

JURNAL ILMU PENDIDIKAN DAN EKONOMI

Halaman Jurnal: <https://journal.staidenpasar.ac.id/index.php/wb>

Halaman Utama Jurnal : <https://journal.staidenpasar.ac.id/index.php>

ANALISIS BUTIR SOAL PENILAIAN AKHIR SEMESTER GENAP MATA PELAJARAN PENDIDIKAN AGAMA ISLAM DI SMP BINTANG PERSADA TABANAN-BALI

Rusmayani

Sekolah Tinggi Agama Islam Denpasar Bali

e-mail: rusmayani88@yahoo.co.id

Abstrak

Penelitian evaluatif ini bertujuan untuk mengetahui validitas, reliabilitas, daya beda, dan taraf kesukaran pada soal Penilaian Akhir Semester Genap dengan bentuk soal pilihan ganda pada mata pelajaran PAI Kelas VII Semester Genap di SMP Bintang Persada Tabanan-Bali Tahun Pelajaran 2018-2019. Objeknya adalah 40 butir soal bentuk pilihan ganda yang dibuat oleh guru MGMP PAI di Kabupaten Tabanan. Jumlah responden 39 siswa, data didapatkan melalui dokumentasi hasil tes dan wawancara. Analisis data yang digunakan adalah analisis campuran, yakni menyajikan data dalam bentuk numerik selanjutnya diinterpretasikan ke dalam kriteria yang telah ditetapkan. Hasil penelitian diperoleh, bahwa: 1) Dari 40 butir soal, terdapat 15 butir soal yang drop, dan 25 butir soal valid, 2) Nilai koefisien reliabilitas sebesar 0.855 dengan kategori sangat tinggi, 3) Terdapat 11 butir kategori sedang dan 14 butir soal kategori mudah, hal ini menunjukkan bahwa soal buatan guru terbukti belum memenuhi syarat taraf kesukaran butir soal yang proporsional, 4) Terdapat 8 butir soal memiliki indeks daya beda > 0.20 dengan kategori sedang, dan 17 butir soal memiliki indeks daya beda < 0.20 dengan kategori lemah, sehingga dari 25 butir soal, sebagian besar butir tersebut belum mampu membedakan antara siswa yang berada pada kelompok atas dengan siswa yang berada pada kelompok bawah.

Kata Kunci : soal pilihan ganda, validitas, reliabilitas, daya beda, tingkat kesukaran tes

Abstract

This evaluative study aims to determine the validity, reliability, differentiating the strength of questions, the level of difficulty of the questions in the Final Assessment with multiple choice forms in odd semester on PAI subjects VII grade SMP Bintang Persada Tabanan-Bali in 2018-2019 Academic Year. The research object is 40 multiple choice questions submitted by MGMP PAI teachers in Tabanan Regency. The number of respondents was 39 students, data obtained through documentation of test results and interviews. Analysis of the data used is a mixed analysis, which presents data in numerical form then interpreted into predetermined criteria. The results obtained are: 1) Of the 40 items, there were 15 items that went down, and 25 valid items, 2) The reliability coefficient value was 0.855 with a very high category, 3) There were 11 items in the media category and 14 items in the easy category, this shows that the questions made by the teacher proved to not meet the difficulty level of proportional item requirements, 4) There were 8 items that had different power indices > 0.20 in the medium category, and 17 items had different strength index < 0.20 with weak categories, so from the 25 items, most of these items have not been able to distinguish between students in the upper group and students in the lower group.

Keywords: multiple choice, validity, reliability, distinguish power of question, level of difficulty of question

PENDAHULUAN

Dalam kurikulum 2013 terbaru terdapat beberapa perubahan istilah salah satunya yakni perubahan istilah Ujian Akhir Semester menjadi Penilaian Akhir Semester¹. Penilaian Akhir Semester (PAS) dilakukan di semester 1 dan Penilaian Akhir Tahun (PAT) dilakukan di semester 2. Penilaian akhir semester dilakukan untuk mengukur seberapa jauh pemahaman siswa dalam pembelajaran di awal semester, hal ini sesuai dengan pernyataan yang berbunyi bahwa “Penilaian adalah proses pengumpulan dan pengolahan informasi untuk mengukur pencapaian hasil belajar peserta didik”². Penilaian akhir semester dilaksanakan dengan memberikan soal kepada siswa, bentuk soal yang iasa digunakan saat penilaian akhir semester adalah soal dalam bentuk pilihan ganda dan soal dalam bentuk uraian.

Penilaian akhir semester Genap Tahun Pelajaran 2018-2019 untuk mata pelajaran Pendidikan Agama Islam di Sekolah Menengah Pertama Umum di Kabupaten Tabanan, disusun oleh guru PAI yang sebelumnya melaksanakan rapat, dengan seluruh guru PAI dalam MGMP (Musyawarah Guru Mata Pelajaran) di Kota Tabanan untuk menyusun soal tersebut. Sehingga disepakati bahwa bentuk soal untuk penilaian akhir semester adalah sebanyak 40 butir soal bentuk pilihan ganda dan 5 butir soal bentuk uraian. Penggunaan kedua bentuk tes ini bertujuan agar kelemahan tes uraian dapat ditutupi dengan kelebihan dari tes pilihan ganda begitupula sebaliknya, kelemahan dari tes pilihan ganda dapat ditutupi dengan kelebihan dari tes uraian.

Adapun beberapa kelebihan dari tes bentuk pilihan ganda adalah menghasilkan skor yang sama meskipun dinilai oleh orang yang berbeda, jawabannya pendek dan ringkas³, mudah dalam menilai jawabannya, bahan yang dapat dicakup dalam tes lebih luas⁴, dan lebih menghemat waktu, yakni untuk tes yang berlangsung 60 menit dapat diberikan 30 – 40 butir soal⁵. Sedangkan kelemahannya adalah membutuhkan ketelitian dan tingkat kesulitan yang lebih tinggi dalam menyusun butir tes daripada tes esay, kesempatan untuk menebak-nebak jawaban lebih besar, dan peluang kerjasama antar siswa pada waktu mengerjakan soal tes lebih terbuka⁶.

Untuk soal bentuk uraian juga memiliki kelebihan dan kekurangan tersendiri, kelebihanannya yakni meningkatkan daya nalar siswa, melatih siswa untuk dapat memberikan jawaban dengan menggunakan kata-kata sendiri terkait dengan soal yang diujikan sehingga siswa sukar untuk melakukan tebak-tebakan dalam menjawab soal dengan tipe ini. Kelemahannya yakni sulitnya dalam memberikan skor dari jawaban tiap siswa, membutuhkan waktu koreksi yang cukup lama, dan dalam mengkoreksi jawaban siswa banyak dipengaruhi oleh unsur subyektifitas sehingga membutuhkan pertimbangan-pertimbangan pribadi dalam menilai jawaban siswa⁷. Fakta dilapangan menunjukkan bahwa dalam melakukan penilaian hasil belajar guru cenderung menggunakan tipe soal campuran yakni bentuk pilihan ganda dan soal uraian, dalam hal ini untuk mata pelajaran PAI SMP di Kota Tabanan berdasarkan hasil keputusan rapat jumlah soalnya 40 untuk pilihan ganda dan 5 untuk soal uraian, untuk di KKMA (Kelompok Kerja Madrasah Aliyah) untuk Aliyah jumlah soal sebanyak 50 dengan bentuk pilihan ganda semua, sedangkan untuk pelajaran ilmu pasti/eksak jumlah butir soalnya lebih sedikit dibandingkan dengan mata pelajaran umum lainnya yakni 30 soal pilihan ganda dan 5 soal uraian⁸.

⁴ Sudjana, Penilaian Hasil Belajar, Proses Belajar Mengajar, (Bandung: Remaja Rosdakarya, 1995), hal.44

⁵ Arikunto, Suharsimi. Dasar-dasar Evaluasi Pendidikan, (Jakarta: Bumi Aksara, 1997), hal. 166⁶ Ibid, Arikunto hal. 166

⁷ Arikunto, Suharsimi. Dasar-dasar Evaluasi Pendidikan Edisi Revisi, (Jakarta: Bumi Aksara, 1999), hal. 163

⁸ Hasil wawancara dengan Guru PAI SMP Bintang Persada Tabanan, Bp. Haris Nursyah Arifin, S.Si., M.Pd, 13 April 2020

Dalam menyusun soal tes penilaian tengah semester harus berpedoman pada langkah- langkah menyusun soal dengan benar agar benar- mampu benar mengukur tingkat pengetahuan dan pemahaman siswa sesuai dengan tujuan pembelajaran, adapun langkah-langkah penulisan soal dari segi materi dalam bentuk pilihan ganda adalah soal harus sesuai dengan indikator, pengecoh harus berfungsi dan setiap soal harus mempunyai satu jawaban yang benar sedangkan untuk tes uraian soal harus sesuai dengan indikator, setiap pertanyaan harus diberikan batasan jawaban yang diharapkan, materi yang ditanyakan harus sesuai dengan tujuan pengukuran, dan sesuai dengan jenjang dan jenis sekolah atau tingkat kelas⁹.

Mata Pelajaran Pendidikan Agama Islam di sekolah umum di Bali dalam ini Sekolah Menengah Pertama Bintang Persada Tabanan dilaksanakan diluar jam pelajaran yang artinya pelajaran ini tidak termuat dalam jadual pelajaran guru PAI hanya datang ke sekolah disaat jam mengajar saja, sehingga guru PAI tidak menganalisis soal karna sifatnya diluar jam sekolah¹⁰. Oleh karena itu perlu kiranya dilakukan evaluasi hasil penilaian tengah semester melalui butir-butir soal yang telah diberikan kepada siswa, untuk mengetahui apakah soal tersebut telah memenuhi kriteria butir soal yang baik atau belum serta memberikan hasil yang maksimal dalam mengukur tingkat pemahaman dan pengetahuan siswa terhadap materi yang telah diajarkan. Analisis kualitas soal dalam hal ini adalah mengetahui validitas, dan reliabilitas soal yang dilanjutkan dengan uji taraf kesukaran butir, dan daya beda butir.

A. Validitas

Validitas atau kesahihan adalah menunjukkan sejauhmana suatu alat ukur mampu mengukur apa yang hendak diukur (*valid measure if it seccesfully measure the phenomenon*). Dapat didefinisikan bahwa validitas instrument/tes menyangkut ketepatan inferensi yang dibuat peneliti/guru berdasarkan data yang dikumpulkan menggunakan instrument/tes tersebut. Sebagai contoh, tes untuk mgnukur prestasi belajar matematika harus menjamin bahwa tes tersebut dapat menggali data yang dapat digunakan untuk menarik kesimpulan tentang prestasi belajar matematika tersebut¹². Ada beberapa jenis ukuran validitas yang umum digunakan, antara lain validitas isi, validitas kriterium, dan validitas konstruk. Namun dalam hal ini yang hanya dicari sebatas validitas konstruk saja.

Validitas konstruk adalah validitas yang berkaitan dengan kesanggupan suatu alat ukur dalam mengukur pengertian suatu konsep yang diukurnya. Validitas konstruk merupakan yang terluas cakupannya dibandingkan dengan validasi yang lain¹³ karena melibatkan banyak prosedur termasuk validasi isi dan validasi kriteria. Teknik korelasi yang digunakan untuk menentukan validitas butir tes tergantung kepada jenis tes yang akan diuji. Perhitungan validitas tes berupa tes objektif yakni menggunakan teknik korelasi yang berbeda dengan perhitungan validitas instrument berupa tes uraian atau angket¹⁴.

Butir tes objektif sering disebut butir tes dikotomi karena skor tes objektif berupa skala dikotomi yakni 1 (satu) atau 0 (nol). Skor 1 diberikan untuk butir jawaban yang benar ,sedangkan skor 0 (nol) diberikan pada butir jawaban soal yang salah. Validitas butir tes dihitung dengan mengkorelasikan skor butir dengan skor total yang diperoleh responden. Skor butir tes objektif berupa skala dikotomi sedangkan skor totalnya berupa skala interval. Validitas butir soal objektif menggunakan teknik korelasi *point biserial* (γ_{pbi}) dengan rumus:

$$\gamma_{pbi} = \frac{M_p - M_t}{S_t} \sqrt{\frac{p}{q}}$$

Koefisien korelasi dalam perhitungan validitas terdapat pada rentang -1,00 sampai +1,00. Untuk pengujian validitas dengan bantuan SPSS maka kriteria pengujiannya adalah: 1) Jika $\gamma_{pbi} > r_{tabel}$, maka butir tersebut dinyatakan valid, sebaliknya 2) Jika $\gamma_{pbi} \leq r_{tabel}$, maka butir tersebut dinyatakan tidak valid/drop.

B. Reliabilitas

Reliabilitas adalah suatu uji untuk mengetahui sejauh mana hasil pengukuran tetap konsisten, apabila dilakukan pengukuran dua kali atau lebih terhadap gejala dan alat ukur yang samapula¹⁵. Ada beberapa teknik yang dapat digunakan untuk mengukur reliabilitas suatu instrument/tes, tergantung dari skala yang digunakan. Untuk data yang bersifat dikotomi dihitung dengan menggunakan rumus rumus KR 21 dengan asumsi bahwa tingkat kesukaran tes homogen, adapun rumusnya sebagai berikut¹⁶:

$$r_{11} = \left(\frac{n}{n-1} \right) \left(\frac{S_{t-n\bar{p}\bar{q}}^2}{S_t^2} \right)$$

Harga koefisien reliabilitas tersebut dilihat pada rentangan koefisien korelasi pada tabel Guildford antara 0,00-1,00, dimana rentang koefisien korelasi 0.40-0,60 berada pada derajat reliabilitas sedang.

C. Taraf Kesukaran Butir Tes

Instrumen/tes yang diuji coba merupakan prestasi (*achievement test*) berbentuk objektif. Oleh karena itu, setelah validitas dan reliabilitas tes diketahui harus dilanjutkan dengan uji taraf (indeks) kesukaran butir. Umumnya uji taraf kesukaran butir dilakukan dengan memperhitungkan banyak responden yang menjawab butir tersebut dengan benar. Adapun rumus yang digunakan adalah¹⁷:

$$I = \left(\frac{B}{N} \right)$$

Kriteria untuk indeks kesukaran butir tes adalah berada pada rentang 0,00 – 1,00. Untuk butir dengan P 0,00–0,30 berada pada kategori sukar, butir 0.31-0,70 berada pada kategori sedang, dan 0,71-1,00 berada pada kategori mudah. Butir tes dengan indeks kesukaran 0.00 merupakan butir tes yang amat sukar, karena tidak ada satupun peserta tes yang mampu menjawab soal tersebut dengan benar, sebaliknya butir tes yang memiliki indeks 1.00 merupakan butir tes yang amat mudah, karena semua peserta tes mampu menjawab butir tes tersebut dengan benar.

Indeks kesukaran butir tes bermanfaat untuk memilih antara butir tes yang akan dipertahankan dan dipergunakan kembali, atau butir tes yang akan digugurkan. Butir soal yang memiliki rentangan indeks kesukaran 0.30 – 0.70 masih bisa diterima untuk butir tes yang patut dipilih. Indeks kesukaran butir tidak mampu memberikan informasi mengenai perbedaan kemampuan peserta tes, apabila butir tes tidak memenuhi kriteria, maka pilihannya butir tersebut digugurkan atau direvisi.

D. Daya Beda Butir

Analisis daya beda adalah mengkaji butir-butir soal untuk mengetahui kesanggupan butir-butir tes tersebut membedakan siswa yang tergolong mampu dan tidak mampu dalam menguasai suatu materi tertentu. Butir tes yang baik harus mampu membedakan antara peserta tes yang mampu dan peserta tes yang tidak mampu, dengan kata lain, butir tes yang baik seharusnya dijawab benar oleh peserta tes yang mampu dan dijawab salah oleh peserta tes yang tidak mampu. Indeks daya beda butir dinyatakan dengan d dan dihitung dengan rumus sebagai berikut¹⁸:

$$d = \frac{U - L}{N}$$

Terdapat beberapa indikator dalam daya beda butir, yakni secara langsung dan tidak langsung. Secara tidak langsung dengan mencari koefisien korelasi biserial antara skor butir dengan skor total sedangkan secara langsung dengan menggunakan perbandingan kinerja tes pada kelompok atas (*upper region*) dan kelompok bawah (*lower region*) pada skor total.

Kelompok atas dan kelompok bawah ditentukan berdasarkan skor total dengan rentangan 10 sampai 33 persen. Jika skor total berdistribusi normal, maka diambil 27 persen skor tertinggi dan 27 persen skor terendah, sehingga masing-masing akan menjadi kelompok atas dan kelompok bawah. Jika kondisi skor total lebih landai (datar) dari kurva normal, maka persentase yang diambil lebih besar dari 27 persen yakni mendekati atau 33 persen.

Indeks daya beda butir memiliki rentang -

1.0 sampai dengan +1.0. Indeks daya beda negatif menunjukkan bahwa kelompok bawah lebih banyak menjawab butir tes dengan benar daripada kelompok atas, kondisi ini harus dihindari dalam pengukuran hasil belajar. Indeks daya beda yang diharapkan adalah indeks daya beda yang bernilai positif, semakin mendekati +1.0 maka butir tes tersebut memiliki daya beda yang semakin baik¹⁹. Butir soal yang memiliki indeks daya beda sama dengan 0.0 bermakna bahwa peserta tes yang menjawab benar butir tes pada kelompok bawah sama dengan peserta tes yang menjawab benar pada kelompok atas sehingga butir tes tersebut harus direvisi atau digugurkan, karena butir tes tersebut tidak mampu membedakan kemampuan peserta tes.

Adapun kriteria tingkat daya beda tiap butir tes adalah pada rentang 0 – 0.2 berkategori lemah, 0.21 – 0.4 berkategori sedang, 0.41 – 0.70 berkategori baik, 0.71 – 1.00 berkategori sangat kuat, dan bertanda negatif berkategori sangat lemah/jelek.

METODE PENELITIAN

Penelitian ini merupakan penelitian evaluatif. Penelitian evaluatif merupakan rangkaian kegiatan pengumpulan data atau informasi yang kemudian dibandingkan dengan suatu kriteria untuk mengambil kesimpulan²⁰. Kegiatan evaluatif ini dilaksanakan dengan menghitung validitas, reliabilitas, tingkat kesukaran, dan daya beda butir tes buatan guru PAI, kemudian diinterpretasikan ke dalam kriteria yang telah ditetapkan sebelumnya. Setiap analisis dalam penelitian evaluatif ini memiliki tujuan untuk mengetahui seberapa banyak butir tes yang valid, tingkat reliabilitasnya, serta besar kecilnya tingkat kesukaran dan daya beda butir tes pilihan ganda kelas VII yang telah disusun oleh guru PAI di SMP Bintang Persada Tabanan, Bali. Penentuan lokasi penelitian dilakukan secara sengaja yakni dengan pertimbangan-pertimbangan tertentu oleh si peneliti.

Objek dari penelitian ini adalah 40 butir soal pilihan ganda yang disusun oleh guru MGMP PAI di Kabupaten Tabanan. Jumlah responden berjumlah 39 siswa. Teknik pengumpulan data dengan wawancara dan dokumentasi hasil Penilaian Akhir Semester Genap pada mata pelajaran Pendidikan Agama Islam. Analisis data yang digunakan adalah analisis campuran (kualitatif dan kuantitatif) yakni menyajikan data validitas, reliabilitas, taraf kesukaran dan daya beda butir soal dalam bentuk numerik selanjutnya diinterpretasikan ke dalam kriteria yang telah ditetapkan²¹.

HASIL DAN PEMBAHASAN

Hasil Penelitian ini menyajikan dan memaparkan hasil perhitungan validitas, reliabilitas, tingkat kesukaran, dan daya beda butir soal pilihan ganda yang disusun oleh guru MGMP PAI di Kabupaten Tabanan, Bali.

1. Validitas

Validitas butir soal didasarkan pada koefisien korelasi *point biserial* antara skor butir dengan skor total dibandingkan dengan harga r_{tabel} , jika $\gamma_{pbi} > r_{tabel}$ maka skor butir soal tersebut berkorelasi secara signifikan dengan skor total, sehingga butir soal tersebut dinyatakan valid begitupula sebaliknya jika $\gamma_{pbi} \leq r_{tabel}$ maka skor butir soal tersebut tidak berkorelasi secara signifikan dengan skor total, sehingga butir soal tersebut dinyatakan drop/tidak valid.

Dengan jumlah responden (N) sebanyak 39 siswa maka diketahui r_{tabel} sebesar 0.316 pada taraf signifikansi 0.05. Dari hasil perhitungan di dapatkan butir soal yang drop sebanyak 15 butir, sedangkan yang valid sebanyak 25 butir. Berikut rekap butir soal yang valid dan drop.

Tabel 1. Rekap Validitas Butir Soal

No	Status	Butir Soal	Jumlah
1	Valid	25	2,3,6,7,9,10,11, 12,13,14,15,16, 17,19,21,26, 27,28,29,31,34, 35,36,38,40
2	Drop	15	1,4,5,8,18,20,22, 23,24,25,30, 32,33,37,39
Total			40

Data Primer yang Diolah

Dari tabel diatas di dapatkan informasi bahwa sebanyak 25 butir soal atau sebesar 62.5% menyatakan valid, sisanya sebanyak 15 butir soal atau sebesar 37.5% menyatakan butir tersebut tidak valid/drop. Butir soal yang valid sebanyak 25 butir soal selanjutnya dipergunakan untuk menghitung derajat reliabilitas butir tes.

2. Reliabilitas

Reliabilitas tes mengacu pada konsistensi hasil pengukuran yang ditunjukkan oleh tes tersebut. Reliabilitas tes bentuk dikotomi (pilihan ganda), dihitung dengan rumus KR-20, reliabilitas dihitung hanya untuk butir-butir soal yang valid saja (2, 3, 6, 7, 9, 10, 11, 12, 13, 14, 15, 16, 17, 19, 21, 26, 27, 28, 29, 31, 34, 35, 36, 38, dan 40). Adapun hasil perhitungan reliabilitas tes PAS Genap dapat dilihat pada tabel dibawah ini. menunjukkan bahwa reliabilitas soal akan semakin tinggi jika butir soal di dominasi oleh butir soal dengan kategori sedang²².

3. Taraf Kesukaran Butir Tes

Taraf kesukaran butir tes merupakan gambaran apakah butir soal tersebut mudah, sedang atau sulit untuk dikerjakan oleh peserta tes. Taraf kesukaran butir didapatkan melalui pembagian banyak siswa yang menjawab butir tersebut dengan benar dengan jumlah responden. Adapun rekapitulasi taraf kesukaran butir tes dapat dilihat pada tabel dibawah ini.

Tabel 3. Taraf Kesukaran Butir Tes

No	No Item	Jml	Ket
1	2,9,13,14,15, 16,21,27, 29,38,40	11	Sedang
2	3,6,7,10,11, 12,17,19,26, 28,31,34,35, 36	14	Mudah
Total			25

Data Primer yang Diolah

Dari tabel tersebut diperoleh informasi bahwa sebanyak 11 butir soal atau sebesar 44% berada pada kategori sedang, sisanya sebanyak

14 butir soal atau sebesar 56% berada pada kategori mudah. Dengan demikian, butir tes dengan rentangan indeks kesukaran 0.3 – 0.7 atau berkategori sedang masih ditolerir sebagai butir tes yang patut dipilih, atau dengan kata lain butir soal yang baik adalah butir soal yang tidak terlalu sukar dan tidak terlalu mudah.

Tabel 2. Tabel Reliabilitas Tes

Jumlah Butir	KR-20	Kategori
25	0.855	Sangat Tinggi

Data Primer yang Diolah

Dari perhitungan didapatkan nilai reliabilitas tes sebesar 0.855 dengan kriteria derajat reliabilitas sangat tinggi. Beberapa penelitian Berdasarkan tabel didapatkan informasi butir soal yang patut dipilih hanya sebanyak 11 butir dari 25 butir soal atau sebesar 44%, sisanya sebanyak 14 butir soal atau sebesar 56% direvisi atau digugurkan. Salah satu acuan dasar dalam menentukan proporsi jumlah soal kategori sukar, sedang, dan mudah adalah adanya keseimbangan yang didasarkan atas arah kurva normal, hal ini memiliki arti proporsi jumlah soal sukar dan mudah sama²⁴. Perbandingan ideal antara soal yang sukar, sedang, dan mudah adalah 3-4-3, atau 30% soal sukar, 40% soal sedang, dan 30% soal mudah. Jika mengikuti perbandingan ideal tersebut maka sebaiknya dari 25 soal Penilaian Akhir Semester tersebut terdiri dari 7 butir soal sukar, 10 butir soal sedang, dan 8 butir soal mudah.

4. Daya Beda Butir Tes

Berdasarkan tabel kerja reliabilitas yang terdiri dari butir soal yang valid saja selanjutnya mencari kelompok atas dan kelompok bawah dengan mengurutkan skor total tertinggi sampai terendah dari masing-masing peserta, kemudian diambil 33% dari kelompok atas (13 orang dari atas) dan 33% dari kelompok bawah (13 orang dari bawah) sisanya no urut 14-27 tidak digunakan. Kelompok atas dan bawah ditentukan 33% karena distribusi sebaran skornya tidak mengikuti arah distribusi normal, hal ini sesuai dengan pernyataan yang menyebutkan bahwa “apabila distribusi skor total lebih datar (landai) daripada kurva normal, maka persentase yang diambil lebih besar, yakni mendekati 33%”²⁵.

Selanjutnya mencari nilai total masing-masing butir soal dari 13 peserta kelompok atas dan kelompok bawah tersebut. Kemudian membuat tabel kerja kembali dengan urutan no butir, U (nilai total kelompok atas), L (nilai total kelompok bawah), $d = (U-L)/N$, dan keterangan. Berikut disajikan tabel kerja perhitungan daya beda butir soal terbatas pada 10 butir soal.

Tabel 4. Tabel Kerja Perhitungan Daya BedaButir

No Butir	U	L	$d = (U-L)/N$	Keterangan
2	11	2	0.23	Sedang
3	13	8	0.13	Lemah
6	12	8	0.10	Lemah
7	13	9	0.10	Lemah
9	13	5	0.21	Sedang
10	17	9	0.21	Sedang
11	13	5	0.21	Sedang
12	13	7	0.15	Lemah
13	9	1	0.21	Sedang
14	12	6	0.15	Lemah

Data Primer yang Diolah

Hasil rekap daya beda butir dapat dilihat padatabel di bawah ini:

Tabel 5. Rekap Daya Beda Butir Soal

No	No Item	Jml	Ket
1	2,9,10,11,13,15,1 9,27	8	Sedang
2	3,6,7,12,14,16,17 ,21,26,28,29,31,3 4,35,36,38,40	17	Lemah
Total		25	

Data Primer yang Diolah

Dari tabel tersebut menunjukkan bahwadari 25 butir soal, 8 butir soal atau sebesar 32% berada pada kategori sedang, dan 17 butir soal atau sebesar 68% berada pada kategori lemah. Hal ini berarti terdapat 8 butir soal yang berkategori ideal dan masih mampu dalam membedakan kemampuan antara siswa yang memiliki kemampuan tinggi dan siswa yang memiliki kemampuan rendah. Dari kedelapan butir soal tersebut jumlah siswa kelompok atas atau yang memiliki kemampuan tinggi lebih banyak dalam menjawab betul butir soal tersebut daripada siswa yang berada pada kelompok bawah atau siswa yang memiliki kemampuan rendah. Sedangkan sisanya sebesar 17 butir soal memiliki rentangan daya beda kurang dari 0.20 dengan kategori lemah/jelek sehingga butir soal tersebut digugurkan/diganti²⁶. Ketujuh belas butir soal yang berada dalam kategori lemah/jelek tersebut belum mampu membedakan kemampuan antara siswa yang berada pada kelompok atas dengan siswa yang berada pada kelompok bawah, hal ini bisa dilihat dari jumlah siswa kelompok atas yang menjawab betul ketujuh belas butir soal tersebut hampir sama dengan jumlah siswa yang berada pada kelompok bawah dalam menjawab betul pada soal-soal tersebut.

KESIMPULAN DAN SARAN

Dari pembahasan tersebut dapat disimpulkan bahwa dari 40 butir soal, terdapat 15 butir soal yang drop, dan 25 butir soal valid, nilai koefisien reliabilitas sebesar 0.855 dengan kategori sangat tinggi, beberapa penelitian menunjukkan bahwa reliabilitas soal akan semakin tinggi jika butir soal didominasi oleh butir soal dengan kategori sedang, butir soal buatan guru terbukti belum memenuhi syarat taraf kesukaran butir soal yang proporsional. Ditemukan proporsi soal hanya berkisar sedang dan mudah, dengan jumlah 11 butir soal sedang dan 14 butir soal mudah, hal ini belum sesuai dengan proporsi soal yang ideal yakni pertimbangan keseimbangan dan mengikuti arah kurva normal (3-4-3). Selanjutnya dari 25 butir soal, sebagian besar butir tersebut belum mampu membedakan antara siswa yang berada pada kelompok atas dengan siswa yang berada pada kelompok bawah, dari perhitungan didapatkan 8 butir soal memiliki indeks daya beda lebih dari 0.20 dengan kategori sedang, dan 17 butir soal memiliki indeks daya beda kurang dari 0.20 dengan kategori lemah/jelek sehingga butir soal buatan guru tersebut masih banyak yang perlu diperbaiki. Berdasarkan hasil temuan tersebut diatas diharapkan dapat memberikan manfaat kepada pihak-pihak terkait dalam penyusunan butir soal dengan lebih baik lagi sehingga dalam menilai hasil belajar siswa didapatkan hasil yang lebih akurat. Hasil temuan ini juga diharapkan sebagai salah satu rujukan bagi peneliti yang lain dalam mengembangkan penelitian serupa.

DAFTAR REFERENSI

- Anwar, Syafri. 2009. Penilaian Berbasis Kompetensi. Padang: UNP Press.
- Arikunto, Suharsimi. 1997. Dasar-dasar Evaluasi Pendidikan. Jakarta: Bumi Aksara
- _____. 1999. Dasar-dasar Evaluasi Pendidikan (edisi revisi). Jakarta: Bumi Aksara
- Candiasa, I Made. 2011. Pengujian Instrumen Penelitian Disertai Aplikasi ITEM dan BIGSTEPS. Singaraja: Undiksha Press
- Departemen Pendidikan Nasional, Direktorat Jenderal Manajemen Pendidikan Dasar dan Menengah. 2008. Panduan Penulisan Butir Soal. Jakarta: Depdiknas
- Kuswati, Eni. 2017. <https://enikuswati.gurusiana.id/article/2017/7/daftar-perubahan-istilah-dalam-k-13-revisi-2017-1041540>, diakses tanggal 11 April 2020
- Marhaeni, A.A.I.N. 2012. Evaluasi Program Pendidikan. Buku Ajar Mahasiswa Program Pasca Sarjana, Singaraja, Program Pasca Sarjana UNDIKSHA Singaraja
- Nurhasanah, Herlina Ahmad. 2017. Analisis Soal Ujian Tengah Semester Mata Kuliah Geometri. Jurnal Pendidikan PEPATUDZU Media Pendidikan dan Sosial Kemasyarakata Vol. 13, No. 1
- Peraturan Pemerintah RI Nomor 19 Tahun 2005. tentang Standar Nasional Pendidikan. Jakarta: Depdiknas
- Purwanto, M. Ngalim. 1994. Prinsip-prinsip dan Teknik Evaluasi Pengajaran. Bandung: Remaja Rosdakarya
- Sudjana, Nana. 2012. Penilaian Hasil Belajar, Proses Belajar Mengajar. Bandung: Remaja Rosdakarya
- Siregar, Syofian. 2010. Statistika Deskriptif Untuk Penelitian. Jakarta: Raja Grafindo Persad