

ANALISIS HASIL UJI KOMPETENSI PELAJARAN BAHASA INGGRIS DENGAN MODEL LOGISTIK

Nur Hidayanto

FBS Universitas Negeri Yogyakarta

email: nur_hidayanto@uny.ac.id

Abstrak

Penelitian ini bertujuan untuk menganalisis hasil uji kompetensi pelajaran bahasa Inggris dengan menggunakan Model Logistik 1 Parameter, 2 Parameter, dan 3 Parameter logistik. Penelitian ini diawali dengan pengumpulan hasil uji coba tes dari seluruh peserta di Indonesia dengan menggunakan dokumentasi. Data yang terkumpul dianalisis dengan software Komputer Bilog MG. Program ini memberikan informasi yang cukup untuk menganalisis butir soal dengan *IRT* (Teori Respon Butir). Hasil penelitian menunjukkan bahwa Model Logistik 3 Parameter (3PL) merupakan model yang paling sesuai untuk memberikan informasi mengenai estimasi tingkat kemampuan peserta, kesalahan pengukuran, kecocokan data dengan model, serta korelasi daya beda butir soal karena model ini mampu menunjukkan lebih banyak butir soal yang menunjukkan informasi tersebut.

Kata kunci: uji kompetensi, pelajaran bahasa inggris, model logistik

ANALYSIS OF COMPETENCE TEST RESULTS OF ENGLISH SUBJECTS USING THE LOGISTICS MODEL

Abstract

The study is aimed at analyzing the results of the competence assessment of English subjects using the logistics model using the logistic parameter 1, parameter 2, and parameter 3. The study begins with the collection of the results of the test try-outs of the participants through out Indonesia using the documentation technique. The collected data are analyzed using the Computer Bilog MG software. This program gives adequate information to analyze test items using the IRT (Item Response Theory). Research results show that the Parameter 3 Logistics Model (3PL) turns out to be the most suitable model to give information on the estimation of the participants levels of competence, standard error of measurement, compatibility of data and model, and correlation of item discriminating power. The findings exert the Parameter 3 Logistics Model (3PL) test can show more information about test items.

Keywords: competency testing, english subjects, logistics model

PENDAHULUAN

Tantangan dunia global yang sarat dengan muatan persaingan mengisyaratkan bahwa seseorang yang ingin berhasil dalam memerangi dunia nyata perlu memiliki kemahiran yang diakui oleh dunia global. Terkait dengan hal ini, prestasi seseorang juga mesti diukur dengan cara dan hasil yang

dapat diakui oleh dunia global. Mengikuti alur berpikir ini, kemahiran berbahasa Inggris juga perlu diukur dengan cara yang tidak hanya diakui di Indonesia tetapi juga di manca negara sehingga kesempatan untuk melanjutkan *study* dan atau mencari kerja tidak hanya di negeri sendiri, tetapi juga di mancanegara.

Selama ini, alat ukur yang digunakan untuk mendapatkan informasi tentang kemahiran berbahasa Inggris adalah tes yang dibuat oleh lembaga asing, seperti TOEFL, IELTS, dan TOEIC. Namun, biaya yang harus dibayar ke lembaga asing, yang berarti ke negara asing, untuk menempuh tes-tes tersebut tentu saja kurang menguntungkan. Untuk memberi kesempatan kepada siswa/lulusan SMA mengukur kemahiran berbahasa Inggris/kompetensi komunikatif serta mengurangi ketergantungan dunia pendidikan Indonesia pada pihak asing dalam hal pengukuran kemahiran berbahasa Inggris, sejak tahun 2007 sampai dengan 2011, Direktorat Pendidikan SMA telah mengembangkan tes profisiensi bahasa Inggris (*Test of English Proficiency* atau TOEP) untuk mengukur kemahiran menggunakan bahasa Inggris. Beberapa paket soal TOEP telah diujicobakan di beberapa SMA di seluruh provinsi Indonesia. Analisis hasil dengan 1 PL (teori tes klasik) juga telah dilakukan sebelumnya. Namun demikian, analisis dengan teori tes klasik masih kurang memberikan informasi yang cukup untuk mengetahui keefektifan butir-butir soal dalam tes tersebut karena terdapat asumsi-asumsi yang tidak terpenuhi jika butir-butir tes tersebut hanya diukur dengan teori tes klasik. Hasil analisis dengan teori tes klasik hanya memberikan informasi tentang tingkat kesulitan soal, daya beda, serta keberfungsian pengecoh.

Terkait dengan hal tersebut di atas, diperlukan suatu analisis lanjut dengan analisis 2 PL dan 3 PL untuk memperoleh data *output* yang memberikan informasi dalam mengestimasi tingkat kemampuan peserta (*measure of difficulty*), kesalahan pengukuran (*standard error of measurement*), kecocokan data dengan model (*infit dan outfit*), serta korelasi daya beda butir soal (*point biserial*). Hasil analisis tersebut diharapkan dapat memberikan informasi mengenai karakteristik butir soal tersebut yang nantinya dapat digunakan dalam penyusunan bank soal

ataupun *Computer Adapted Testing*. Dengan demikian, tujuan penelitian ini adalah untuk memperoleh informasi hasil *Test of English Proficiency* (TOEP) dalam mengestimasi tingkat kemampuan peserta (*measure of difficulty*), kesalahan pengukuran (*standard error of measurement*), kecocokan data dengan model (*infit dan outfit*), serta korelasi daya beda butir soal (*point biserial*) jika dianalisis dengan model logistik 1 Parameter, 2 Parameter, dan 3 Parameter.

TOEP merupakan sebuah tes kemampuan Berbahasa Inggris yang dikembangkan oleh Kementerian Pendidikan Nasional sejak tahun 2007. Tes ini disusun oleh pakar tes yang berasal dari beberapa universitas di Indonesia. Tes tersebut disusun dalam beberapa paket. Setiap paket soal dikembangkan sesuai dengan standar minimal kemampuan berbahasa Inggris yang dikuasai oleh lulusan sekolah menengah atas (SMA). Setiap paket terdiri atas beberapa bagian, yakni *Listening Comprehension* dan *Reading Comprehension*.

Bagian tes ini digunakan untuk mengukur kemampuan dasar peserta tes dalam hal menyimak dan merespon percakapan pendek, percakapan yang agak panjang, serta monolog. Jumlah butir soal menyimak dalam setiap paket adalah 50 butir. Setiap butir disusun sesuai dengan SK dan KD yang ditetapkan untuk SMA.

Bagian tes ini dimaksudkan untuk mengukur kemampuan peserta tes dalam hal memahami makna bacaan serta mendapatkan berbagai informasi rinci dari sebuah teks. Setiap paket terdiri dari 50 soal *reading comprehension*. Jenis bacaan yang digunakan dalam setiap paket tes bermacam-macam, seperti pengumuman, memo, sampai dengan teks akademik pendek.

Analisis butir soal dengan teori respon butir pada dasarnya merupakan perkembangan dari teori tes klasik. Dalam perkembangannya, Teori Tes Klasik dianggap memiliki kekurangan antara lain hasil pengukuran terikat pada sampel dan keberadaan skor

murni. Untuk mengakomodasi kekurangan tersebut, munculah teori modern atau lebih dikenal dengan Teori Respon Butir. Pemikiran munculnya teori ini adalah karena kemungkinan menjawab benar tidak dipengaruhi oleh sampelnya, akan tetapi oleh kondisi subjek pada saat pengukuran sehingga hasil pengukuran dapat diprediksi berdasarkan kemampuan atau dikenal dengan *trait* yang dimiliki oleh subjek. Teori ini menganut beberapa asumsi dasar, antara lain unidimensi, independensi lokal, dan invariansi parameter.

Unidimensionalitas sebuah tes dapat diartikan bahwa butir-butir soal dalam perangkat tes tersebut hanya mengukur satu kemampuan saja. Namun dalam kenyataannya, asumsi ini sulit dipenuhi sehingga unidimensi dapat dilihat dari item yang dominan. Asumsi independensi lokal digunakan untuk mengukur kemampuan bahwa respon terhadap suatu butir tidak akan mempengaruhi ataupun dipengaruhi oleh respon terhadap butir yang lain. Asumsi ketiga menyangkut parameter. Yang dimaksud dengan parameter di sini adalah ciri-ciri butir atau karakteristik butir. Seperti halnya dalam Teori Tes Klasik, dalam Teori Respon Butir dikenal beberapa parameter butir antara lain daya beda dan tingkat kesulitan. Dalam Teori Respon Butir juga dikenal adanya *pseudoguessing* atau kemungkinan jawaban tebakan yang dilakukan oleh peserta tes. *Pseudoguessing* ini juga dapat dimaknai sebagai kondisi butir dijawab benar bukan karena kemampuan tetapi hanya tebakan semata.

Menurut Mardapi (2008:135) *trait* merupakan dimensi kemampuan seseorang seperti kemampuan verbal, kemampuan psikomotor, kemampuan kognitif dan sebagainya. Postulat kedua adalah bahwa hubungan antara kemampuan subjek pada suatu item dan perangkat kemampuan laten yang mendasarinya dapat digambarkan melalui kurva karakteristik butir (*Item Characteristic Curve*). Model-model kurva karakteristik

butir tergantung pada bentuk matematis fungsi karakteristik butirnya dan pada banyaknya parameter yang dilibatkan dalam model yang dilibatkan (Azwar, 2008b: 82).

Suatu kurva karakteristik butir adalah suatu rumusan matematis yang menghubungkan probabilitas keberhasilan menjawab dengan benar pada suatu butir soal dengan kemampuan yang diukur oleh tes tersebut dan karakteristik soal bersangkutan. Model-model kurva karakteristik butir tergantung pada bentuk matematis fungsi karakteristik butirnya dan banyaknya parameter yang dilibatkan dalam model yang digunakan. Hal ini dikarenakan tidak semua model respon butir cocok untuk perangkat data tes yang lain. Maka langkah pertama yang harus dilakukan dalam menganalisis dengan menggunakan teori respon butir adalah menentukan kecocokan antara model dengan perangkat data tes yang hendak dianalisis. Jika hal ini telah dilakukan, hasil yang akan didapat adalah *item independence* dan *subject independence*. *Item independence* yaitu bahwa estimasi kemampuan butir tidak tergantung pada butir. *Subject independence* bermakna estimasi kemampuan subjek tidak tergantung pada tes. Selain hal-hal yang telah disebutkan di atas, berbeda dengan teori tes klasik yang tidak dapat mengestimasi kekeliruan kemampuan individual, teori respon butir dapat memberikan perkiraan tentang *standard error* untuk setiap estimasi kemampuan individu dan bukan estimasi kekeliruan untuk semua peserta tes seperti dalam teori tes klasik.

Model ogive normal pada awalnya adalah model yang paling dominan digunakan dalam pengembangan teori respon butir. Namun model tersebut saat ini sudah jarang digunakan dan digantikan dengan model-model logistik. Model logistik lebih sering digunakan karena prosedur komputasinya lebih mudah dan sederhana dibandingkan model ogive normal. Ada tiga model logistik yang sering digunakan saat ini, yaitu (1) model

logistik satu parameter, (2) model logistik dua parameter dan (3) model logistik tiga parameter (Hambleton & Swaminathan, 1985, Hambleton et al, 1991). Pemilihan model yang digunakan tergantung pada asumsi yang cocok bagi perangkat data yang akan dianalisis serta selera pada pemakainya.

Menurut Naga (1992:255) dalam Teori Respon Butir terdapat dua jenis parameter yang satu sama lain saling berhubungan. Dalam hal ini parameter ciri peserta dapat diketahui bila parameter ciri butir telah diketahui atau dikenal dengan estimasi model logistik. Estimasi model inilah yang kemudian berkembang menjadi model logistik 1 sampai 3 parameter, dan sebaliknya, parameter ciri butir dapat diukur bila parameter ciri peserta diketahui atau dikenal dengan estimasi *maximum likelihood* atau estimasi kebolehjadian maksimum. Pada awal perkembangan teori ini digunakan distribusi normal, namun pada masa sekarang berkembang menjadi distribusi logistik karena lebih sederhana analisis matematikanya. Kurva karakteristik butir soal dinyatakan dalam tiga fungsi dan menjadi model logistik (Azwar, 2008a: 88-94).

Model logistik satu parameter (model Rasch)

Asumsi utama model ini adalah bahwa tingkat kesulitan butir soal menunjukkan fungsi kemampuan seseorang. Persamaan matematik model logistik satu parameter, yaitu:

$$P_i(\theta) = \frac{e^{(\theta-b_i)}}{1 + e^{(\theta-b_i)}} \quad i = 1,2,3, \dots, n \dots \dots \dots (8)$$

- $P_i(\theta)$ = probabilitas seorang subjek yang memiliki kemampuan untuk menjawab butir i dengan benar
- θ = tingkat kemampuan
- b_i = parameter tingkat kesulitan butir i
- n = banyaknya butir dalam tes
- e = angka transendental yang bernilai 2,718

Sebagai salah satu model yang paling banyak dikenal, model ini hanya menggunakan parameter tingkat kesulitan butir soal sebagai karakteristik butir. Tingkat kesulitan butir soal didefinisikan sebagai kemungkinan seorang *testee (subjek)* mendapatkan 50% jawaban benar dari butir i jika level *trait testee (subjek)* sesuai dengan tingkat kesulitan butir soal. Model ini mengasumsikan bahwa daya beda semua butir soal dalam tes yang diujikan adalah sama.

Model logistik dua parameter

Model ini dikembangkan berdasarkan pada distribusi normal kumulatif. Persamaan matematik model logistik dua parameter, yaitu:

$$P_i(\theta) = \frac{e^{Dai(\theta-b_i)}}{1 + e^{Dai(\theta-b_i)}} \quad i = 1,2,3, \dots, n \dots \dots \dots (9)$$

- $P_i(\theta)$ = probabilitas seorang subjek yang memiliki kemampuan untuk menjawab butir i dengan benar
- θ = tingkat kemampuan
- b_i = parameter tingkat kesulitan butir i
- n = banyaknya butir dalam tes
- e = angka transendental yang bernilai 2,718
- D = faktor penyekalan yang diikutkan untuk menjadikan fungsi logistik semirip mungkin dengan fungsi ogive normal
- ai = parameter daya beda atau daya diskriminasi

Selain menggunakan parameter kesulitan butir soal, model ini juga menggunakan parameter daya diskriminasi. Parameter daya beda ini sejalan dengan kecuraman/*slope* ICC di titik b_i pada skala kemampuan.

Model logistik tiga parameter

Model ini dikembangkan dari dua model sebelumnya dengan menambahkan satu parameter lagi yakni parameter *pseudo-chance level* atau lebih dikenal dengan

probabilitas *testee* (*subjek*) menjawab benar secara kebetulan.

Persamaan matematik model logistik tiga parameter, yaitu:

$$P_i(\theta) = ci + (1 - ci) \frac{e^{(\theta - b_i)}}{1 + e^{(\theta - b_i)}} \quad i = 1, 2, 3, \dots, n \dots \dots \dots (10)$$

- $P_i(\theta)$ = probabilitas seorang subjek yang memiliki kemampuan untuk menjawab butir i dengan benar
- θ = tingkat kemampuan
- b_i = parameter tingkat kesulitan butir i
- n = banyaknya butir dalam tes
- e = angka transendental yang bernilai 2,718
- ci = parameter *pseudo-chance level* atau probabilitas menebak dengan benar

Asumsi yang mendasari model ini adalah bahwa subjek dengan kemampuan paling rendahpun masih memiliki kemungkinan untuk menjawab benar melalui tebakan atau *guessing*. Asumsi inilah yang digunakan oleh pengembang model ini untuk memasukkan parameter probabilitas *testee* (*subjek*) menjawab benar secara kebetulan.

Seperti telah diungkap sebelumnya, estimasi parameter butir dapat dilakukan jika parameter kemampuan peserta telah diketahui. Untuk menghitung parameter ciri butir, atau *maximum likelihood*, yang didapatkan dari respon peserta yang dilambangkan dengan 0 untuk jawaban salah (Q) dan 1 untuk jawaban benar (P).

Rumus untuk menghitung *maximum likelihood* adalah sebagai berikut:

$$\ln L(X/0) = \sum_{i=1}^M [X_i \ln P_i(\theta) + (1 - X_i) \ln Q_i(\theta)] \quad \dots \dots \dots (11)$$

Dimana:

- L = *maximum likelihood*
- P = probabilitas menjawab benar
- θ = parameter kemampuan peserta
- Q = probabilitas menjawab salah

Program komputer *Bilog MG* digunakan dalam penelitian ini untuk memperoleh data *output* yang memberikan informasi untuk mengetahui estimasi tingkat kemampuan peserta (*measure of difficulty*), kesalahan pengukuran (*standard error of measurement*), kecocokan data dengan model (*infit dan outfit*), serta korelasi daya beda butir soal (*point biserial*). Program ini memberikan informasi yang cukup untuk menganalisis butir soal dengan *IRT* (Teori Respon Butir).

METODE

Penelitian ini dilaksanakan di Universitas Negeri Yogyakarta dan dilaksanakan dari bulan September–November 2012. Penelitian ini merupakan penelitian *ex-post facto* yang menggunakan pendekatan kuantitatif deskriptif untuk mendeskripsikan hasil tes uji kompetensi pelajaran bahasa Inggris TOEP dengan metode 1 parameter, 2 parameter dan 3 parameter. Analisis kuantitatif dilakukan dengan teori respon butir (Program Komputer *Bilog MG*). Dari hasil analisis didapatkan data *output* yang memberikan informasi mengenai model yang paling sesuai untuk mengetahui estimasi tingkat kemampuan peserta (*measure of difficulty*), kesalahan pengukuran (*standard error of measurement*), kecocokan data dengan model (*infit dan outfit*), serta korelasi daya beda butir soal (*point biserial*).

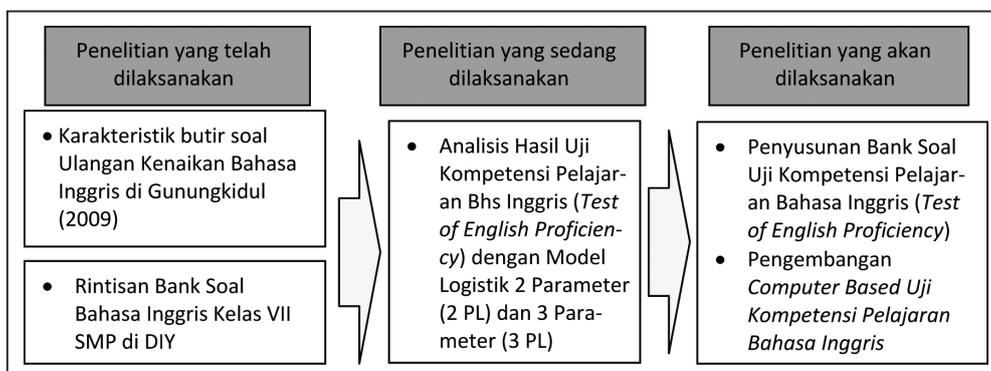
Asumsi yang digunakan dalam penelitian ini adalah bahwa pelaksanaan uji tes kemampuan bahasa Inggris terlaksana dengan baik dan bersih serta sesuai standar yang telah ditetapkan.

Penelitian ini merupakan penelitian tindak lanjut dari penelitian sebelumnya yang dilakukan oleh peneliti, yakni tentang karakteristik butir soal ulangan kenaikan

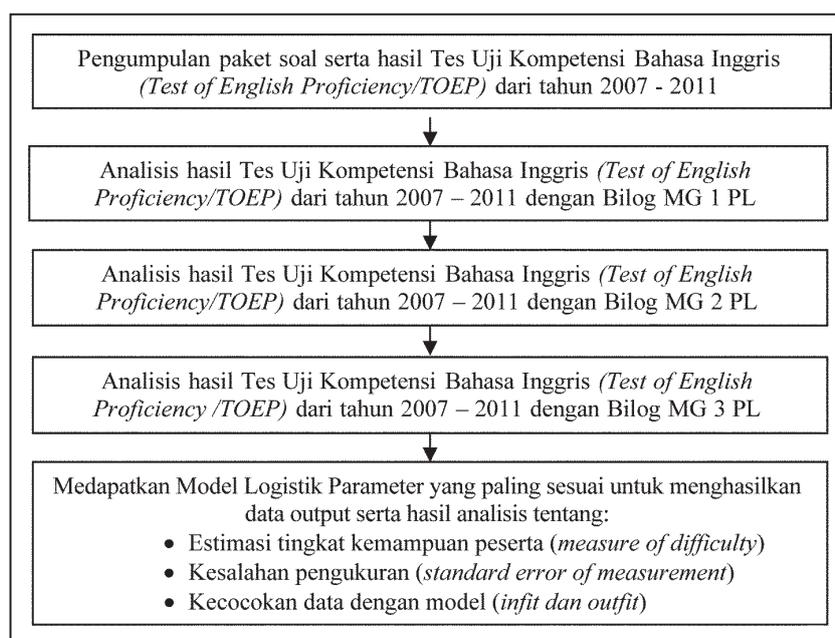
kelas Bahasa Inggris di Gunungkidul pada tahun 2009 serta rintisan bank soal Bahasa Inggris kelas VII Sekolah Menengah Pertama di Daerah Istimewa Yogyakarta pada tahun 2010. Selain itu, penelitian ini juga merupakan kelanjutan analisis sebelumnya yang dilakukan oleh Madya et al (2011) tentang analisis butir TOEP dengan 1 PL. Hasil dari penelitian ini diharapkan dapat digunakan untuk penyusunan bank soal uji kompetensi pelajaran Bahasa Inggris (*Test of English Proficiency*) serta pengembangan *computer based* uji kompetensi pelajaran Bahasa Inggris. Alir penelitian dapat dilihat

pada Gambar 1:

Prosedur yang digunakan dalam penelitian ini dapat dilihat pada Gambar 2. Seperti yang dilihat pada Gambar 2 pada prosedur penelitian, penelitian ini diawali dengan pengumpulan data berupa soal serta hasil jawaban peserta TOEP dalam kurun waktu 2007 – 2011. Selanjutnya, data tersebut akan dimasukkan dan diolah dengan program komputer *Bilog MG* untuk mengetahui model parameter logistic yang paling sesuai untuk estimasi tingkat kemampuan peserta (*measure of difficulty*), kesalahan pengukuran (*standard error of measurement*), kecocokan data



Gambar 1. Roadmap Penelitian



Gambar 2. Alur Penelitian yang Direncanakan untuk Dilakukan

dengan model (*infit dan outfit*), serta korelasi daya beda butir soal (*point bisserial*).

HASIL PENELITIAN DAN PEMBAHASAN

Untuk memastikan kelancaran pelaksanaan penelitian ini, tim peneliti melakukan pertemuan rutin tiap 2 minggu sekali untuk memantau kemajuan atau *progress check* sebagai persiapan penelitian. Pertemuan-pertemuan tersebut membahas tentang jadwal pertemuan berikutnya serta persiapan pengumpulan data penelitian. Selain itu, tim peneliti juga melakukan pertemuan dengan ahli pengukuran, untuk mengonsultasikan langkah pengumpulan data, proses menganalisis data, serta pemaknaan hasil analisis.

Pelaksanaan Kegiatan

Seperti yang diungkapkan pada bagian design penelitian pada bab sebelumnya, penelitian ini diawali dengan pengumpulan paket soal serta hasil Tes Uji Kompetensi Bahasa Inggris (*Test of English Proficiency/TOEP*) dari tahun 2007-2011. Sebelum terjun untuk keperluan pengambilan data, tim penelitian terlebih dahulu melakukan pembagian tugas. Setelah itu, kemudian tim peneliti bersama-sama mengambil data dari petugas Direktorat Pendidikan Sekolah Menengah Atas yang ditunjuk untuk menyimpan hasil uji coba TOEP. Namun demikian, sumber yang memiliki data hanya mengizinkan peneliti untuk mengambil sampel dari salah satu tahun. Untuk itu tim peneliti kemudian memutuskan untuk mengambil data ujicoba tahun 2008 yang terdiri dari 4 paket soal dengan masing –masing 100 butir soal.

Langkah selanjutnya adalah analisis dengan hasil Tes Uji Kompetensi Bahasa Inggris (*Test of English Proficiency/TOEP*). Pada langkah ini, tim peneliti menganalisis data yang telah terkumpul kemudian menganalisisnya dengan Bilog MG untuk 1, 2 dan 3 parameter Logistik (PL).

Tabel 1. Hasil Analisis Bilog

Paket Soal	1 PL	2 PL	3 PL
2A11	6	44	49
2B11	6	37	55
3A11	6	43	51
3B11	5	42	52
	23	166	207

Tabel di atas menunjukkan hasil analisis data penelitian dengan menggunakan Bilog MG 1, 2, serta 3 Parameter Logistik. Dari table tersebut diperoleh informasi bahwa model 3 Parameter merupakan model yang paling baik dibandingkan dengan model 1 dan 2 parameter. Hal ini dapat dilihat dari jumlah butir soal yang informasi estimasi tingkat kemampuan peserta (*measure of difficulty*), kesalahan pengukuran (*standard error of measurement*), kecocokan data dengan model (*infit dan outfit*), serta korelasi daya beda butir soal (*point bisserial*) dari model logistik 3 parameter (207 butir) paling banyak dibanding dengan model logistik 1 (23 butir) dan 2 (166 butir) parameter. Hasil analisis tersebut selanjutnya digunakan untuk dapat memberikan informasi mengenai karakteristik butir soal tersebut yang nantinya ditindaklanjuti dengan pelaksanaan penelitian lanjutan yang bertujuan untuk penyusunan Bank Soal ataupun *Computer Adapted Testing*.

Untuk keperluan penelitian selanjutnya, penelitian ini telah memberikan informasi yang cukup untuk menyatakan bahwa model 3 PL nantinya dapat digunakan sebagai dasar pengambilan keputusan bahwa untuk pengambilan butir yang sesuai untuk pengembangan bank soal dan CAT nantinya dilakukan dengan model 3 PL. Hal ini senada dengan pernyataan Andreswari (2007:3), yang menyatakan bahwa model 3 PL dianggap lebih baik dan juga lebih rumit, namun demikian model ini dianggap memberikan perhitungan estimasi yang lebih teliti. Selain itu, model ini merupakan model yang paling memenuhi untuk soal pilihan ganda.

Model 3 PL ini memberikan informasi mengenai parameter pembeda soal, tingkat kesulitan dan juga kemungkinan menebak dengan benar. Parameter pembeda soal (a) adalah diskriminasi atau daya beda butir soal. Daya beda soal adalah kemampuan suatu soal untuk membedakan antara peserta tes yang berkemampuan tinggi dan berkemampuan rendah. Daya beda soal dapat diketahui dengan melihat besar kecilnya indeks diskriminasi atau angka yang menunjukkan besar kecilnya daya beda. Fungsi daya beda soal adalah mendeteksi perbedaan individual yang sekecil-kecilnya diantara para peserta tes.

Parameter tingkat kesulitan soal (b) adalah satu titik pada skala kemampuan dimana kemungkinan untuk menjawab benar sebesar 0,5. Semakin besar parameter b semakin besar pula kemampuan yang dituntut dari seorang subjek untuk memperoleh 50% peluang menjawab dengan benar. Parameter tebakan soal (c) yaitu kemungkinan untuk menjawab benar secara kebetulan yang biasanya dikenal dengan *pseudo - chance level*. Dengan demikian dalam model logistik tiga parameter juga terdapat satu asumsi dimana seorang subjek yang memiliki kemampuan rendah pun bisa menjawab butir dengan benar. Hal ini biasanya berlaku untuk format tes pilihan ganda. Harga c biasanya diasumsikan akan lebih kecil daripada harga yang akan diperoleh bila subjek menjawab dengan tebakan secara acak. Seperti di ungkap di awal, ketiga karakteristik butir yang dihasilkan dari penelitian ini nantinya dapat digunakan sebagai dasar untuk penyusunan Bank Soal Bahasa Inggris SMA dan juga CAT yang rencananya akan dilakukan setelah penelitian ini selesai.

SIMPULAN

Berdasarkan hasil penelitian dan juga pembahasan pada bab sebelumnya, dapat ditarik beberapa simpulan. Pertama, di antara model logistik 1 Parameter (1 PL), 2 Parameter (2 PL) dan 3 Parameter (3 PL) yang

digunakan untuk menganalisis Hasil Uji Coba Tes Kompetensi Bahasa Inggris (TOEP), model logistik yang paling sesuai untuk mengukur kesesuaian butir dengan model adalah model logistik 3 Parameter (3PL) yang melibatkan tingkat kesulitan daya beda serta *pseudoguessing*. Kedua, ketiga karakteristik butir yang dihasilkan dari analisis dengan 3 PL dari penelitian ini nantinya dapat digunakan sebagai dasar untuk penyusunan Bank Soal Bahasa Inggris SMA dan juga CAT yang rencananya akan dilakukan setelah penelitian ini selesai.

DAFTAR PUSTAKA

- Andreswari, D. 2007. *Penerapan Metode Item Response Theory Three Parameter dalam Tes Potensi Akademik Berbasis Cat (Computer Adaptive Test)*. Bengkulu: Fakultas Teknik Universitas Bengkulu.
- Azwar, S. 2008a. *Dasar-dasar Psikometri*. Yogyakarta: Pustaka Pelajar.
- Azwar, S. 2008b. *Reliabilitas dan Validitas*. Yogyakarta: Pustaka Pelajar.
- Naga, D.S. 1992. *Pengantar Teori Skor*. Jakarta: Gunadarma.
- Mardapi, Djemari. 2008. *Teknik Penyusunan Instrumen Tes dan Nontes*. Yogyakarta: Mitra Cendikia.
- Hambleton, R.K. & Swaminathan, H. 1985. *Item Response Theory*. Boston, MA: Kluwer Inc.
- Hambleton, R.K., Swaminathan, H., & Rogers, H.J. 1991. *Fundamentals of Item Response Theory*. Newbury Park, CA: Sage Publications, Inc.
- Madya, Suwarsih, Ali Saukah, Heri Retnawati, dan Suharso. 2008. "Pengembangan Tes *Test of English Proficiency* (TOEP) untuk Memenuhi Kebutuhan Indonesia akan Tes Baku Kemahiran Bahasa Inggris". *Laporan Tidak Dipublikasikan*. Jakarta: Depdiknas.