

**KARAKTERISTIK SOAL UASBN MATA PELAJARAN BAHASA INDONESIA  
DI DAERAH ISTIMEWA YOGYAKARTA PADA TAHUN PELAJARAN 2008/2009**

***CHARACTERISTICS OF THE TEST ITEMS IN THE NATIONAL STANDARD SCHOOL  
FINAL EXAMINATION OF THE INDONESIAN LANGUAGE SUBJECT  
IN THE YOGYAKARTA SPECIAL TERRITORY IN THE ACADEMIC YEAR OF 2008/2009***

Deri Anggraini, Pujiati Suyata  
Universitas PGRI Yogyakarta, Universitas Negeri Yogyakarta  
derianggraini@yahoo.com, pujiati\_suyata@uny.ac.id

**Abstrak**

Penelitian ini bertujuan menggambarkan karakteristik soal UASBN Bahasa Indonesia di DIY tahun 2008/2009. Perangkat tes ditelaah dari aspek materi, konstruksi, dan bahasa. Analisis kuantitatif berdasarkan teori tes klasik menggunakan *MicroCAT IteMan* versi 3.00 dan teori tes modern menggunakan *MicroCAT Bigsteps* versi 2.30. Hasil penelitian menunjukkan bahwa: (1) secara teoretis, 100% butir soal berkualitas baik dari aspek materi, konstruksi, dan bahasa, (2) secara empiris dengan teori tes klasik, 11 butir soal berkualitas baik (22%) dengan rerata tingkat kesukaran 0,759 (mudah), rerata daya beda 0,701 (sangat baik), distribusi respons yang berfungsi 32%, dan indeks reliabilitas 0,933 (handal), (3) secara empiris dengan teori respons butir, 39 butir soal berkualitas baik (78%) dengan rerata tingkat kemampuan peserta 1,92 (tinggi), 49 butir soal cocok, fungsi informasi tes 87,76%, dan 44 butir (89,80%) butir soal dengan tingkat kesukaran sedang, (4) secara teoretis dan empiris, 8 butir soal berkualitas baik (16%), (5) tingkat konsistensi hasil analisis teori tes klasik dan teori respons butir sebesar -0,068 (inkonsisten).

**Kata kunci:** karakteristik soal, UASBN, analisis kualitatif, IteMan, Bigsteps

**Abstract**

*This study aims to reveal characteristics of the test items in the National Standard School Final Examination (NSSFE) of the Indonesian language subject in the Yogyakarta Special Territory in the academic year of 2008/2009. The test set was reviewed in terms of the relevance level from the aspects of materials, construction, and language. The students' responses to the test set were quantitatively analyzed in accordance with classical test theory using the programs of MicroCAT IteMan version 3.0 and modern test theory using the programs of MicroCAT Bigsteps version 2.30. The results of the analysis show that: (1) From the theoretical aspect as a whole the test items have good characteristics because, of the 50 test items tested, all or 100% are in the good category in terms of the aspects of materials, construction, and language. (2) Based on the classical test theory analysis using the IteMan program, of the 50 test items, 11 test items have good quality (22%) with an average difficulty index of 0.759 (easy), an average discrimination index of 0.701 (very good), functional response distribution (32%), and a reliability index of 0.933 (reliable). (3) Based on the item response theory analysis with 1 parameter (Rasch Model) using the Bigsteps program, of the 50 test items, 39 test items have good quality (78%) with testees' average ability of 1.92 (high), 49 test items fit Rasch Model, 43 test items (87.76%) of 49 items have good item information, and 4 test items (89.80%) have moderate difficulty indices. (4) Based on the theoretical aspect, classical test theory, and item response theory of the 50 test items, 8 test items have good quality (16%). (5) The number of the poor test items based on classical test theory is higher than that based on item response theory; the proportion is 39 to 11, with a consistency level of 0.068 (inconsistent).*

**Keywords:** item characteristics, NSSFE, qualitative analysis, IteMan, Bigsteps

## Pendahuluan

Pendidikan pada semua jalur, jenjang, dan jenis diselenggarakan sebagai satu kesatuan yang sistemik dengan sistem yang terbuka dan multimakna; diselenggarakan sebagai suatu proses pembudayaan dan pemberdayaan peserta didik yang berlangsung sepanjang hayat; diselenggarakan dengan memberi keteladanan, membangun kemauan, dan mengembangkan kreativitas peserta didik dalam proses pembelajaran; diselenggarakan dengan mengembangkan budaya membaca, menulis, dan berhitung bagi segenap warga masyarakat; dan diselenggarakan dengan memberdayakan semua komponen masyarakat melalui peran serta dalam penyelenggaraan dan pengendalian mutu layanan pendidikan. Pemerintah, pemerintah daerah, dewan pendidikan, dan komite sekolah/madrasah melakukan pengawasan atas penyelenggaraan pendidikan sesuai dengan kewenangan masing-masing. Untuk menjamin mutu pendidikan nasional, pemerintah menentukan kebijakan nasional dan Standar Nasional Pendidikan.

Dalam Undang-Undang Nomor 20 Tahun 2003 tentang Sistem Pendidikan Nasional, evaluasi diatur dalam Bab XVI Pasal 57, 58, dan 59. Evaluasi pada dasarnya merupakan penafsiran atau interpretasi yang sering bersumber pada data kuantitatif (Sudijono, 2008, p.5). Pelaksanaan evaluasi bertujuan untuk mengukur dan mengendalikan mutu pendidikan. Evaluasi dilakukan untuk memantau proses, kemajuan, dan perbaikan hasil belajar peserta didik secara berkesinambungan. Penjabaran pelaksanaan evaluasi dinyatakan dalam Peraturan Pemerintah Nomor 19 Tahun 2005 tentang Standar Nasional Pendidikan. Standar Nasional Pendidikan terdiri atas standar isi, proses, kompetensi lulusan, tenaga kependidikan, sarana dan prasarana, pengelolaan, pembiayaan, dan penilaian pendidikan yang harus ditingkatkan secara berencana dan berkala. Pada pasal 63 ayat (1) disebutkan bahwa penilaian pendidikan pada jenjang pendidikan dasar dan menengah terdiri atas: (a) penilaian hasil belajar oleh pendidik, (b) penilaian hasil belajar oleh satuan pendidikan, dan (c) penilaian hasil belajar oleh pemerintah.

Evaluasi mencakup dua kegiatan yaitu pengukuran dan penilaian. Pengukuran merupakan suatu proses pemberian angka kepada individu-individu secara sistematis sebagai sebuah alat untuk merepresentasikan properti individu-

individu tersebut (Allen & Yen, 1979, p.2). Pengukuran dibatasi pada kegiatan membandingkan sesuatu dengan alat ukur yang sesuai secara sistematis. Dalam kegiatan pengukuran terdapat seperangkat aturan yang harus dipenuhi untuk mendapatkan hasil ukur yang tepat (Thorndike, 2005, p.25). Penilaian adalah pengambilan keputusan berdasarkan terhadap hasil pengukuran, sehingga penilaian bersifat kualitatif. Penilaian dalam proses pembelajaran antara lain sebagai kegiatan menghimpun fakta-fakta dan dokumen belajar peserta didik yang dapat dipercaya untuk melakukan perbaikan program (Surapranata, 2007, p.2). Penilaian hasil belajar oleh pendidik dan satuan pendidikan merupakan bentuk evaluasi internal (*internal evaluation*), sedangkan penilaian hasil belajar oleh pemerintah merupakan bentuk evaluasi eksternal (*external evaluation*). Penilaian hasil belajar peserta didik pada jenjang Sekolah Dasar (SD), pada tahun pelajaran 2008/2009 ditetapkan sistem penilaiannya dalam bentuk Ujian Akhir Sekolah Berstandar Nasional (UASBN) sebagaimana ditetapkan dalam Peraturan Menteri Pendidikan Nasional Nomor 82 Tahun 2008.

UASBN merupakan salah satu kegiatan pengukuran dan penilaian kompetensi peserta didik secara nasional untuk jenjang SD. UASBN dilaksanakan secara terintegrasi dalam pelaksanaan ujian sekolah/madrasah untuk SD/MI/SDLB. Dalam UASBN tahun pelajaran 2008/2009 ada tiga mata pelajaran yang diujikan secara nasional yaitu Bahasa Indonesia, IPA, dan Matematika. UASBN SD/MI/SDLB Tahun 2008/2009 diatur dalam Peraturan Menteri Pendidikan Nasional Nomor 82 Tahun 2008. Dalam Peraturan Menteri Pendidikan Nasional Nomor 82 Tahun 2008 tentang UASBN untuk SD/MI/SDLB Tahun 2008/2009 Pasal 3 disebutkan bahwa tujuan UASBN adalah (a) menilai pencapaian kompetensi lulusan secara nasional pada mata pelajaran Bahasa Indonesia, Matematika, dan IPA; dan (b) mendorong tercapainya target wajib belajar pendidikan dasar yang bermutu.

Materi UASBN adalah materi ujian yang sesuai dengan Standar Kompetensi Lulusan UASBN yang 25% soalnya dipersiapkan oleh Pusat, dan 75% soal disiapkan oleh. Penyusunan soal UASBN tahun 2008/2009 berdasarkan pada (a) Peraturan Menteri Pendidikan Nasional Republik Indonesia Nomor 23 Tahun 2006 tentang Standar Kompetensi Lulusan untuk Satuan Pendidikan Dasar dan Mene-

ngah dan (b) kisi-kisi soal UASBN yang terlampir dalam Peraturan Menteri Pendidikan Nasional Nomor 82 Tahun 2008. Soal UASBN berbentuk pilihan ganda. Tes bentuk pilihan ganda adalah tes yang jawabannya dapat diperoleh dengan memilih alternatif jawaban yang telah disediakan. Pengembangan tes prestasi belajar mengikuti langkah-langkah standar dalam konstruksi tes. Karakteristik tes yang baik, yaitu (1) valid, (2) reliabel, (3) objektif, (4) praktis, dan (5) ekonomis.

Dalam Peraturan Menteri Pendidikan Nasional Nomor 82 Tahun 2008 (Depdiknas, 2008) disebutkan bahwa hasil UASBN digunakan sebagai salah satu pertimbangan untuk: (a) pemetaan mutu satuan pendidikan, (b) dasar seleksi masuk jenjang pendidikan berikutnya, (c) penentuan kelulusan peserta didik dari satuan pendidikan, dan (d) dasar pembinaan dan pemberian bantuan kepada satuan pendidikan dalam upaya peningkatan mutu pendidikan. Dari Seksi Data dan Teknologi Informasi, Dinas Pendidikan, Pemuda, dan Olahraga, Daerah Istimewa Yogyakarta diperoleh informasi bahwa sekolah yang mempunyai rata-rata nilai UASBN pada tahun 2008/2009 dari 48.463 peserta pada tiga mata pelajaran adalah 7,3. Rata-rata masing-masing mata pelajaran adalah 7,8 (Bahasa Indonesia), 6,9 (Matematika), dan 7,3 (IPA). Rata-rata nilai Bahasa Indonesia tertinggi 9,28 dan terendah 4,75.

Deskripsi data nilai UASBN di Daerah Istimewa Yogyakarta menunjukkan bahwa hasil UASBN pada Mata Pelajaran Bahasa Indonesia di beberapa sekolah di Daerah Istimewa Yogyakarta pada tahun pelajaran 2008/2009 masih rendah. Rendahnya hasil UASBN dapat disebabkan oleh kemampuan mengajar guru rendah, kemampuan belajar siswa yang rendah, fasilitas belajar yang kurang, kualitas soal ujian yang kurang baik, acuan ujian yang tidak baku, atau gabungan dari kelima faktor tersebut. Kelima faktor tersebut disebut oleh Mardapi (1999) sebagai faktor penyebab rendahnya Ebtanas. Faktor kemampuan mengajar guru, kemampuan belajar siswa, dan fasilitas belajar siswa merupakan faktor yang berpengaruh dalam proses pembelajaran pra-UASBN. Acuan ujian yang tidak baku merupakan faktor penyebab rendahnya hasil UASBN yang berkaitan dengan kebijakan penyelenggaraan ujian. Faktor penyebab yang berkaitan dengan pelaksanaan ujian dan menarik untuk diteliti adalah kualitas soal ujian.

Soal UASBN Mata Pelajaran Bahasa Indonesia di Daerah Istimewa Yogyakarta pada

tahun pelajaran 2008/2009 perlu diketahui kualitasnya secara keseluruhan, baik yang dibuat oleh pusat maupun wilayah. Perangkat tes harus benar-benar berkualitas, baik secara kualitatif (teoretis) maupun kuantitatif (empiris). Kualitas tes secara kualitatif dapat dilihat dari sisi materi, konstruksi, dan bahasa. Kualitas tes secara kuantitatif dapat ditinjau berdasarkan dua teknik yaitu dengan menggunakan teori tes klasik dan teori tes modern.

Teori Tes Klasik (TTK) atau *Classical Test Theory* (CTT) merupakan teori yang menggunakan asumsi-asumsi yang cukup sederhana dan mudah dipahami (Mardapi, 1999, p.6). Asumsi-asumsi dalam teori tes klasik menurut Allen & Yen (1979, pp.57–60) sebagai berikut: (1) terdapat hubungan antara skor tampak (*observed score*) atau  $X$ , skor murni (*true score*) atau  $T$ , dan skor kesalahan (*error score*) yang dilambangkan dengan  $E$ , (2) nilai harapan ( $\epsilon(X)$ ) adalah skor murni ( $T$ ), (3) tidak ada korelasi antara skor kesalahan dan skor murni yang diperoleh dari populasi subjek, (4) tidak ada korelasi antara skor kesalahan tes pertama dan skor kesalahan tes kedua, (5) tidak ada korelasi antara skor kesalahan tes pertama dan skor murni tes kedua, (6) dua perangkat tes merupakan tes yang paralel jika populasi subjek yang menempuh kedua tes tersebut mendapat skor murni yang sama ( $T = T'$ ) dan varian skor-skor kesalahannya sama ( $=$ ), dan (6) jika dua perangkat tes mempunyai skor tampak ( $X_1$  dan  $X_2$ ) yang memenuhi asumsi 1 sampai 5 dan apabila untuk setiap populasi subjek,  $T_1 = T_2 + c_{12}$ ,  $c_{12}$  adalah sebuah bilangan konstanta, maka kedua tes itu disebut tes yang setara.

*Item Response Theory* (IRT) atau *Latent Trait Theory* (LTT), atau *Characteristics Curve Theory* (CCT) merupakan suatu teori yang dikembangkan atas dasar dua postulat yaitu: (1) performansi peserta tes pada suatu butir soal dapat diprediksikan (atau dijelaskan) oleh seperangkat faktor yang disebut *trait*, *latent trait*, atau kemampuan dan (2) hubungan performansi peserta tes pada suatu butir soal dan perangkat *trait* yang mendasarinya dapat digambarkan oleh fungsi naik monoton yang disebut fungsi karakteristik butir atau kurva karakteristik butir (Hambleton, Swaminathan, & Rogers, 1991, p.7). IRT dilandasi tiga asumsi yaitu (1) unidimensi, (2) independensi lokal, dan (3) fungsi karakteristik butir atau kurva karakteristik butir.

Menurut Hambleton, Swaminathan, & Rogers (1991, p.12), tiga macam model logistik

dalam teori respons butir yaitu model logistik satu parameter, model logistik dua parameter, dan model logistik tiga parameter. Model logistik satu parameter (Model Rasch) menggunakan satu parameter untuk menganalisis data yaitu parameter tingkat kesukaran soal. Kelebihan IRT dibandingkan CTT yaitu (1) IRT tidak berdasarkan *group dependent*, (2) skor siswa dideskripsikan bukan *test dependent*, (3) IRT menekankan pada tingkat butir soal bukan tes, (4) IRT tidak memerlukan paralel tes untuk menentukan reliabilitas tes, (5) IRT memerlukan suatu pengukuran ketepatan untuk setiap skor tingkat kemampuan.

Analisis kuantitatif dibedakan secara klasik dan modern. Aspek analisis butir soal secara klasik adalah validitas, reliabilitas, tingkat kesukaran butir, daya pembeda butir, penyebaran pilihan jawaban (untuk soal bentuk objektif), dan kesalahan pengukuran. Analisis kuantitatif butir soal secara klasik dapat dilakukan dengan menggunakan program komputer *Item and Test Analysis (Iteman)* versi 3.00. Hayat (1997, p.1) menerangkan bahwa program *Iteman* dibuat dengan pendekatan klasik yang dapat digunakan untuk menganalisis file data dengan format ASCII (*American Standard Code for Information Interchange*). Program ini mampu menganalisis maksimal 250 butir soal dengan 3000 responden dan memberikan informasi tentang statistik butir soal, pilihan jawaban, dan tes.

Analisis kuantitatif butir soal secara modern dengan teori respons butir bersifat umum karena komposisi sampel dapat mengestimasi parameter dan tingkat kesukaran soal tanpa bias. Analisis dengan model Rasch diarahkan kepada kecocokan butir soal dengan model yang dipilih (*Rasch Model*) dan estimasi tingkat kesukaran butir soal, sehingga dapat memberikan informasi apakah suatu butir soal cocok dengan model Rasch (*fit statistic*). Kriteria data sesuai dengan model Rasch adalah apabila hasil korelasi *pointbiserial* positif dan *outfit*-nya  $< 2,00$  baik *outfit* butir soal maupun *outfit* peserta tes. Dilihat dari daya pembeda, butir soal yang mempunyai daya beda (*Ptbis*) positif berarti cocok dengan model Rasch, sedangkan butir soal yang mempunyai daya beda (*Ptbis*) negatif berarti tidak cocok dengan model Rasch. *Outfit* butir soal merupakan statistik butir soal yang menunjukkan bagaimana perilaku yang tidak diharapkan dari orang yang mempunyai kemampuan lebih dengan tingkat kesukaran yang bersangkutan. *Outfit* peserta tes merupakan statistik orang yang menunjukkan

bagaimana perilaku yang tidak diharapkan pada butir soal yang mempunyai tingkat kesukaran jauh dengan kemampuan orang yang bersangkutan. Butir soal yang cocok berarti berperilaku secara konsisten dengan apa yang diharapkan oleh model. Butir soal yang memiliki *out fit*  $> 2,00$  dan daya beda (*Ptbis*) negatif ditolak dan dinyatakan tidak cocok dengan model Rasch. Teori respons butir menggunakan istilah informasi untuk menyatakan reliabilitas dan validitas butir soal ataupun tes

Rumusan masalah penelitian adalah sebagai berikut: (1) Bagaimanakah karakteristik butir soal UASBN Mata Pelajaran Bahasa Indonesia di Daerah Istimewa Yogyakarta pada tahun pelajaran 2008/2009 ditinjau dari aspek teoretis? (2) Bagaimanakah karakteristik butir soal UASBN Mata Pelajaran Bahasa Indonesia di Daerah Istimewa Yogyakarta pada tahun pelajaran 2008/2009 ditinjau dari aspek empiris?

Penelitian ini bertujuan untuk mendeskripsikan karakteristik butir soal UASBN Mata Pelajaran Bahasa Indonesia di Daerah Istimewa Yogyakarta pada tahun pelajaran 2008/2009 ditinjau dari aspek teoretis dan empiris. Kontribusi penelitian ini bagi dunia pendidikan adalah dapat memberikan gambaran tentang karakteristik butir soal UASBN Mata Pelajaran Bahasa Indonesia di Daerah Istimewa Yogyakarta pada tahun pelajaran 2008/2009, baik secara teoretis maupun empiris. Selain itu, secara praktis, hasil penelitian dapat digunakan sebagai masukan guna peningkatan kualitas soal dan peningkatan kontrol terhadap alat evaluasi yang digunakan pada ujian nasional masa mendatang.

## Metode

Penelitian ini merupakan penelitian kuantitatif noneksperimental dengan menggunakan metode deskriptif. Desain penelitian ini adalah *cross sectional*. Penelitian ini dilaksanakan di Daerah Istimewa Yogyakarta. Subjek penelitian ini adalah butir soal UASBN Bahasa Indonesia buatan tim penyusun soal dan sumber datanya adalah lembar jawaban siswa atas 50 butir soal pilihan ganda. Populasi dalam penelitian ini adalah respons siswa peserta tes UASBN mata pelajaran Bahasa Indonesia tahun pelajaran 2008/2009 di Daerah Istimewa Yogyakarta sebanyak 48.463 siswa dari 2009 sekolah (1514 SD Negeri dan 495 SD Swasta). Populasi terdiri atas 11.360 respons siswa dari 368 SD di Kabupaten Bantul, 10.585 respons

siswa dari 558 SD di Kabupaten Gunungkidul, 6.086 respons siswa dari 380 SD di Kabupaten Kulon Progo, 13.200 respons siswa dari 512 SD di Kabupaten Sleman, dan 7.232 respons siswa dari 191 SD di Kotamadya Yogyakarta.

Sampel respons siswa sejumlah 655 ditentukan dengan teknik *stratified sampling* dari 48.463 siswa di 2009 sekolah. Variabel utama penelitian ini adalah karakteristik butir soal Bahasa Indonesia yang digunakan pada UASBN di Daerah Istimewa Yogyakarta. Sub-variabel penelitian ini adalah tingkat kesukaran butir, daya beda butir, distribusi respons, validitas tes, indeks keandalan tes, fungsi informasi, tingkat kecocokan butir tes dengan model, dan kesalahan baku pengukuran.

**Hasil Penelitian dan Pembahasan**

Hasil analisis kualitatif menunjukkan bahwa semua (100%) butir soal UASBN Mata Pelajaran Bahasa Indonesia di Daerah Istimewa Yogyakarta pada tahun ajaran 2008/2009 yang terdiri dari lima puluh butir soal memenuhi kriteria dari aspek materi, konstruksi, dan bahasa. Hasil tersebut menunjukkan bahwa lima puluh butir soal UASBN Mata Pelajaran Bahasa Indonesia di Daerah Istimewa Yogyakarta pada tahun ajaran 2008/2009 berkategori baik.

Hasil analisis kuantitatif dengan *Iteman* menunjukkan bahwa butir soal UASBN Mata Pelajaran Bahasa Indonesia di Daerah Istimewa Yogyakarta pada tahun ajaran 2008/2009 yang berkualitas baik dari aspek tingkat kesukaran, daya beda, dan distribusi jawaban sebanyak butir soal yang berkualitas baik ada 11 butir (22%) dan 39 butir soal (78%) tidak baik. Enam belas butir atau 32% butir soal memenuhi kriteria yang ditentukan pada aspek tingkat kesukaran, yaitu butir soal nomor 1, 5, 12, 20, 28, 29, 30, 31, 34, 39, 41, 42, 43, 44, 47, dan 49. Sebanyak 49 butir atau 98% butir soal memenuhi kriteria yang ditentukan pada aspek daya beda. Yang tidak memenuhi daya beda adalah butir soal nomor 6. Sebanyak 16 butir atau 32% memiliki distribusi jawaban yang baik.

Berdasarkan kriteria tingkat kesukaran butir soal yang baik dari 0,30 sampai dengan 0,70, jumlah butir yang baik untuk perangkat tes UASBN Mata Pelajaran Bahasa Indonesia di Daerah Istimewa Yogyakarta pada tahun ajaran 2008/2009 sebagai berikut: terdapat 33 butir soal (66%) memiliki tingkat kesukaran lebih besar dari 0,70 (kategori mudah), 16 butir soal (32%) memiliki tingkat kesukaran antara

0,30 dan 0,70 (sedang), dan 1 butir soal (2%) memiliki tingkat kesukaran lebih kecil dari 0,30 (kategori sukar). Tingkat kesukaran disajikan dalam Tabel 1.

Tabel 1. Tingkat Kesukaran Butir Soal

Kategori	Jumlah	%	Butir Soal
Mudah $p > 0,70$	33	66	2, 3, 4, 6, 7, 8, 9, 10, 11, 13, 14, 15, 16, 17, 18, 19, 21, 22, 23, 24, 25, 26, 27, 32, 33, 35, 36, 37, 38, 45, 46, 48, 50
Sedang $0,30 \leq p \leq 0,70$	16	32	1, 5, 12, 20, 28, 29, 30, 31, 34, 39, 41, 42, 43, 44, 47, 49
Sukar $(p < 0,30)$	1	2	40

Dengan menggunakan kriteria indeks daya beda soal yang dinyatakan dengan koefisien *biserial* dengan menggunakan kriteria batas minimum daya beda yang diterima adalah  $\geq 0,40$  ada 29 butir (58%) soal UASBN mata pelajaran Bahasa Indonesia di Daerah Istimewa Yogyakarta pada tahun ajaran 2008/2009 memiliki kategori sangat baik, 20 butir (40%) kategori baik, 1 butir (2%) kategori cukup, dan tidak ada butir yang tidak baik. Daya pembeda disajikan dalam Tabel 2.

Tabel 2. Daya Pembeda Butir Soal

Kategori	Jumlah	%	Butir Soal
Sangat Baik $0,70 \leq D \leq 1,00$	29	58	1, 2, 3, 4, 5, 8, 9, 10, 13, 15, 18, 20, 24, 26, 27, 30, 31, 32, 33, 34, 36, 37, 38, 39, 41, 42, 45, 46, 48
Baik $0,40 \leq D \leq 0,69$	20	40	7, 11, 12, 14, 16, 17, 19, 21, 22, 23, 25, 28, 29, 35, 40, 43, 44, 47, 49, 50
Cukup Baik $0,20 \leq D \leq 0,39$	1	2	6
Tidak Baik $D < 0,20$	0	0	-

Pada butir soal UASBN Mata Pelajaran Bahasa Indonesia di Daerah Istimewa Yogyakarta pada tahun ajaran 2008/2009 terdapat distraktor yang belum berfungsi baik yaitu soal nomor 2, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 17, 18, 19, 21, 22, 23, 25, 26, 27, 29, 32, 33, 35, 36, 37, 38, 41, 45, 46, 48, 49, dan 50. Hasil analisis menunjukkan bahwa 68% soal UASBN Mata Pelajaran Bahasa Indonesia di Daerah Istimewa Yogyakarta pada tahun ajaran 2008/2009 distribusi responsnya kurang baik dan perlu direvisi.

Tabel 3. Statistik Tes dengan Program *Iteman*

Aspek	Hasil
<i>N of Items</i>	50
<i>N of Examinees</i>	655
<i>Mean</i>	37,945
<i>Variance</i>	93,744
<i>Std. Deviasi</i>	9,682
<i>Skew</i>	-0,368
<i>Kurtosis</i>	-1,204
<i>Minimum</i>	10,000
<i>Maximum</i>	50,000
<i>Median</i>	36,000
<i>Alpha</i>	0,933
<i>SEM</i>	2,502
<i>Mean P</i>	0,759
<i>Mean Item-Tot.</i>	0,477
<i>Mean Biserial</i>	0,701

Berdasar Tabel 3 dapat dinyatakan bahwa dari 50 jumlah butir soal yang dikerjakan oleh 655 peserta tes, hasilnya menunjukkan rata-rata skor 37,945 dengan sebaran skor 93,744 dan simpangan baku 9,682. Kemiringan distribusi skor -0,368 (kemiringan negatif) menunjukkan bahwa sebagian besar skor berada pada bagian atas (skor tinggi) dari distribusi skor. Kurtosis bernilai -1,204 menunjukkan distribusi skor landai (merata). Skor minimum 10,000 dan maksimum 50,00. Indeks reliabilitas tes 0,933 berarti handal ( $0,933 > 0,700$ ), rerata tingkat kesukaran 0,759 berarti mudah (tidak baik) karena tingkat kesukaran tes yang baik (sedang) adalah 0,30-0,70 dengan kesalahan baku pengukuran 2,502. Rerata daya beda 0,701 berarti daya beda butir tes berfungsi/baik ( $0,701 > 0,40$ ).

Hasil analisis butir soal dengan menggunakan *Bigsteps* menunjukkan butir soal UASBN Mata Pelajaran Bahasa Indonesia di Daerah Istimewa Yogyakarta pada tahun ajaran 2008/2009 yang berkualitas baik ada 39 butir (78%) yaitu butir soal nomor 1, 3, 4, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 23, 24,

25, 26, 29, 30, 31, 32, 34, 35, 36, 37, 38, 39, 41, 42, 45, 46, 47, 48, dan 49. Butir soal UASBN Mata Pelajaran Bahasa Indonesia di Daerah Istimewa Yogyakarta pada tahun ajaran 2008/2009 yang berkualitas tidak baik ada 11 butir soal (22%) yaitu nomor 2, 5, 6, 22, 27, 28, 33, 40, 43, 44, dan 50.

Tabel 4. Statistik Variabel Analisis dengan *Bigsteps*

No	Statistik	
A. Statistik Peserta Tes Peserta		
1	Rerata Tingkat Kemampuan	1,92
2	Deviasi Standar	1,63
3	Rerata Kesalahan Estimasi	0,50
4	Deviasi Standar Kesalahan Estimasi	0,21
5	Rerata Skor	37,6
6	Deviasi Standar Skor	9,6
7	Fungsi Informasi	0,89
B. Statistik Butir Tes		
1	Rerata Tingkat Kemampuan	0,00
2	Deviasi Standar	1,32
3	Rerata Kesalahan Estimasi	0,13
4	Deviasi Standar Kesalahan Estimasi	0,05
5	Rerata Skor	481,1
6	Deviasi Standar Skor	98,8
7	Fungsi Informasi	0,99

Berdasarkan hasil analisis program *Bigsteps* diperoleh soal sebanyak 49 butir (98%) yang cocok dengan model (*fit*) dan yang tidak cocok dengan model sebanyak 1 butir (2%). Berdasarkan kriteria yang telah ditentukan, perangkat soal UASBN Mata Pelajaran Bahasa Indonesia di Daerah Istimewa Yogyakarta pada tahun ajaran 2008/2009 memiliki kualitas baik.

Tabel 5. Kecocokan Butir Soal dengan Model Rasch

Kriteria	Nomor Soal	Jumlah	%
Outfit < 2,00	1, 2, 3, 4, 5, 7, 8,	49	98
Atau Ptbis (+)	9, 10, 11, 12, 13,		
(cocok)	14, 15, 16, 17, 18,		
	19, 20, 21, 22, 23,		
	24, 25, 26, 27, 28,		
	29, 30, 31, 32, 33,		
	34, 35, 36, 37, 38,		
	39, 40, 41, 42, 43,		
	44, 45, 46, 47, 48,		
	49, 50		
Outfit > 2,00	6	1	2
Atau Ptbis (-)			
(tidak cocok)			
Jumlah	50	50	100

Tingkat kesukaran dalam model Rasch dikelompokkan dalam tiga kategori yaitu (a) kategori Sukar, dengan indeks kesukaran >2,00, (2) sedang, dengan indeks kesukaran -2,00 sampai dengan 2,00, dan (3) mudah, dengan indeks kesukaran < -2,00. Butir soal yang baik adalah yang memiliki tingkat kesukaran sedang sejumlah 44 (89,90%) berada antara -2,00-2,00.

Tabel 6. Distribusi Tingkat Kesukaran Butir Soal yang Cocok (Model Rasch)

Kriteria	Nomor Soal	Jumlah	%
Mudah < -2,00	33,50	2	4,08
Sedang -2,00 – 2,00	1, 2, 3, 4, 5, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 34, 35, 36, 37, 38, 39, 41, 42, 45, 46, 47, 48, 49	44	89,8
Sukar > 2,00	40, 43, 44	3	6,12
Jumlah		49	100%

Tingkat kemampuan tertinggi peserta UASBN Mata Pelajaran Bahasa Indonesia di Daerah Istimewa Yogyakarta pada tahun ajaran 2008/2009 dalam merespon 50 soal adalah 5, 44 dimiliki oleh 11 peserta yaitu, *Testee 394, Testee 395, Testee 396, Testee 397, Testee 399, Testee 401, Testee 402, Testee 403, Testee 404, Testee 405, dan Testee 407* dengan perolehan skor 50 dan tingkat kemampuan terendah adalah -1,79 dimiliki oleh *Testee 262* perolehan skor 10. Tidak ada butir soal yang mempunyai daya beda (*Ptbis*) yang negatif sehingga semua butir tersebut cocok dengan model Rasch karena memenuhi kriteria butir soal yang baik.

Tabel 7. Butir yang Ditolak dan Diterima pada Teori Tes Klasik dan Teori Respons Butir

Aspek	Pendekatan Analisis			
	Teori Tes Klasik		Teori Respons Butir	
	Ditolak	Diterima	Ditolak	Diterima
Nomor Butir	2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 21, 22, 23, 24, 25, 26, 27, 29, 32, 33, 35, 36, 37, 38, 40, 41, 45, 46, 48, 49, 50	1, 20, 28, 30, 31, 34, 39, 42, 43, 44, 47	2, 5, 6, 22, 27, 28, 33, 40, 43, 44, 50	1, 3, 4, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 23, 24, 25, 26, 29, 30, 31, 32, 34, 35, 36, 37, 38, 39, 41, 42, 45, 46, 47, 48, 49
Jumlah	39	11	11	39

Harga fungsi informasi tes dihitung dengan menggunakan program Microsoft Excel 2007. Harga fungsi informasi butir yang memiliki nilai maksimum > 0,7 ada 43 dari 49 butir yang cocok dengan model Rasch atau 87,76%, dan yang memiliki nilai maksimum < 0,7 ada 6 atau 12,24%. Butir soal yang memiliki nilai maksimum < 0,7 yaitu butir soal nomor 2 (0,6908), nomor 15 (0,6932), nomor 22 (0,6908), nomor 27 (0,6978), nomor 28 (0,6956), dan nomor 44 (0,6978).

Hasil analisis butir soal dengan menggunakan *Bigsteps* menunjukkan butir soal UASBN Mata Pelajaran Bahasa Indonesia di Daerah Istimewa Yogyakarta pada tahun ajaran 2008/2009 yang berkualitas baik ada 39 butir (78%) yaitu butir soal nomor 1, 3, 4, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 23, 24, 25, 26, 29, 30, 31, 32, 34, 35, 36, 37, 38, 39, 41, 42, 45, 46, 47, 48, dan 49. Butir soal UASBN Mata Pelajaran Bahasa Indonesia di Daerah Istimewa Yogyakarta pada tahun ajaran 2008/2009 yang berkualitas tidak baik ada 11 butir soal (22%) yaitu nomor 2, 5, 6, 22, 27, 28, 33, 40, 43, 44, dan 50.

Hasil analisis menunjukkan bahwa tingkat konsistensi ( $\Phi$ ) perangkat tes UASBN Bahasa Indonesia di Daerah Istimewa Yogyakarta pada tahun ajaran 2008/2009 adalah sebesar -0,068. Angka tersebut menunjukkan bahwa butir yang baik atau tidak baik menurut teori tes klasik dan modern tidak konsisten. Banyak butir yang ditolak menurut teori klasik, diterima menurut teori respons butir. Nomor dan jumlah butir yang ditolak dan diterima dengan analisis teori tes klasik dan teori respons butir tersaji dalam Tabel 7. Estimasi Tabel *Phi* tersaji dalam Tabel 8.

Tabel 8. Estimasi Tabel *Phi* Naskah UASBN Mata Pelajaran Bahasa Indonesia di Daerah Istimewa Yogyakarta pada Tahun Ajaran 2008/2009

		<i>Teori Tes Klasik</i>		
		0	1	
<i>Teori Respons Butir</i>	0	2, 5, 6, 22, 27, 33, 40, 50	28, 43, 44	<i>a+b</i>
		<i>a=8</i>	<i>b=3</i>	<i>11</i>
	1	3, 4, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 21, 23, 24, 25, 26, 29, 32, 35, 36, 37, 38, 41, 45, 46, 48, 49	1, 20, 30, 31, 34, 39, 42, 47	<i>c+d</i>
		<i>c=31</i>	<i>d=8</i>	<i>39</i>
		<i>a+c=39</i>	<i>b+d=11</i>	<i>50</i>

Butir soal yang baik dalam penelitian ini adalah butir soal UASBN Bahasa Indonesia di Daerah Istimewa Yogyakarta tahun ajaran 2008/2009 yang memenuhi persyaratan secara kuantitatif dan kualitatif. Butir soal yang berkualitas baik ada 8 butir (16%) dan yang tidak baik ada 42 butir soal (84%).

Tabel 9. Butir Soal yang Baik Berdasarkan Analisis Kuantitatif dan Kualitatif

Butir Soal	Jumlah	%
1, 20, 30, 31, 34, 39, 42, 47	8	16

## Simpulan dan Saran

### Simpulan

Berdasarkan hasil penelitian dan pembahasan, dapat disimpulkan sebagai berikut: (1) perangkat tes UASBN Bahasa Indonesia tahun pelajaran 2008/2009 di Daerah Istimewa Yogyakarta ditinjau dari aspek teoretis secara keseluruhan memiliki karakteristik yang baik, karena dari 50 butir soal yang diujikan semuanya termasuk kategori baik atau 100%, baik dari aspek materi, konstruksi, maupun bahasa. (2) Perangkat tes UASBN Bahasa Indonesia tahun pelajaran 2008/2009 di Daerah Istimewa Yogyakarta berdasarkan analisis teori tes klasik dengan program *Iteman* dari 50 butir soal yaitu 11 butir soal kualitasnya baik (22%). Hal ini dilihat dari rerata tingkat kesukaran 0,759 (mudah), rerata daya beda 0,701 (sangat baik), distribusi respon yang berfungsi 32%, dan indeks reliabilitas 0,933 (handal). (3) Perangkat tes UASBN Bahasa Indonesia tahun pelajaran 2008/2009 di Daerah Istimewa Yogyakarta berdasarkan analisis teori respons butir 1 parameter (*Rasch model*) dengan program *Bigsteps* menunjukkan bahwa dari 50 butir soal yaitu 39 butir soal kualitasnya baik (78%). Hal ini dilihat dari rerata tingkat kemampuan peserta 1,92 (tinggi), 49 butir soal cocok dengan model

*Rasch*, fungsi informasi butir yang baik sebanyak 43 dari 49 butir soal (87,76%), dan butir soal yang memiliki tingkat kesukaran sedang ada 44 butir (89,80%). (4) Perangkat tes UASBN Bahasa Indonesia tahun pelajaran 2008/2009 di Daerah Istimewa Yogyakarta yang tidak baik menurut teori tes klasik lebih banyak dibanding teori respons butir yaitu 39 berbanding 11. Tingkat konsistensi sebesar -0,068 (tidak konsisten).

### Saran

Berdasarkan beberapa simpulan penelitian di atas, dapat dikemukakan saran-saran sebagai berikut: (1) Kemampuan tim penyusun soal UASBN Daerah Istimewa Yogyakarta secara keseluruhan sudah baik. Meskipun demikian, masih tetap perlu ditingkatkan karena kesalahan teknik penulisan alternatif jawaban yang sama secara berulang masih dijumpai pada soal UASBN Bahasa Indonesia tahun 2008/2009. (2) Peningkatan kemampuan guru di sekolah dalam membuat soal agar tingkat kesukaran, daya beda, distraktor butir soal pilihan ganda, dan reliabilitasnya baik, dapat dilakukan dengan pelatihan penyusunan soal dan analisisnya dengan menggunakan program *iteman* serta pembinaan rutin di KKG maupun pertemuan gugus. Hal ini menjadi tanggung jawab Dinas Pendidikan, Pemuda, dan Olahraga Daerah Istimewa Yogyakarta dengan difasilitasi oleh LPMP selaku penjamin mutu pendidikan. (3) Peningkatan kemampuan guru di sekolah atau tim penyusun soal dalam membuat soal agar butir soal cocok dengan model, tingkat kesukarannya sedang, dan fungsi informasi butirnya lebih baik, dapat dilakukan dengan pelatihan penyusunan soal dan analisisnya menggunakan program *Bigsteps* serta pembinaan rutin di KKG maupun pertemuan gugus. Hal ini juga menjadi tanggung jawab Dinas Pendidikan, Pemuda, dan Olahraga Daerah Istimewa Yogyakarta dengan difasilitasi oleh LPMP selaku pen-

jamin mutu pendidikan. (4) Referensi analisis kuantitatif dengan pendekatan teori tes modern, khususnya *Bigsteps* masih terbatas. Mengingat pentingnya kegiatan analisis soal dengan pendekatan teori tes modern, para praktisi pendidikan yang berkompoten dapat terus menyumbangkan ilmu dan pengetahuannya, baik melalui media cetak maupun elektronik. (5) Rendahnya konsistensi soal UASBN Bahasa Indonesia tahun pelajaran 2008/2009 di Daerah Istimewa Yogyakarta disebabkan butir yang diterima dan ditolak pada teori tes klasik berbeda dengan teori respons butir. Guru perlu melatih keterampilan menyusun soal, baik soal untuk tes formatif, maupun sumatif.

#### Daftar Pustaka

- Allen, M. J., & Yen, W. M. (1979). *Introduction to measurement theory*. Monterey: Brooks/Cole Publishing Company.
- BSNP (2008) *Keputusan Badan Standar Pendidikan Nasional Nomor 1514/BSNP-XII/2008 tentang Prosedur Operasi Standar UASBN untuk SD/MI/SDLB Tahun Pelajaran 2008/2009*. Diambil pada tanggal 16 Januari 2009, dari <http://www.bsnp-indonesia.org>.
- Depdiknas. (2003). *Undang-Undang Republik Indonesia Nomor 20, Tahun 2003, tentang Sistem Pendidikan Nasional*.
- Depdiknas. (2005). *Peraturan Pemerintah Republik Indonesia Nomor 19, Tahun 2005, tentang Standar Nasional Pendidikan*.
- Depdiknas. (2006). *Peraturan Pemerintah Republik Indonesia Nomor 23, Tahun 2006, tentang Standar Kompetensi Lulusan untuk Satuan Pendidikan Dasar dan Menengah*.
- Depdiknas. (2008). *Peraturan menteri pendidikan nasional nomor 82, Tahun 2008, tentang UASBN untuk SD/MI/SDLB Tahun 2008/2009*.
- Hambleton, R.K., & Swaminathan, H., & Rogers, H.J. (1993). *Fundamentals of item response theory*. Newbury Park, CA: Sage Publications, The International Professional Publishers.
- Hartono. (2009). *Statistik untuk penelitian*. Yogyakarta: Pustaka Pelajar.
- Iskandarwassid & Sunendar, D. (2008). *Strategi pembelajaran bahasa*. Bandung: Remaja Rosdakarya.
- Mardapi, D. (1999). *Estimasi kesalahan pengukuran dalam bidang pendidikan dan implikasinya pada ujian nasional*. Pidato Pengukuhan Guru Besar, di Universitas Negeri Yogyakarta.
- Sudijono, A. (2008). *Pengantar evaluasi pendidikan*. Jakarta: PT RajaGrafindo Persada.
- Surapranata, S. (2007). *Panduan penulisan tes tertulis: implementasi kurikulum 2004*. Bandung: PT Remaja Rosdakarya.
- Thorndike, R. M. (2005). *measurement and evaluation in psychology and education (7<sup>th</sup> Ed.)*. New York: Merrill Prentice Hall.