

**VALIDITAS DAN RELIABILITAS SOAL TENGAH SEMESTER GENAP
KAITANNYA DENGAN KETERCAPAIAN TUJUAN PEMBELAJARAN
BAHASA INDONESIA KELAS VIII A SMP NEGERI 2 BANYUDONO
TAHUN PELAJARAN 2013/2014**

Herika Ambar Tri Handani dan Harun Joko Prayitno
Prodi Pendidikan Bahasa dan Sastra Indonesia
Fakultas Keguruan dan Ilmu Pendidikan, Universitas Muhammadiyah Surakarta
Jalan A. Yani, Tromol Pos 1, Pabelan, Kartasura Surakarta 57102
Alamat Email:
herikaath@gmail.com dan harun.prayitno@ums.ac.id

Abstrak

Studi ini bertujuan untuk (1) mendeskripsikan validitas soal tengah semester genap bahasa Indonesia kelas VIII A SMP Negeri 2 Banyudono tahun pelajaran 2013/2014, (2) mendeskripsikan reliabilitas soal tengah semester genap bahasa Indonesia kelas VIII A SMP Negeri 2 Banyudono tahun pelajaran 2013/2014, dan (3) mengidentifikasi keterkaitan validitas dan realibilitas soal tengah semester genap terhadap ketercapaian tujuan pembelajaran bahasa Indonesia kelas VIII A SMP Negeri 2 Banyudono tahun pelajaran 2013/2014. Penelitian ini mengambil data di SMP Negeri 2 Banyudono. Bentuk penelitian ini adalah penelitian deskriptif kuantitatif terkait dengan interpretasi validitas dan reliabilitas hasil UTS genap kelas VIII A bahasa Indonesia dengan bantuan rumus statistik. Data penelitian berupa butir soal dan hasil jawaban siswa. Adapun sumber data diperoleh dari dokumentasi naskah soal UTS, kunci jawaban, 32 lembar jawab siswa, dan hasil kuesioner guru. Teknik analisis data menggunakan metode kuantitatif yang dipadukan dengan model Miles and Huberman meliputi reduksi data, penyajian data, dan penarikan kesimpulan. Data yang diperoleh dari hasil perhitungan validitas dan realibilitas kemudian diinterpretasikan untuk mengetahui ketercapaian tujuan pembelajaran. Hasil penelitian ini menunjukkan tiga temuan. Pertama, hasil uji validitas butir dengan perhitungan korelasi point biserial (rpbis) dan pearson product-moment dari 45 butir pilihan ganda dan 5 butir uraian menunjukkan bahwa 25 butir soal tergolong valid dan 25 butir soal tidak valid. Interpretasi hasil perhitungan validitas tersebut merujuk pada signifikansi 5% sejumlah 32 siswa menghasilkan t-tabel sebesar 0,296. Kedua, hasil uji reliabilitas butir pilihan ganda dan uraian dengan rumus KR 20 dan *alpha cronbach* diperoleh koefisien sebesar -0,117 dan -0,495. Interpretasi hasil perhitungan tersebut menunjukkan bahwa butir soal UTS tergolong tidak reliabel karena tidak memenuhi batas reliabel sebesar 0,6. Ketiga, ketercapaian tujuan pembelajaran dikaitkan dengan perolehan jumlah butir valid pada soal pilihan ganda diperoleh sebesar 51,1% dan pada soal uraian diperoleh sebesar 40%. Adapun persentase ketercapaian tujuan pembelajaran juga dilakukan dengan cara penskoran jumlah jawaban benar pada setiap butir soal dibagi jumlah siswa dikali dengan 100%. Hasil perhitungan persentase ini menunjukkan bahwa berturut-turut 68,8% dan 100% pada soal pilihan ganda dan uraian dapat dikatakan memenuhi ketercapaian tujuan pembelajaran.

Kata kunci: *validitas, reliabilitas, ketercapaian tujuan pembelajaran.*

A. Pendahuluan

Interaksi belajar-mengajar diprioritaskan pada profesionalisme guru dan prestasi belajar siswa dengan menitikberatkan pada kebermutuan sumber ajar. Proses transfer ilmu tersebut

mebutuhkan kesatuan komponen-komponen pembelajaran dalam bersinergi membentuk kompleksitas ranah, baik kognitif, afektif, maupun psikomotorik sebagai pondasi pembentukan pengetahuan siswa. Keterkaitan antarkomponen

pembelajaran dijalankan selaras dengan konteks pembelajaran dan berpedoman pada tujuan yang akan dicapai. Tujuan pembelajaran yang telah dirancang, sesuai dengan silabus, menjadi tolok ukur keberhasilan belajar siswa.

Selaras dengan pencapaian tujuan pembelajaran, diperlukan evaluasi secara periodik terhadap perkembangan hasil belajar siswa. Hasil belajar siswa sebagai bahan evaluasi untuk mengukur sejauh mana tingkat penguasaan siswa terhadap materi ajar yang telah disampaikan. Evaluasi sebagai proses penilaian pendidikan secara keseluruhan mencakup segala ketercapaian satuan pendidikan menghasilkan keberhasilan usaha yang ditempuh sesuai dengan tujuan pendidikan, yakni menghasilkan *output* selaras dengan bidang yang dipelajari. Salah satu bentuk evaluasi pendidikan yang bersifat konkret dan numerikal dapat diketahui dari hasil belajar siswa. Hasil belajar siswa diperoleh melalui penilaian.

Penilaian sebagai suatu proses untuk mengetahui apakah proses dan hasil dari suatu program kegiatan telah sesuai dengan tujuan atau kriteria yang telah ditetapkan (Suwandi, 2009:6). Tidak terlepas dari bentuk evaluasi, penilaian berhubungan erat dengan pengukuran. Pengukuran menghasilkan data untuk proses penilaian. Seperti yang dikemukakan Suwandi (2009:6), aspek kuantitatif pada penilaian diperoleh melalui pengukuran, sedangkan aspek kualitatifnya berupa penafsiran dan pertimbangan terhadap data kuantitatif hasil pengukuran tersebut. Hasil pengukuran menghasilkan data deskriptif berdasarkan penafsiran sesuai kriteria penilaian yang telah ditetapkan.

Penilaian sebagai sebuah proses membutuhkan cara untuk mengimplementasikan di kelas. Penilaian dapat dilakukan baik dengan tes maupun non-tes. Tes sebagai alat pengukur penguasaan siswa dalam kurun belajar yang telah ditetapkan. Berdasarkan jenis tes sebagai pengukur keberhasilan atau kemampuan siswa, Ulangan Tengah Semester maupun Ulangan Akhir Semester

baik jenjang pendidikan SD/ SMP/ SMA maupun sederajat tergolong pada tes sumatif. Tes sumatif dilakukan setelah selesai semua kegiatan belajar mengajar atau seluruh program yang direncanakan (Suwandi, 2009:37). Materi yang diujikan pada tes sumatif mencakup semua kompetensi dasar yang telah diajarkan dalam setengah/ satu semester. Dengan demikian, tes sumatif dilaksanakan untuk mencapai tujuan umum pembelajaran sesuai dengan standar kompetensi yang telah ditetapkan.

Pertimbangan kriteria soal yang berkategori baik diharapkan mampu memberikan informasi yang dapat dipertanggungjawabkan. Dengan kata lain, guru dituntut mampu mempersiapkan dan melakukan penilaian dengan baik sehingga tujuan pembelajaran yang telah ditetapkan dapat dicapai secara optimal. Di akhir pembelajaran, guru diharapkan mampu menyusun alat tes yang dapat dipertanggungjawabkan. Seperti yang diungkapkan oleh Tuckman (dalam Nurgiyantoro, 2010:150) bahwa alat tes mesti dapat dipertanggungjawabkan dari segi kelayakan (*appropriateness*), validitas, reliabilitas, ketertafsiran (*interpretability*), dan kebergunaan (*usability*). Jadi, tujuan utama kegiatan penilaian digunakan untuk mengetahui sejauh mana kompetensi dasar yang dikuasai siswa setelah mengikuti serangkaian pembelajaran.

Seperti yang dikemukakan Tuckman, Purwanto (2011:114) juga sependapat bahwa sebagai sebuah alat ukur maka THB (Tes Hasil Belajar) harus memenuhi syarat sebagai alat ukur yang baik. Alat ukur yang baik harus memenuhi dua syarat yaitu validitas dan reliabilitas. Purwanto menjelaskan bahwa THB yang valid adalah THB yang mengukur dengan tepat keadaan yang ingin diukur. Sebaliknya, THB dikatakan tidak valid bila digunakan untuk mengukur suatu keadaan yang tidak tepat diukur dengan THB tersebut. Pengujian validitas dapat digolongkan menjadi dua jenis. *Pertama*, validitas tes meliputi: validitas isi, validitas konstruk, validitas ramalan, dan validitas bandingan. *Kedua*, pengujian validitas butir dapat

dilakukan dengan korelasi *point biserial* dan korelasi *pearson product-moment*. Pengujian validitas hasil tes UTS semester genap kelas VIII A SMP N 2 Banyudono tahun pelajaran 2013/2014 akan dilakukan dengan cara pengujian validitas butir. Soal pilihan ganda diuji dengan perhitungan korelasi *point biserial*. Adapun pengujian validitas soal uraian dihitung dengan korelasi *pearson product-moment*.

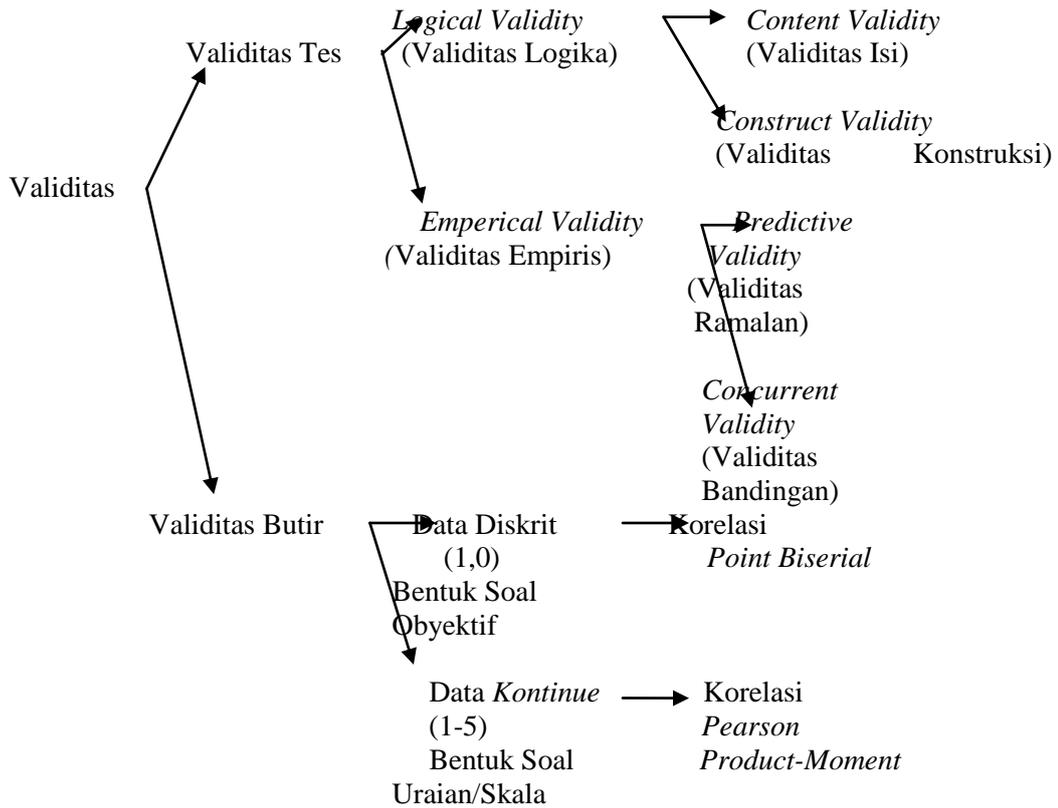
Interpretasi validitas tersebut diujikan pada hasil tes bukan sekadar alat tes. Pedoman pengujian data ini didasarkan pada pernyataan Gronlund (Nurgiyantoro, 2010:152) yang menyebutkan salah satu pertimbangan kaitan antara validitas dengan tes, yakni validitas menunjuk pada kelayakan interpretasi yang dibuat berdasarkan skor hasil tes yang berkaitan dengan penggunaan tertentu dan *bukan* terhadap instrumennya itu sendiri. Gronlund berpendapat bahwa penggunaan istilah yang lazim terdengar “validitas alat tes” sebenarnya yang lebih tepat adalah validitas interpretasi terhadap hasil tes. Jadi, pengujian validitas butir pada penelitian ini akan dilakukan pada alat tes (soal) dan hasil tes (jawaban siswa).

Validitas berhubungan erat dengan reliabilitas. Reliabilitas atau konsistensi pengukuran dibutuhkan untuk memperoleh hasil yang valid, tetapi reliabilitas dapat diperoleh tanpa harus valid (Nurgiyantoro, 2010:150). Jika validitas berkaitan dengan

kelayakan penafsiran hasil tes, maka reliabilitas berkaitan dengan konsistensi hasil pengujian tes. Pengujian hasil tes yang relatif tetap dapat dikatakan bahwa hasil tes tersebut reliabel/ dapat dipercaya, dalam arti kompetensi yang diujikan selaras dengan penguasaan siswa. Penelitian ini akan melakukan pengujian reliabilitas internal. Berkaitan dengan butir soal pada UTS Bahasa Indonesia kelas VIII A SMP N 2 Banyudono berjumlah ganjil, maka peneliti menggunakan metode *Kuder-Richardson* untuk pengujian reliabilitas soal pilihan ganda dan metode *Alpha Cronbach* untuk pengujian reliabilitas soal uraian.

Purwanto (2011:115) menjelaskan bahwa THB yang valid adalah THB yang mengukur dengan tepat keadaan yang ingin diukur. Sebaliknya, THB dikatakan tidak valid bila digunakan untuk mengukur suatu keadaan yang tidak tepat diukur dengan THB tersebut. Adapun menurut Gronlund dalam Nurgiyantoro (2010:150) mengemukakan bahwa validitas menunjuk pada kelayakan interpretasi yang dibuat berdasarkan skor hasil tes yang berkaitan dengan penggunaan tertentu. Berdasarkan uraian pengertian di atas, dapat disimpulkan bahwa validitas menunjuk pada ketepatan, kelayakan, dan kesahihan interpretasi terhadap hasil tes berdasarkan kriteria yang telah dirumuskan pada tujuan pembelajaran.

Adapun pembagian jenis validitas menurut Iskandar (2012) sebagai berikut.



Keterkaitan keduanya terletak pada reliabilitas atau konsistensi pengukuran dibutuhkan untuk memperoleh hasil yang valid, tetapi reliabilitas dapat diperoleh tanpa harus valid. Butir tes yang valid dan reliabel menunjukkan kesesuaian kompetensi yang harus dikuasai sesuai dengan tujuan pembelajaran yang telah dirumuskan. Dalam kaitan ini sebagaimana ditegaskan Nurgiyantoro (2010:150) menegaskan bahwa validitas berhubungan erat dengan reliabilitas. Jika validitas lebih berurusan dengan kelayakan penafsiran terhadap skor hasil tes, reliabilitas berurusan dengan keajegan skor hasil tes.

B. Metode Penelitian

Penelitian ini merupakan jenis kuantitatif deskriptif. Data penelitian ini berupa butir soal UTS genap mata pelajaran bahasa Indonesia kelas VIII SMP Negeri 2 Banyudono tahun pelajaran 2013/2014 dan jawaban siswa. Sumber data berupa dokumentasi naskah soal UTS genap mata pelajaran bahasa Indonesia kelas VIII A SMP Negeri 2 Banyudono tahun ajaran

2013/2014, kunci jawaban soal UTS, 32 lembar jawaban siswa, dan hasil kuesioner guru. Pengumpulan data dengan menggunakan teknik dokumentasi dan kuesioner. Teknik analisis data penelitian ini dilakukan secara kuantitatif dengan dipadukan *model Miles and Huberman*. Miles dan Huberman (dalam Sugiyono, 2009:246) mengemukakan bahwa aktivitas dalam analisis data kualitatif dilakukan secara interaktif dan berlangsung secara terus menerus sampai tuntas, sehingga datanya sudah jenuh. Analisis data dalam penelitian ini meliputi tiga langkah meliputi reduksi data, penyajian data, dan penarikan kesimpulan. Adapun keabsahan data diukur dengan menggunakan teknik triangulasi teori.

Selanjutnya untuk mengukur validitas butir dilakukan dengan bantuan perhitungan statistik korelasi. Berkaitan dengan bentuk soal berupa pilihan ganda, pengujian validitas menggunakan perhitungan korelasi *point biserial*. Adapun bentuk soal uraian diuji dengan perhitungan korelasi *pearson product-moment*.

Berikut rumus rpbis (korelasi *point biserial*).

$$r_{pbis} = \sqrt{\frac{M_p - M_t}{S_{dt}} \frac{p}{q}}$$

(Iskandar, 2012)

Keterangan :

r_{pbis} : koefisien korelasi *point biserial*

M_p : skor rata-rata hitung untuk butir yang dijawab betul

M_t : skor rata-rata dari skor total

S_{dt} : standar deviasi skor total

p : proporsi siswa yang menjawab betul pada butir yang diuji validitasnya

q : proporsi siswa yang menjawab salah pada butir yang diuji validitasnya

Adapun perhitungan validitas uraian menggunakan rumus korelasi *pearson – product moment*. Berikut rumusnya.

$$\text{Pearson } r = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{\{N \sum X^2 - (\sum X)^2\}} \sqrt{\{N \sum Y^2 - (\sum Y)^2\}}}$$

(Iskandar, 2012)

Keterangan :

r : koefisien korelasi *Pearson*

$\sum XY$: jumlah hasil kali skor X dan Y

$\sum X$: jumlah skor X

$\sum Y$: jumlah skor Y

$\sum X^2$: jumlah kuadrat skor X

$\sum Y^2$: jumlah kuadrat skor Y

N : jumlah peserta

Pengujian signifikansi korelasi dilakukan dengan membandingkan antara korelasi hitung (r_{xy}) dengan r pada tabel (r tabel). Pada korelasi positif, bila $r_{xy} > r$ tabel maka dapat disimpulkan bahwa xy mempunyai korelasi positif secara signifikan.

Uji reliabilitas tes dilakukan dengan tujuan untuk memberikan antara lain: (1) tes dapat memberikan hasil yang relatif tetap terhadap sesuatu yang diukur, (2) jawaban peserta didik terhadap butir-butir tes secara relatif tetap, dan (3) hasil tes diperiksa oleh siapa pun juga akan menghasilkan skor yang kurang lebih sama. Adapun Purwanto (2011:153-154) mengemukakan bahwa keandalan (*reliability*) berasal dari kata *rely* yang artinya percaya dan reliabel yang artinya dapat dipercaya. Keterpercayaan

berhubungan dengan ketetapan dan konsistensi. THB dikatakan dapat dipercaya apabila memberikan hasil pengukuran hasil belajar yang relatif tetap secara konsisten. Berdasarkan beberapa pengertian di atas, dapat disimpulkan bahwa reliabilitas menunjuk pada keterpercayaan/ konsistensi hasil tes yang relatif tetap, baik dari segi jawaban siswa maupun diperiksa oleh penilai yang berbeda, skor yang dihasilkan akan kurang lebih sama sesuai dengan kriteria penilaian yang ditetapkan.

Purwanto (2011:155-156) menyebutkan bahwa metode pengujian reliabilitas dibedakan menjadi dua kelompok, yakni *external stability* dan *internal consistency*. Reliabilitas sebagai stabilitas eksternal ini memandang bahwa THB dikatakan reliabel apabila diujikan

beberapa kali akan memberikan hasil pengukuran yang relatif konsisten. Metode yang tergolong dalam kelompok ini adalah metode tes ulang dan metode paralel. Adapun THB dikatakan reliabel apabila di antara butir THB memberikan hasil pengukuran yang konsisten. Metode ini dikelompokkan menjadi dua berdasarkan jumlah butirnya. Metode yang termasuk pada klasifikasi jumlah butir genap, antara lain: metode belah dua, *Flanagan* dan *Rulon*. Adapun metode yang termasuk pada

klasifikasi jumlah butir ganjil, antara lain: metode *Kuder-Richardson*, *Hoyt*, dan *Alpha Cronbach*.

Berdasarkan pembagian metode pengujian reliabilitas tersebut, interpretasi penilaian dalam penelitian ini menggunakan reliabilitas internal. Berkaitan dengan butir tes yang akan diuji berjumlah ganjil, maka peneliti menggunakan metode *Kuder-Richardson 20* untuk uji reliabilitas soal pilihan ganda dan metode *Alpha-Cronbach* untuk uji reliabilitas soal uraian.

Berikut rumus KR20 :

$$r_{KR20} = \frac{k}{k-1} \left(1 - \frac{\sum pq}{s^2} \right)$$

(Iskandar, 2012)

Keterangan :

r_{KR20} : koefisien korelasi dengan KR20

k : jumlah butir soal

p : proporsi jawaban benar pada butir tertentu

q : proporsi jawaban salah pada butir tertentu
($q = 1-p$)

s^2 : varians skor total

Adapun berikut ini rumus korelasi *alpha cronbach*.

$$r_{11} = \left[\frac{n}{n-1} \right] \left[1 - \frac{\sum si^2}{\sum st^2} \right]$$

(Purwanto, 2011:175)

Keterangan :

r_{11} : koefisien reliabilitas tes

n : jumlah butir

si^2 : varians butir

st^2 : varians total

Batas kriteria reliabilitas ditunjukkan pada tabel korelasi. Bila r hitung $>$ r tabel maka kedua skor hasil pengukuran THB berkorelasi signifikan. Signifikansi korelasi menunjukkan adanya konsistensi sehingga THB telah dapat dikatakan reliabel (Purwanto, 2011:180).

C. Hasil dan Pembahasan

1. Hasil

Penelitian ini mengukur hasil tes UTS genap mata pelajaran bahasa Indonesia kelas VIII A SMP Negeri 2 Banyudono dengan uji validitas dan reliabilitas untuk mengukur ketercapaian tujuan pembelajaran. Purwanto (2011:114) menerangkan bahwa

alat ukur yang baik harus memenuhi dua syarat yaitu validitas dan reliabilitas. Berkaitan dengan hal tersebut, peneliti menganalisis THB dengan cara pengujian signifikansi korelasi (:membandingkan antara korelasi hitung (r-hitung) dengan r pada tabel (r-tabel)). Hasil uji korelasi tersebut menunjukkan bahwa THB tergolong alat ukur yang valid dan atau reliabel/ tidak.

**Validitas Soal Tengah Semester Genap
Bahasa Indonesia Kelas VIII A SMP**

**Negeri 2 Banyudono Tahun Pelajaran
2013/2014**

Peneliti menghitung validitas butir soal pilihan ganda dengan rumus rpbis (korelasi *point biserial*). Adapun, validitas soal uraian dihitung dengan rumus korelasi *pearson product moment*. Berikut hasil perhitungan validitas butir soal pilihan ganda.

Tabel 1. Validitas Butir Soal Pilihan Ganda

Butir Soal	r-tabel	r-hitung	Hasil
1	0,296	0,254	TV
2	0,296	0,319	V
3	0,296	1,066	V
4	0,296	0	TV
5	0,296	0	TV
6	0,296	0,862	V
7	0,296	0	TV
8	0,296	0	TV
9	0,296	0,415	V
10	0,296	0,716	V
11	0,296	0,164	TV
12	0,296	0,764	V
13	0,296	0,250	TV
14	0,296	0	TV
15	0,296	0,505	V
16	0,296	0,316	V
17	0,296	0,554	V
18	0,296	0	TV
19	0,296	0,104	TV
20	0,296	0,553	V
21	0,296	0	TV
22	0,296	0	TV
23	0,296	0,899	V
24	0,296	1,027	V
25	0,296	0	TV
26	0,296	0	TV
27	0,296	0,874	V
28	0,296	0	TV
29	0,296	0,366	V
30	0,296	0	TV
31	0,296	0,571	V
32	0,296	0,662	V
33	0,296	0,681	V
34	0,296	0,452	V
35	0,296	0,486	V

36	0,296	0	TV
37	0,296	0,681	V
38	0,296	0	TV
39	0,296	0	TV
40	0,296	0	TV
41	0,296	0,296	V
42	0,296	0,260	TV
43	0,296	0,636	V
44	0,296	0,630	V
45	0,296	0	TV

Keterangan:

V : Valid

TV : Tidak Valid

Berdasarkan perhitungan validitas pada signifikansi 5% sejumlah 32 siswa dengan tabel korelasi sebesar 0,296, validitas butir soal pilihan ganda hasil UTS genap mata pelajaran bahasa Indonesia kelas VIII A SMP Negeri 2 Banyudono tahun pelajaran

2013/2014 menunjukkan sejumlah 23 butir soal valid dan 22 butir soal tidak valid.

Berdasarkan perhitungan korelasi *pearson - product moment*, adapun keputusan validitas butir soal uraian sebagai berikut.

Tabel 2. Validitas Butir Soal Uraian

Butir Soal	r - tabel	r - hitung	Hasil
1	0,296	0,006	TV
2	0,296	0	TV
3	0,296	0,640	V
4	0,296	0	TV
5	0,296	0,572	V

Keterangan :

V : valid

TV : tidak valid

Berdasarkan perhitungan validitas pada signifikansi 5% sejumlah 32 siswa dengan tabel korelasi sebesar 0,296, validitas butir soal uraian hasil UTS genap mata pelajaran bahasa Indonesia kelas VIII A SMP Negeri 2 Banyudono tahun pelajaran 2013/2014 menunjukkan sejumlah 2 butir soal valid dan 3 butir soal tidak valid.

Reliabilitas Soal Tengah Semester Genap Bahasa Indonesia Kelas VIII A SMP Negeri 2 Banyudono Tahun Pelajaran 2013/2014

THB yang reliabel adalah THB yang memberikan hasil pengukuran hasil belajar yang relatif tetap secara konsisten. Sebaliknya, THB yang tidak reliabel adalah

THB yang tidak tetap dalam memberikan hasil pengukuran hasil belajar secara konsisten. Jadi, untuk mengetahui THB reliabel/ tidak, diperlukan pengujian reliabilitas hasil tes. Peneliti menghitung reliabilitas butir soal pilihan ganda dengan metode *Kuder-Richardson* (KR20). Adapun, reliabilitas soal uraian dihitung dengan metode *Alpha Cronbach*.

THB dapat dinyatakan reliabel apabila koefisien yang diperoleh melalui perhitungan menggunakan metode pengujian reliabilitas tertentu lebih besar dibandingkan dengan batas keputusan reliabilitas (Purwanto, 2011:177). Angka koefisien batas reliabel yang digunakan sebagai dasar untuk menentukan koefisien reliabilitas hasil

perhitungan adalah sebesar 0,6. Semakin tinggi angka koefisien hasil perhitungan dari batas 0,6, maka THB tersebut tergolong sangat reliabel. Berdasarkan perhitungan yang dilakukan, diperoleh koefisien reliabilitas soal pilihan ganda sebesar -0,117 dan soal uraian sebesar -0,495.

Ketercapaian Tujuan Pembelajaran Bahasa Indonesia Kelas VIII A SMP Negeri 2 Banyudono Tahun Pelajaran 2013/2014

Pengukuran ketercapaian tujuan pembelajaran butir soal ini dikaitkan dengan pemerolehan jumlah soal yang tergolong valid dan pemerolehan persentase hasil dari jumlah jawaban benar setiap butir dibagi jumlah siswa. Dalam konteks perhitungan validitas dan reliabilitas soal pilihan ganda yang telah dilakukan, dapat diketahui bahwa terdapat 23 butir soal yang memenuhi ketercapaian tujuan pembelajaran. Adapun, soal uraian ditemukan 2 butir soal yang memenuhi ketercapaian tujuan pembelajaran. Dalam bentuk persentase, ketercapaian tujuan pembelajaran berdasarkan keterkaitan valid dan reliabel menunjukkan pada pilihan ganda sebesar 51,1% dan pada uraian sebesar 40% soal yang memenuhi kompetensi yang diajarkan.

Merujuk perhitungan penskoran jumlah jawaban benar menunjukkan bahwa terdapat 31 butir soal dapat dikatakan tujuan pembelajaran pada materi tersebut sudah tercapai. Adapun dalam bentuk persentase menunjukkan ketercapaian tujuan pembelajaran soal pilihan ganda sebesar 68,8% dan soal uraian sebesar 100% memenuhi kompetensi yang diajarkan.

2. Pembahasan

Seperti dikemukakan oleh Purwanto (2011:47) bahwa hasil belajar perlu dievaluasi. Evaluasi dimaksudkan sebagai cermin untuk melihat kembali apakah tujuan yang ditetapkan telah tercapai dan apakah proses belajar mengajar telah berlangsung efektif untuk memperoleh hasil belajar. Oleh karena itu, evaluasi tes hasil belajar dilakukan untuk mengetahui bahwa THB itu memenuhi syarat sebagai alat ukur yang baik.

Berikut interpretasi validitas butir soal hasil pengujian signifikansi korelasi yang dilakukan dengan membandingkan antara korelasi hitung (r -hitung) dengan r pada tabel (r -tabel). Purwanto (2011:118) mengungkapkan bahwa pada korelasi positif, bila $r_{xy} > r_{tabel}$ maka dapat disimpulkan bahwa xy mempunyai korelasi positif secara signifikan. Adapun, pada korelasi negatif bila $r_{xy} < r_{tabel}$ maka dapat disimpulkan bahwa xy mempunyai korelasi negatif secara signifikan. Pengambilan keputusan valid atau tidak valid butir soal ini didasarkan pada rujukan teori Purwanto tersebut.

Hasil uji validitas butir menunjukkan bahwa soal pilihan ganda terdiri atas 23 butir soal valid dan 22 butir soal tidak valid. Adapun soal uraian terdiri atas 2 soal valid dan 3 soal tidak valid. Jumlah benar yang berhasil dijawab siswa mempengaruhi batas valid setiap butir soal. Semakin besar skor total dari ke-32 siswa pada tiap butir soal maka batas valid akan tercapai. Penentuan keputusan batas valid koefisien yang dicapai bergantung pada perolehan nilai M_p (rata-rata hitung dari skor total yang dijawab benar). Dengan kata lain, skor total siswa yang menjawab benar butir soal mempengaruhi hasil nilai M_p .

Berdasarkan hasil perhitungan reliabilitas butir soal pilihan ganda diperoleh koefisien sebesar -0,117. Merujuk pada batas reliabel sebesar 0,6, maka dapat dikatakan bahwa butir soal pilihan ganda termasuk kategori instrumen yang kurang baik karena tidak reliabel. Adapun, hasil perhitungan reliabilitas butir soal uraian menghasilkan koefisien sebesar -0,495. Dikaitkan dengan batas reliabel sebesar 0,6, maka dapat dikatakan bahwa butir soal uraian tidak reliabel. Soal uraian tersebut termasuk kategori instrumen yang kurang baik pula.

Alat tes yang baik akan menghasilkan hasil tes yang baik pula. Butir soal yang valid dan reliabel diharapkan akan mampu dijawab benar oleh siswa sehingga hasil tes memenuhi batas nilai yang ditentukan. Jadi, hasil belajar siswa sesuai dengan tujuan pembelajaran yang telah dirancang.

Keterkaitan validitas dan reliabilitas terletak pada reliabilitas atau konsistensi pengukuran dibutuhkan untuk memperoleh hasil yang valid, tetapi reliabilitas dapat diperoleh tanpa harus valid. Hasil analisis menunjukkan bahwa perolehan hasil valid pada soal pilihan ganda sebanyak 23 soal dan 2 soal uraian tergolong dalam kategori soal yang reliabel. Ketercapaian tujuan pembelajaran berdasarkan keterkaitan valid dan reliabel menunjukkan sebesar 51,1% untuk pilihan ganda dan 40% untuk uraian. Adapun ketercapaian tujuan pembelajaran berdasarkan skor jawaban benar menghasilkan persentase sebesar 68,8% untuk pilihan ganda dan 100% untuk uraian.

Dikaitkan dengan hasil temuan terdahulu. Peneliti mengaitkan hasil analisis yang diperoleh dengan tiga hasil temuan terdahulu. Pertama, penelitian oleh Adiputra (2011) dengan judul "Analisis Butir Soal Tes Ulangan Akhir Semester IPS Terpadu Buatan MGMP IPS Kabupaten Gianyar Kelas VII Semester 1 Tahun Pelajaran 2011-2012". Analisis yang dilakukan bertujuan untuk mengetahui kualitas butir soal tes tersebut ditinjau dari relevansi antara kisi-kisi, SK, KD, dan Indikator mata pelajaran, validitas isi tes, validitas butir soal, reliabilitas tes, taraf kesukaran butir soal, daya beda butir soal, dan efektivitas pengecoh butir soal.

Hasil penelitian Adiputra menunjukkan relevansi antara Standar Kompetensi, Kompetensi Dasar, dan Indikator dengan butir soal dari 60 butir soal terdapat 56 (93%) butir soal relevan dan 4 butir soal (7%) tidak relevan dengan indikator soal. Validitas isi yang diuji dengan Uji Gregory menunjukkan validitas yang sangat tinggi (0,933). Bila ditinjau dari validitas butir didapatkan 53 butir soal (88%) dinyatakan valid dan 7 butir soal (8%) dinyatakan tidak valid. Reabilitas tes didapatkan $r_{11} = 0,860$ dengan formula KR 20. Jika ditinjau dari taraf kesukaran tes didapatkan 23% butir soal kategori soal mudah, 62 % butir kategori soal sedang, dan 15 % butir kategori soal sukar. Dilihat dari daya beda, 82 % memiliki daya beda yang dapat diterima, yang memiliki daya beda

kurang baik 13 % dan yang buruk harus dibuang 5%. Dan jika ditinjau dari efektivitas pengecoh 85% memiliki pengecoh yang berfungsi dengan baik dan 15% tidak dapat menjalankan fungsinya dengan baik.

Berdasarkan hasil penelitian Adiputra yang dipaparkan di atas, berikut persamaan dan perbedaan dengan temuan penelitian ini. Persamaannya yakni sama-sama menganalisis butir soal tes. Selain itu, penelitian Adiputra menguji kualitas tes juga dengan cara validitas dan reliabilitas hasil tes. Adapun perbedaannya yakni: (1) Adiputra menguji butir tes pada mata pelajaran IPS terpadu, sedangkan penelitian ini pada mata pelajaran bahasa Indonesia dan (2) penelitian Adiputra menganalisis butir tes secara komprehensif, mulai dari validitas tes, reliabilitas tes, tingkat kesukaran soal, daya pembeda, dan efektivitas pengecoh butir soal yang direlevansikan dengan kisi-kisi, SK, KD, dan indikator, sedangkan penelitian ini memfokuskan analisis butir soal berdasarkan validitas dan reliabilitas hasil tes untuk mengetahui ketercapaian tujuan pembelajaran. Ketercapaian tujuan pembelajaran penelitian ini juga dikaitkan dengan KD yang telah dirumuskan.

Dikaitkan dengan hasil temuan, persentase validitas butir yang diperoleh oleh Adiputra tergolong tinggi, yakni sebesar 88%. Adapun hasil temuan penelitian ini memperoleh persentase valid yang rendah. Persentase validitas butir pilihan ganda sebesar 51,1%, sedangkan perolehan validitas butir uraian sebesar 40%. Selaras dengan hal itu, reliabilitas hasil perhitungan Adiputra menunjukkan koefisien sebesar 0,860. Perolehan koefisien tersebut tergolong tinggi sehingga dapat dikatakan instrumen tes tersebut tergolong reliabel. Berbeda halnya dengan perolehan koefisien reliabilitas dalam penelitian ini. Soal pilihan ganda dan uraian sama-sama menunjukkan kualitas soal yang kurang baik, dengan perolehan koefisien di bawah batas reliabel (0,6) yakni -0,117 dan -0,495.

Kedua, penelitian Dwipayani (2011) dengan judul "Analisis Validitas dan Reliabilitas Butir Soal Ulangan Akhir

Semester Bidang Studi Bahasa Indonesia Kelas X.D SMA N 1 Bangli terhadap Pencapaian Kompetensi". Penelitian ini bertujuan mengetahui dan mendeskripsikan (1) kualitas butir soal ulangan akhir semester ganjil bidang studi bahasa Indonesia kelas X.d SMA N 1 Bangli berdasarkan uji validitas isi, tingkat kesukaran, analisis daya beda, dan efektivitas opsi, (2) kualitas perangkat soal ulangan akhir semester ganjil bidang studi bahasa Indonesia kelas X.d SMA N 1 Bangli ditinjau dari analisis reliabilitas, (3) tingkat ketercapaian siswa dalam menjawab benar dan salah tes/soal ulangan akhir semester ganjil bidang studi bahasa Indonesia kelas X.d di SMA N 1 Bangli.

Hasil penelitian Dwipayani menunjukkan bahwa validitas isi soal tidak valid atau tergolong rendah. Hasil tingkat kesukaran butir soal yang termasuk dalam kriteria soal mudah 24%, soal sedang 56%, dan soal sulit 20%. Butir soal yang termasuk dalam kriteria daya pembeda jelek 40%, daya pembeda cukup 32%, daya pembeda baik 20%, dan daya pembeda sangat baik 8%. Analisis efektivitas opsi kunci butir soal yang berfungsi secara efektif 72%, sedangkan opsi kunci butir soal yang tidak berfungsi efektif 28%, untuk opsi pengecoh butir soal yang berfungsi efektif 76%, sedangkan opsi pengecoh butir soal yang tidak berfungsi efektif 24%. Hasil uji reliabilitas siswa pada saat mengikuti tes I berbeda secara signifikan dengan hasil tes II. Rata-rata persentase tingkat ketercapaian siswa yang menjawab benar pada soal ulangan akhir semester ganjil mata pelajaran bahasa Indonesia kelas X.d SMA N 1 Bangli sebesar 52% sedangkan rata-rata persentase tingkat ketercapaian siswa yang menjawab salah sebesar 48%.

Berdasarkan hasil temuan analisis Dwipayani di atas, perbedaan dengan hasil penelitian ini sebagai berikut. *Pertama*, penggunaan metode uji validitas. Dwipayani menggunakan metode validitas isi. Uji validitas ini dapat dilakukan tanpa pengujian statistika. Berbeda halnya dengan penelitian ini menggunakan perhitungan statistika melalui uji korelasi. Pengujian validitas soal

pilihan ganda pada penelitian ini menggunakan rumus rpbis (korelasi *point biserial*). Adapun pengujian soal uraian penelitian ini menggunakan korelasi *pearson product moment*. *Kedua*, penelitian Dwipayani juga menganalisis butir soal berdasarkan tingkat kesukaran soal, daya pembeda, efektivitas opsi kunci soal, dan efektivitas opsi pengecoh butir soal. Adapun, penelitian ini memfokuskan analisis butir soal berdasarkan pengujian validitas dan reliabilitas soal. *Ketiga*, penelitian Dwipayani menguji reliabilitas soal dengan uji ulang tes. Adapun pada penelitian ini menggunakan hasil tes pada UTS. Hal ini dilakukan oleh peneliti karena selaras dengan metode uji reliabilitas yang digunakan. Peneliti menghitung reliabilitas tanpa melakukan korelasi dari dua pengukuran. Metode ini dilakukan dengan menghubungkan setiap butir dalam satu tes dengan butir-butir lainnya dalam tes itu sendiri secara keseluruhan. Pada soal pilihan ganda, peneliti menggunakan rumus *kuder-richardson* (KR 20), sedangkan pada soal uraian digunakan metode *alpha cronbach*.

Di samping beberapa persamaan yang ditemukan di atas, terdapat pula persamaan hasil temuan antara hasil penelitian ini dengan penelitian Dwipayani. *Pertama*, perolehan hasil valid butir soal sama-sama rendah. Dwipayani memperoleh hasil validitas soal yang tergolong rendah. Adapun hasil penelitian ini memperoleh jumlah butir soal valid yang berbeda tipis, yakni 23 soal valid dan 22 soal tidak valid (pilihan ganda) serta 2 soal valid dan 3 soal tidak valid (uraian). *Kedua*, persamaan terletak pada perhitungan ketercapaian tujuan pembelajaran. Ketercapaian ini sama-sama diukur dengan cara perhitungan persentase jawaban benar. Perolehan angka ketercapaian pun sama-sama berbeda tipis. Penelitian Dwipayani memperoleh sebesar 52% butir soal yang mencapai tujuan pembelajaran. Adapun penelitian ini memperoleh persentase ketercapaian sebesar 68,8% untuk soal pilihan ganda dan 100% untuk soal uraian. Selain perhitungan jumlah benar tersebut, penelitian ini mengukur ketercapaian tujuan pembelajaran dengan

cara menghitung persentase jumlah soal valid.

Ketiga, penelitian oleh Nuratni (2011) dengan judul “Analisis Butir-Butir Soal Ulangan Umum Bahasa Indonesia Kelas X SMAK Harapan Semester Ganjil Tahun Pelajaran 2011/2012”. Penelitian ini merupakan penelitian deskriptif kuantitatif-kualitatif yang bertujuan untuk mendeskripsikan (1) validitas butir-butir soal ulangan umum bahasa Indonesia kelas X, tentang validitas isi, konstruk dan analisis butir-butir soal; (2) taraf kesukaran butir-butir soal ulangan umum bahasa Indonesia kelas X; (3) daya beda butir-butir soal ulangan umum bahasa Indonesia kelas X; dan (4) reliabilitas butir-butir soal ulangan umum bahasa Indonesia kelas X SMA.

Hasil penelitian terhadap butir-butir soal ulangan umum bahasa Indonesia kelas X SMAK Harapan menunjukkan bahwa (1) validitas isi yang mencakup keterampilan membaca dan menulis tergolong baik; validitas konstruk tergolong baik yaitu 82% soal yang sesuai dengan kisi-kisi soal, dan 18% soal tidak sesuai dengan kisi-kisi soal; dan validitas butir-butir soal tergolong kurang baik yaitu 68% soal valid dan 32% tidak valid. (2) Tingkat kesukaran butir-butir soal tergolong kurang baik. (3) Daya beda butir-butir soal tergolong baik. (4) Reliabilitas tergolong tinggi.

Berdasarkan hasil penelitian Nuratni di atas, terdapat persamaan dan perbedaan dengan hasil temuan penelitian ini. Persamaannya yakni: (1) sama-sama menganalisis butir soal tes; (2) sama-sama menggunakan uji validitas dan reliabilitas tes; (3) sama-sama diujikan pada mata pelajaran bahasa Indonesia; dan (4) metode pengumpulan data menggunakan teknik dokumentasi. Adapun perbedaannya antara lain: (1) penelitian Nuratni lebih kompleks, meliputi: uji validitas isi, uji validitas konstruk, uji validitas butir, uji reliabilitas, tingkat kesukaran soal, dan daya beda; (2) jenis penelitian Nuratni yaitu penelitian semi kualitatif-kuantitatif, dan (3) teknik analisis data Nuratni juga semi kualitatif-kuantitatif. Teknik kualitatif digunakan pada saat mendeskripsikan data, sedangkan teknik

kuantitatif digunakan pada saat penghitungan skor melalui program iteman. Adapun penelitian ini menggunakan teknik analisis data kuantitatif yang dipadukan dengan model *Miles and Huberman*. Model analisis penelitian ini dilakukan dengan urutan reduksi data – penyediaan data – penarikan kesimpulan.

Dikaitkan dengan hasil temuan, perolehan uji validitas isi dan konstruk penelitian Nuratni tergolong baik. Akan tetapi, berbeda dengan perolehan hasil uji validitas butir penelitian Nuratni tergolong kurang baik dengan persentase sebesar 68% butir tergolong valid. Sama halnya dengan temuan hasil penelitian ini diperoleh persentase validitas sebesar 51,1% pilihan ganda dan 40% uraian. Selaras dengan ketercapaian soal valid, penelitian Nuratni menghasilkan reliabilitas yang tinggi, sedangkan penelitian ini menghasilkan reliabilitas rendah di bawah batas reliabel 0,6.

Berdasarkan hasil membandingkan temuan hasil analisis, penelitian analisis butir soal ini diharapkan menjadi langkah wajib yang semestinya ditempuh bagi guru maupun tim penyusun soal untuk menyusun soal yang berkualitas. Mengacu pada teori Purwanto, syarat alat tes harus memenuhi kedua syarat yakni valid dan reliabel. Salah satu analisis butir soal dengan cara pengujian validitas dan reliabilitas soal dapat ditempuh untuk mengetahui ketercapaian tujuan pembelajaran. Hasil perhitungan yang valid akan mempengaruhi kualitas keterbacaan soal pada siswa. Oleh karena itu, soal yang memenuhi kompetensi sesuai dengan tujuan pembelajaran akan membantu siswa menjawab dengan benar. Dengan kata lain, hasil belajar siswa akan memperoleh hasil maksimal apabila kualitas soal yang diujikan telah memenuhi syarat sebagai alat tes yang baik.

D. Simpulan

Dalam rangka menjawab tiga rumusan masalah yang diajukan pada bagian awal penelitian, pada bagian ini dipaparkan simpulan dari serangkaian analisis yang telah dilakukan. *Pertama*, hasil uji validitas butir

dengan perhitungan korelasi *point biserial* (rpbis) dan *pearson product-moment* dari 45 butir pilihan ganda dan 5 butir uraian menunjukkan bahwa 25 butir soal tergolong valid dan 25 butir soal tidak valid. Interpretasi hasil perhitungan validitas tersebut merujuk pada signifikansi 5% sejumlah 32 siswa dengan tabel korelasi sebesar 0,296. *Kedua*, hasil uji reliabilitas butir pilihan ganda dan uraian dengan rumus KR 20 dan *alpha cronbach* diperoleh koefisien sebesar -0,117 dan -0,495. Interpretasi hasil perhitungan tersebut menunjukkan bahwa butir soal UTS tergolong tidak reliabel karena tidak memenuhi batas reliabel sebesar 0,6. *Ketiga*, ketercapaian tujuan pembelajaran dikaitkan dengan perolehan jumlah butir valid pada soal pilihan ganda diperoleh sebesar 51,1% dan pada soal uraian diperoleh sebesar 40%. Adapun persentase ketercapaian tujuan pembelajaran juga dilakukan dengan cara penskoran jumlah jawaban benar pada setiap butir soal dibagi jumlah siswa dikali dengan 100%. Hasil perhitungan persentase ini menunjukkan bahwa berturut-turut 68,8% dan 100% pada soal pilihan ganda dan uraian dapat dikatakan memenuhi ketercapaian tujuan pembelajaran.

E. Saran

Pertama, guru bidang studi maupun tim MGMP ikut serta melakukan uji validitas dan reliabilitas soal UTS sebelum pelaksanaan UTS tersebut. Sependapat dengan teori Nurgiyantoro, uji validitas dan reliabilitas dilakukan oleh guru sebelum melaksanakan tes. Hasil uji di awal akan memberikan pertimbangan baik untuk perbaikan dan penyempurnaan instrumen tes tersebut sehingga layak untuk diujikan.

Kedua, sebaiknya setelah UTS selesai, perlu adanya evaluasi terhadap soal tes, baik bersama siswa maupun bersama guru bidang studi bahasa Indonesia yang mengampu kelas VIII. Hal ini dimaksudkan untuk mengoreksi kesesuaian soal dengan jawaban siswa. Selain itu, kompetensi yang belum tercapai sebagai target guru untuk memperbaiki penerimaan kompetensi tersebut kepada siswa. Evaluasi di akhir

pelaksanaan tes memberikan penyempurnaan pemahaman siswa terhadap materi yang dikuasai.

Ketiga, guru menjelaskan kisi-kisi secara terperinci tentang semua kompetensi yang telah dipelajari. Kisi-kisi soal disesuaikan dengan naskah soal yang akan diujikan. Pembelajaran yang mengacu pada Lembar Kerja Siswa (LKS) dan diselaraskan dengan silabus perlu dikelola dengan baik agar dapat menghindari materi yang tidak mempunyai relevansi dengan kompetensi yang harus dikuasai. Selain itu, kecermatan perlu ditingkatkan untuk memperhatikan indikator pada LKS yang diselaraskan dengan silabus.

Keempat, guru mempunyai antisipasi terhadap komposisi soal. Keberagaman tingkat kesulitan soal yang disusun oleh tim MGMP perlu diidentifikasi guru untuk mengetahui kategori kesulitan soal. Dengan demikian, sebaiknya guru telah merancang batasan kompetensi yang tergolong sulit, sedang, ataupun mudah sehingga diharapkan secara keseluruhan keterbacaan soal dapat diterima baik bagi siswa.

Kelima, tim MGMP menyusun kisi-kisi untuk soal UTS pada semua bidang studi. Kisi-kisi tidak hanya diperlukan untuk pelaksanaan UAS. Kisi-kisi sebagai pegangan guru dan siswa untuk mempersiapkan pendalaman materi yang akan diujikan.

Keenam, tim MGMP/ penyusun soal memperhatikan komposisi tingkat kesulitan soal. Sesuai kesepakatan dan ketentuan, 60% soal sedang, 20% soal sulit, dan 20% soal mudah. Keberadaan anak didik yang jauh dari pusat kota menjadi bahan pertimbangan penyusunan soal secara komprehensif. Diusahakan keterbacaan soal dapat diterima dengan baik bagi anak didik di seluruh lapisan sekolah.

Daftar Pustaka

- Adiputra, Ida Bagus Ragita. "Analisis Butir Soal Tes Ulangan Akhir Semester IPS Terpadu Buatan MGMPIPS Kabupaten Gianyar Kelas VII Semester 1 Tahun Pelajaran 2011-2012". *Jurnal*. Universitas Pendidikan Ganesha. Vol. 2, No. 1, Th. 2012.
- Dwipayani, Sri. 2011. "Analisis Validitas dan Reliabilitas Butir Soal Ulangan Akhir Semester Bidang Studi Bahasa Indonesia Kelas X.D SMA N 1 terhadap Pencapaian Kompetensi". *Jurnal*. Universitas Pendidikan Ganesha.
- Iskandar, Akbar. 2012. "Cara Manual Analisis Validitas Butir Soal Bentuk Objektif dan Uraian". [Http: www.akbar-iskandar.blogspot.com](http://www.akbar-iskandar.blogspot.com). Diakses pada Tanggal 3 November 2014. Pukul 13.05 WIB.
- Nuratni, Ni Putu. 2011. "Analisis Butir-Butir Soal Ulangan Umum Bahasa Indonesia Kelas X SMAK Harapan Semester Ganjil Tahun Pelajaran 2011/2012". *Jurnal*. Universitas Pendidikan Ganesha.
- Nurgiyantoro, Burhan. 2010. *Penilaian Pembelajaran Bahasa Berbasis Kompetensi*. Yogyakarta: BPF.
- Purwanto. 2011. *Evaluasi Hasil Belajar*. Yogyakarta: Pustaka Pelajar.
- Sugiyono. 2009. *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Bandung: Alfabeta.
- Suwandi, Sarwiji. 2009. *Modul Penilaian Kelas dalam Pembelajaran Bahasa Indonesia*. Surakarta: Universitas Sebelas Maret.