

ANALISIS KUALITAS TES SUMATIF BAHASA INDONESIA KELAS VIII SMPN8 DENPASAR SEMESTER GENAP TAHUN 2012

N.M Darini¹, N. Martha², G. Artawan³

^{1,2,3}Program Studi Pendidikan Bahasa, Program Pasca Sarjana
Universitas Pendidikan Ganesha, Singaraja, Indonesia

Email: made.darini@pasca.undiksha.ac.id;nengahmartha@yahoo.com;
gartawan@yahoo.com

Abstrak

Penelitian ini bertujuan mengetahui kualitas tes sumatif bahasa Indonesia kelas VIII SMPN 8 Denpasar. Rancangan penelitian bersifat ex-post facto. Subjek berjumlah 349 orang. Objeknya yaitu tes dan hasil tes sumatif. Metode dokumen digunakan untuk mengumpulkan data. Teknik analisis, yaitu statistik deskriptif. Hasil penelitian menunjukkan: (1) analisis validitas isi diperoleh validitas yang tinggi. (2) Analisis reliabilitas tes memiliki reliabilitas tinggi. (3) Analisis daya beda ditemukan 1 soal daya bedanya sangat baik, 6 soal daya bedanya baik, 29 soal daya bedanya sedang, 13 soal daya bedanya kurang baik, dan 1 soal daya bedanya tidak baik. (4) Analisis tingkat kesukaran diperoleh 42 soal klasifikasi sedang, 8 soal klasifikasi mudah. (5) Analisis kemampuan pengecoh ditemukan 49 soal pengecohnya berfungsi efektif dan 1 soal tidak berfungsi efektif.

Kata Kunci : Kualitas tes sumatif, bahasa Indonesia.

Abstract

This study aims to determine the quality of Indonesian language summative tests class VIII SMPN 8 Denpasar. The research design is ex-post factor. Subjects totaling 349 students. The object are summative test and the result summative test. Document method was used for collecting data. The data analyzed by descriptive statistic technique. The result showed: (1) content validity analyzed is obtained high content validity. (2) Test reliability analysis has high reliability. (3) In power difference analysis found a question has very well, 6 questions its power difference is good, 29 questions its power difference is medium, 13 questions its power difference is not so good, a question its power difference not good. (4) Analysis of difficulty level gained 42 medium classification questions, 8 easy classification questions. (5) Analysis capabilities rapscaillon found 49 questions the rapscaillon functioning effectively and a question is not effectively.

Keywords : Summative test quality, Indonesian language

PENDAHULUAN

Peningkatan mutu pendidikan adalah bagian terpadu dari upaya peningkatan kualitas manusia, baik aspek kemampuan, kepribadian maupun tanggung jawab sebagai warga negara. ahli-ahli pendidikan telah menyadari bahwa mutu pendidikan sangat bergantung kepada kualitas guru dan praktek pembelajarannya sehingga

peningkatan kualitas pembelajaran merupakan isu mendasar bagi peningkatan mutu pendidikan secara nasional.

Guru merupakan kunci dalam peningkatan mutu pendidikan dan mereka berada pada titik sentral dari setiap usaha reformasi pendidikan yang diarahkan pada perubahan-perubahan kualitatif. Menurut PPRI No. 19 Tahun 2005 (Pasal 28 ayat

3) tentang Standar Pendidikan Nasional menyatakan bahwa guru diharapkan memiliki 4 kompetensi. Pertama, kompetensi pedagogik, yaitu kemampuan guru untuk mengelola pembelajaran yang meliputi pemahaman terhadap peserta didik, perancangan dan pelaksanaan pembelajaran, evaluasi hasil belajar, dan pengembangan peserta didik untuk mengaktualisasikan berbagai potensi. Kedua, kompetensi keperibadian, yaitu keperibadian yang mantap, stabil, dewasa, arif, dan berwibawa menjadi teladan bagi peserta didik dan berakhlak mulia. Ketiga, kompetensi profesional, yaitu kemampuan penguasaan materi pembelajaran secara luas dan mendalam yang memungkinkannya membimbing peserta didik memenuhi standar kompetensi yang ditetapkan dalam Standar Nasional Pendidikan. Keempat, kompetensi sosial, yaitu kemampuan pendidik sebagai bagian dari masyarakat untuk berkomunikasi dan bergaul secara aktif dengan peserta didik, semua pendidik, tenaga kependidikan, orang tua atau wali dan masyarakat umum.

kompetensi pedagogik, menjelaskan bahwa seorang guru harus memiliki kemampuan untuk mengadakan evaluasi hasil belajar. Evaluasi hasil belajar dilaksanakan untuk mengukur kompetensi yaitu: kognitif, afektif, dan psikomotor peserta didik. kompetensi ini terbentuk secara berkesinambungan dalam setiap kegiatan pembelajaran. Terbentuknya kompetensi dalam setiap kegiatan pembelajaran menandakan bahwa kegiatan penilaian dilaksanakan pada setiap kegiatan pembelajaran. Penilaian akan berhasil bila dibandingkan dengan tujuan pembelajaran yang telah ditetapkan, sehingga dapat diketahui, bahwa proses belajar mengajar yang dilakukan cukup efektif memberikan hasil yang baik dan memuaskan, atau sebaliknya. Berhasil tidaknya proses pembelajaran selalu didasarkan atas kurikulum yang berlaku. Dalam kurikulum tingkat satuan pendidikan penilaian yang dilakukan adalah penilaian berbasis kelas. Penilaian berbasis kelas merupakan suatu sistem penilaian yang dilaksanakan secara terpadu dengan

kegiatan pembelajaran. Namun apa yang terjadi di lapangan masih banyak guru yang belum melaksanakan penilaian ini, penilaian hasil belajar dilaksanakan dengan tes objektif. Hal itu terbukti dengan penelitian yang dilakukan oleh Made Gosong, dkk (2005). mengungkapkan bahwa dalam pembelajaran bidang studi bahasa Indonesia (khususnya SMP), sebagian besar guru belum melakukan pembelajaran dan penilaian portofolio (*portfolio assessment*), penilaian diri (*self-assessment*) sebagai bentuk-bentuk penilaian autentik (*autentic assessment*), padahal kurikulum 2004 sudah resmi dilaksanakan sejak tahun 2004. Uji coba sudah dilaksanakan sejak tahun 2002 di sekolah-sekolah pilot project yang ditunjuk oleh Dirjen pendidikan Dasar dan Menengah. Selain hal tersebut penilaian hasil belajar yang dilaksanakan dengan tes objektif sangat merugikan siswa karena masih banyak kompetensi siswa yang belum terungkap dengan penilaian ini.

Dalam penelitian ini permasalahan yang diajukan adalah (1) bagaimanakah validitas (validitas isi) tes sumatif bahasa Indonesia kelas VIII SMP Negeri 8 Denpasar semester genap tahun pelajaran 2011/2012?, (2) bagaimanakah reliabilitas tes sumatif bahasa Indonesia kelas VIII SMP Negeri 8 Denpasar semester genap tahun pelajaran 2011/2012?, (3) bagaimanakah daya beda tes sumatif bahasa Indonesia kelas VIII SMP Negeri 8 Denpasar semester genap tahun pelajaran 2011/2012?, (4) bagaimanakah tingkat kesukaran tes sumatif bahasa Indonesia kelas VIII SMP Negeri 8 Denpasar semester genap tahun pelajaran 2011/2012? dan (5) bagaimanakah kemampuan pengecoh tes sumatif bahasa Indonesia kelas VIII SMP Negeri 8 Denpasar semester genap tahun pelajaran 2011/2012?.

Sesuai dengan rumusan masalah yang telah dirumuskan pada bagian sebelumnya, maka tujuan penelitian ini untuk mengetahui kualitas tes sumatif bahasa Indonesia kelas VIII Sekolah Menengah Pertama Negeri 8 Denpasar, yang berhubungan dengan (1) validitas (validitas isi) tes sumatif bahasa Indonesia

kelas VIII SMP Negeri 8 Denpasar semester genap tahun pelajaran 2011/2012, (2) reliabilitas tes sumatif bahasa Indonesia kelas VIII SMP Negeri 8 Denpasar semester genap tahun pelajaran 2011/2012, (3) daya beda tes sumatif bahasa Indonesia kelas VIII SMP Negeri 8 Denpasar semester genap tahun pelajaran 2011/2012, (4) tingkat kesukaran tes sumatif bahasa Indonesia kelas VIII SMP Negeri 8 Denpasar semester genap tahun pelajaran 2011/2012 dan (5) kemampuan pengecoh tes sumatif bahasa Indonesia kelas VIII SMP Negeri 8 Denpasar semester genap tahun pelajaran 2011/2012.

Keberhasilan merupakan harapan setiap orang, demikian pula bagi guru yang sedang melakukan pembelajaran. Keberhasilan tersebut akan dapat diketahui dengan melakukan evaluasi. Evaluasi merupakan suatu kegiatan yang tidak mungkin dipisahkan dari kegiatan pendidikan dan pembelajaran secara umum. Semua kegiatan pembelajaran yang dilakukan harus diikuti dan disertai dengan kegiatan evaluasi. Tanpa evaluasi, guru tidak mungkin dapat menilai dan melaporkan proses dan hasil belajar siswa secara objektif. Evaluasi mencakup dua kegiatan yaitu, mencakup pengukuran dan penilaian. Evaluasi selalu didahului dengan kegiatan pengukuran dan penilaian, sebab lewat proses tersebut evaluasi dapat dibuat dengan baik.. Dalam dunia pendidikan khususnya pembelajaran, evaluasi adalah suatu kegiatan terencana, sistematis, dan terarah berdasarkan atas tujuan yang jelas; untuk mengetahui keadaan dan hasil pembelajaran dengan menggunakan alat ukur dan hasilnya dibandingkan dengan kriteria norma untuk memperoleh keputusan-keputusan atau simpulan-simpulan (Wahyuni, 2011:10). Penilaian berarti menilai sesuatu. Sedangkan menilai itu mengandung arti mengambil keputusan terhadap sesuatu dengan mendasarkan diri atau berpegang pada ukuran baik atau buruk, sehat atau sakit, pandai atau bodoh dan sebagainya. Pengukuran yang bersifat kuantitatif itu, dapat dibedakan menjadi tiga macam, yaitu : (1) pengukuran yang dilakukan

bukan untuk menguji sesuatu, misalnya : pengukuran yang dilakukan oleh seorang penjahit pakaian, (2) pengukuran yang dilakukan untuk menguji sesuatu, misalnya : pengukuran untuk menguji daya tahan per baja terhadap tekanan berat. (3) pengukuran untuk menilai, yang dilakukan dengan jalan menguji sesuatu, misalnya : mengukur kemajuan belajar peserta didik dalam rangka mengisi nilai rapor yang dilakukan dengan menguji mereka dalam bentuk tes hasil belajar (Anas, 1995 : 4).

Djiwandono (2011:12) mendefinisikan bahwa tes adalah alat, prosedur, atau rangkaian kegiatan yang digunakan untuk memperoleh contoh tingkah laku seseorang yang memberikan gambaran tentang kemampuannya dalam suatu bidang ajaran tertentu. Dengan demikian, tes merupakan suatu bentuk pemberian tugas atau pertanyaan yang harus dikerjakan oleh siswa yang sedang diberi tes. Jawaban yang diberikan siswa terhadap pertanyaan-pertanyaan itu dianggap sebagai informasi terpercaya yang mencerminkan kemampuannya. Kegiatan tes dapat terlaksana, jika tersedia suatu perangkat tugas, pertanyaan atau latihan. Perangkat tugas, latihan, dan pertanyaan itulah yang dikenal sebagai alat tes atau instrument tes (Suwandi Sarwiji, 2011 : 48)

Setelah tes diselenggarakan, pekerjaan peserta tes pertama-tama diskor oleh kolektor yang bertugas untuk menentukan kebenaran jawaban yang diberikan oleh peserta tes terhadap semua butir tes yang telah dikerjakan. Atas dasar itu kolektor menuliskan skor, yaitu jumlah jawaban benar. Penskoran ini dapat dilakukan menurut salah satu dari dua cara yang secara mendasar berbeda berdasarkan tingkat objektivitas penskoran, yaitu (1) tes objektif, dan tes subjektif. Menurut Djiwandono (2011 : 36) Tes Objektif adalah tes yang penskorannya dapat dilakukan dengan tingkat objektivitas yang tinggi. Skor yang dihasilkan akhir penskoran tidak akan berbeda dan akan sama seandainya dilakukan oleh dua atau lebih korektor pada waktu yang berlainan. Penskoran tes objektif secara objektif

karena tersedia dan digunakannya kunci jawaban berisi daftar jawaban benar sebagai pegangan penskoran. Tes objektif dapat dituangkan dalam bentuk (a) tes menjodohkan, (b) tes benar salah, dan tes pilihan ganda. Tes menjodohkan memberi tugas kepada peserta tes untuk menjodohkan atau mencocokkan (*matching*) dua bagian tes yang dari segi isi atau arti merupakan dua bagian yang secara nalar saling berkaitan. Tes benar salah terdiri dari sejumlah butir tes, masing –masing berupa pernyataan. Beberapa diantara pernyataan itu benar dalam arti sesuai dengan yang seharusnya, beberapa yang lain berupa pernyataan yang salah, yaitu tidak sesuai atau bertentangan dengan yang seharusnya. Tes pilihan ganda adalah sejenis tes objektif yang masing-masing butir tesnya memiliki lebih dari dua pilihan jawaban. Dari semua pilihan yang tersedia, hanya satu yang benar yang merupakan jawaban kunci. Kunci jawaban adalah pilihan yang merupakan jawaban paling tepat. Pilihan-pilihan diluar jawaban yang benar itu disebut pengecoh. Sedangkan Tes dikategorikan sebagai tes subjektif apabila penskoran pekerjaan peserta tes tidak mungkin dilakukan secara objektif dan hanya dapat dilakukan secara subjektif. Pertanyaan dan tugas yang diberikan dalam tes itu dirumuskan sedemikian rupa sehingga mengundang jawaban dan pelaksanaan tugas peserta tes yang beragam dalam fokus isi, susunan kata-kata, dan panjang pendeknya jawaban.

Berdasarkan fungsi dalam penyelenggaraan pembelajaran tes dapat dibedakan menjadi; (1) tes formatif yang memberikan informasi tentang pelaksanaan sebagian dari penyelenggaraan kegiatan yang direncanakan, (2) tes sumatif yang memberikan informasi tentang pelaksanaan seluruh kegiatan pembelajaran yang direncanakan. (3) pre tes yang bertujuan untuk melakukan pengukuran terhadap tingkat kemampuan awal menjelang dimulainya kegiatan pembelajaran, dan (4) pos tes yang diselenggarakan menjelang atau pada akhir program pembelajaran untuk memastikan adanya perbedaan tingkat kemampuan pembelajaran sebagai

dampak dari seluruh kegiatan pembelajaran yang telah diselesaikan.

Dalam penyelenggaraan pembelajaran bahasa, evaluasi ditujukan terhadap tingkat kemampuan bahasa sebagai pencerminan hasil pembelajaran. Secara tidak langsung hasil evaluasi juga dimaksudkan sebagai umpan balik bagi seluruh penyelenggaraan pembelajaran yang meliputi kurikulum yang dijadikan landasan pembelajaran, bahan ajar yang digunakan, latihan dan tugas-tugas yang diberikan kepada pembelajar, juga umpan balik bagi pengajar terhadap cara mengajar, pilihan dan kesesuaian bahan ajar, tes yang digunakan. Oleh karena itu evaluasi perlu dilaksanakan dengan baik menggunakan tes sebagai alat yang bermutu serta dapat dipertanggung-jawabkan dan memiliki ciri-ciri yang dipersyaratkan, seperti : (1) validitas, (2) reliabilitas, (3) kepraktisan dan beban keuangan (Djiwandono Soenardi, 2011: 163).

Menurut Djiwandono Soenardi (2011:164). validitas itu dapat diartikan kesesuaian tes sebagai alat ukur dengan sasaran pokok yang perlu diukur. Tes bahasa yang valid sebagai alat ukur kemampuan bahasa memusatkan pengukurannya pada kemampuan bahasa peserta tes, bukan kemampuan lain, seperti pengetahuan tentang sejarah perkembangan bahasa. Hasil tes bahasa yang valid memberikan informasi tentang kemampuan bahasa peserta tesnya, bukan tingkat kemampuan lain seperti berhitung atau menyanyi. Tes kemampuan membaca akan valid, relevan, cocok, sesuai, untuk pengukuran kemampuan membaca, dan tidak untuk kemampuan berbicara atau kemampuan lainnya. Menurut jenisnya validitas dapat dibedakan menjadi: (1) validitas isi (*content validity*), (2) validitas konstruk (*construct validity*), (3) validitas konvergen, (4) validitas prediktif, (5) validitas ukuran, (6) validitas sejalan, (7) validitas butir soal. Validitas isi atau *content validity*, atau validitas tes mempersoalkan apakah isi butir teks yang diujikan mencerminkan isi kurikulum yang seharusnya diukur atau tidak, hal ini diungkapkan pula oleh Mimin

(2007:82) bahwa dalam sistem penilaian yang berbasis kompetensi di sekolah menekankan pada kesahihan isi/validitas isi yaitu menunjukkan seberapa jauh kesesuaian materi ujian dengan kompetensi dasar yang hendak diukur. Suatu alat ukur dikatakan memiliki validitas isi apabila isi materi / bahan alat ukur tersebut betul-betul merupakan bahan yang representatif terhadap bahan pelajaran yang diberikan. Artinya isi alat ukur diperkirakan sesuai dengan apa yang telah diajarkan berdasarkan kurikulum.

Reliabilitas merupakan kriteria ukuran apakah suatu alat ukur dapat mengukur secara konsisten sesuatu yang akan diukur dari waktu ke waktu. Reliabilitas merujuk pada derajat keajegan (*consistency*) alat tersebut dalam mengukur apa saja yang diukurnya. Suatu alat ukur seperti tes dikatakan memiliki reliabilitas atau keterandalan bilamana tes tersebut dipakai mengukur berulang-ulang hasilnya sama. Hal ini sesuai dengan pendapat Djiwandono (2011:170), yang mengatakan reliabilitas sebagai ciri tes yang memiliki kemampuan untuk menghasilkan pengukuran yang ajeg, tidak berubah-ubah seandainya digunakan secara berulang-ulang pada sasaran yang sama.

METODE PENELITIAN

Penelitian ini termasuk penelitian evaluasi, dengan rancangan bersifat *ex-post facto*, karena tidak melakukan manipulasi terhadap gejala yang diteliti dan gejalanya sudah ada di lapangan. Metode dokumen yang digunakan untuk mengumpulkan data. Data yang terkumpul dianalisis dengan teknik statistik deskriptif. Statistik deskriptif ini memiliki beberapa tujuan diantaranya : a) menggambarkan data atau fenomena yang diperoleh dari lapangan dalam bentuk numerik, b) meringkas data dalam bentuk gambaran, dan c) menampilkan data dalam bentuk tabel (Sukardi, 2008: 145). Sedangkan Analisis deskriptif dilakukan dengan menelaah kesesuaian isi butir-butir tes dengan isi kurikulum. Metode dokumen yang digunakan untuk mengumpulkan data ditinjau dari segi, validitas, reabilitas,

daya beda, tingkat kesulitan, dan kemampuan pengecoh. Subjek uji coba diambil dari seluruh siswa kelas VIII Sekolah Menengah Pertama Negeri 8 Denpasar tahun pelajaran 2011/2012. Metode dokumen yang digunakan untuk mengumpulkan data, selanjutnya data dianalisis. Validitas isi dianalisis dengan tabel kerja *inter observers agreement* yaitu penilaian beberapa ahli (dalam hal ini guru bahasa Indonesia) terhadap kesesuaian isi tes dengan kurikulum (standar kompetensi). Reliabilitas tes dianalisis dengan tes ulang atau tes retest dengan rumus *Product Moment Pearson*, tingkat kesukaran dianalisis dengan rumus $IF = \frac{FL+FL}{N}$, daya beda dianalisis dengan rumus $ID = \frac{FH-FL}{N}$, dan keefektifan option juga menggunakan pengelompokan atas dan bawah.

HASIL PENELITIAN DAN PEMBAHASAN

Ada lima temuan dalam penelitian ini (1) Ditemukan 4 keterampilan berbahasa sesuai dengan kemampuan yang diuji dalam kisi-kisi soal, yaitu keterampilan menyimak, membaca, menulis, dan berbicara. Berdasarkan tabel kerja *inter observers agreement* terdapat 12 soal yang tidak valid (24%) dan 38 (76 %) soal valid. Validitas isi tes sumatif bahasa Indonesia kelas VIII semester genap tahun pelajaran 2011/2012 adalah tinggi karena penyimpangan butir-butir soal sangat rendah (1,76%). (2) Terdapat hubungan yang signifikan antara hasil tes pertama dengan hasil tes kedua. Dengan metode tes retest reliabilitas butir-butir instrument tes yang diuji coba adalah 1.00 ada pada batasan reliabilitas $0.80 < r \leq 1.00$, artinya instrument butir-butir tes sumatif bahasa Indonesia di SMP Negeri 8 Denpasar memiliki reliabilitas tinggi.. (3) Dari analisis daya beda butir tes terdapat 1 butir soal (2%) daya bedanya sangat baik, 6 butir soal (12%) daya bedanya baik, 29 butir soal (58%) daya bedanya sedang, 13 soal (26%) daya bedanya kurang baik, dan 1 soal (2%) daya bedanya tidak baik/ perlu diganti. Adapun 13 butir soal yang merupakan soal kurang baik, karena $ID <$

0.19 berarti perolehan nilai dari kelompok atas kecil dikurangi perolehan nilai dari kelompok bawah yang besar, sehingga hasil menjadi kecil. Ini berarti soal tersebut daya bedanya kecil karena tidak mampu membedakan siswa yang mampu/ pandai dan siswa yang tidak mampu / kurang pandai. Demikian pula 1 soal (2%) soal yang tidak baik atau perlu direvisi/ diganti, yaitu soal nomor 26. Hal ini terjadi karena pengurangan perolehan nilai antara kelompok atas dan bawah sangat kecil $ID < 0.00$ (negatif), Ini berarti, soal tersebut tidak bisa membedakan siswa yang pandai dan tidak pandai. (4) Berdasarkan analisis tingkat kesukaran terdapat klasifikasi soal sedang 42 butir soal, klasifikasi soal mudah 8 butir soal, dan tidak ada soal yang terlalu sukar bagi siswa.

Berdasarkan hasil analisis kualitas tes sumatif bahasa Indonesia kelas VIII Sekolah Menengah Pertama Negeri 8 Denpasar semester genap tahun pelajaran 2011-2012, diperoleh temuan (1) validitas isi mencakup: identifikasi kemampuan yang diuji dalam kisi-kisi soal, mengidentifikasi 4 macam keterampilan yaitu keterampilan mendengarkan, membaca, menulis, dan berbicara. Jika dicermati dalam mata pelajaran bahasa Indonesia, maka ada 4 ruang lingkup keterampilan yang mencakup keterampilan mendengarkan, keterampilan membaca, keterampilan menulis, dan keterampilan berbicara. Sehubungan dengan adanya 4 keterampilan berbahasa yang mendukung butir-butir tes sumatif bahasa Indonesia kelas VIII SMP Negeri 8 Denpasar semester genap tahun pelajaran 2011/2012, maka terlihat jelas butir-butir tes telah memproporsikan keseimbangan 4 ruang lingkup keterampilan. Hal ini dapat dilihat dalam pemetaan SK-KD mata pelajaran bahasa Indonesia kelas VIII, mencakup 4 ruang lingkup keterampilan. Walaupun, jumlah butir-butir tes antara keterampilan mendengarkan, membaca, menulis dan berbicara tidak seimbang. Dapat disimpulkan bahwa butir-butir tes sumatif bahasa Indonesia kelas VIII semester genap di SMP Negeri 8 Denpasar tahun pelajaran 2011/2012,

sudah memenuhi kriteria soal yang baik karena mencakup 4 ruang lingkup keterampilan. Hal ini sesuai dengan paparan pada silabus bahasa Indonesia yaitu penyusunan soal yang baik sekurang-kurangnya harus mencakup 50% ruang lingkup dari jumlah ruang lingkup keterampilan yang ditetapkan. Adapun ruang lingkup keterampilan pada mata pelajaran bahasa Indonesia, yaitu : keterampilan membaca, keterampilan menulis, keterampilan menyimak, dan keterampilan berbicara. (2) Kesesuaian kemampuan yang diuji dengan standar kompetensi diperoleh kesesuaian dari 50 soal ada 38 soal (76 %), Sedangkan soal-soal yang tidak sesuai 12 soal (24%). Sehubungan dengan penilaian tiga (3) observers diperoleh hasil dari 50 butir instrumen berupa tes, diperoleh 97 yang mengatakan “ ya´ artinya observer setuju, bahwa tes tersebut sesuai dengan kurikulum / standar kompetensi yang telah ditetapkan sebelumnya, dan 53 menjawab “tidak”, artinya observer menilai bahwa tes tersebut tidak sesuai dengan standar kompetensi yang telah ditetapkan sebelumnya. Berdasarkan penilaian itu 12 soal yang tidak valid artinya penyimpangan yang terjadi 24 %, maka dapat disimpulkan bahwa validitas isi tes sumatif bahasa Indonesia kelas VIII SMP Negeri 8 Denpasar semester genap tahun pelajaran 2011/2012 dapat dikatakan tinggi, karena penyimpangan butir-butir soal yang tidak sesuai dengan standar kompetensi dalam kisi-kisi soal adalah 12 (24 %). Seharusnya tidak ada soal yang tidak sesuai dengan kisi-kisi mengingat peruntukan butir-butir soal sebagai tes sumatif. Hasil tes sumatif untuk menentukan keberhasilan belajar siswa dengan kriteria ketuntasan minimal (KKM) yang telah ditentukan dan tercantum dalam KTSP SMP Negeri 8 Denpasar. Hal ini sesuai dengan Peraturan Kementrian Pendidikan dan Kebudayaan nomor 20 tahun 2007 tentang Standar Penilaian Pendidikan. (2) Sesuai dengan kriteria reliabilitas instrumen, maka 1.00 ada pada batasan reliabilitas $0.80 < r \leq 1.00$, artinya instrument butir-butir tes sumatif bahasa Indonesia di SMP Negeri 8

Denpasar memiliki reliabilitas tinggi. Ini berarti bahwa tes yang disusun guru sudah baik dan dapat dipergunakan lagi pada siswa lain yang mempunyai kondisi yang relatif sama. Hal ini sejalan dengan yang dikatakan Candiasa (dalam Nuratni 2011:151) bahwa soal yang baik adalah soal yang reliabilitasnya tinggi. Walaupun tes ini sudah baik, dapat dipergunakan kembali pada siswa lain yang mempunyai kemampuan yang relatif sama, perlu ditingkatkan kembali kemampuan dan pemahaman Bapak/ibu guru SMP Negeri 8 Denpasar tentang standar kompetensi dan kompetensi dasar yang dijabarkan kembali dalam kisi-kisi soal sehingga soal yang dibuat sesuai dengan standar kompetensi yang telah ditetapkan sebelumnya, mengingat soal buatan guru yang digunakan untuk menentukan hasil belajar siswa. Hasil belajar siswa ini digunakan untuk membuat laporan hasil belajar, diselain itu guru adalah orang yang berwenang membuat soal sesuai dengan kemampuannya dan bertanggung jawab terhadap kemajuan pendidikan siswa-siswinya, baik secara individual maupun klasikal. Dalam hal ini guru harus memahami kompetensi guru dan segala seluk beluknya. Kompetensi tersebut merupakan modal bagi guru untuk merencanakan, melaksanakan, dan mengevaluasi proses pembelajaran sehingga tujuan pendidikan dapat tercapai dengan baik. (3) Dalam Analisis daya beda terdapat 1 butir soal (2%) daya bedanya sangat baik, 6 butir soal (12%) daya bedanya baik, 29 butir soal (58%) daya bedanya sedang, 13 soal (26%) daya bedanya kurang baik, dan 1 soal (2%) daya bedanya tidak baik/ perlu diganti. Adapun 13 butir soal yang merupakan soal kurang baik karena $ID < 0.19$. Perolehan nilai dari kelompok atas kecil dikurangi perolehan nilai dari kelompok bawah yang besar, sehingga hasil menjadi kecil. Ini berarti soal tersebut daya bedanya kecil karena tidak mampu membedakan siswa yang mampu atau pandai dan siswa yang tidak mampu atau kurang pandai. Demikian pula 1 soal (2%) soal yang tidak baik atau perlu direvisi/ diganti, yaitu soal nomor 26. Hal ini terjadi

karena pengurangan perolehan nilai antara kelompok atas dan bawah sangat kecil $ID < 0.00$ (negatif), Ini berarti, soal tersebut tidak bisa membedakan siswa yang pandai dan siswa yang tidak pandai. (4) Analisis tingkat kesukaran terdapat klasifikasi soal sedang 42 butir soal, klasifikasi soal mudah 8 butir soal, dan tidak ada soal yang terlalu sukar bagi siswa. Jika dicermati, maka terdapat beberapa kelemahan, yaitu : Proporsi soal mudah (16%), soal katagori sedang (84%), dan soal sulit (0%), tidak menunjukkan proporsi yang seimbang. Kurangnya pengetahuan guru tentang proporsi soal mudah, sedang, dan sukar. Menurut Sri Wahyunu, (2012 : 51) proporsi soal yang mudah, sedang, dan sukar berkisar 30%, 50%, dan 20%. Bila ditinjau kembali ketentuan standar proses pada KTSP SMP Negeri 8 Denpasar, bahwa siswa dapat mengikuti proses pembelajaran pada semester berikutnya, apabila memenuhi standar KKM (KKM pada proses pembelajaran semester bersangkutan tuntas), bila kurang atau tidak tuntas diupayakan mengikuti remidi. Ini berarti secara tidak langsung terjadi pengkatrolan nilai sehingga nilai meningkat sesuai batas KKM (peningkatan kualitas yang semu). Pengkatrolan nilai, dalam penilaian ini sangat tidak adil karena hanya berdasarkan atas patokan mutlak, yaitu nilai yang didasarkan atas jumlah soal yang dapat dijawab benar oleh siswa, tanpa memperhatikan kompetensi lain yang dimiliki siswa dalam proses pembelajaran (penilaian autentik assessment). (5) Analisis kemampuan / keefektifan pengecoh ditemukan dari 50 soal, ada 49 butir-butir soal yang pengecohnya sudah berfungsi dengan efektif, dan 1 butir soal pengecohnya tidak berfungsi efektif. Ini terjadi karena jumlah siswa kelompok atas dan siswa kelompok bawah tidak ada yang benar memilih option tersebut (0%), sehingga soal tersebut tidak bisa membedakan siswa yang mampu atau pandai dan tidak mampu atau kurang pandai. Berdasarkan perbandingan hasil interpretasi itu, maka tes sumatif bahasa Indonesia semester genap tahun pelajaran 2011 / 2012 secara

umum berfungsi dengan efektif. Namun ironis, ternyata tes sumatif yang dirancang dan dipersiapkan berdasarkan kisi-kisi, terdapat pilihan jawaban/pengecoh yang tidak baik. Hal ini memunculkan keinginan siswa untuk berspekulasi bekerja sama dengan siswa yang lain.

PENUTUP

Berdasarkan hasil analisis dapat disimpulkan sebagai berikut: (1) Validitas isi tes sumatif bahasa Indonesia kelas VIII semester genap di SMP Negeri 8 Denpasar, masih ada yang perlu direvisi. Hal ini dapat dilihat dari tabel kerja yang menyatakan kemampuan yang diuji dalam kisi-kisi / standar kompetensi ada 12 soal (24%) yang tidak sesuai dengan standar kompetensi. Sedangkan 38 soal (76%) sesuai dengan standar kompetensi (valid), (2) Reliabilitas tes sumatif bahasa Indonesia kelas VIII di SMP Negeri 8 Denpasar reliabilitasnya tinggi, sehubungan dengan menggunakan metode tes retest / tes ulang dengan rumus Product Moment Pearson menghasilkan reliabilitas tes $0,80 < r \leq 1,00$, memiliki reliabilitas tinggi atau keajegan yang tinggi, (3) Daya beda butir-butir tes sumatif bahasa Indonesia kelas VIII di SMP Negeri 8 Denpasar sudah baik karena dari 50 butir tes, terdapat 1 soal (2%) daya bedanya sangat baik, 6 tes (12%) daya bedanya baik, 29 soal (58%) daya bedanya sedang, 13 soal (26%) daya bedanya kurang baik/ perlu perbaikan, dan 1 soal (2%) daya bedanya tidak baik perlu diganti (4) Tingkat kesukaran tes sumatif bahasa Indonesia kelas VIII di SMP Negeri 8 Denpasar, diperoleh proporsi sedang 42 soal (82%), proporsi mudah 8 soal (16%), dan tidak ditemukan soal klasifikasi sukar. (5) Kemampuan pengecoh tes sumatif bahasa Indonesia di kelas VIII SMP Negeri 8 Denpasar, sudah cukup baik, karena dari 50 butir soal hanya terdapat 1 soal (2%) yang pengecohnya tidak berfungsi dengan efektif karena jumlah pemilih dari kelompok atas dan kelompok bawah sama sehingga tidak dapat membedakan siswa yang pintar dan bodoh/ kurang pintar.

Terkait dengan simpulan terhadap temuan-temuan dalam penelitian ini, disarankan kepada pakar bahasa dan pakar pendidikan bahasa agar memanfaatkan temuan penelitian ini sebagai sumbangan dalam usaha memperkaya khasanah ilmu pengetahuan terutama dalam bidang evaluasi yang terkait dengan kualitas tes, khususnya tentang validitas isi, reliabilitas, daya beda, tingkat kesukaran, dan kemampuan pengecoh. Temuan ini diharapkan juga dapat digunakan bagi guru sebagai standarisasi untuk mengkaji butir-butir tes. Guru diharapkan dapat menetapkan kriteria objektif dalam melakukan penilaian hasil belajar peserta didik, khususnya dalam pembelajaran bahasa Indonesia, dan bagi sekolah untuk meningkatkan hasil belajar, dan pedoman untuk mengetahui apakah yang dilakukan sekolah sudah memenuhi standar atau belum. Pemenuhan standar akan terlihat dari nilai-nilai evaluasi yang diperoleh siswa. Sedangkan bagi praktisi pendidikan untuk menyempurnakan populasi dan sampel yang akan diteliti, sehingga diperoleh sampel yang lebih valid dan bermanfaat bagi pengembangan penilaian atau evaluasi hasil belajar

DAFTAR PUSTAKA

- Arikunto, Suharsini. 2001. *Prosedur Penilaian Suatu Pendekatan Praktek*. Jakarta: Rineka Cipta.
- Akhadiah, Sabarti. 1988. *Evaluasi Dalam Pengajaran Bahasa*. Jakarta: Depdikbud Dirjen Pendidikan Tinggi Proyek Pengembangan Lembaga Pendidikan Tenaga Pendidikan.
- Bachman, Lyle F. 1990. *Fundamental Consideration in Language testing*. Oxford: Oxford University Press.
- Badan Standar Nasional Pendidikan. 2006. *Panduan Penyusunan KTSP : Jenjang Pendidikan Dasar dan Menengah*. Jakarta : BNSP.

- Candiasa, I Made. 2010. *Pengujian Instrumen Penelitian Disertai Aplikasi Iteman dan Bigsteps*. Singaraja : Unit Penerbit Undhiksa.
- Depiknas. 2004. *Kurikulum Tingkat Satuan Pendidikan*. Jakarta : Depdiknas.
- _____. 2004. *Pedoman Khusus Pengembangan Silabus dan Penilaian*. Jakarta: Depdiknas.
- Djiwandono, M Sunardi. 1996. *Tes Bahasa Dalam pengajaran*. Bandung:Tarsito.
- Djiwandono, M Soenardi. 2011. *Tes Bahasa Pegangan Bagi Pengajar Bahasa*. Jakarta: Indeks.
- Haryati, Mimin. 2007. *Model & Teknik Penilaian pada Tingkat Satuan Pendidikan*. Jakarta: GP PRESS.
- Mahadi, Ni Made. 2007. *Penerapan Evaluasi Diri dalam Meningkatkan Keterampilan Menulis Teks Pidato Siswa Kelas X₄ Semester II SMA Negeri 1 Kediri Tahun Pelajaran 2006/2007*. Laporan Penelitian (tidak diterbitkan). Kediri.
- Margono. 2000. *Metodologi Penelitian Pendidikan*. Jakarta: Rineka Cipta.
- Martha, I Ngh. 2012 "Analisis Tes Secara Emperis Melalui Pretesting". Makalah disampaikan dalam seminar Guru-Guru SD, Denpasar.
- _____. 2007. " *Penilaian Autentik (Authentic Assessment) Dalam Proses Belajar Mengajar pada Pendidikan Dasar dan Menengah*" Orasi (tidak diterbitkan) Singaraja: Undiksha.
- Nababan, Sri Utari Subiakto. 1993. *Metodologi pengajaran Bahasa*. Jakarta: PT Gramedia Pustaka Utama
- Nuratni, Ni Putu. 2012. *Analisis Butir-Butir Soal Ulangan Umum Bahasa Indonesia Kelas X SMK Harapan Denpasar Semester Ganjil Tahun Pelajaran 2011/2012*. Tesis (tidak diterbitkan). Program Pascasarjana
- _____. 2012. *Peraturan Menteri Pendidikan Nasional Nomor 23 Tahun 2006 tentang Standar Kompetensi Lulusan*. Jakarta: Dinas Pendidikan Nasional.
- _____. 2012. *Peraturan Menteri Pendidikan Nasional Nomor 22 Tahun 2006 tentang Standar Isi*. Jakarta : Dinas Pendidikan Nasional.
- _____. 2012. *Peraturan Menteri Pendidikan Nasional Nomor 20 Tahun 2007 tentang Standar Penilaian Pendidikan*. Jakarta : Dinas Pendidikan Nasional.
- _____. 2005. *Peraturan Pemerintah Nomor 19 Tahun 2005 tentang Standar Nasional Pendidikan*. Jakarta : Dinas Pendidikan Nasional.
- Oller, Jhon. 1979. *Language test*. London: Longman Group Limited.
- Pardini, Ni Putu. 2006. *Analisis Tes Bahasa Indonesia Siswa kelas I Sekolah Menengah Atas Negeri Ubud Tahun 2004/2005 dari segi Pendekatan Komunikatif dan Integratif*. Tesis (tidak diterbitkan). Program Pascasarjana
- Purwanto. 2011. *Evaluasi hasil Belajar*. Surakarta : Pustaka Pelajar.
- Ridwan, Akdon. 2008. *Rumus dan Data dalam Analisis Statistika*. Bandung. alfabeta.
- Setyosari, Punaji. 2010. *Metode Penelitian Pendidikan dan Pengembangan*. Malang: Kencana.
- Sirat, Bistok. 1989. *Bahan Pengajaran Untuk Mata Kuliah Evaluasi Hasil Belajar Siswa*. Jakarta: Departemen Pendidikan dan Kebudayaan Direktorat Jendral Pendidikan Tinggi Proyek Pengembangan Lembaga Pendidikan Tenaga Kependidikan.

Sudijono, Anas. 1995. *Pengantar Evaluasi Pendidikan*. Yogyakarta. PT Raja Grafindo.

Sukardi. 2003. *Metodologi Penelitian Pendidikan*. Yogyakarta. Bumi Aksara.

_____. 2008. *Evaluasi Pendidikan Prinsip & Operasionalnya*. Yogyakarta: Bumi Aksara.

Suwandi, Sarwiji. 2011. *Model-Model Asesmen Dalam Pembelajaran*. Surakarta: Yuma Pustaka.

Suwarno, Bambang. 2008. *Rumus dan Data dalam Analisis Statistika*. Bandung: Alfabeta.

Thoha, M. Chabib. 2003. *Teknik Evaluasi Pendidikan*. Jakarta : Divisi Buku Perguruan Tinggi PT Raja Grafindo Persada.

Wahyuni, Ibrahim. 2012. *Assesmen Pembelajaran Bahasa*. Malang: Aditama.