

BAB III

METODE PENELITIAN

A. Metode Penelitian

1. Pendekatan

Penelitian ini tergolong dalam penelitian deskriptif kuantitatif dimana tujuan dari penelitian ini adalah untuk mencari informasi dan data yang dapat digunakan untuk mendeskripsikan kualitas tes. Pendekatan yang digunakan dalam penelitian ini adalah pendekatan kuantitatif karena data yang diperoleh di lapangan adalah dalam bentuk angka-angka dan dianalisis dengan statistic menggunakan program *ANBUSO Version 8.0*.

2. Variabel penelitian

Variabel penelitian ini adalah analisis butir soal yang dilihat dari aspek validitas, reliabilitas, daya pembeda, tingkat kesukaran, efektivitas pengecoh serta kaidah isi dan kaidah penulisan soal.

3. Populasi dan sampel, lokasi dan objek penelitian

Populasi dalam penelitian ini adalah seluruh siswa kelas VII lebih tepatnya kelas A dan E SMP Muhammadiyah 3 Yogyakarta. Lokasi penelitian ini dilaksanakan di SMP Muhammadiyah 3 Yogyakarta Tahun Ajaran 2019/2020. Waktu penelitian adalah pada semester I (ganjil) tahun ajaran 2019/2020 dimulai pada bulan November 2019

tepatnya setelah ujian tengah semester Mata Pelajaran Aqidah Akhlak. Objek Penelitian ini adalah hasil ujian tengah semester peserta didik kelas VII A dan E SMP Muhammadiyah 3 Yogyakarta Tahun Ajaran 2019/2020 yang berjumlah 57 peserta didik.

4. Teknik pengumpulan data

Pada penelitian ini, teknik pengumpulan data yang dipakai adalah metode dokumentasi. Sehingga metode dokumentasi disini mengumpulkan data dengan cara mengutip dari data yang ada atau keterangan yang ada atau bisa juga mempelajari data yang tertulis serta tercatat dan sesuai dengan masalah yang diteliti. Metode ini digunakan untuk mendapatkan hasil data butir – butir soal, kunci jawaban peserta didik dan soal Formatif mata pelajaran Aqidah Akhlak.

5. Validitas dan Reabilitas

a. Validitas Tes

Validitas ialah tingkat ketepatan suatu tes. Penentuan kategori validitas bisa dilakukan dari segi kesesuaian isi soal dengan kisi-kisi soal dan pengukuran oleh butir soal. Validitas butir soal dapat dihitung dengan menggunakan korelasi *point biserial* untuk soal pilihan ganda dan korelasi *product moment* untuk soal uraian. Setelah diperoleh hasil dari perhitungan dengan menggunakan korelasi *product moment* dan korelasi *point biserial*. Kemudian dikonsultasikan dengan r_{tabel} pada taraf signifikansi 5% sesuai dengan jumlah siswa yang diteliti. Suatu tes hasil belajar dapat

dikatakan valid apabila diukur dengan tepat sesuai dengan tujuan pengukuran. Sebaliknya, tes hasil belajar dapat dikatakan tidak valid apabila tidak sesuai dengan tujuan pengukuran atau tepat dalam pengukurannya.

a. Reliabilitas Tes

Reliabilitas merupakan tingkat atau derajat konsistensi dari suatu instrumen. Suatu tes dapat dikatakan memiliki Reliabilitas yang tinggi apabila selalu memberikan hasil yang sama bila diteskan pada kelompok yang sama pada waktu atau kesempatan yang berbeda. Reliabilitas untuk soal pilihan ganda dapat dihitung dengan menggunakan rumus K.R-20 dan untuk soal uraian dapat dihitung dengan rumus Alpha. Soal tes dikatakan memiliki Reliabilitas tinggi apabila koefisien Reliabilitas soal (r_{11}) sama dengan atau lebih besar dari 0,70.

6. Analisis data

a. Validitas

Analisis validitas bertujuan untuk mengetahui apakah suatu tes sudah tepat digunakan sebagai alat ukur. Validitas item dihitung dengan rumus korelasi point biserial sebagai berikut:

$$r_{pb1} = \frac{M_r - M_t}{S_t} \sqrt{\frac{p}{q}}$$

Keterangan:

Y_{rhi} : Koefisien korelasi biserial

M_p : Rerata skor dari subjek yang menjawab betul

M_t : Rerata skor total

S_t : Standar deviasi dari skor total

p : Proporsi siswa yang menjawab benar

(p = banyaknya siswa yang benar/jumlah semua siswa)

q : Proporsi siswa yang menjawab salah

$$(q = 1 - p)$$

Indeks korelasi (Y_{rhi}) yang diperoleh merupakan hasil dari perhitungan dikonsultasikan dengan r table pada taraf signifikan 5% sesuai dari siswa yang diteliti. Apabila $Y_{rhi} > r_{tabel}$ maka butir soal valid

b. Reliabilitas

Reliabilitas untuk soal bentuk pilihan ganda dapat dihitung dengan rumus kr20 sebagai berikut:

$$r_{11} = \left(\frac{n}{n-1} \right) \left(\frac{S^2 - \sum pq}{S^2} \right)$$

Keterangan:

r_{11} : reliabilitas tes secara keseluruhan

p : proporsi subjek yang menjawab item dengan benar

q : proporsi subjek yang menjawab item dengan salah ($q = 1 - p$)

$\sum pq$: jumlah hasil perkalian antara p dan q

n : banyaknya item

S : standar deviasi dari tes (standar deviasi adalah akar devians)

Reliabilitas untuk soal bentuk uraian dapat dihitung dengan rumus Alpha sebagai berikut:

$$r_{11} = \left(\frac{n}{(n-1)} \right) \left(1 - \frac{\sum \sigma^2}{\sigma_t^2} \right)$$

Keterangan:

r_{11} = Reliabilitas yang dicari

$\sum \sigma^2$ = Jumlah varians skor tiap-tiap item

σ_t^2 = Varians total

Kemudian pada pemberian interpretasi terhadap koefisien tes (r_{11}) pada umumnya memiliki patokans sebagai berikut:

- 1) Ketika r_{11} sama dengan atau lebih besar daripada 0,70 bisa diartikan tes hasil belajar yang diuji reliabilitasnya dinyatakan telah memiliki reliabilitas yang tinggi (= *reliable*).
- 2) Apabila r_{11} lebih kecil dari pada angka 0,70 berarti bahwa tes hasil belajar yang sedang diuji reliabilitasnya dinyatakan belum memiliki reliabilitas yang tinggi (*un-reliable*).

c. Tingkat Kesukaran

Rumus mencari Tingkat Kesukaran (P) sebagai berikut:

$$P = B / JS$$

Keterangan:

P : Indeks kesukaran

B : Siswa yang menjawab dengan betul

JS : Jumlah seluruh siswa peserta tes

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

P : 0,00 – 0,30 : sukar/sulit

P : 0,31 – 0,70 : sedang

P : 0,71 – 1,00 : mudah

d. Daya Pembeda

Untuk menghitung Daya Pembeda perlu dibedakan antara kelompok kecil (kurang dari 100) dan kelompok besar (lebih dari 100).

1) Untuk kelompok kecil

Seluruh kelompok testee dibagi dua sama besar, 50% kelompok atas (J_A) dan kelompok bawah (J_B). Seluruh pengikut tes, dideretan mulai dari skor teratas sampai terbawah lalu dibagi dua.

2) Untuk kelompok besar

Mengingat biaya dan waktu untuk menganalisis, maka untuk kelompok besar biasanya hanya diambil kedua kutubnya saja, yaitu 27% skor teratas sebagai kelompok atas (J_A) dan 27% skor terbawah sebagai kelompok bawah (J_B).

Untuk mencari Daya Pembeda dapat digunakan rumus sebagai berikut:

$$D = \frac{B_A}{J_A} - \frac{B_B}{J_B} \text{ atau } P_A - P_B$$

Keterangan:	
D	: Jumlah peserta tes
J_A	: Banyaknya peserta kelompok atas
J_B	: Banyaknya peserta kelompok bawah
B_A	: Banyaknya kelompok atas yang menjawab benar
B_B	: banyaknya kelompok bawah yang menjawab benar
$P_A = \frac{B_A}{J_A}$: proporsi kelompok atas yang menjawab benar
$P_B = \frac{B_B}{J_B}$: proporsi kelompok bawah yang menjawab benar

Selanjutnya daya Pembeda akan diklasifikasikan sesuai dengan pengklasifikasiannya untuk mengetahui klasifikasi butir soal tersebut. Klasifikasi Daya Pembeda:

D : 0,00 – 0,19 : jelek (*poor*)

D : 0,20 – 0,29 : cukup (*marginal*)

D : 0,30 – 0,39 : baik (*good*)

D : 0,40 ke atas : sangat baik (*very good*)

e. Efektivitas Pengecoh/*Distractor*

Efektivitas penggunaan pengecoh/*distractor* dapat dilihat dari pola sebaran siswa. Pola sebaran jawaban diperoleh dengan menghitung banyaknya *testee* yang memilih pilihan jawaban atau yang tidak memilih apapun. Dari pola sebaran jawaban dapat ditentukan apakah pengecoh dapat berfungsi atau tidak. Suatu butir soal dapat dikategorikan sebagai soal yang baik apabila pengecoh atau *distractor* dapat berfungsi dengan baik. Pengecoh yang baik ditandai dengan dipilih oleh sedikitnya 5% dari peserta tes. Sebuah pengecoh atau *distractor* dapat dikatakan berfungsi dengan baik apabila pengecoh tersebut mempunyai daya tarik yang besar bagi peserta tes yang kurang memahami konsep atau kurang menguasai materi.

f. Kualitas Soal

Setelah dianalisis menurut masing-masing kriteria, butir-butir soal kemudian dianalisis secara keseluruhan berdasarkan kriteria validitas, reliabilitas, tingkat kesukaran, daya pembeda dan

efektivitas pengecoh/*distractor* untuk menentukan kualitas soal yang digunakan dalam evaluasi. Penentuan kualitas soal antara soal yang berkualitas baik, cukup baik dan tidak baik didasarkan pada beberapa pertimbangan sebagai berikut:

- 1) Butir soal dikatakan memiliki kualitas yang baik, apabila soal tersebut memenuhi empat kriteria yaitu validitas, tingkat kesukaran, daya pembeda dan efektivitas pengecoh/*distractor*.
- 2) Butir soal dikatakan memiliki kualitas yang cukup baik, apabila soal tersebut hanya memenuhi tiga dari empat kriteria.
- 3) Butir soal dikatakan memiliki kualitas yang tidak baik, apabila soal tersebut tidak memenuhi dua atau lebih kriteria butir soal yang baik.

