh dua sumber varian kesalahan:

- a) Pencuplikas isi (sebagaimana dalam bentuk alternatif dan reliabilitas belah separuh)
- b) Heterognis dari dominan perilaku yang disampelkan. Semakin homogen, semakin tinggilah konsistensi antar personal.

5) Reliabilitas Pemberi Skor

Reliabilitas pemberi skor dapat ditemukan dengan memiliki sampel lembaran tes yang diskor secara terpisah oleh dua penguji. Dengan demikian, dua skor yang didapatkan oleh masing-masing peserta tes ini kemudian dikorelasikan dengan cara biasa dan koefisien korelasi yang dihasilkan adalah ukuran reliabilitas pemberi skor.

c. **Tingkat Kesukaran**

Indeks kesukaran butir adalah bilangan yang menunjukkan sukar dan mudahnya soal. Semakin tinggi indeks kesukaran butir maka soal semakin mudah. Soal yang baik adalah soal tidak terlalu mudah atau tidak terlalu sukar. Analisis tingkat kesukaran soal adalah mengkaji soal-soal dari segi kesulitannya sehingga dapat diperoleh soal-soal mana yang termasuk mudah, sedang, dan sukar. (Ata Nayla Amalia and Ani Widayati, 2012:8)

Menurut Witherington dalam Anas Sudijono (2012: 371) angka indeks kesukaran butir itu besarnya berkisar antara 0,00 sampai dengan 1,00. Semakin besar angka indeks kesukaran maka soal semakin mudah. Jika seluruh peserta ujian menjawab dengan salah butir tersebut maka soal tersebut sangat sukar dengan angka kesukaran 0,00 dan jika angka kesukaran 1,00 maka soal sangat mudah karena dijawab dengan benar oleh seluruh peserta tes

Tes terdiri dari dua bentuk yaitu tes objektif dan tes uraian, maka dalam melakukan perhitungan tingkat kesukaran digunakan cara yang berbeda. Untuk tes bentuk objektif dalam menghitung tingkat kesukaran dapat dilakukan dengan menggunakan rumus sebagai berikut:

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

Keterangan:

P = Angka indeks kesukaran item

B = Banyaknya siswa yang menjawab soal dengan betul

JS = Jumlah seluruh siswa peserta tes

(Anas Sudijono, 2012: 371)

d. Daya Pembeda

Menurut Anas Sudijono (2012: 385), daya pembeda item adalah kemampuan suatu butir item tes hasil belajar untuk dapat membedakan antara *testee* yang berkemampuan tinggi dengan *testee* yang berkemampuan rendah. Mengetahui daya pembeda item sangat penting, sebab salah satu dasar pegangan untuk menyusun butir tes hasil belajar adalah adanya anggapan bahwa kemampuan antara *testee* yang satu dengan *testee* yang lain berbeda-beda. Selain itu, butir tes hasil belajar harus mampu memberikan hasil tes yang mencerminkan adanya perbedaan kemampuan yang terdapat di kalangan *testee* tersebut.

Daya pembeda item dapat diketahui dengan melihat besar kecilnya angka indeks diskriminasi item. Angka indeks diskriminasi

item adalah sebuah angka atau bilangan yang menunjukkan besar kecilnya daya pembeda (*discrimination power*) yang dimiliki oleh sebutir item. Sama halnya dengan menganalisis tingkat kesukaran, dalam menganalisis daya pembeda soal bentuk objektif dan bentuk uraian dilakukan dengan cara yang berbeda. Tes bentuk objektif dalam menghitung daya pembeda dapat dilakukan dengan menggunakan rumus sebagai berikut:

$$D = P_A - P_B$$

Keterangan:

D = angka indeks diskriminasi

$P_A = \frac{B_A}{J_A}$ = Proporsi peserta kelompok atas yang menjawab benar

$P_B = \frac{B_B}{J_B}$ = Proporsi peserta kelompok bawah yang menjawab benar

(Anas Sudijono, 2012: 390)

e. Fungsi Pengecoh/ Distractor

Berbeda dengan soal bentuk uraian, pada soal pilihan ganda telah dilengkapi beberapa pilihan jawaban. Di antara pilihan jawaban

yang ada, hanya satu yang benar. Selain jawaban yang benar tersebut, adalah jawaban yang salah. Jawaban yang salah itulah yang dikenal dengan *distractor* (pengecoh). Butir soal yang baik, pengecohnya akan dipilih secara merata oleh peserta didik yang menjawab salah. Sebaliknya, butir soal yang kurang baik, pengecohnya akan dipilih secara tidak merata oleh peserta didik.

Tujuan utama dari pemasangan *distractor* pada setiap butir item adalah agar dari sekian banyak peserta tes yang mengikuti tes hasil belajar ada yang tertarik untuk memilihnya. *Distractor* akan mengecoh peserta didik yang kurang mampu untuk dapat dibedakan dengan yang mampu. *Distractor* yang baik adalah yang dapat dihindari oleh peserta didik yang pandai dan akan dipilih oleh peserta didik yang kurang pandai. Dengan demikian *distractor* baru dapat dikatakan telah berfungsi dengan baik apabila *distractor* tersebut telah memiliki daya rangsang atau daya tarik yang baik.

Menurut Anas Sudijono (2012: 411), mengungkapkan bahwa *distractor* telah dapat menjalankan fungsinya dengan baik apabila *distractor* tersebut telah dipilih sekurangnya 5% dari seluruh peserta tes *Distractor* yang telah menjalankan fungsinya dengan baik dapat digunakan kembali pada tes yang akan datang. Dengan demikian, efektivitas *distractor* adalah seberapa baik pilihan yang

salah tersebut dapat mengecoh peserta tes yang memang tidak mengetahui kunci jawaban yang tersedia. Semakin banyak peserta tes yang memilih *distractor* tersebut, maka *distractor* itu dapat menjalankan fungsinya dengan baik. Jika peserta tes mengabaikan semua option (tidak memilih) disebut omit. Dilihat dari segi omit, sebuah item dikatakan baik jika omitnya tidak lebih dari 10 % pengikut tes.

3. Ulangan Tengah Semester

Ulangan adalah proses yang dilakukan untuk mengukur pencapaian kompetensi peserta didik secara berkelanjutan dalam proses pembelajaran, untuk memataui kemajuan, melakukan perbaikan pembelajaran, dan menentukan keberhasilan belajar peserta didik. Ulangan bisa dilakukan guru pada setiap tahap pembelajaran, bisa meliputi ulangan harian, ulangan tengah semester, ulangan akhir semester dan ulangan kenaikan kelas. Ulangan tengah semester adalah kegiatan yang dilakukan oleh pendidik untuk mengukur pencapaian kompetensi peserta didik di pertengahan semester. Cakupan ulangan meliputi seluruh indikator yang mempresentasikan semua kompetensi dasar pada semester tersebut.

Tes yang dilakukan di sekolah, mempunyai fungsi untuk mengukur kemampuan siswa dalam menguasai mata pelajaran tertentu, yang nantinya hasil dari penilaian ini akan dipergunakan sebagai bahan evaluasi dan penentuan program pembelajaran yang akan diambil. Penilaian Ulangan Akhir Semester di SMA/SMK/MA dan sederajat

4. Anates

Anates merupakan sebuah program aplikasi komputer yang bertujuan untuk menganalisis butir soal. Program ini sangat bermanfaat khususnya bagi para guru umumnya para pemerhati evaluasi pendidikan. Program ini dikembangkan oleh Bapak Karno, seorang dosen Psikologi di UPI dan Bapak Yudi Wibisono, seorang Konsultan komputer. Fasilitas yang ada dalam program anates yaitu:

- a. Penyekoran Data, meliputi:
 - 1) Memasukan skor data hasil tes
 - 2) Membobot skor data sesuai yang dibutuhkan

- b. Pengolahan Data, meliputi:
 - 1) Reliabilitas
 - 2) Kelompok unggul dan asor

- 3) Daya Pembeda
- 4) Tingkat Kesukaran Soal
- 5) Korelasi skor butir soal dengan skor total
- 6) Kualitas pengecoh

5. Mata Pelajaran PAI-2

Mata pelajaran PAI-2 adalah mata pelajaran gabungan dari tiga mata pelajaran agama yaitu Tarikh, Kemuhammadiyah dan Akhlak. Kebijakan dari sekolah sehingga mapel yang tadinya terbagi-bagi dijadikan satu mapel, buku pegangan siswa dan guru mapel PAI-2 menggunakan Al-Islam dan Kemuhammadiyah diterbitkan oleh mentari pustaka.

Tarikh dalam pengertiannya adalah ilmu untuk mengetahui sejarah atau kejadian masa lampau, dimana siswa akan mengerti akan kejadian-kejadian atau peristiwa penting yang terjadi dalam perkembangan Islam dari zaman Rasulullah sampai sekarang. Kemuhammadiyah guna mengenalkan siswa sejarah perkembangan dan pembentukan muhammadiyah serta mengenal tokoh-tokoh penting yang memajukan muhammadiyah dan akhlak sendiri guna membentuk sifat dan sikap selayaknya muslim yang beriman kepada Allah.