

ANALISIS BUTIR TES FORMATIF SOAL PILIHAN GANDA MATA PELAJARAN AQIDAH AKHLAK KELAS VII DI SMP MUHAMMADIYAH 3 YOGYAKARTA

Sri Devi Monica ; Nurul Aisyah, S.Pd.I., M.Pd.

*Universitas Muhammadiyah Yogyakarta, Fakultas Agama Islam, Program Studi Pendidikan
Agama Islam*

E-mail : sridevimonica2122@gmail.com

ABSTRAK

Penelitian ini bertujuan untuk mengetahui kualitas butir Tes Formatif soal Pilihan Ganda Mata Pelajaran Aqidah Akhlak Kelas VII di SMP Muhammadiyah 3 Yogyakarta tahun ajaran 2019/2020 yang ditinjau dari validitas, reliabilitas, tingkat kesukaran, daya pembeda dan efektifitas pengecoh. Penelitian ini merupakan penelitian deskriptif kuantitatif. Objek dalam penelitian ini adalah hasil ujian Tengah Semester peserta didik kelas VII A dan E SMP Muhammadiyah 3 Yogyakarta. Teknik Pengumpulan data dilakukan dengan metode dokumentasi. Data yang diperoleh dianalisis menggunakan program *Anbuso 8.0* dan perhitungan manual menggunakan *Ms.Excel*.

Hasil dari penelitian menunjukkan bahwa kualitas soal Aqidah Akhlak: (1) Segi Validitas, 26 soal atau 65% kategori valid dan 14 soal atau 35% kategori tidak valid; (2) Reliabilitas soal menunjukkan dari 40 butir soal termasuk reliabel koefisien tingkat tinggi, karena dari analisis menghasilkan angka 0,754; (3) Indeks kesukaran 25 butir soal sekitar 62,5% kategori mudah, 12 soal sekitar 30% kategori sedang, 3 soal sekitar 7,5% kategori sukar; (4) Daya Pembeda 17 butir soal sekitar 42,5% kategori tidak baik, 9 soal sekitar 22,5% kategori cukup baik, 14 soal sekitar 35% kategori baik; (5) Efektifitas pengecoh terdapat 10 butir soal sekitar 25% kategori sangat baik, 9 butir soal sekitar 22,5% kategori baik, 17 soal sekitar 42,5% cukup baik dan 4 soal sekitar 10% tidak baik dan harus direvisi ataupun ditingkatkan.

Kata kunci: Analisis Butir Soal, Tes Formatif, Aqidah Akhlak

PENDAHULUAN

Evaluasi merupakan hal yang perlu adanya dilaksanakan ataupun dilakukan mengingat pentingnya pendidikan dengan adanya evaluasi. Evaluasi memiliki berupa rangkaian dari sistem manajemen yaitu perencanaan, pelaksanaan, monitoring, serta evaluasi. Evaluasi merupakan hal yang harus ada dalam sebuah sistem manajemen. Tanpa adanya evaluasi maka akan terasa sulit diketahui sejauh mana tingkat keberhasilan ataupun tujuan dari manajemen tersebut telah tercapai. Kegiatan evaluasi dalam proses pembelajaran sangat erat kaitannya dengan bagaimana cara mengetahui segala tujuan atau yang telah dicapai peserta didik. Langkah evaluasi sangat jelas dan diatur sesuai undang-undang Republik Indonesia No. 20 Tahun 2003 tentang Sistem Pendidikan Nasional BAB XVI pasal 58 ayat 1 yang berbunyi :

“Evaluasi hasil peserta didik dilakukan untuk memantau proses, kemajuan, dan perbaikan hasil belajar peserta didik secara berkesinambungan”.

Dari pasal tersebut, bisa ditarik kesimpulan bahwa salah satu tujuan evaluasi hasil pembelajaran atau belajar yaitu memberikan nilai dari keberhasilan peserta didik serta bisa memperbaiki dari proses pembelajaran yang telah dilaksanakan. Penilaian merupakan bagian paling mendasar untuk mengukur ataupun memperbaiki suatu proses dalam pembelajaran secara keseluruhan yang telah dilaksanakan maupun secara berkesinambungan. Suatu upaya untuk meningkatkan kualitas dari belajar mengajar yaitu adanya penilaian dari hasil belajar. Tes merupakan salah satu yang digunakan untuk mengukur kemampuan siswa yang didalamnya berupa pertanyaan-pertanyaan.

Bagi seorang guru, mengevaluasi dalam pembelajaran adalah kegiatan yang wajib dilakukan, dengan mengadakan kegiatan evaluasi, maka akan memudahkan guru dalam mengamati kemampuan siswa. Pemberian tes terhadap siswa merupakan salah satu cara seorang pendidik mengadakan kegiatan evaluasi hasil belajar.

Realita dilapangan masih banyak pendidik yang belum melakukan analisis terhadap item soal yang akan dijadikan sebagai instrument evaluasi belajar peserta didik, adapun kendala seorang pendidik belum melakukan analisis butir soal adalah karena proses analisis tersebut memakan waktu yang lama, dalam pelaksanaannya ada beberapa tahapan yang harus dilakukan, dalam menganalisis item tes pendidik harus memiliki

kemampuan teknik analisis butir soal sedangkan belum semua guru mengetahui teknik analisis tersebut (Oktanin dan Sukirno, 2015:37).

Tidak adanya analisis item tes maka berdampak pada kualitas soal yang belum diketahui, adanya ketidak jelasan kualitas soal tersebut maka akan menimbulkan informasi yang salah terkait dengan hasil evaluasi belajar siswa sehingga pendidik tidak bisa mengukur kemampuan dan hasil belajar peserta didik.

MANFAAT PENELITIAN

Adapun manfaat diadakannya penelitian ini secara teoritis yaitu memberikan sumbangan terutama dalam dunia pendidikan terkhusus dalam evaluasi belajar mengajar. Manfaat bagi guru adalah hasil penelitian ini bisa digunakan ataupun dipakai sebagai alat ukur untuk mengetahui hasil belajar peserta didik dan sebagai bahan masukan supaya soal yang dibuat oleh guru semakin berkualitas. Adapun kebermanfaatannya bagi pihak sekolah yaitu agar bisa menjadi motivasi lembaga pendidikan untuk mampu mengadakan pelatihan kepada guru dalam pembuatan dan analisis instrumen evaluasi yang baik.

TUJUAN PENELITIAN

1. Mengetahui Validitas dan Reliabilitas Tes Formatif soal pilihan ganda mata pelajaran Aqidah Akhlak kelas VII di SMP Muhammadiyah 3 Yogyakarta tahun ajaran 2019/2020.
2. Menganalisis Indeks Tingkat Kesukaran, Daya Pembeda dan Efektifitas Pengecoh/*Distractor* Tes formatif soal pilihan ganda Mata Pelajaran Aqidah Akhlak Kelas VII di SMP Muhammadiyah 3 Yogyakarta Tahun Ajaran 2019/2020.

LANDASAN TEORI

Evaluasi meliputi pengukuran dan penilaian. Pengukuran berkaitan dengan ukuran kuantitatif, sedangkan penilaian terkait dengan kualitas. Perencanaan evaluasi pembelajaran berarti persiapan atau pemikiran guru tentang pengukuran dan penilaian proses dan hasil kegiatan belajar mengajar (Divayana dan Sugiharni, 2017:161).

Evaluasi didasarkan memiliki tujuan untuk mengetahui tingkat keberhasilan suatu program. Dalam kaitannya Anas menyatakan bahwa tujuan evaluasi dapat dikelompokkan menjadi dua, yaitu : (1) untuk memperoleh data yang mendukung tingkat ketercapaian kompetensi dan tingkat keberhasilan siswa dalam pencapaian tujuan-tujuan kurikuler setelah mereka mengikuti proses pembelajaran dalam jangka waktu tertentu, dan (2)

untuk mengetahui tingkat efektivitas metode - metode pengajaran yang telah digunakan oleh pengajar.

Daryanto (2008: 28) stated that evaluation techniques are divided into two, test and non-test. Non-test techniques are techniques for evaluating student learning outcomes without testing students, but observing, interviewing, questionnaires and checking documentation (Aisyah, 2019:495).

Analisis butir soal merupakan kegiatan yang berfungsi untuk mengetahui tingkat kualitas item soal yang diujikan, serta untuk meningkatkan mutu dan kualitas soal yang telah dibuat. Nana Sudjana mendefinisikan “analisis item soal merupakan sebuah kegiatan mengkaji pertanyaan-pertanyaan yang ada dalam soal tes agar diperoleh soal dengan kualitas baik”. Analisis butir soal dapat dilakukan dengan dua cara, yaitu tanpa uji coba soal dan uji coba soal. Instrument soal dengan uji coba dianalisis dari aspek substansi, kontraksi, dan bahasa. Instrument soal dengan uji coba soal dapat dianalisis dari aspek validitas, reliabilitas, daya beda, tingkat kesukaran dan fungsi *distractor* untuk pilihan ganda dan lainnya (Latip, 2018:166).

a. Validitas

Validitas soal akan menginformasikan kepada pembuat soal tentang ketelitian dan akurasi dalam instrument berupa tes berperan seperti perangkat untuk mengukur perolehan belajar peserta didik. Instrument tes akan dikatakan valid jika tes itu mampu menghitung objek yang diukur dan sesuai dengan patokan serta instrument tersebut dapat menjalankan fungsi ukurnya dan memberikan hasil pengukuran yang sesuai dengan maksud dilakukannya pengukuran tersebut. Tes yang memiliki validitas tinggi maka akan menghasilkan data yang relevan dengan tujuan pengukuran sebaliknya tes yang memiliki validitas rendah maka akan menghasilkan data yang tidak sesuai dengan tujuan. Ada dua unsur penting dalam validitas. Pertama, validitas menunjukkan suatu derajat, ada yang sempurna, ada yang sedang, dan ada pula yang rendah. Kedua, validitas selalu dihubungkan dengan suatu keputusan atau tujuan yang spesifik (Arifin, 2016:247).

b. Reliabilitas

Reliabilitas adalah tingkat atau derajat konsistensi dari suatu instrumen. Isjoni mengungkapkan bahwa suatu soal dikatakan reliabilitas ketika soal dapat dipercayai, konsisten, stabil, serta produktif. Dikatakan reliabelnya suatu tes ketika tes diujikan dimana

saja dan waktu atau kesempatan berbeda hasilnya akan sama. KR 20 merupakan sama halnya dengan rumus r_{11} Untuk mencari reliabilitas tes bentuk objektif bisa dilakukan dengan menggunakan rumus:

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

Keterangan :

r_{11} = reliable tes keseluruhan

p = proporsi subjek yang menjawab item benar

q = proporsi subjek yang menjawab item dengan salah

$\sum pq$ = jumlah perkalian dari p dan q

n = banyaknya item tes

S = standar deviasi tes

Reliabilitas dikatakan juga tes dengan tingkat kesamaan konsisten. Tes yang reliabel biasanya koefisien reliabilitas tinggi dan kesalahan baku pengukurannya (*standard error of measurement*) rendah. Tes dikatakan mempunyai tafaat kepercayaan yang tinggi ketika tes tersebut jika diujikan kembali akan menghasilkan hasil yang sama. Adapun kriteria instrument reliabilitas yakni sebagai berikut: $0,80 < r_{11} \leq 1,00$ = Sangat Tinggi, $0,60 < r_{11} \leq 0,80$ = Tinggi, $0,40 < r_{11} \leq 0,60$ = Sedang, $0,20 < r_{11} \leq 0,40$ = Rendah (Pasi dan Yusrizal, 2018 : 197).

c. Tingkat Kesukaran

Analisis tingkat kesukaran soal adalah mengkaji soal-soal dari segi kesulitannya sehingga dapat diperoleh soal-soal mana yang termasuk mudah, sedang, dan sukar. Untuk soal berupa tes Objektif dapat dilakukan dengan menggunakan rumus sebagai berikut :

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

Keterangan:

P = Angka indeks kesukaran item

B = Banyaknya siswa/peserta tes menjawab dengan benar

JS = Jumlah seluruh siswa peserta tes

Menurut Witherington dalam Anas Sudijono (2011:371) dan Arikunto (2009:207) angka indeks kesukaran butir itu besarnya antara 0,00 sampai dengan

1,00. Sehingga jika angka indeks semakin besar maka soal semakin mudah (Amalia & Widayati, 2012 : 8). Hal ini mengisyaratkan jika seluruh peserta ujian menjawab dengan salah maka soal dikatakan sukar dengan angka kesukaran 0,00 dan jika angka kesukaran 1,00 berlaku kebalikan atau soal sangat mudah karena dijawab dengan benar oleh semua peserta. Menurut Mudjijo (1995: 62) tes yang baik memiliki distribusi secara normal atau memiliki proporsi butir soal dengan tingkat kesukaran seimbang. Mengingat distribusi normal ini, maka dapat dijadikan pedoman bahwa proporsi tingkat kesukaran soal yang mudah, sedang dan sukar masing – masing 30%,50% dan 20%.

d. Daya Pembeda

Menurut Anas Sudijono, daya pembeda item adalah kemampuan suatu butir item tes hasil belajar untuk dapat membedakan antara *testee* yang berkemampuan tinggi dengan *testee* yang berkemampuan rendah. 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.

Menurut Daryanto, Daya Pembeda soal adalah kemampuan suatu soal untuk membedakan antara siswa yang pandai (berkemampuan tinggi) dengan siswa yang bodoh (berkemampuan rendah). Klasifikasi daya pembeda, yaitu D : 0,00 - 0,20 kategori jelek (poor), D : 0,20 - 0,40 kategori cukup (satisfactory), D : 0,40 - 0,70 kategori baik (good), D : 0,70 - 1,00 kategori baik sekali (excellent), D : negatif, semuanya tidak baik. Jadi semua butir soal yang menghasilkan nilai D negatif sebaiknya dibuang saja (Srika dan Yusrizal, 2018 : 197).

e. Efektifitas pengecoh/*Distractor*

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.

TINJAUAN PUSTAKA

Pertama, Penelitian yang dilakukan oleh Dwi Cahya Sulistyadi dengan judul “Analisis Kualitas Butir Soal Pilihan Ganda Pada Ulangan Akhir Semester Mata Pelajaran Penjasorkes Kelas X SMKN 1 Ngasem Kabupaten Kediri Tahun Ajaran 2017/2018” dan hasil dari penelitian tersebut menghasilkan indeks tingkat kesukaran soal kategori sulit mencapai 10% dari 4 butir, tingkat sedang 32,5% 13 butir soal dan mudah 57,5% beraral dari 23 butir. Indeks dayaabeda menunjukkan yang layak dengan angka 62,5% atau 25 butir, kemudian tidak layak 37,5% (15 butir). Suatu efektivitas pengecoh yang layak mencapai 2,5% (1 butir), direvisi 97,5% (39 butir) dan dibuang 0%. Hasil data bisa disimpulkan bahwa kualitas soal masih kurang baik dan perlunya direvisi.

Kedua, penelitian yang dilakukan oleh Kusnani, Nuri Dewi Muldayanti, Hanum Mukti Rahayu dengan judul “Analisis Butir Soal Ulangan Akhir Semester Ganjil Pada mata pelajaran Biologi Kelas X Miasmanegeri 1 Sungai Rayatahun Ajaran 2014/2015” dan hasil penelitiannya dari segi validitas item yaitu soal ulangan akhir semester ganjil kelas X MIA yang valid sebanyak 15 soal (30%), reliabilitas yaitu 0,69. Tingkat kesukaran item menunjukkan soal yang mudah 6 soal (12%), sedang 41 soal (82%) dan sukar 3 soal (6%). Daya pembeda item menunjukkan soal yang jelek 17 soal (34%), sedang 23 soal (46%) dan baik 10 soal (20%). Soal ulangan akhir semester ganjil dapat disimpulkan yaitu memiliki validitas yang rendah, reliabilitas cukup, tingkat kesukaran baik dan daya pembeda baik (Saputri, Dewi dan Setiadi, 2016 : 42).

Ketiga, penelitian yang dilakukan oleh Insar Damopolli dengan judul penelitian “Analisis Butir Soal Ujian Tengah Semester Mata Kuliah Biologi Umum Di Universitas Papua” Hasil analisis menunjukkan bahwa kualitas butir soal ujian tengah semester Biologi Umum di Universitas Papua memiliki tingkat kesukaran cukup/sedang, daya pembeda cukup, kualitas pengecoh sangat baik, terdapat 70 butir soal yang valid dan reliabilitas 0,79 dengan kategori tinggi. Secara keseluruhan butir soal yang diterima sebesar 7 %, butir soal yang direvisi sebesar 63 % dan butir soal yang tidak digunakan sebesar 30 %. Dari hasil diatas bisa disimpulkan bahwa butir soal ujian tengah semester

mata kuliah Biologi Umum di Universitas Papua perlu dilakukan revisi untuk memperbaiki kualitas butir soal (Damopolli, 2016 : 145).

METODE PENELITIAN

Penelitian ini tergolong dalam penelitian deskriptif kuantitatif dengan menggunakan pendekatan kuantitatif karena data yang diperoleh di lapangan adalah dalam bentuk angka-angka dan dianalisis dengan statistic menggunakan program *ANBUSO Version 8.0*. 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.

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.

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.

HASIL DAN PEMBAHASAN

Hasil serta pembahasan dalam penelitian ini yaitu

(1) Validitas,

Penganalisan validitas menghasilkan bahwa butir soal Tes Fomatif mata pelajaran Aqidah Akhlak pada penelitian ini untuk taraf signifikan 5% hanya terdapat 26 butir soal yang valid dan 14 butir soal yang tidak valid. Jika dipersentasekan 65% dari 40 soal yang valid dan 35% dari 40 soal Aqidah akhlak yang tidak valid. Ini

menunjukkan bahwa 65% butir soalnya yang layak diujikan kepada peserta didik dan 35% butir soal yang tidak layak diujikan.

Hasil analisis inipun diperkuat dengan adanya teori yang disampaikan oleh Anas Sudijono bahwa pada tiap butir soal yang memiliki validitas yang tinggi mencerminkan soal tersebut telah memiliki kehandalan dan tidak perlu diragukan ketepatannya dalam mengukur kemampuan peserta didik. Dalam hal ini untuk soal yang memiliki validitas yang rendah mencerminkan soal tersebut tidak valid sehingga soal tersebut perlu dilakukan tindakan.

Melihat dari hasil penelitian di atas, tingkat validitas tes standar Analisis butir tes Formatif soal pilihan Ganda mata pelajaran Aqidah Akhlak kelas VII (tujuh) SMP Muhammadiyah 3 Yogyakarta tahun ajaran 2019/2020 tergolong baik terhadap validitasnya. Hal tersebut dilihat dari persentase butir soal lebih tinggi soal yang valid dibandingkan soal yang tidak valid. Adapun butir soal pada tingkat tidak layak ataupun tidak valid alangkah baik jika diperbaiki dengan cara meningkatkan pada bagian penguasaan teknis tentang penyusunan tiap butir soal.

(2) Reliabilitas

Analisis soal secara keseluruhan dapat dilihat berdasarkan besar atau kecilnya indeks reliabilitas soal. Reliabilitas soal dihitung dengan menggunakan rumus KR-20. Interpretasi koefisien reliabilitas adalah $r_{11} \geq 0,70$ maka soal yang diujikan memiliki reliabilitas tinggi, tetapi apabila $r_{11} < 0,70$ maka soal yang diujika memiliki reliabilitas yang rendah atau tidak reliabel. Hasil penelitian menunjukkan bahwa butir tes Formatif soal pilihan Ganda mata pelajaran Aqidah Akhlak kelas VII SMP Muhammadiyah 3 Yogyakarta tahun ajaran 2019/2020 menunjukkan bahwa butir soal dinyatakan reliabilitas tinggi karena hasil menunjukkan angka di atas 0,7 yaitu 0,752 hal ini telah diuji dengan rumus r_{11} . Hasil penelitian ini didukung dengan adanya teori dari Anas Sudijono bahwa apabila Sebuah tes hasil belajar dikatakan reliable apabila $r_{11} \geq 0,70$ maka soal yang diujikan memiliki reliabilitas tinggi, tetapi apabila $r_{11} < 0,70$

(3) Tingkat Kesukaran,

Berdasarkan hasil tingkat kesukaran butir tes formatif soal pilihan ganda yang dianalisis menggunakan program *Anbuso 8.0* terhadap 40 butir soal yang

dianalisis, dapat disimpulkan bahwa sebanyak 25 butir soal termasuk dalam kategori mudah, yaitu butir item nomor 1, 2, 3, 4, 5, 8, 9, 10, 15, 16, 17, 18, 19, 20, 21, 22, 23, 28, 29, 31, 33, 34, 35, 39, dan 40 yaitu sekitar 62,5% termasuk tingkat kesukaran mudah. Dan 12 butir soal termasuk dalam kategori sedang atau cukup (tidak terlalu mudah dan tidak terlalu sukar) yaitu butir item nomor 6, 7, 11, 12, 13, 24, 26, 27, 30, 32, 36, dan 37 atau sekitar 30% termasuk tingkat kategori sedang atau cukup. Dan terdapat 3 butir soal termasuk dalam kategori sulit atau sukar yaitu butir item nomor 14, 25, dan 38 atau sekitar 7,5% termasuk tingkat kategori sulit atau sukar. Berdasarkan teori tingkat kesukaran soal butir item tes hasil belajar dapat dikatakan sebagai butir item tes yang baik apabila butir-butir tes tersebut tidak terlalu sukar dan tidak pula terlalu mudah.

Witherington dan Anas Sudijono (2011:371) serta Arikunto (2009:207) berpendapat angka indeks kesukaran butir itu besarnya berkisar antara 0,00 sampai 1,00. Semakin besar angka indeks kesukaran maka soal semakin mudah. Butir soal yang terlalu mudah tidak akan merangsang peserta didik untuk berfikir. Namun sebaliknya jika butir soal terlalu sukar atau sulit bisa menyebabkan peserta didik menjadi putus asa sehingga tidak mempunyai semangat untuk mencoba kembali karena soal yang baik itu memiliki indeks antara 30%-70%. Butir soal yang memiliki indeks dibawah 30% dikatakan sukar. Butir soal yang memiliki indeks 30%-70% dikatakan sedang, sedangkan butir soal yang memiliki indeks diatas 70% dikatakan terlalu mudah.

Dari hasil penelitian tingkat kesukaran soal menunjukkan persentase dengan perbandingan 62,5% : 30% : 7,5% atau mudah : sedang : sukar. Namun, hal ini berbeda dengan teori yang sampaikan dari Mudjijo (1995:62) ia berpendapat bahwa proporsi tingkat kesukaran soal mudah, sedang, sukar masing-masing 30%, 50% dan 20%. Bisa disimpulkan bahwa analisis butir tes formatif soal pilihan ganda mata pelajaran Aqidah Akhlak di smp Muhammadiyah 3 Yogyakarta tahun ajaran 2019/2020 kategori kesukaran soal belum bisa dikatakan Baik. Karena, dari hasil menunjukkan bahwa tidak seimbang antara soal yang mudah dan sukar.

(4) Daya Pembeda,

Daya Pembeda merupakan kemampuan suatu butir item tes hasil belajar untuk dapat membedakan antara peserta didik yang memiliki kemampuan tinggi dengan peserta didik yang memiliki kemampuan rendah. Daya pembeda item dapat diketahui dengan melihat besar kecilnya angka indeks diskriminasi item. Berdasarkan analisis daya pembeda terhadap 40 butir soal pilihan ganda, maka dapat disimpulkan bahwa gambaran daya pembeda terdiri dari 17 butir soal memiliki indeks daya pembeda dalam kategori sangat jelek (tidak baik), yaitu butir item nomor 4, 5, 8, 9, 10, 14, 15, 16, 17, 19, 20, 22, 31, 32, 38, 39 dan 40 atau sekitar 42,5%, 9 butir soal memiliki indeks daya pembeda dalam kategori cukup baik, yaitu item soal nomor 1, 3, 11, 12, 18, 24, 25, 35, dan 37 atau sekitar 22,5%, 14 butir soal memiliki indeks daya pembeda dalam kategori baik, yaitu item soal nomor 2, 6, 7, 13, 21, 23, 26, 27, 28, 29, 30, 33, 34, dan 36 atau sekitar 35%.

Dari hasil analisis butir tes formatif soal pilihan ganda dalam kategori daya pembeda menunjukkan bahwa nilai pesentase akhir yaitu 42,5% termasuk dalam kategori tidak baik (sangat jelek), 22,5% dalam kategori cukup baik dan 35 % butir soal dalam kategori baik. Kesimpulan dari analisis butir tes formatif soal pilihan ganda mata pelajaran Aqidah Akhlak kelas VII di smp Muhammadiyah 3 Yogyakarta tahun ajaran 2019/2020 menghasilkan daya pembeda yang positif.

Hal inipun dipertegas teori Daryanto dalam Srika dan Yusrizal (2018:197) bahwa jika jumlah kelompok atas lebih banyak dibandingkan kelompok bawah, maka ini menunjukkan bahwa daya pembeda yang positif karena item bisa membedakan antara kelompok bawah dan kelompok atas.

(5) Efektifitas pengecoh/*distractor*

Sebuah pengecoh atau distractor dapat dikatakan berfungsi dengan baik apabila distractor tersebut mempunyai daya tarik yang besar bagi peserta didik yang kurang memahami konsep atau kurang menguasai bahan. Suatu distractor dapat dikatakan berfungsi dengan baik jika paling sedikit dipilih oleh 5% pengikut tes.

Berdasarkan analisis efektifitas pengecoh dari 40 butir soal terdapat hasil yang beragam dari masing-masing butir soal memiliki indeks tersendiri misalnya dari

4 butir soal memiliki indeks *distractor* termasuk dalam kategori sangat jelek (tidak baik), yaitu butir item nomor 1, 2, 18, dan 37 atau sekitar 10%, 17 butir soal memiliki indeks *distractor* termasuk dalam kategori jelek (cukup baik), yaitu item soal nomor 4, 5, 8, 9, 10, 14, 15, 16, 17, 19, 20, 22, 31, 32, 38, 39, dan 40 atau sekitar 42,5%, 9 butir soal memiliki indeks *distractor* termasuk dalam kategori baik, yaitu item soal nomor 3, 21, 23, 25, 28, 29, 33, 34, dan 35 atau sekitar 22,5%, 10 butir soal memiliki indeks *distractor* termasuk dalam kategori sangat baik, yaitu item soal nomor 6, 7, 11, 12, 13, 24, 26, 27, 30, dan 36 atau sekitar 25%.

Hasil penelitian ini diperkuat dengan Menurut Anas Sudjana, *distractor* dapat berjalan fungsinya dengan baik jika *distractor* tersebut telah dipilih sekurang-kurangnya 5% dari seluruh peserta tes atau peserta didik.

Dari data di atas bisa disimpulkan bahwa analisis butir tes formatif soal pilihan ganda kelas VII di Smp Muhammadiyah 3 Yogyakarta tahun ajaran 2019/2020 menunjukkan bahwa besarnya butir soal dengan kualitas *distractor* yang kurang baik mengidentifikasi bahwa pengecoh belum dapat berfungsi dengan baik, pengecoh terlalu mencolok dan cenderung heterogen. Sehingga efektifitas pengecoh pada butir tes perlu adanya revisi.

DAFTAR PUSTAKA

- Aisyah, N. (2019). *An Analysis of Summative Test on the 8 th Grade of Muhammadiyah Junior High School of Yogyakarta*. 353(IcoSIHESS)
- Amalia, A. N., & Widayati, A. (2012). Analisis Butir Soal Tes Kendali Mutu Kelas XII Sma Mata Pelajaran Ekonomi Akuntansi Di Kota Yogyakarta Tahun 2012. *Jurnal Pendidikan Akuntansi Indonesia*. <https://doi.org/10.21831/jpai.v10i1.919>
- Arifin, Z. (2016). *Evaluasi Pembelajaran*. Bandung: PT Remaja Rosdakarya.
- Asep Ediana Latip, M. (2018). *Evaluasi Pembelajaran di SD dan MI*. Bandung: PT Remaja Rosdakarya.
- Damopoli, I. (2016). Analisis Butir Soal Ujian Tengah Semester Mata Kuliah Biologi Umum Di Universitas Papua. *Jurnal Pendidikan Vol.2 No.1*.
- Divayana, D. G. H., & Sugiharni, G. A. D. (2017). Evaluasi Program Sertifikasi Komputer Pada Universitas Teknologi Indonesia Menggunakan Model Cse-Ucla. *JPI (Jurnal Pendidikan Indonesia)*, 5(2), 865. <https://doi.org/10.23887/jpi-undiksha.v5i2.8586>

- Oktanin, W. S., & Sukirno, S. (2015). Analisis Butir Soal Ujian Akhir Mata Pelajaran Ekonomi Akuntansi. *Jurnal Pendidikan Akuntansi Indonesia*, 13(1). <https://doi.org/10.21831/jpai.v13i1.5183>
- Saputri, L. A., Dewi, N. M., & Setiadi, A. eka. (2016). Jurnal Biologi Education, Vol. 3, No.2, Agustus 2016 53. *Jurnal Biologi Education*, 3(2), 53–62.
- Yusrizal, Srika. *Analisis Butir Soal Ujian Bahasa Indonesia Buatan Guru Mtsn Di Kabupaten Aceh Besar*195. (4), 195–202.