

## **KARAKTERISTIK DAN NILAI FUNGSI INFORMASI BUTIR SOAL UJIAN AKHIR SEKOLAH DASAR DI KOTA BENGKULU**

*Oleh:*

*Nyayu Masyita Ariani*

### **Abstrak**

Pendidikan di sekolah dasar (SD), merupakan dasar bagi jenjang pendidikan selanjutnya. Sebagai *mastery test* bagi calon tamatan SD, soal ujian akhir sekolah dasar (UASD) harus berkarakteristik baik dan dapat dipertanggungjawabkan. Penelitian ini bertujuan untuk mendeskripsikan karakteristik soal serta nilai fungsi informasi butir soal UASD di Kota Bengkulu tahun 2004 .

Dengan teknik *purposive sampling*, dipilih soal tes mata uji Matematika, bahasa Indonesia, dan Ilmu Pengetahuan Sosial (IPS) sebagai sampel soal beserta respons peserta ujian terhadap soal tersebut. Tiap mata uji tersebut terdiri dari 35 soal pilihan ganda, 10 soal isian singkat, dan 5 soal uraian. Respons peserta ujian terhadap soal pilihan ganda dianalisis dengan pendekatan teori respons butir 3 parameter, sedangkan soal isian singkat dan uraian dianalisis dengan pendekatan teori tes klasik. Sampel respons terhadap soal UASD diambil dari respons peserta ujian dari gradasi sekolah yang berbeda.

Hasil penelitian mengungkapkan bahwa karakteristik butir soal UASD yang tergolong baik pada bentuk soal pilihan ganda untuk mata uji matematika, bahasa Indonesia, dan IPS, secara berturut-turut adalah: 24, 5, dan 1 butir; pada bentuk isian singkat, soal yang berkarakteristik baik ada 5, 4, dan 5 butir; sedangkan pada soal bentuk uraian ada 4, 1, dan 2 butir soal.

**Kata kunci:** *soal, karakteristik, fungsi-informasi.*

## **Pendahuluan**

Penyelenggaraan ujian akhir pada akhir program pendidikan di Indonesia telah beberapa kali mengalami perubahan format ujian. Mulai dari ujian yang diselenggarakan sendiri oleh sekolah pada era sebelum tahun 80-an, dan pada era tahun 80-an diselenggarakan ujian yang bersifat nasional yang kita kenal dengan Ebtanas (evaluasi belajar tahap akhir nasional). Masing-masing penyelenggaraan ujian ini dilatar-belakangi oleh dasar pemikiran yang sangat mendasar pada eranya.

Apapun, kenyataannya penyelenggaraan Ebtanas Sekolah Dasar (SD) sejak tahun 2002 telah dihapus melalui SK Mendiknas No. 011/U/2002. Sebagai penggantinya, diselenggarakanlah Ujian Akhir Sekolah Dasar (UASD). Hal ini tertuang dalam SK Mendiknas No.012/U/2002. Penyelenggaraan UASD sebagai pengganti Ebtanas ini dimaksudkan agar penilaian yang selama ini bersifat sentralistik dan kurang memberdayakan sekolah dapat lebih dipercayakan pada pihak yang berada paling depan dalam dunia pendidikan, yaitu guru dan sekolah. Pemberian kepercayaan dan kewenangan kepada guru dan pihak sekolah untuk melaksanakan penilaian, merupakan konsekuensi dari agenda reformasi pendidikan nasional.

Penyelenggaraan UASD bertujuan untuk menilai hasil belajar siswa di sekolah dan untuk mempertanggungjawabkan penyelenggaraan pendidikan kepada masyarakat (Depdiknas, 2002). Untuk mencapai tujuan tersebut, tentunya penyelenggaraan UASD harus dikelola dengan manajemen yang baik, terutama dalam penyediaan bahan ujian, yang harus dapat dipertanggungjawabkan kualitasnya.

Namun, kenyataannya tidak semua daerah dan sekolah siap melaksanakan UASD. Ini karena beberapa hal, antara lain kondisi daerah yang berbeda-beda. Kualitas dan kuantitas sumber daya manusia yang masih terbatas. Sarana dan prasarana di daerah yang belum mendukung. Belum lagi ditambah dengan biaya yang diperlukan untuk penyelenggaraan UASD yang tidak sedikit. Berkaitan dengan ini, bagaimanakah penyelenggaraan UASD di Kota Bengkulu? Berdasarkan wawancara awal dengan beberapa sumber di Dinas Pendidikan Nasional Kota Bengkulu, diperoleh informasi bahwa penyelenggaraan UASD

dapat dilaksanakan, namun dalam penyusunan soal masih dirasakan berat, hal ini karena belum tersedianya bank soal UASD, sumber daya manusia, serta sarana dan prasarana yang belum memadai. Oleh sebab itu penyelenggaraan UASD di Kota Bengkulu belum dikelola oleh masing-masing sekolah, namun oleh Dinas Pendidikan Nasional Kota Bengkulu. Alasan lain adalah untuk mengukur mutu pendidikan SD di tingkat Kota Bengkulu. Soal-soal untuk keperluan UASD dibuat oleh panitia penyusun soal. Persoalannya apakah soal-soal UASD tersebut telah memiliki karakteristik yang baik dan dapat dipertanggung-jawabkan?

Berdasarkan hal ini, maka perlu kiranya diteliti tentang karakteristik dan nilai fungsi informasi butir soal UASD yang digunakan. Hal ini diharapkan dapat memberikan masukan yang konseptual kepada pihak yang terkait dalam rangka penyempurnaan penyajian soal UASD pada masa datang. Berdasarkan ini, maka rumusan masalah di dalam penelitian ini adalah, bagaimanakah karakteristik butir serta nilai informasi butir soal UASD di Kota Bengkulu?

Tujuan pendidikan dasar sebagaimana yang tertuang di dalam Pasal 3 Peraturan Pemerintah Nomor 28 tahun 1990 tentang Pendidikan Dasar, adalah untuk memberikan bekal kemampuan dasar kepada siswa untuk mengembangkan kehidupannya sebagai pribadi, anggota masyarakat, warga negara dan anggota umat manusia serta mempersiapkan siswa untuk mengikuti pendidikan menengah. Adapun tujuan pendidikan dasar yang diselenggarakan di sekolah dasar adalah bertujuan memberikan bekal kemampuan dasar “Baca-Tulis-Hitung”, pengetahuan dan keterampilan dasar yang bermanfaat bagi siswa sesuai dengan tingkat perkembangannya serta mempersiapkan mereka untuk mengikuti pendidikan di Sekolah Lanjutan Tingkat Pertama (Depdikbud, 1994/1995:3). Di samping kemampuan dasar tadi, pengetahuan sosial juga sangat diperlukan, sebagaimana di sampaikan Djemari Mardapi (1999:22), “Pada jenjang sekolah dasar, kemampuan utama yang dibutuhkan adalah membaca, menulis dan berhitung. Selain itu pengembangan wawasan peserta didik dalam bidang sosial sangat dibutuhkan.” Ketiga pengetahuan ini diperoleh siswa SD melalui mata pelajaran matematika, bahasa Indonesia, dan IPS ini, merupakan pengetahuan minimal yang harus dimiliki siswa SD. Berdasarkan hal ini,

selanjutnya soal UASD tiga mata pelajaran ini yang dibahas di dalam penelitian ini.

Pada akhir suatu penyelenggaraan program pendidikan, termasuk di SD, diselenggarakan ujian akhir. Ujian akhir sekolah (UAS) dimaksudkan antara lain untuk menilai hasil belajar siswa di sekolah, mempertanggung-jawabkan penyelenggaraan pendidikan pada masyarakat, dan untuk mengetahui mutu pendidikan pada sekolah. Selain itu hasil penilaian juga berfungsi sebagai bahan pertimbangan di dalam menentukan kelulusan siswa. Untuk itu, alat penilaian yang berupa soal-soal yang digunakan untuk keperluan UASD harus disusun melalui prosedur yang sistematis dan harus memiliki karakteristik yang memenuhi syarat dan dapat dipertanggung-jawabkan.

Adanya persyaratan prosedur yang sistematis maupun aturan tertentu dalam mengobservasi perilaku seseorang termasuk menilai hasil belajar seseorang, dinyatakan oleh Mehrens & Lehmann (1973:6) berikut, "*... a systematic procedure for observing a persons's behavior and describing it with the aid of a numerical scale or a category-system.* Oleh sebab itu, berdasarkan hal ini, suatu tes harus disusun secara sistematis dengan aturan-aturan yang jelas, sehingga informasi yang diperoleh berupa hasil pengukuran melalui tes dapat akurat, atau paling tidak mendekati keadaan sesungguhnya dan dapat dipertanggungjawabkan.

Di dalam dunia pendidikan, tujuan diberikan tes adalah untuk mengukur penguasaan peserta didik terhadap pelajaran yang telah diajarkan. Tes semacam ini selanjutnya dikenal dengan tes prestasi (Tim Pusiasjian, 1997). Menilik fungsinya, tes prestasi belajar ditempatkan sebagai alat bantu untuk mengambil keputusan pendidikan. Sebagai tes yang berfungsi untuk sertifikasi, ujian akhir tentu berbeda dengan tes untuk fungsi lainnya seperti tes untuk dignosa, atau tes penempatan . Di dalam penyajiannya, secara tertulis soal-soal tes prestasi belajar biasanya disajikan dalam bentuk objektif, dan uraian. Hal ini tergantung dari sisi kemampuan yang mana yang akan diukur dari peserta tes.

Di dalam dunia pengukuran pendidikan dikenal dua teori tes, yaitu teori tes klasik dan teori respons butir. Masing-masing teori ini memiliki keunggulan dan keterbatasan yang khas. Penerapan teori inipun sesuai kebutuhan yang disesuaikan dengan keunggulan dan keterbatasannya. Keunggulan teori tes

klasik terletak pada kemudahan pemahaman konsep dan penggunaannya, sehingga sering digunakan. Namun kelemahan teori ini terletak pada hasil estimasi parameter yang tergantung pada karakteristik peserta ujian (*group dependent*), dan hasil estimasi kemampuan peserta tergantung pada karakteristik butir (*item dependent*). Butir yang dirasakan sukar bagi seorang peserta ujian mungkin saja dirasakan tidak sukar oleh peserta ujian yang lain (Hambleton, et. al., 1991; Azwar, 2001). Berdasarkan uraian ini jelas bahwa teori tes klasik memiliki ciri bahwa karakteristik butir tes tidak dapat dipisahkan dari karakteristik peserta tes. Bila tes yang sama dikerjakan oleh peserta tes yang berbeda, maka karakteristik soal itu pada umumnya berubah (Naga, 1992:4).

Di dalam teori respons butir, tidak dikenal kelemahan-kelemahan yang disebutkan di atas, karakteristik soal ujian tidak tergantung peserta ujian. Soal ujian tetap mempunyai karakteristik yang sama walaupun peserta ujiannya berbeda. Menurut Hambleton, et. al. (1991:7), hal yang mendasar dari teori respons butir adalah pertama, respons seorang subyek (peserta tes) pada suatu butir dapat memprediksi seperangkat faktor yang disebut *trait* (kemampuan) yang dimilikinya. Kedua, hubungan antara kemampuan subjek pada suatu butir (hasil tes) dan kemampuan yang mendasarinya dapat digambarkan oleh suatu fungsi yang menaik secara monoton, yang disebut kurva karakteristik butir (*item characteristic curve (ICC)*). Hal yang mendasar dari teori respons butir adalah fungsi matematika serta hubungan probabilitas respons terhadap butir oleh subyek untuk menentukan karakteristik dari subyek dan butir.

Untuk mengestimasi kemampuan peserta dibutuhkan parameter yang membentang sepanjang garis kontinum variabel yang diukur. Garis kontinum ini bergerak dari sebelah kiri yang menunjukkan kemampuan kurang, menuju ke sebelah kanan yang menunjukkan kemampuan tinggi, sedangkan untuk mengestimasi butir membutuhkan parameter butir. Parameter butir ini terdiri dari  $a_i$  (daya beda),  $b_i$  (tingkat kesukaran) dan  $c_i$  (*pseudo guessing*) (Naga, 1992:12). Tentang besarnya parameter  $c_i$  (peluang tebakan) ini tergantung dari banyaknya pilihan jawaban pada butir soal tersebut. Jika banyaknya pilihan jawaban dalam satu butir soal dimisalkan  $k$ , maka nilai  $c_i$  maksimal adalah sebesar  $1/k$  (Mehrens & Lehmann, 1973: 280).

Selanjutnya, secara matematis, model logistik 3-parameter (3P) dapat dinyatakan sebagai berikut (Hambleton, R. K & Swaminathan, H., 1985:49)

$$P_i(\theta) = c_i + (1 - c_i) \frac{e^{D a_i (\theta - b_i)}}{1 + e^{D a_i (\theta - b_i)}} \dots\dots\dots (1)$$

- dengan  $i = 1, 2, 3, \dots, n$ ,
- $D$  : Konstanta yang bernilai 1,7
- $a_i$  : Parameter daya beda butir ke- $i$
- $b_i$  : Parameter tingkat kesukaran butir ke- $i$
- $c_i$  : *Pseudo guessing* butir ke- $i$
- $e$  : Bilangan transendental yang nilainya 2,718
- $n$  : Banyaknya butir soal dalam tes

Dimana Fungsi Informasi. Di dalam teori tes klasik, kehandalan tes ditunjukkan dengan koefisien reliabilitas, dimana nilai koefisien reliabilitas melibatkan total skor tes. Berbeda dengan teori tes klasik, teori respons butir menggunakan istilah informasi untuk menyatakan kehandalan suatu butir soal. Nilai fungsi informasi butir menyatakan kekuatan atau sumbangan butir soal dalam mengungkap *latent trait* yang diukur oleh suatu tes yang memuat butir soal tersebut. Dengan pendekatan teori respons butir, kehandalan suatu butir soal tes dapat diestimasi secara independen. Butir-butir pembangun tes masing-masing dapat dihitung nilai fungsi informasinya. Selanjutnya kehandalan tes dapat diketahui dari kontribusi nilai-nilai fungsi informasi butir soal penyusun tes tersebut (Lord, 1980:22).

Untuk soal dengan skor dikotomi dalam teori respons butir 3P, fungsi informasi butir ( $I_i(\theta)$ ) dinyatakan sebagai berikut.

$$I_i(\theta) = \frac{2.89 a_i^2 (1 - c_i)}{\left[ c_i + e^{1.7 a_i (\theta - b_i)} \right] \left[ 1 + e^{-1.7 a_i (\theta - b_i)} \right]^2} \dots\dots\dots (2)$$

Dimana  $\theta$  adalah parameter kemampuan peserta tes (Hambleton dan Swaminathan, 1991: 91).

Jika  $c_i = 0$ , fungsi informasi maksimum terjadi pada  $e_i = b_i$ . Namun jika  $c_i > 0$ , maka fungsi informasi butir akan maksimum pada  $e_i$  sedikit lebih besar dari nilai parameter  $b_i$  yaitu,

$$(\hat{e}_i) = b_i + \frac{1}{Da_i} \ln \left[ 0,5 \left( 1 + \sqrt{1 + 8c_i} \right) \right] \dots\dots\dots$$

(3)

### Metode Penelitian

Penelitian ini dilakukan di Kota Bengkulu tahun 2004. saat penelitian ini dilakukan, UASD di Kota Bengkulu diselenggarakan oleh Dinas Pendidikan Nasional Kota Bengkulu. Seluruh siswa peserta UASD di Kota Bengkulu mengerjakan soal yang hanya disusun dan dikeluarkan oleh Dinas Pendidikan Nasional Kota Bengkulu. Untuk mendiskripsikan karakteristik butir soal UASD, digunakan teknik dokumentasi untuk pengumpulan data yang berupa naskah soal beserta respons peserta ujian (siswa) terhadap soal tersebut. Naskah soal diperlukan untuk ditelaah/dianalisis secara kualitatif, yang pada prinsipnya adalah untuk mengetahui validitas isi masing-masing soal. Sementara respons peserta ujian terhadap soal dilakukan analisis kuantitatif untuk diketahui karakteristik butir soalnya. Tiap mata uji UASD terdiri dari 3 bentuk soal, yaitu, pilihan ganda, isian singkat dan uraian. Mata uji yang dikaji adalah matematika, bahasa Indonesia, dan IPS. Ketiga mata uji dianggap telah ini mewakili kelompok mata pelajaran yang diajarkan di SD.

Analisis butir soal pilihan ganda dilakukan dengan pendekatan teori respons butir 3P, sedangkan soal isian singkat dan uraian dianalisis dengan pendekatan teori tes klasik. Sampel respons peserta ujian diambil dari lembar jawab siswa (LJS) peserta UASD dari sekolah dengan gradasi yang berbeda. Hal ini dilakukan agar tingkat kesulitan soal sesuai dengan kualitas peserta UASD. Dengan demikian dapat dihindari data yang hanya berasal dari satu kelompok peserta ujian saja. Teknik *purposive sampling* ini dapat digunakan dengan memperhatikan ciri kelompok yang dapat diduga sebagai sampel. Hal ini sesuai yang dikemukakan Kerlinger (2004:206) bahwa, "Ciri sampling ini adalah penilaian dan upaya cermat untuk memperoleh sampel representatif

dengan cara memilih sampel yang meliputi wilayah-wilayah atau kelompok-kelompok yang diduga sebagai anggota sampelnya”.

Analisis butir dengan pendekatan teori respons butir 3P terhadap respons soal pilihan ganda dilakukan untuk mengestimasi parameter butir soal. Untuk itu analisis butir soal dilakukan dengan bantuan komputer program BILOG versi 3,07 pada fase II. Pada *output* (keluaran) analisis ini diperoleh nilai estimasi parameter  $a_i$  (*item discriminating/ slope*),  $b_i$  (*item difficulty level/ threshold*), dan  $c_i$  (*pseudo guessing/ lower asymptot*), serta nilai  $\chi^2$  dan harga derajat kebebasan (DF) masing-masing butir soal yang dianalisis. Kecocokan butir soal dengan model 3P dapat diketahui dengan membandingkan nilai  $\chi^2$  dengan harga kritik  $\chi^2$  pada derajat kebebasannya. Butir-butir soal pilihan ganda ini dapat dikatakan berkarakteristik baik jika memiliki nilai parameter  $a_i$  pada interval (0,2), parameter  $b_i$  pada [-2,2], serta parameter  $c_i$  pada (0, 0,25] secara bersama-sama.

Setelah diperoleh nilai parameter butir dan kecocokan dengan model masing-masing butir soal, selanjutnya butir-butir soal yang memiliki karakteristik butir baik menurut pendekatan teori respons butir 3P, diestimasi nilai fungsi informasi butirnya dengan rumus (1), dan parameter kemampuan peserta ( $\theta$ ) yang sesuai diestimasi dengan rumus (3). Nilai fungsi informasi butir soal diestimasi dengan rumus (2).

Pada soal bentuk isian singkat dan uraian dilakukan analisis butir dengan pendekatan teori tes klasik. Dua hal yang paling penting yang harus diketahui tentang karakteristik butir soal melalui analisis butir soal adalah tingkat kesukaran ( $p$ ) dan indeks daya beda butir ( $d$ ). Di dalam penelitian ini hal ini dilakukan. Hal ini sesuai yang diutarakan oleh Barnard (1999:196) berikut ini, *“Two most basic statistics computed and examined during item analysis are the items’ difficulty and discrimination values.”* Tingkat kesukaran soal, adalah proporsi menjawab benar yang besarnya antara 0 dan 1, sementara indeks daya beda soal adalah kemampuan soal tersebut membedakan peserta tes yang berkemampuan tinggi dan rendah. Ada beberapa cara dalam menghitung (estimasi) indeks daya beda soal, dengan menggunakan *point biserial*, *biserial*, mengelompokkan peserta menurut skor yang diperoleh ke dalam kelompok atas dan bawah, dan cara-cara yang lain.



Penghitungan indeks daya beda dengan *biserial* dipandang lebih teliti, namun dalam menggunakannya diperlukan persyaratan analisis yaitu antara lain, jenis datanya harus dikotomi, sampel respons yang relatif besar, dan datanya harus berdistribusi normal (Glass & Hopkins, 1984:100-102). Persyaratan tentang besarnya sampel adalah hal yang penting. Bila sampel yang digunakan relatif kecil, maka hasil estimasi indeks daya beda akan kasar, sebagaimana dinyatakan Glass & Hopkins (1984:102), "*When n is small,  $r_{bis}$  is a very crude estimate of r; indeed it is probably best not to use  $r_{bis}$  when n is small, especially when p differs considerably from .5.*"

Berdasarkan ini, karena tidak terpenuhinya persyaratan analisis, terutama besarnya sampel dan jenis data, maka estimasi indeks daya beda soal isian singkat dan uraian di dalam penelitian ini tidak dapat menggunakan *biserial* ataupun *point-biserial*. Oleh sebab itu, digunakan cara pengelompokan peserta tes ke dalam kelompok atas dan bawah. Untuk menghitung indeks daya beda butir soal ini digunakan rumus (4), dan tingkat kesukaran butir soal dengan rumus (5) (Allen & Yen, 1979 : 122).

$$d_i = \frac{U_i}{n_{iU}} - \frac{L_i}{n_{iL}} \dots\dots\dots (4)$$

dimana,

$U_i$  : Jumlah seluruh nilai yang diperoleh peserta kelompok atas

$L_i$  : Jumlah seluruh nilai yang diperoleh peserta kelompok bawah

$n_{iU}$  : (Jumlah seluruh peserta kelompok atas) x (nilai maksimum butir-i)

$n_{iL}$  : (Jumlah seluruh peserta kelompok bawah) x (nilai maksimum butir-i)

$$p_i = \frac{\text{jumlah seluruh nilai yang diperoleh peserta}}{(\text{nilai maksimum butir ke - i}) \times (\text{jumlah seluruh peserta})} \dots\dots\dots (5)$$

Suatu butir soal dikatakan baik jika butir soal tersebut memiliki indeks daya beda ( $d_i$ ) bernilai  $0,2d''d_i < 1$  (Djemari, 1999 : 2), dan tingkat kesukaran ( $p_i$ ) bernilai  $0,3d''p_i d''0,7$  (Allen & Yen, 1979: 121) secara bersama-sama.

## Hasil Penelitian dan Pembahasan

### Telaah Butir

Penelaahan butir soal ini dilakukan terhadap naskah soal. Berdasarkan telaah butir yang dilakukan terhadap ketiga mata uji UASD diperoleh informasi bahwa masing-masing mata uji yang diteliti telah sesuai dengan indikator yang telah ditetapkan, dan memiliki validitas isi yang baik.

### Karakteristik Soal Pilihan Ganda dan Fungsi Informasi Butir

Mata Uji Matematika. Berdasarkan *output* hasil analisis, bahwa pada semua butir soal mata uji matematika yang dianalisis dengan program BILOG versi 3.07 cocok dengan model 3P, dan memiliki nilai parameter  $a_i$  yang berada pada interval (0,2). Ini berarti butir-butir soal ini memiliki cukup kemampuan untuk membedakan tingkat kemampuan peserta. Namun, ada butir soal yang memiliki nilai parameter  $b_i$  yang terletak di luar interval [-2,2]. Ada 5 butir soal (17,14%) yang memiliki nilai parameter  $b_i < -2$  (sangat mudah) yaitu nomor 8, 23, 26, 28, 30, dan nomor 34. Selain itu ternyata ada 1 butir (2,86%) yang memiliki nilai parameter  $c_i$  (*pseudo guessing*) sebesar 0,363 yaitu butir soal nomor 21. Artinya soal ini memberikan peluang sebesar 36,3% kepada peserta dengan kemampuan rendah menjawab benar butir ini.

Jadi dari hasil analisis butir dengan pendekatan teori respons butir 3P, ada 24 butir (68,57%) soal pilihan ganda matematika UASD yang tergolong baik. Butir-butir soal tersebut adalah nomor 1, 2, 4, 5, 6, 7, 9, 11, 13, 15, 16, 17, 18, 19, 20, 24, 25, 27, 29, 32, 31, 33, dan 35. Jika ingin mengembangkan bank soal, butir soal ini dapat dipergunakan.

Bahasa Indonesia. Soal pilihan ganda mata uji bahasa Indonesia yang dinyatakan 'bonus' ada 4 butir yaitu 14, 21, 28, dan 29. Hasil analisis terhadap soal yang bukan bonus, menunjukkan bahwa terdapat 1 butir soal (2,867%) yang tidak cocok model, yaitu butir soal nomor 5, karena harga  $c^2$  tidak lebih besar dari  $c^2_{tab}$ . Di samping itu, terdapat 2 butir soal (5,74%) yang tidak konvergen, yaitu butir soal nomor 17 dan 19. Bila ditinjau dari statistik klasikal butir tampak bahwa butir soal nomor 17 dan 19 ini memiliki statistik klasikal *percent* masing-masing sebesar 0,988 dan 0,986; atau dengan kata lain bahwa hampir 99% peserta ujian dapat menjawab benar soal ini. Secara klasik kedua

butir ini digolongkan ke dalam butir yang sangat mudah. Hal ini juga dibuktikan dengan nilai parameter  $b_i < -2$  yaitu berturut-turut sebesar -5,137 dan -3,000. kedua butir soal ini tidak baik (sangat mudah). Kecuali 2 butir yang tidak konvergen, dan 1 butir yang tidak cocok model, butir-butir soal yang lainnya (92,43%) cocok dengan model 3P.

Seluruh butir yang cocok dengan model 3P ini, memiliki nilai parameter  $a_i$  yang berada pada interval (0,2) yang berarti butir-butir tersebut memiliki cukup kemampuan untuk membedakan tingkat kemampuan peserta ujian. Namun, dari hasil estimasi parameter  $b_i$ , ternyata ada 5 butir soal (14,29%) yaitu butir soal nomor 4, 5, 12, 20, dan 31 yang memiliki nilai parameter  $b_i < -2$  (sangat mudah). Selain itu ternyata ada 2 butir soal (5,71%) yang memiliki parameter  $b_i > 2$  (sulit), yaitu butir soal nomor 33 dan 34.

Menilik hasil estimasi parameter  $c_i$  ternyata 20 butir soal (57,14%) memiliki nilai parameter  $c_i$  yang tinggi, yaitu berkisar antara 0,252 dan 0,382. Artinya butir-butir soal tersebut memberikan sumbangan peluang menjawab benar sebesar 0,252 sampai 0,382 kepada peserta yang memiliki kemampuan rendah. Hanya terdapat 5 butir soal (14,29%) yaitu soal nomor 1, 10, 26, 27, dan nomor 35, yang memiliki nilai parameter  $c_i$  yang baik.

Berdasarkan kenyataan ini, maka butir soal pilihan ganda bahasa Indonesia UASD Kota Bengkulu yang memiliki nilai parameter butir  $a_i$ ,  $b_i$ , dan  $c_i$  yang baik secara bersama-sama hanya 5 butir (14,29%), yaitu butir-butir soal nomor 1, 10, 26, 27, dan nomor 35. Jika ingin mengembangkan bank soal pilihan ganda bahasa Indonesia, butir soal ini dapat digunakan.

Ilmu Pengetahuan Sosial (IPS). Ada 2 butir soal yang dinyatakan sebagai 'bonus' yaitu soal nomor 27 dan 30. Berdasarkan hasil analisis terhadap soal yang bukan 'bonus', ada 2 butir soal (5,71%) yang tidak cocok dengan model 3P, yaitu butir soal nomor 5 dan 20. Sementara 29 butir soal (82,9%) lainnya cocok model. Pada butir-butir soal yang cocok model, seluruhnya (29 butir) atau 82,86% butir soal memiliki nilai parameter  $a_i$  yang baik yaitu berada pada interval (0,2).

Bila ditinjau dari nilai parameter  $b_i$ , ada 4 butir soal (11,43%) yang tidak baik, yaitu nomor 1, 3, 12, dan 17 karena nilai  $b_i < -2$  (sangat mudah). Sementara

ada 2 butir soal (5,71%) yaitu soal nomor 12 dan 31 yang juga tidak baik, karena memiliki nilai  $b_1 > 2$  (sulit). Butir soal ini tergolong sulit, karena membutuhkan nilai parameter kemampuan peserta masing-masing sebesar 2,151 dan 3,208 untuk menjawab benar butir soal tersebut.

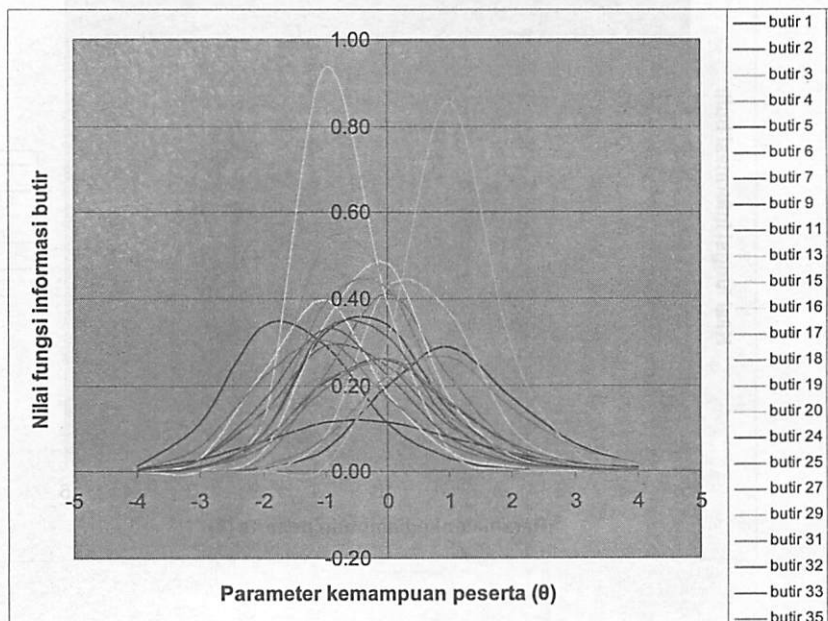
Secara keseluruhan, butir soal pilihan ganda IPS sebagian besar (77,14%) memiliki nilai parameter  $c_1$  yang tinggi yaitu berkisar antara 0,251 dan 0,441. Artinya butir-butir soal ini memberikan sumbangan peluang sebesar 0,251 sampai 0,441 kepada peserta yang memiliki kemampuan rendah untuk menjawab benar butir-butir soal ini. Sementara ada hanya 2 butir soal (5,71%) yang memiliki parameter  $c_1$  yang baik yaitu soal nomor 2 dan 12. Dengan demikian, maka soal pilihan ganda IPS yang berkarakteristik baik hanya ada 1 butir soal (2,86%), yaitu soal nomor 2. Jika akan mengembangkan soal, soal ini dapat dipergunakan.

Nilai Fungsi Informasi Butir. Terhadap butir-butir soal pilihan ganda yang memiliki karakteristik yang baik, dihitung nilai informasi butirnya dengan rumus (2), dan parameter kemampuan peserta yang sesuai dengan rumus (3). Butir soal dengan nilai fungsi informasi butir yang tinggi, memiliki kesalahan pengukuran yang kecil, dan dapat memberikan kontribusi yang tinggi pula kepada nilai fungsi informasi tes. Grafik nilai fungsi informasi butir soal dengan parameter kemampuan peserta yang sesuai untuk tiap mata uji disajikan dalam gambar 1, gambar 2, dan gambar 3.

Secara mudah dapat dilihat dari gambar grafik, tinggi atau rendahnya nilai informasi butir serta parameter kemampuan peserta tes yang sesuai. Bila suatu butir soal dikerjakan oleh peserta dengan kemampuan yang sesuai, maka akan memberikan nilai fungsi informasi yang maksimum. Dengan grafik fungsi informasi butir soal ini, kita dapat terbantu dalam seleksi butir sesuai dengan tingkat kesukarannya. Dalam hal ini tingkat kesukaran soal berhubungan erat dengan parameter kemampuan peserta.

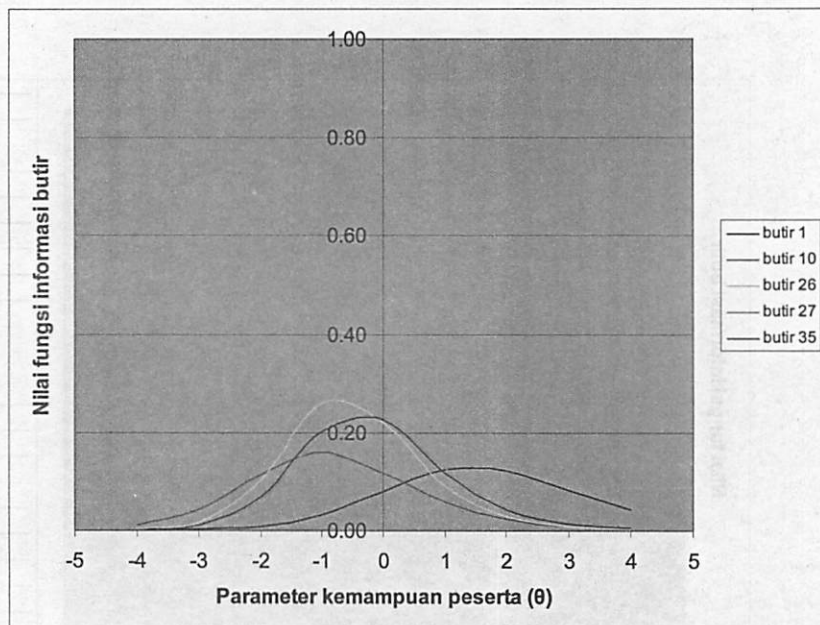
Pada gambar 1, dapat dilihat bahwa fungsi informasi mencapai nilai maksimum pada parameter kemampuan peserta  $\hat{e} < 0$ . Artinya dibutuhkan nilai parameter  $\hat{e} < 0$  untuk menjawab benar soal-soal pilihan ganda mata uji matematika ini, dengan kata lain soal-soal ini mudah hingga sedang. Tiga soal

yang paling baik sesuai parameter kemampuan peserta yang sesuai adalah soal yang memiliki puncak yang tinggi. Soal tersebut bernomor 15, 18, dan 20.



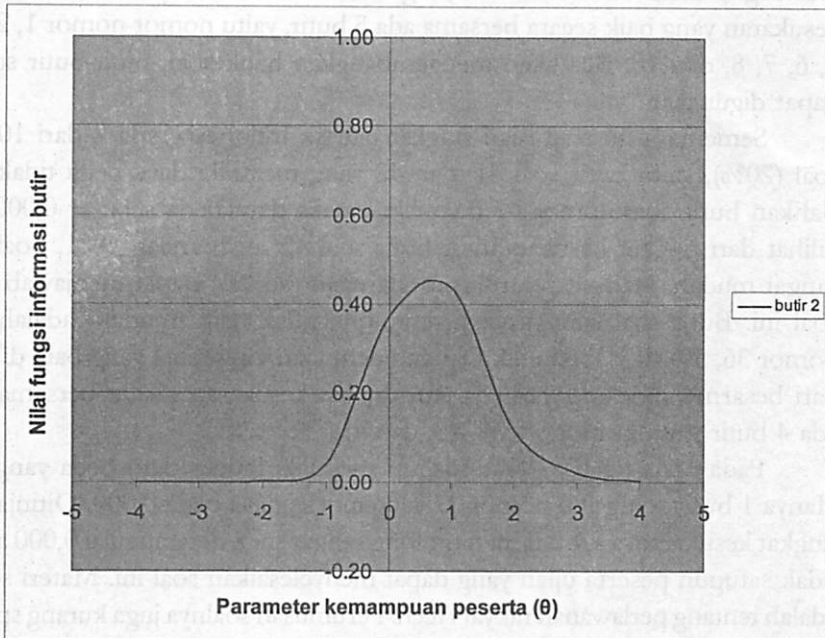
Gambar 1. Grafik Nilai Fungsi Informasi Butir Soal Pilihan Ganda Matematika

Lima butir soal pilihan ganda bahasa Indonesia yang telah diestimasi nilai fungsi informasi butirnya, memiliki grafik nilai fungsi informasi butir seperti pada gambar 2. Pada gambar 2 tersebut terlihat bahwa tinggi kelima puncak grafik fungsi terletak di bawah 0,3. Empat dari lima butir soal ini memiliki ordinat titik puncak yang kurang dari 0,3, namun satu butir lainnya, yaitu butir soal nomor 1, memiliki ordinat lebih dari 0 yang berarti butir ini cukup sulit .



Gambar 2. Grafik Nilai Fungsi Informasi Butir Soal Pilihan Ganda Bahasa Indonesia

Pada mata uji IPS, satu-satunya butir soal yang memiliki karakteristik butir soal yang baik adalah nomor 2. Butir soal ini memiliki tinggi puncak grafik nilai fungsi informasi butir lebih dari 0,4, dan ordinat titik puncak grafik lebih dari 0. Butir soal ini memberikan nilai fungsi informasi butir yang maksimum sebesar 0,528 pada parameter kemampuan peserta sebesar 0,557.



Gambar 3. Grafik Nilai Fungsi Informasi Butir Soal Pilihan Ganda IPS

### Bentuk Soal Isian Singkat

Sementara itu untuk karakteristik soal isian singkat dan uraian, dilakukan analisis terhadap respons peserta ujian dengan menggunakan pendekatan teori tes klasik. Tiap mata uji terdapat 10 butir soal isian singkat. Pada mata uji matematika, dari hasil analisis, 9 butir soal isian singkat (90%) memiliki indeks daya beda yang baik. Butir tersebut adalah nomor 1, 2, 3, 4, 5, 6, 7, 8, dan 10. Artinya butir-butir soal ini mampu membedakan peserta ujian yang berkemampuan tinggi dan rendah. Bila dilihat dari besarnya tingkat

kesukarannya, hanya 5 butir (50%) tergolong baik, yaitu nomor 1, 2, 3, 5, dan 6. Dengan demikian butir-butir yang memiliki indeks daya beda dan tingkat kesukaran yang baik secara bersama ada 5 butir, yaitu nomor nomor 1, 2, 3, 4, 5, 6, 7, 8, dan 10. Bila akan mengembangkan bank soal, butir-butir soal ini dapat digunakan.

Sementara itu soal isian singkat bahasa Indonesia, ada 2 dari 10 butir soal (20%), yaitu butir soal 41, dan 42 yang memiliki daya beda tidak baik. Bahkan butir soal nomor 42 memiliki indeks daya beda sebesar 0,000. Bila dilihat dari tingkat kesukarannya, butir soal 42 ini bernilai 0,981. Soal yang sangat mudah. Hampir seluruh peserta ujian (98,1%) dapat menjawab benar soal ini. Butir soal lainnya yang tergolong tidak baik (mudah) adalah butir nomor 36, 37, 40, 41, dan 44. Dengan demikian, butir soal yang baik di tinjau dari besarnya indeks daya beda dan tingkat kesukaran secara bersama-sama ada 4 butir yaitu nomor 38, 39, 43, dan 45.

Pada mata uji IPS, 90% soalnya memiliki indeks daya beda yang baik. Hanya 1 butir soal yaitu nomor 37 yang indeks daya beda 0,000. Ditinjau dari tingkat kesukarannya, butir ini tergolong sangat sulit, dengan nilai 0,000 artinya tidak satupun peserta ujian yang dapat menyelesaikan soal ini. Materi soal ini adalah tentang perlawanan rakyat Aceh. Perumusan soalnya juga kurang spesifik, dan dapat menimbulkan penafsiran ganda. Butir soal ini sangat tidak baik. 3 butir soal lainnya yaitu nomor 37, 41, dan 44, juga tidak baik (sulit). Sementara ada 2 butir yang tergolong mudah, yaitu soal nomor 39 dan 45. Dengan demikian, butir soal isian singkat IPS yang terbukti berkarakteristik baik dan dapat digunakan untuk pengembangan soal ada 5 butir (50%), yaitu nomor 36, 38, 40, 42, dan 43.

### **Bentuk Soal Uraian**

Tiap mata uji yang dianalisis, terdapat 5 butir soal bentuk uraian. Pada mata uji matematika, semua butir soal uraian memiliki indeks daya beda yang baik. Namun ada 1 butir soal yang memiliki tingkat kesukaran ( $p$ ) < 0,3. Soal ini tergolong tidak baik (sulit). Sementara 4 butir soal lainnya yang tergolong baik. Dengan demikian soal uraian matematika yang berkarakteristik ada 4



butir (80%) yaitu butir soal nomor 2, 3, 4, dan 5.

Pada mata uji bahasa Indonesia, 3 dari 5 butir soal uraiannya (60%) memiliki indeks daya beda yang tidak baik, yaitu butir soal nomor 46, 47, dan 49. Besarnya indeks daya beda tersebut berkisar antara 0,061 dan 0,091. Bila ditinjau dari tingkat kesukarannya ada 4 butir soal (80%) tidak baik (mudah). Butir tersebut adalah nomor 46, 47, 49, dan 50. Dengan demikian hanya 1 butir soal (20%) uraian bahasa Indonesia yang berkarak-teristik baik, yaitu nomor 48.

Pada mata uji IPS, semua butir soalnya (100%) memiliki indeks daya beda yang baik. Namun bila ditinjau dari tingkat kesukarannya, ada 4 butir soal (80%) yang tidak baik karena memiliki tingkat kesukaran ( $p_i$ ) > 0,7 (mudah). Butir soal tersebut adalah nomor 46, 47, 49, dan 50. Dengan demikian butir soal uraian IPS yang berkarakteristik baik ada 1 butir, yaitu nomor 48.

## **Simpulan**

Berdasarkan uraian di atas, dapat disimpulkan, bahwa soal-soal pada mata uji yang ditelaah, telah sesuai dengan indikator yang telah ditetapkan. Sementara itu berdasarkan analisis butir soal pilihan ganda dengan pendekatan teori respons butir, butir soal yang berkarakteristik baik pada mata uji matematika ada 24 butir (68,57%), bahasa Indonesia 5 butir (14,29%), dan IPS 1 butir (2,86%). Berdasarkan hasil analisis butir soal dengan pendekatan teori tes klasik, diperoleh kesimpulan bahwa soal yang berkarakteristik baik pada bentuk isian singkat mata uji matematika 5 butir (50%), bahasa Indonesia 4 butir (40%), dan IPS 5 butir (50%). Sementara pada soal bentuk uraian, butir soal yang berkarakteristik baik pada mata uji matematika ada 4 butir (80%), bahasa Indonesia ada 1 butir (20%), dan IPS ada 2 butir (40%).

Berdasarkan hasil penelitian ini, terbukti bahwa soal yang digunakan untuk ujian akhir sekolah dasar tidak semuanya berkarakteristik baik. Padahal untuk keperluan ujian akhir, karakteristik soal harus baik dan dapat dipertanggungjawabkan. Untuk itu, disarankan agar para penyusun, dapat meningkatkan pengetahuannya dalam penyusun soal. Selanjutnya hal yang terpenting adalah, sebelum soal disajikan untuk keperluan ujian, terutama ujian

akhir, soal harus dianalisis terlebih dahulu agar diketahui karakteristiknya. Hanya soal yang telah ditelaah dan berkarakteristik baik yang disajikan. Dengan demikian hasil pengukuran akan akurat, dapat mencerminkan kemampuan peserta ujian, dan dapat dipertanggungjawabkan.

### **Daftar Pustaka**

- Allen, M. J. & Yen, W. M. (1979). *Introduction to measurement theory*. Belmont, CA: Brooks/Cole Publishing Company.
- Azwar, S. (2001). *Tes prestasi fungsi dan pengembangan pengukuran prestasi Belajar*. Edisi II. Yogyakarta: Pustaka Pelajar.
- Barnard, J. J. (1999). Item analysis in test construction. Dalam Masters, G. N. & Keeves, J. P. (Eds.), *Advance in measurement in educational research and assessment* (pp.195-206). Amsterdam: Pergamon, An imprint of Elsevier Science.
- Depdikbud.(1994/1995). *Kurikulum pendidikan dasar landasan, program dan pengembangan*. Dirjendikti Bagian Proyek Pengembangan Pendidikan Guru Sekolah Dasar.
- Hambleton, R.K. & Swaminathan, H. (1985). *Item response theory: Principles and applications*. Kluwer: Nijhoff Publishing.
- Hambleton, R.K. et al. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage Publications.
- Kerlinger, F. N. (2004). *Asas-asas penelitian behavioral*. (Terjemahan Landung R. Simatupang). Yogyakarta: Gajah Mada University Press.
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. New Jersey: Lawrence Erlbaum Associates, Publishers.
- Mehrens, W. A. & Lehmann, I. J. (1973). *Measurement and evaluation in education and psychology*. Nem York: Hold, Rinehart and Winston, Inc.
- Mislevy, R.J. (1990). *Bilog 3, item analysis and test scoring with binary logistic models* (2<sup>nd</sup> ed). Chicago: Mooresville
- Naga, D. S. (1992). *Pengantar teori skor pada pengukuran pendidikan*. Jakarta: Gunadarma.