

Kualitas Soal Penilaian Tengah Semester Mata Pelajaran Bahasa Sunda di SMPN 2 Leuwisadeng Bogor

Vika Destivani¹, Usep Kuswari², Yatun Romdonah²

¹SMA Negeri 1 Cibungbulang, ²Universitas Pendidikan Indonesia
vikadestivani@gmail.com

Sejarah Artikel: Diterima (01 Februari 2020); Diperbaiki (13 Maret 2020); Disetujui (04 April 2020); Published (30 April 2020)

Bagaimana mengutip artikel ini (dalam gaya APA): Destivani, V., Kuswari, U., Romdonah, Y. (2020). Struktur Kalimat Bahasa Sunda dalam Artikel Siswa. *Lokabasa*, 11(1), 73-81. doi: <https://doi.org/10.17509/jlb.v11i1.25200>

Abstrak: Penelitian ini dilatarbelakangi oleh adanya kualitas soal penilaian tengah semester mata pelajaran bahasa Sunda yang dibuat oleh guru masih belum memenuhi kriteria tingkat validitas, reliabilitas, dan kelayakan soal penilaian tengah semester. Penelitian ini bertujuan untuk mendeskripsikan validitas, baik validitas isi maupun validitas empiris (tingkat kesukaran dan daya pembeda), tingkat reliabilitas, dan tingkat kelayakan soal-soal penilaian tengah semester basa Sunda di SMPN 2 Leuwisadeng Bogor. Sumber data dalam penelitian ini didapat dari pengumpulan soal-soal penilaian tengah semester bahasa Sunda dan lembar jawaban siswa semester 1 dan semester 2 taun ajar 2018/2019. Berdasarkan hasil penelitian 11 aspek yang telah dianalisis ini yaitu validitas isi soal penilaian tengah semester dari 110 soal yaitu 44 soal valid, 21 soal harus direvisi, dan 45 soal harus diganti. Validitas empiris, berdasarkan kesukaran soal ada 62 soal yang layak, 40 soal direvisi, dan 8 soal harus diganti. Berdasarkan tingkat daya pembeda ada 49 soal layak, 31 soal direvisi, dan 30 soal harus diganti. Berdasarkan tingkat reliabilitas soal Penilaian Tengah Semester ada 50% kategori tinggi, 30% kategori sedang, dan 20% kategori rendah. Oleh sebab itu, berdasarkan validitas, baik validitas isi maupun validitas empiris (tingkat kesukaran soal dan daya pembeda soal), serta tingkat reliabilitas terdapat kelayakan soal yaitu ada 35 soal yang layak (32%), 50 soal harus direvisi (45%), dan 25 soal harus diganti (23%). Jadi, soal-soal penilaian tengah semester bahasa Sunda di SMPN 2 Leuwisadeng Bogor mempunyai tingkat kelayakan soal yang “cukup”.

Kata Kunci: penilaian tengah semester; validitas; reliabilitas

Mid-Semester Test Questions Quality of Sundanese Language Subject in SMPN 2 Leuwisadeng Bogor

Abstract: This research based on the quality of mid-semester test questions in Sundanese language subject which is proposed by the teacher is still not meeting the criteria of validity, reliability, and questions properness. This research aims to describe the validity level, content validity, and empirical validity (difficulty level and differentiator power), reliability, and questions appropriateness level of Sundanese language subject in SMPN 2 Leuwisadeng Bogor. This research uses descriptive method, specifically, descriptive and inferential statistics. The data for this research are Sundanese subject mid-term questions and the answer sheets of the first and second semester student year 2018/2019. Based on the 11 analyzed aspects, the result shows that mid-semester test question content validity of 110 questions that 44 questions are valid, 21 questions are obliged to be revised, and 45 questions should be changed. Empirical Validity based on difficulty level shows that 62 questions are proper, 40 questions should be revised and 8 questions should be changed. Based on the differentiator power shows that 49 questions are proper, 31 questions should be revised, and 30 questions should be substituted. Based on the reliability level, 50% is in good category, 30% is in middle category, and 20% is in the bad category. Therefore, based on validity, either content validity or empirical validity (questions' difficulty level and differentiator power), as well as reliability level, there are 35 proper questions (32%), 50 questions should be revised (45%), and 25 questions should be changed (23%). The conclusion is mid-semester test questions of Sundanese subject in SMPN 2 Leuwisadeng Bogor has 'good' standard of questions appropriateness.

Keywords: mid-Semester test; validity; reliability

PENDAHULUAN

Evaluasi merupakan proses menilai kemampuan siswa dalam proses belajar-mengajar. Tujuan dasar diadakan evaluasi yaitu untuk mendapatkan informasi hasil pembelajaran. Informasi hasil belajar tersebut dapat dilakukan dengan melakukan pengolahan hasil tes yang baik (Nugraha et al., 2018). Selain itu, Sukardi (2012, hlm. 4) mengatakan evaluasi juga mempunyai macam-macam fungsi dalam proses belajar mengajar yaitu untuk mengetahui aspek-aspek kekurangan siswa dalam melaksanakan kegiatan belajar, tingkat ketercapaian siswa dalam kegiatan belajar, dan sebagai alat untuk mengetahui tercapai atau tidaknya belajar siswa.

Salah satu alat evaluasi yang kerap digunakan yaitu tes, bagaimana yang telah dipaparkan oleh Sudjana (2009, hlm. 35) yaitu tes selaku alat penilaian yang isinya berupa pertanyaan-pertanyaan yang diberikan kepada siswa dalam bentuk lisan (tes lisan), dan dalam bentuk tulisan, yang berguna untuk mengukur hasil belajar siswa, tentunya dari hasil belajar yang sama dengan tujuan pembelajaran.

Tes hasil pembelajaran bahasa Sunda bisa dilaksanakan dalam bentuk tertulis yaitu pilihan ganda ataupun esay. Tes harus memberikan nilai yang objektif serta bisa menjelaskan kemampuan siswa dalam menguasai materi pembelajaran. Menurut Tim Direktorat Pembinaan SMP (2017, hlm. 7) terdapat beberapa tes yang diadakan di sekolah, seperti: penilaian harian, penilaian tengah semester, dan penilaian akhir taun.

Penilaian Tengah Semester (PTS) termasuk jenis evaluasi yang digunakan di sekolah untuk mengukur tercapai tidaknya kompetensi dasar mata pelajaran, setelah pembelajaran berlangsung selama 8-9 minggu dalam Tim Direktorat Pembinaan SMP (2017, hlm. 8). Bagi guru, hasil penilaian tengah semester bisa dipakai untuk mengetahui meningkat atau tidaknya hasil belajar siswa, dan apakah teknik pembelajaran harus dikoreksi atau diganti. Untuk siswa, hasil penilaian tengah semester ber-

fungsi untuk menguasai atau tidaknya materi yang telah diberikan oleh guru.

Soal PTS di sekolah belum tepat dari bentuk penyusunannya. Guru seharusnya guru bisa menyusun soal yang baik, sehingga bisa memberikan solusi kepada siswa untuk meningkatkan pembelajaran dan bisa memberikan informasi tentang kompetensi yang telah tercapai oleh siswa. Jika guru tidak bisa membuat soal yang baik, dapat menjadikan pengetahuan siswa tidak terbahas, sehingga dapat membuat nilai mutu siswa rendah.

Permasalahan yang terdapat dalam penelitian ini adalah soal PTS yang dibuat oleh guru di sekolah belum diuji kualitas soalnya. Seharusnya PTS bisa menampilkan hasil belajar siswa yang baik, sehingga tujuan pembelajaran terlihat telah tercapai atau tidak. Oleh karena itu, untuk memecahkan masalah tersebut dilakukan kajian ini.

Taib (2014, hlm. 117) mengatakan dalam pelaksanaan evaluasi, aspek materi diperlukan untuk mengupas pemahaman siswa yang sama dengan pembelajaran. Untuk mengetahui kualitas soal bisa tercapai dengan cara meneliti pada kaedah nulis soal dari aspek materi yang telah ditentukan.

Cara menulis soal merupakan proses untuk menentukan kualitas butir soal yang baik, jadi harus dilakukan secara baik. Tim Direktorat Pembinaan SMP (2017, hlm. 1) mengatakan untuk menentukan kualitas tes yang baik harus mengembangkan tes melalui beberapa langkah. Langkah-langkah yang dilakukan untuk menyusun tes, menentukan tujuan yang dipakai (kriteria atau norma), menyusun kisi-kisi, memilih soal dari kumpulan soal yang sudah ada dan sama dengan kisi-kisi.

Untuk mendapatkan kualitas soal yang baik, penulis soal harus memperhatikan kaidah-kaidah dengan cara melihat materi, konstruksi dan bahasa kebudayaan. Tim Direktorat Pembinaan SMP (2017, hlm. 13).

Dalam kegiatan evaluasi pasti membutuhkan alat atau teknik evaluasi, sehingga dalam pelaksanaannya bisa lebih jelas. Alat evaluasi PTS tentu akan berbeda dengan

Penilaian Harian dan Penilaian Akhir Semester.

Kualitas alat evaluasi yang baik menurut Sudjana (2009, hlm. 12) yaitu jika alat itu mempunyai dua, yaitu validitas dan reliabilitas.

Validitas dan reliabilitas soal sangat penting dalam mengukur kemampuan siswa, karena soal yang baik yaitu soal yang mempunyai kriteria valid dan reliabel. Purwanti (2014, hlm. 87) mengatakan soal yang dikatakan valid yaitu soal yang bisa mengukur apa yang akan diukur atau alat ukur yang sama dengan indikator serta materi ajarnya. Sedangkan menurut Koni (2018, hlm. 153) reliabel yaitu tes yang bisa menghasilkan data yang sama (ajeg) jika dipakai beberapa kali untuk mengukur objek yang sama.

Menurut Sukardi (2012, hlm. 32) secara *metodologis*, validitas dibagi menjadi empat macam, yaitu validitas isi, validitas konstruk, validitas *konkure* dan validitas prediksi. Keempat validitas, dikelompokkan menjadi dua macam yaitu validitas logis dan validitas empiris.

Walaupun sama dengan penelitian sebelumnya, yaitu membahas tentang validitas dan reliabilitas soal-soal PTS berdasarkan tingkat validitas dan tingkat reliabilitas, tapi penelitian ini ada perbedaannya. Dalam penelitian ini terfokus pada analisis tingkat validitas, tingkat reliabilitas, dan kelayakan soal yang baik untuk soal PTS.

Dalam penelitian ini yang menjadi fokus masalah yaitu bagaimana kualitas soal penilaian tengah semester dan lembar jawaban siswa melalui tingkat validitas, tingkat reliabilitas dan tingkat kelayakan soal penilaian tengah semester yang dibuat oleh guru. Kualitas soal penilaian tengah semester yang baik seharusnya berdasarkan tingkat validitas yang akan diuji yaitu validitas isi, dalam penelitian ini yaitu soal dianalisis berdasarkan pada jelas tidaknya petunjuk soal, jelas tidaknya rumusan indikator, terkait antara jenis penilaian dan indikator, jelas tidaknya instrumen dengan materi yang diujikan, jelas tidaknya bahan penilaian dengan alat penilaian yang digunakan, bahasa yang digunakan, rumusan bahasa komuni-

katif, menggunakan kata dan kalimat yang mudah dipahami, huruf yang digunakan, ukuran huruf yang digunakan, dan format atau layout. Sedangkan validitas empiris yaitu apakah soal yang dipakai mudah atau sulit, serta bagaimana daya pembeda soal. Bagaimana reliabilitas soal yaitu apakah soal sama atau tidak. Serta bagaimana kelayakan soal PTS.

Berdasarkan pembahasan di atas, tulisan ini mempunyai tujuan untuk mendeskripsikan validitas, validitas isi maupun validitas empiris, tingkat reliabilitas dan kelayakan soal-soal PTS bahasa Sunda di SMPN 2 Leuwisadeng Bogor.

METODE

Penelitian ini menggunakan metode deskriptif dengan didukung oleh studi pustaka berdasarkan dengan Setyosari (2012, hlm. 39).

Metode deskriptif dipakai untuk menggambarkan dan mendeskripsikan data-data yang berkaitan dengan tes, khususnya soal penilaian tengah semester bahasa Sunda dengan hasil data yang telah dijawab oleh siswa kelas VII dan kelas VIII SMPN 2 Leuwisadeng Bogor.

Teknik yang digunakan yaitu tehnik studi dokumentasi.

Sumber data dalam penelitian ini yaitu soal-soal PTS bahasa Sunda dengan lembar jawaban siswa semester 1 dan semester 2 kelas VII dan VIII di SMPN 2 Leuwisadeng Bogor tahun ajaran 2018/2019. Data diolah melalui analisis validitas dan analisis tingkat reliabilitas soal.

HASIL DAN PEMBAHASAN

Berdasarkan pada hasil analisis data, di dapat tujuan-tujuan penelitian yang mencakup validitas isi, validitas empiris (kesukaran soal dan daya pembeda soal), reliabilitas dan tingkat kelayakan soal. Sama dengan Rasyid dan Mansur dalam Kholis (2017, hlm. 96) mengatakan jika dalam menguji satu instrumen dalam wanda soal terdapat empat yaitu a) Uji Validitas, b) Uji Reliabilitas, c) Analisis Daya Pembeda, dan d) Analisis Tingkat Kesulitan Soal.

Validitas Isi

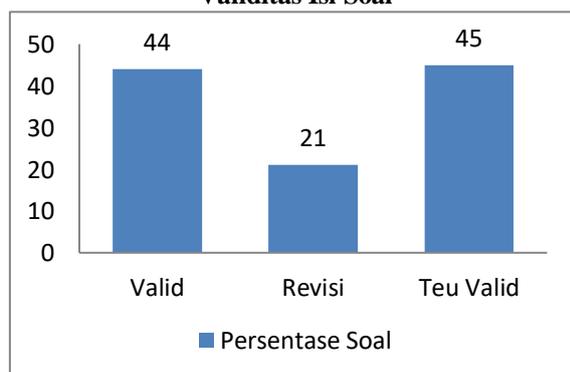
Menurut Koni (2018, hlm. 151) mengatakan yaitu validitas berhubungan dengan ketepatan pada apa harus diukur oleh tes serta seberapa cermat tes melaksanakan pengukuran.

Dalam menganalisis soal dan lembar jawaban siswa berdasarkan pada validitas isi diukur mengenai: jelas tidaknya petunjuk soal, jelas tidaknya rumusan indikator, terkait antara jenis penilaian dengan indikator, jelas tidaknya instrumen dengan materi yang diujikan, jelas tidaknya bahan penilaian dengan alat penilaian yang digunakan, bahasa yang digunakan, rumusan bahasa komunikatif, menggunakan kata dan kalimat yang mudah dipahami, huruf yang digunakan, ukuran huruf yang digunakan, dan format atau layout yang ditelaah oleh praktisi 1 dan praktisi 2.

Syarat validnya suatu soal minimal harus memenuhi 4 aspek yang penting yaitu: jelas tidaknya rumusan indikator, terkait antara jenis penilaian dengan indikator, jelas tidaknya instrumen dengan materi yang diujikan, jelas tidaknya bahan penilaian dengan alat penilaian yang digunakan. Seumpama 4 aspek itu sudah dipenuhi, soal itu sudah termasuk dalam kategori valid. Sedangkan aspek yang lainnya merupakan pelengkap, artinya kalau buruk masih bisa direvisi.

Setelah dianalisis berdasarkan pada validitas isi soal-soal penilaian tengah semester bahasa unda di SMPN 2 Leuwisadeng Bogor taun ajar 2018/2019 memiliki validitas isi yang “cukup” lebih jelasnya dijelaskan dalam grafik dibawah ini.

Grafik 1
Validitas Isi Soal



Berdasarkan grafik diatas, soal yang valid ada 44 soal (40%), soal yang harus direvisi ada 21 soal (19%), dan soal yang tidak valid ada 45 soal (41%).

44 Soal yang valid yaitu soal yang melewati 11 aspek yang telah ditentukan termasuk 4 aspek yang penting, yaitu jelas tidaknya rumusan indikator, terkait antara jenis penilaian dengan indikator, jelas tidaknya instrumen dengan materi yang diujikan, jelas tidaknya bahan penilaian dengan alat penilaian yang digunakan. Dari 21 soal direvisi, ada 7 aspek yang harus direvisi yaitu jelas tidaknya petunjuk soal, bahasa yang digunakan, rumusan bahasa komunikatif, menggunakan kata dan kalimat yang mudah dipahami, huruf yang digunakan, ukuran huruf yang digunakan, dan format atau layout) yang telah diteliti oleh praktisi 1 dan praktisi 2. Serta 45 soal tidak valid yaitu soal yang harus diganti.

Validitas Empiris

Istilah “validitas empiris” terdapat kata “empiris” yang artinya “pengalaman”. Suatu instrumen bisa disebut memiliki validitas empiris apabila udah dinilai dari pengalaman (Arikunto, 2013, hlm. 81). Analisis soal secara kuantitatif menjelaskan pada analisis karakteristik internal tes melalui data empiris. Karakteristik internal secara kuantitatif dimaksudkan pada tingkat kesukaran soal, daya pembeda, dan reliabilitas.

Tingkat kesukaran Soal

Menurut Bagiyono (2017, hlm. 2) meneliti tingkat kesukaran soal artinya meneliti butir-butir soal dari degi tingkat kesukaran soal sehingga bisa didapatkan soal-soal yang mudah, sedang dan sukar.

Soal yang baik yaitu soal yang tidak terlalu mudah atau tidak terlalu sulit (Nugraha, 2018). Soal yang mudah merangsang siswa agar cepat selesai. Sebaliknya soal yang terlalu sulit mengakibatkan siswa jadi putus asa dan tidak mempunyai semangat untuk mencoba lagi karena di luar kemampuannya.

Bilangan yang menunjukkan sulit dan mudahnya soal disebut indeks kesukaran. Besarnya indeks kesukaran yaitu antara 0,00 sampai ke 1,0. Indeks kesukaran menunjukkan tingkat sulitnya soal. Soal yang indeks kesukarannya 0,00 menunjukkan bahwa soal itu terlalu sulit, sebaliknya indeks 1,0 menunjukkan bahwa soalnya telalu mudah (Arikunto, 2013, hlm. 223).

0,00 ————— 1,0
sulit ————— mudah

Klasifikasi indeks kesukaran soal menurut Sudijono dalam Sutarjat (2013, hlm. 35) :

IF 0,00 – 0,10: soal sangat sulit (ganti)

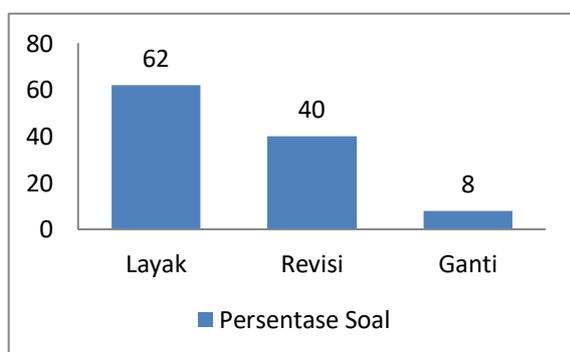
IF 0,11 – 0,30: soal sulit (revisi)

IF 0,31 – 0,70: soal sedang (layak)

IF 0,71 – 0,90: soal mudah (revisi)

IF 0,91 – 1,00: soal sangat mudah (ganti)

Tingkat kesukaran soal penilaian tengah semester bahasa Sunda di SMPN 2 Leuwisadeng Bogor tahun ajaran 2018/2019 bisa dilihat didalam grafik di bawah ini.



Grafik 2
Tingkat Kesukaran Soal

Berdasarkan grafik di atas bisa terlihat bahwa dari seluruh soal yang jumlahnya 110 soal ada 62 soal yang layak (57%), 40 soal yang harus direvisi (36%), dan 8 soal yang harus diganti (7%). Jadi, tingkat kesukaran soal penilaian tengah semester bahasa Sunda di SMPN 2 Leuwisadeng Bogor memiliki tingkat sulit mudah soal yang “tinggi”.

Menurut Sudijono dalam Sutarjat (2013, hlm. 35) klasifikasi indeks kesukaran soal, 62 soal yang layak termasuk pada soal sedang karena termasuk IF 0,31 – 0,70: soal sedang (layak), 40 soal harus direvisi termasuk dalam soal susah dan soal mudah karena IF 0,11 – 0,30: soal sulit (revisi), dan IF 0,71 – 0,90: soal mudah (revisi). f

Sedangkan 8 soal harus diganti termasuk dalam soal sangat sulit dan sangat mudah karena termasuk dalam IF 0,00 – 0,10: soal sangat sulit (ganti) dan IF 0,91 – 1,00: soal sangat mudah (ganti).

Tingkat Daya pembeda Soal

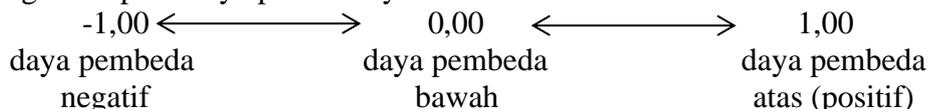
Menurut Sukanti (2013, hlm. 7) daya pembeda soal yaitu kemampuan soal dalam membedakan antara siswa yang menguasai materi yang telah ditanyakan dan siswa yang kurang atau tidak bisa menguasai materi yang telah ditanyakan.

Berbeda dengan Rahayu (2013, hlm. 7) daya pembeda soal yaitu kemampuan suatu soal dalam membedakan antara siswa yang menguasai materi yang ditanyakan dengan siswa yang kurang atau belum menguasai materi yang ditanyakan.

Salah satu tujuan analisis daya pembeda soal adalah untuk menentukan mampu tidaknya suatu butir soal membedakan antara peserta pelatihan yang berkemampuan tinggi dengan peserta pelatihan yang berkemampuan rendah Bagiyono (2017, hlm. 4).

Angka yang menunjukkan besarnya daya pembeda disebut indeks diskriminasi, disingkat D seperti indeks kesukaran, indeks diskriminasi (daya pembeda) antara 0,00 sampai 1,00. Bedanya, indeks kesukaran tidak negatif (-), tapi dalam indeks diskriminasi ada tanda negatif (-). Tanda negatif dalam indeks diskriminasi dipakai seumpama ada soal “terbalik” menunjukkan kualitas *testee*, yaitu siswa pintar disebut siswa kurang paham dan siswa kurang paham disebut pintar.

Jadi ada tiga titik pada daya pembeda yaitu:



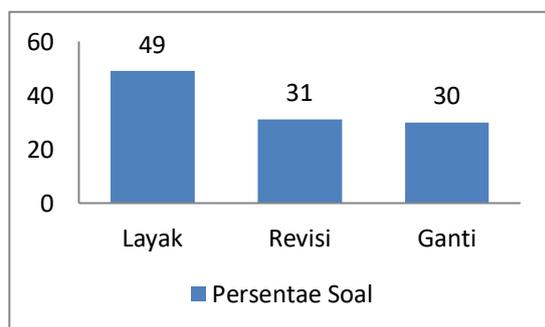
Seluruh siswa yang mengikuti tes dibagi ke dalam dua kelompok, yaitu kelompok pintar atau kelompok tinggi dan kelompok bodoh atau kelompok rendah.

Seandainya seluruh kelompok tinggi bisa menjawab soal dengan benar, serta seluruh kelompok rendah menjawab salah, soal itu memiliki D paling besar, yaitu 1,00. Sebaliknya seumpama seluruh kelompok tinggi menjawab salah, tapi seluruh kelompok rendah menjawab benar, jadi nilai D-nya -1,00. Tapi seumpama siswa kelompok tinggi dan siswa kelompok rendah sama menjawab benar atau sama menjawab salah, soal itu memiliki nilai D-nya 0,00. Karena tidak memiliki daya pembeda Arikunto (2013, hlm. 226).

Adapun menurut Daryanto (2012, hlm. 190) klasifikasi indeks daya pembeda soal di handap ini:

- D 0,00 – 0,20 : buruk (ganti)
- D 0,20 – 0,40 : sedang (revisi)
- D 0,40 – 0,70 : baik (layak)
- D 0,70 – 1,00 : sangat baik (layak)

Tingkat daya pembeda soal penilaian tengah semester bahasa Sunda di SMPN 2 Leuwisadeng Bogor tahun ajar 2018/2019 bisa dilihat dalam grafik di bawah ini.



Grafik 3
Tingkat Daya Pembeda Soal

Berdasarkan grafik di atas, kapaluruh dari 110 soal penilaian tengah semester, ada 49 soal yang layak (45%), 31 soal direvisi (28%), dan 30 soal harus diganti (27%). Jadi,

tingkat daya pembeda soal penilaian tengah semester bahasa Sunda di SMPN 2 Leuwisadeng Bogor mempunyai daya pangbeda soal yang “cukup”.

Menurut Daryanto (2012, hlm. 190) klasifikasi indeks daya pangbeda soal, 49 soal yang layak termasuk dalam soal baik dan sangat baik, karena termasuk dalam D 0,40 – 0,7 : baik (layak), D 0,70 – 1,00 : sangat baik (layak). 31 soal harus direvisi termasuk dalam soal sedeng karena D 0,20 – 0,40 : sedeng (revisi). Sedangkam 30 soal harus diganti termasuk soal buruk karena D 0,00 – 0,20 : buruk (ganti).

Tingkat Reliabilitas Soal

Dari segi bahasa, reliabilitas merupakan penerjemahan dari kata *reliability* yang mempunyai asal kata *rely* dan *ability*. Bila digabungkan, kedua kata tersebut akan mengerucut kepada pemahaman tentang kemampuan alat ukur untuk dapat dipercaya dan menjadi sandaran pengambilan keputusan. Widodo (2006, hlm. 2).

Menurut Kerlinger dalam Arifin (2012, hlm. 258) menyebutkan, “reliabilitas bisa diukur dari tiga kriteria, yaitu *stability*, *dependability*, dan *predictability*.” *Stability* menunjukkan ketetapan suatu tes dalam mengukur gejala yang sama dalam waktu yang beda. *Dependability* menunjukkan kemantapan suatu tes atau seberapa jauh tes itu bisa diandalkan. *Predictability* menunjukkan kemampuan tes untuk meramalkan hasil dalam pengukuran seterusnya. Untuk meningkatkan reliabilitas suatu tes, di antaranya bisa dilakukan dengan memperbanyak butir soal.

Menurut Thorndike dan Hagen dalam Purwanto (2011, hlm. 154) menyebutkan reliabilitas berkaitan dengan akurasi instrumen dalam mengukur apa yang akan diukur, hasil mengukur dikatakan akurat apabila dilakukan pengukuran ulang. Sedangkan menurut Davies dalam Kuswari (2010,

hlm. 62) menyebutkan bahwa reliabilitas penting untuk tes; untuk tes bahasa tertentu yang kemungkinan sulit untuk dicapai.

Berdasarkan beberapa pengertian di atas bisa disimpulkan bahwa reliabilitas berhubungan dengan konsistensi pengukuran, yaitu beberapa konsisten skor tes atau hasil evaluasi dari suatu pengukuran dan pengukuran lainnya.

Tingkat reliabilitas soal bentuk pilihan ganda dan esay menggunakan rumus yang beda. Dalam menentukan tingkat reliabilitas bentuk pilihan ganda menggunakan metode “belah dua”, serta bentuk esay menggunakan rumus *Alpha Cronbach*.

Menurut Arikunto dalam Sunarti (2014 hlm. 99) tingkatan yang dijadikan tolak ukurnya yaitu:

0,80 – 1,00 = sangat tinggi

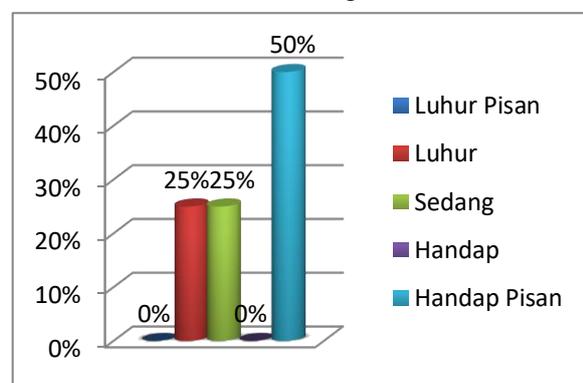
0,60 – 0,79 = tinggi

0,40 – 0,59 = sedang

0,20 – 0,39 = rendah

0,00 – 0,19 = sangat rendah

Tingkat reliabilitas soal penilaian tengah semester bahasa Sunda di SMPN 2 Leuwisadeng Bogor tahun ajar 2018/2019 bisa dilihat dalam bentuk grafik di bawah ini.



Grafik 2
Tingkat Reliabilitas Soal

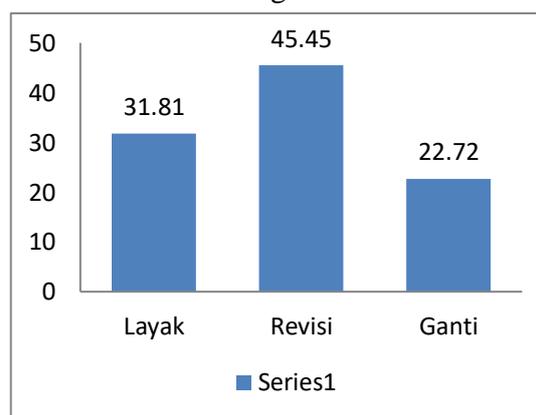
Berdasarkan pada grafik di atas, dari seluruh soal ada 40% yang memiliki reliabilitas tinggi, 30% yang memiliki reliabilitas yang sedang, dan ada 20% yang memiliki reliabilitas rendah. Oleh karena itu, tingkat reliabilitas soal penilaian tengah semester basa Sunda taun ajar 2018/2019 miboga tingkat reliabilitas “tinggi”.

Menurut Arikunto dalam Sunarti (2014, hlm. 99) tingkatan yang dijadikan tolak ukur

reliabilitas, dari semua soal ada 40% yang mempunyai reliabilitas tinggi termasuk dalam sangat tinggi jeung tinggi karena termasuk dalam 0,80 – 1,00 = sangat tinggi dan 0,60 – 0,79 = tinggi. 30% yang mempunyai reliabilitas sedang termasuk dalam soal sedang karena 0,40 – 0,59 = sedang. Sedangkan 20% yang mempunyai reliabilitas rendah termasuk dalam rendah dan sangat rendah karena 0,20 – 0,39 = bawah serta 0,00 – 0,19 = bawah pisan.

Tingkat Kelayakan Soal

Tingkat layak tidaknya suatu soal berdasarkan pada beberapa aspek hasil analisis di antaranya tingkat validitas isi, tingkat validitas empiris (kesukaran soal dan daya pembeda), serta tingkat reliabilitas soal. Tingkat layak tidaknya soal penilaian tengah semester bahasa Sunda di SMPN 2 Leuwisadeng Bogor tahun ajaran 2018/2019 bisa dilihat dalam bentuk grafik di bawah ini.



Grafik 5
Tingkat Kelayakan Soal

Berdasarkan grafik di atas terdapat 35 soal yang layak (31,81%), 50 soal harus direvisi (45,45%), dan 25 soal harus diganti (22,72%). Oleh karena itu, bisa disimpulkan bahwa tingkat validitas dan reliabilitas soal-soal penilaian tengah semester bahasa Sunda di SMPN 2 Leuwisadeng Bogor taun ajar 2018/2019 memiliki tingkat yang “cukup”.

SIMPULAN

Berdasarkan hasil analisis validitas isi (jelas tidaknya petunjuk soal, jelas tidaknya

rumusan indikator, kaitannya antara jenis penilaian dengan indikator, jelas tidaknya instrumen dengan materi yang diujikeun, jelas tidaknya bahan penilaian dengan alat penilaian yang digunakan, bahasa yang digunakan, rumusan bahasa komunikatif, menggunakan kata dan kalimat yang mudah dipahami, huruf yang digunakan, ukuran huruf yang digunakan, dan format atau layout) yang ditelaah oleh praktisi 1 dan praktisi 2 terdapat 44 soal valid (40%), soal yang harus direvisi ada 21 soal (19%), dan soal yang tidak valid 45 (41%).

Berdasarkan validitas empiris, dilihat dari kesukaran soal terdapat 62 soal yang layak (57%), 40 direvisi (36%), dan 8 soal harus diganti (7%), sedangkan dilihat dari tingkat daya pembeda terdapat 49 soal layak (45%), 31 soal revisi (28%), dan 30 soal harus diganti (27%).

Berdasarkan tingkat reliabilitas terdapat 40% mempunyai reliabilitas tinggi, 30% reliabilitas sedang dan 20% reliabilitas rendah.

Oleh sebab itu, berdasarkan validitas, validitas isi maupun validitas empiris (tingkat kesukaran dan daya pembeda soal), serta tingkat reliabilitas. Dari seluruh soal yang jumlahnya 110, terdapat kelayakan soal yaitu 35 soal layak (31,81%), 50 soal harus direvisi (45,45%), dan 25 soal harus diganti (22,72%). Oleh karena itu, bisa disimpulkan bahwa tingkat kelayakan soal PTS bahasa Sunda di SMPN 2 Leuwisadeng Bogor mempunyai tingkat kelayakan soal yang “cukup”.

UCAPAN TERIMA KASIH

Ucapan Syukur kepada Allah SWT, dan terima kasih kepada semua pihak yang telah membantu dalam menyelesaikan tulisan ini.

CATATAN PENULIS

Penulis menyatakan bahwa tidak ada konflik kepentingan terkait publikasi artikel ini. Penulis mengkonfirmasi bahwa data dan artikel ini bebas plagiarisme.

PUSTAKA RUJUKAN

Arifin, Z. (2012). *Evaluasi Pembelajaran*. Bandung: Remaja Rosdakarya.

Arikunto, S. (2013). *Dasar- Dasar Evaluasi Pendidikan* (Edisi 2). Jakarta: Bumi Aksara.

Bagiyono. (2017). Analisis Tingkat Kesukaran dan Daya Pembeda Butir Soal Ujian Pelatihan Radiografi Tingkat 1. *Jurnal Widyanuklida*, 16(1), 1-12. Retrived from <http://jurnal.batan.go.id/index.php/widyanuklida/article/view/4068>

Daryanto. (2012). *Evaluasi Pendidikan*. Jakarta: Bumi Aksara.

Kholis, R. A. N. (2017). Analisis Tingkat Kesulitan (Difficulty Level) Soal Pada Buku Sejarah Kebudayaan Islam Kelas 8 Kurikulum 2013. *Jurnal Penelitian Ilmiah Intaj*, 01(02), 93-113. Retrived from <http://ejournal.alqolam.ac.id0index.php/intaj/article/download/111/109/>

Koni, H. B. uno dan Satria. (2018). *Assesment Pembelajaran*. Jakarta: Bumi Aksara.

Kuswari, U. (2010). *Evaluasi Pembelajaran*. Bandung: Wahana Karya Grafik.

Nugraha, H. S. (2018). Kualitas Soal Ujian Departemen Pendidikan Bahasa Daerah FPBS UPI Tahun Akademik 2016/2017. *Lokabasa*, 9(1), 43-52. <https://doi.org/https://doi.org/10.17509/jlb.v9i1.15617>

Nugraha, H. S., Widyastuti, T., & Sutisna, A. (2018). The Quality of Course Subject Test Items In Sundanese Language Education Program. *Proceedings of the Tenth Conference on Applied Linguistics and the Second English Language Teaching and Technology Conference in Collaboration with the First International Conference on Language, Literature, Culture, and Education (CONAPLIN and ICOLLITE 2)*, 658-663. <https://doi.org/10.5220/0007172806580663>

Purwanti, Muslikah. (2014). *Analisis Butir*

- Soal Ujian Akhir Mata Pelajaran Akuntansi Keuangan Menggunakan Microsoft Office Excel 2010. *Pendidikan Akuntansi Indonesia*, XII(1), 81–94. Retrieved from <https://journal.uny.ac.id/index.php/jpakun/article/view/2710>
- Purwanto. (2011). *Evaluasi Hasil belajar*. Yogyakarta: Pustaka Pelajar.
- Rahayu, R. (2016). Analisis Kualitas Soal Pra Ujian Nasional Mata Pelajaran Ekonomi Akuntansi. *Pendidikan Akuntansi Indonesia*, XIV(1), 85–94. Retrieved from <https://journal.uny.ac.id/index.php/jpakun/article/view/11370>
- Setyosari, P. (2012). *Metode Penelitian Pendidikan dan Pengembangan*. Jakarta: Kencana.
- Sudjana, N. (2009). *Penilaian Hasil Proses Belajar Mengajar*. Bandung: Remaja Rosdakarya.
- Sukanti, dan Eny. P. R. (2013). Analisis Butir Soal Ulangan Kenaikan Kelas Ekonomi Akuntansi. *Pendidikan Akuntansi Indonesia*, XI(1), 67–81. Retrieved from <https://journal.uny.ac.id/index.php/jpakun/article/view/1680>
- Sukardi, M. (2012). *Evaluasi Pendidikan: Prinsip dan Operasionalnya*. Jakarta: Bumi Aksara.
- Sunarti, dan Selly. R. (2014). *Penilaian dalam Kurikulum 2013: Membantu Guru dan Calon guru Mengetahui Langkah-Langkah Penilaian Pembelajaran*. Yogyakarta: Andi Offset.
- Sutarjat, D. (2013). Kualitas Soal Tes Sumatif Pangajaran Basa Sunda di Mts. Sirnamiskin Taun Ajaran 2012/2013. Universitas Pendidikan Indonesia.
- Taib, E. N. (2014). Analisis Kualitas Aspek Materi Butir Soal Buatan Dosen. *Jurnal Biotik*, 2(2), 77–137. Retrieved from <https://jurnal.ar-raniry.ac.id/index.php/biotik/article/view/245%5C>
- Tim Direktorat Pembinaan SMP. (2017). *Panduan Penilaian Oleh Pendidik dan Satuan Pendidik Sekolah Menengah Pertama*. Jakarta: Kementerian Pendidikan dan Kebudayaan.
- Widodo, Prasetyo Budi. 2006. Reliabilitas dan Validitas Konstruk Skala Konsep Diri Untuk Mahasiswa Indonesia. *Jurnal Psikologi Universitas Diponegoro*, 3(1), 1-9. Retrieved from https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=2ahUKEwiFtNi41_zjAhWTinAKHYiNDMQQFjAAegQIBhAC&url=https%3A%2F%2Fjournal.undip.ac.id%2Findex.php%2Fpsikologi%2Farticle%2Fdownload%2F686%2F550&usg=AOvVaw3fEHNxUiWeZ38uaGvoOlt9