

Karakteristik Pengembangan Tes Keterampilan Berpikir Kritis Bumi dan Antariksa Untuk Calon Guru

Taufik Ramlan Ramalis^{a)}, Dadi Rusdiana^{b)}

Departemen Pendidikan Fisika FPMIPA UPI. 3Jl. Dr. Setiabudhi 229 Bandung

Email: ^{a)}taufik_lab.ipba@upi.edu, ^{b)}dadi_rusdiana@upi.edu

Abstract

The purpose of this study was to analyze the characteristics of the f critical thinking skills test and measurement errors according to classical test theory and item response theory. The test is based on the basic concepts and essential principles of the Earth Space, with reference to the core critical thinking skills. The research method used a mixed methods design with embedded experimental models. The analysis in this study is based on 41 responses of students teacher prospective who have taken IPBA courses at the Physics education department of FPMIPA UPI. Based on classical test theory indicates that of 18 test items were analyzed 15 items were good quality, reliability index of 0.89. From information function and standard error measurement of item response theory, the overall conclusion of this evaluation is that the model 2PL is better to use when evaluating the test.

Keywords: critical thinking skills, earth and space, classical test theory, item response theory

Abstrak

Tujuan penelitian ini untuk menganalisis karakteristik tes keterampilan berpikir kritis serta kesalahan pengukuran menurut teori tes klasik dan teori respons item. Tes disusun berdasarkan konsep dasar dan prinsip esensial sains Bumi dan Antariksa, dengan mengacu pada inti keterampilan berpikir kritis. Metode penelitiannya menggunakan *mixed methods* dengan *desain embedded experimental model*. Analisis dalam penelitian ini didasarkan pada respons 41 mahasiswa calon guru yang telah mengambil matakuliah IPBA di departemen pendidikan Fisika FPMIPA UPI. Berdasarkan teori tes klasik menunjukkan bahwa dari 18 item tes yang dianalisis 15 item yang berkualitas baik, indeks reliabilitas 0,89. Dari fungsi informasi dan pengukuran kesalahan baku pada teori respons item, kesimpulan keseluruhan dari evaluasi ini adalah bahwa model 2PL adalah lebih baik untuk digunakan ketika mengevaluasi tes.

Kata-kata kunci: keterampilan berpikir kritis, bumi dan antariksa, teori tes klasik, teori respon item.

PENDAHULUAN

Berpikir kritis merupakan suatu kemampuan yang dapat menciptakan para pemikir tangguh dan pemecah masalah yang handal, hal inilah yang menyebabkan berpikir kritis sangat penting dilatihkan karena kegiatan pembelajaran seharusnya bukan hanya bertujuan mengarahkan siswa dalam rangka memperoleh nilai semata. Berpikir kritis merupakan salah satu berpikir tingkat tinggi yang digunakan untuk pembentukan system konseptual siswa.

Keterampilan Berpikir Kritis (KBK) merupakan bagian penting dalam segala aspek kehidupan seseorang. KBK digunakan dalam berbagai situasi dan kesempatan dalam upaya memecahkan persoalan kehidupan. Oleh karena itu menjadi penting pula seseorang untuk belajar tentang bagaimana berpikir kritis, karena seseorang tidak serta merta mampu berpikir kritis tanpa melalui proses belajar. KBK adalah sebuah ketrampilan yang didapatkan melalui proses. Karena itu perlu upaya untuk mengajarkan bagaimana berpikir kritis kepada peserta didik sedini mungkin (Jurecki & Wander 2012).

Inti dari keterampilan berpikir kritis mencakup: interpretasi, analisis, evaluasi, inferensi, eksplanasi, dan regulasi diri (Facione 2013). Interpretasi diartikan sebagai memahami dan mengungkapkan makna atau arti dari berbagai macam pengalaman, situasi, data, peristiwa, penilaian, konvensi, keyakinan, aturan, prosedur, atau kriteria. Interpretasi meliputi sub-keterampilan kategorisasi, mendekode signifikansi, dan menjelaskan makna.

Analisis diartikan sebagai mengidentifikasi hubungan inferensial dan aktual di antara pernyataan, pertanyaan, konsep, deskripsi, atau bentuk-bentuk representasi yang dimaksudkan untuk mengungkapkan keyakinan, penilaian, pengalaman, alasan, informasi, atau opini. Termasuk dalam keterampilan ini, memeriksa ide-ide, mengidentifikasi dan menganalisis argumen.

Evaluasi diartikan sebagai menilai kredibilitas pernyataan, representasi, atau deskripsi yang berhubungan dengan persepsi, pengalaman, situasi, penilaian, keyakinan, atau pendapat seseorang. Termasuk dalam keterampilan evaluasi ini yakni sub keterampilan menilai kredibilitas klaim dan menilai kualitas argumen.

Inferensi diartikan sebagai mengidentifikasi elemen yang diperlukan untuk menarik kesimpulan logis, membentuk dugaan dan hipotesis, mempertimbangkan informasi yang relevan. Sub keterampilan yang termasuk dalam keterampilan ini adalah menyangsikan bukti, menduga alternatif, dan menarik kesimpulan.

Eksplanasi diartikan sebagai kemampuan penalaran yang koheren, memberikan alasan dan bukti, konseptual, metodologis, dan pertimbangan kontekstual dalam bentuk argumen yang meyakinkan. Sub keterampilan yang termasuk dalam keterampilan ini adalah menetapkan hasil, menjustifikasi prosedur, dan memberikan argumen.

Regulasi diri diartikan sebagai memantau kegiatan kognitif, terutama dengan menerapkan keterampilan analisis, mengkonfirmasi, memvalidasi, atau memperbaiki alasan. Sub keterampilan yang termasuk dalam keterampilan ini adalah monitor diri dan koreksi diri.

Bumi dan Antariksa (BA) merupakan salah satu bagian dari Fisika, yang secara khusus menelaah tentang fenomena alam di Bumi dan benda langit sebagai bagian dari tata surya serta jagat raya secara keseluruhan. Beberapa teori dan hukum fisika dapat digunakan untuk menjelaskan lebih mendalam keadaan bumi serta jagat raya secara keseluruhan.

Pada 15 tahun terakhir, penelitian tentang pendidikan sains BA meningkat relatif lebih tajam dibandingkan dengan penelitian pendidikan sains lainnya, hal ini sejalan dengan issue-isue yang memang sering muncul dan menjadi bahan pembicaraan serta kebutuhan masyarakat di Amerika, bahkan di dunia (Lelliott & Rollnick 2009). Miskonsepsi sains BA pada mahasiswa (Miller & Brewer 2010) serta pada buku teks masih sering ditemukan (Jonh & King 2010).

Fenomena keseharian BA; badai, tsunami, cuaca ekstrim, banjir, letusan gunung berapi, tanah longsor, begitu akrab dengan masyarakat, karena itu beberapa penelitian menyoroiti perlunya meningkatkan literasi BA di sekolah maupun di perguruan tinggi, dan masyarakat secara luas, perlu gerakan yang lebih kohesif untuk mempromosikan pentingnya literasi sistim bumi. melalui pembelajaran yang dikaitkan dengan inkuiri dan pengalaman langsung (LaDue & Clark 2012). Inkuiri dapat menumbuhkan keterampilan berpikir kritis, yang menjadi dasar untuk mengembangkan keterampilan tingkat tinggi lainnya.

Pertanyaan dasar tentang bumi antariksa membantu menjawab pertanyaan mendalam tentang asal-usul alam semesta dan tentang kehidupan itu sendiri. Aspek pengamatan BA akan mempertajam kemampuan anak didik untuk mengamati semua aspek kehidupan kita. Mencatat pengamatan BA akan meningkatkan kemampuan anak didik untuk mendokumentasikan dan menganalisis kejadian di sekitar kita. Pengamatan BA merupakan pengalaman penting dan memiliki implikasi luas.

KBK-BA berarti tidak hanya pengetahuan dan pemahaman tentang bumi dan interaksi dengan benda-benda langit lainnya, tetapi juga pemahaman tentang proses ilmiah yang digunakan untuk menghasilkan konsep-konsep.

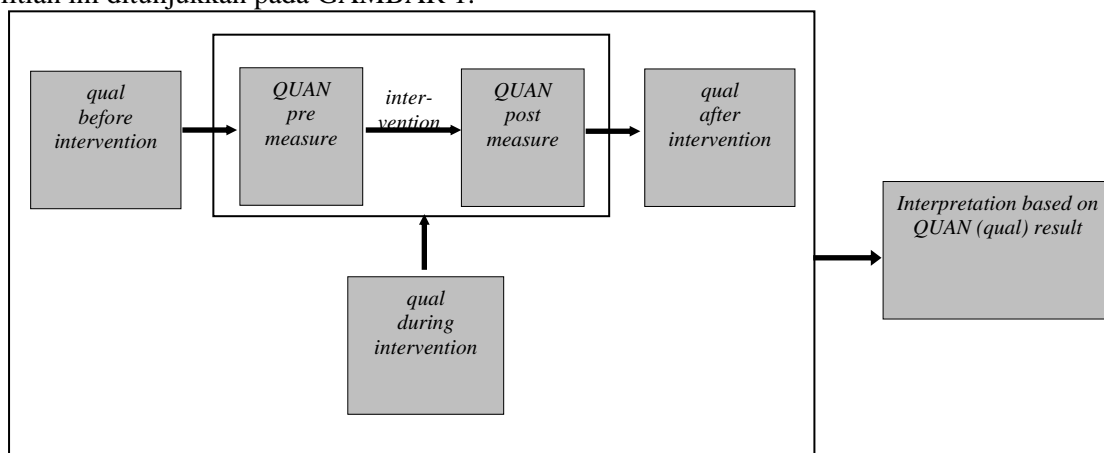
Evaluasi hasil belajar di FPMIPA UPI (UTS dan UAS), menggunakan bentuk tes uraian dan/atau tes pilihan ganda. Mengkontruksi tes pilihan ganda, selain harus memperhatikan daya pembeda dan tingkat kesukaran, juga harus memperhatikan opsi pengecoh. Untuk mengetahui kualitas item tes diperlukan analisis karakteristik item tes, sehingga kekuatan atau kelemahan tes dapat diketahui.

Penelitian ini bertujuan untuk menganalisis karakteristik tes KBK-BA menurut teori tes klasik (*classical test theory*) dan teori respons item (*item response theory*). Analisis dengan teori tes klasik yaitu menentukan standar deviasi, reliabilitas, daya pembeda dan tingkat kesukaran. Analisis dengan teori respon item melalui kurva karakteristik item, fungsi informasi, dan simpangan baku pengukuran.

METODE PENELITIAN

Berdasarkan tujuan, metode penelitian ini dapat dinyatakan sebagai penelitian formatif, yakni penelitian pengembangan atau penelitian tindakan yang dimaksudkan untuk meningkatkan desain praktik atau proses pembelajaran. Penelitian formatif telah direkomendasikan untuk memperluas basis pengetahuan dalam desain teori pembelajaran.

Berdasarkan pada jenis data, model penelitian yang digunakan adalah Mixed Methods (dengan pendekatan Embedded Experimental Model. Model ini memiliki data kualitatif melekat dalam desain eksperimental (Caruth 2013). Prioritas model ini kuantitatif, dengan data kualitatif melekat pada metodologi ini. Desain ini dapat digunakan dengan pendekatan satu fase atau dua fase. Metode penelitian ini ditunjukkan pada GAMBAR 1.



GAMBAR 1. Desain Penelitian Embedded Experimental Model.

Pada qual before intervention, data-data kualitatif prinsip esensial dan konsep dasar BA, diperoleh dari hasil field study, studi literatur, serta analisis dan sintesis jurnal. Bentuk tes berupa pilihan ganda dengan empat opsi pilihan. Konsep dasar dan prinsip penting diperiksa oleh dua orang ahli, dan hasil judgment menjadi masukan untuk revisi. Aspek yang ditelaah oleh ahli mencakup materi dan konstruksi tes. Setiap sub keterampilan KBK terdiri dari tiga item, sehingga jumlah item tes seluruhnya 18 buah. Kisi-kisi tes disusun dalam TABEL 1.

TABEL 1. Kisi-kisi tes KBK-BA

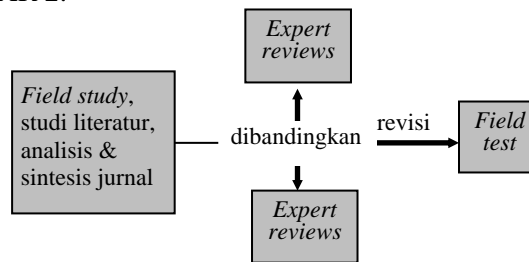
Konsep dasar/ prinsip esensial	Sub KBK, no item
Gerak dan gaya benda langit: hukum Kepler, hukum gravitasi Newton, energi sistem, orbit satelit, gerak semu matahari.	Inferensi, 1; Inferensi, 2; Regulasi diri, 3
Sistem koordinat benda langit: koord. horizon, ekuator, koord. ekliptika.	Analisis, 4
Sistim Bumi-Bulan: fase-fase bulan, gerhana matahari, gerhana bulan, pasang surut.	Evaluasi, 5; Evaluasi, 6; Regulasi diri, 7

TABEL 1. Kisi-kisi tes KBK-BA

Konsep dasar/ prinsip esensial	Sub KBK, no item
Struktur Bumi: bentuk Bumi, rotasi dan revolusi Bumi, lithosfer, magnet bumi, atmosfer, hidrosfer, siklus air.	Analisis, 8; Interpretasi, 9; Eksplanasi, 10
Cuaca dan iklim: unsur cuaca, cuaca dan aktivitas manusia, perubahan iklim.	Regulasi diri, 11; Interpretasi, 12
Tata Surya: sifat fisik planet, klasifikasi planet, konfigurasi planet, model skala tata surya.	Eksplanasi, 13; Evaluasi, 14
Dinamika Bintang: jarak dan gerak bintang, rasi bintang, magnitudo bintang, klasifikasi bintang, siklus hidup bintang,	Interpretasi, 15; Inferensi, 16
Galaksi dan alam semesta: katalog dan klasifikasi galaksi, kosmologi, kuasar, sinar kosmik.	Eksplanasi, 17; Analisis, 18

Pada *QUAN pre measure*, menggali informasi awal KBK mahasiswa melalui tes yang telah disusun. Intervention berupa kegiatan revidu dan revisi berdasarkan perangkat yang telah di buat, sambil digali serta prinsip esensial dan konsep dasar BA (*qual during intervention*).

Setelah *intervention*, tes KBK-BA tersebut kembali dievaluasi pada *QUAN post measure*, dengan teori item klasik dan teori respon item. Hasil analisis selama intervensi menjadi masukan untuk meningkatkan kualialitas tes KBK-BA (*qual after intervention*). Alur pada tahap-tahap tersebut di atas ditunjukkan pada GAMBAR 2.



GAMBAR 2. Tahap penyusunan tes KBK-BA

Setelah seluruh rangkaian tahapan selesai, kemudian dianalisis dan diinterpretasikan pada *Interpretation based on QUAN (qual) result* dengan teori tes klasik dan teori respon item.

Analisis dengan teori item klasik yang dilakukan adalah menentukan standar deviasi, reliabilitas dengan KR-20, daya pembeda dengan koefisien korelasi point biserial (r_{pbis}), dan tingkat kesukaran (p) dengan rata-rata jawaban yang benar.

Teori respon item merupakan metoda untuk item tes berdasarkan kemampuan (latent trait) individu peserta tes, bukan berdasarkan kemampuan kelompok peserta seperti pada teori tes klasik.

Jenis-jenis teori respon item bergantung pada parameter yang digunakan yakni model logistik satu parameter (1PL), dua parameter (2PL) dan tiga parameter (3PL). Masing-masing persamaan ketiga model berturut-turut sebagai berikut (Kohli et al. 2014):

$$P_i(\theta) = \frac{e^{(\theta - b_i)}}{1 + e^{(\theta - b_i)}} \tag{1}$$

$$P_i(\theta) = \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}} \tag{2}$$

$$P_i(\theta) = c_i + (1 - c_i) \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}} \tag{3}$$

dengan

$P_i(\theta)$: probabilitas peserta dengan kemampuan θ yang dapat menjawab item ke i dengan benar.

- θ : tingkat kemampuan peserta tes.
- D : faktor penskalaan yang harganya 1,7
- a_i : daya pembeda item ke-i
- b_i : tingkat kesukaran item ke-i
- c_i : faktor tebakan semu item ke-i
- i : 1, 2, 3 ... n
- n : banyaknya item dalam tes

HASIL DAN PEMBAHASAN

Data kualitatif

Data kualitatif selain konsep dasar dan prinsip esensial BA, juga berupa hasil *judgement* ahli, dengan aspek yang ditelaah seperti pada TABEL 2.

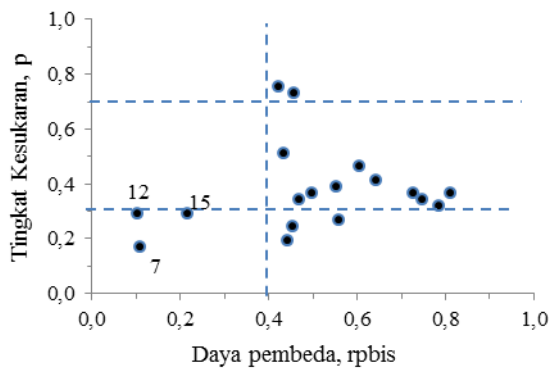
TABEL 2. Aspek tes yang ditelaah

A. Materi	
1.	Tes sesuai dengan indikator
2.	Materi yang ditanyakan sesuai dengan kompetensi
3.	Pilihan jawaban homogen dan logis
4.	Hanya ada satu kunci jawaban
B. Konstruksi	
5.	Tes dirumuskan singkat, jelas, dan tegas
6.	Tes dan pilihan jawaban merupakan pernyataan yang diperlukan saja
7.	Tes tidak memberi petunjuk kunci jawaban
8.	Tes bebas dari pernyataan negatif ganda.
9.	Pilihan jawaban homogen dan logis ditinjau dari segi materi
10.	Gambar, grafik, tabel, diagram, atau sejenisnya jelas dan berfungsi
11.	Panjang pilihan jawaban relatif sama
12.	Pilihan jawaban berbentuk angka disusun dengan urutan besar kecilnya angka.
13.	Item tes tidak bergantung pada jawaban tes sebelumnya

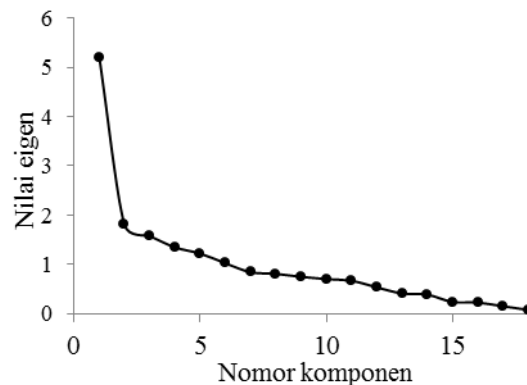
Dari data kualitatif tersebut, ada 5 item tes yang direvisi terkait dengan sub KBK eksplanasi dan regulasi diri.

Teori Tes Klasik

Standar deviasi tes 4,101, reliabilitas KR-20 adalah 0,833, dan kesalahan standar pengukuran 1,676. Daya pembeda r_{pbis} berkisar dari 0,104 sampai 0,798, dan tingkat kesukaran p berkisar dari 0,171 sampai 0,732. Sebaran tingkat kesukaran terhadap daya pembeda ditunjukkan GAMBAR 3.



GAMBAR 3. Sebaran p vs r_{pbis}



GAMBAR 4. Nilai Eigen dari Analisis Faktor

Tingkat kesukaran (p) dikelompokkan menjadi sukar ($p \leq 0,3$), sedang ($0,7 > p > 0,3$), dan mudah ($p \geq 0,7$). Dengan mengacu bahwa item tes yang baik mempunyai $r_{pbis} \geq 0,4$, berdasarkan GAMBAR 3 tersebut tampak bahwa tiga item tes mempunyai daya pembeda yang rendah (item nomor 7, 12, dan 15), dan ketiganya termasuk kelompok sukar. Kemudian berdasarkan tabel 1, konsep dasar (dan prinsip esensial) dan sub KBK, kedua item tersebut masih terwakili oleh item yang lain, sehingga kedua item tersebut tidak digunakan, maka jumlah item menjadi 16 buah. Setelah revisi tersebut di

atas, standar deviasi tes menjadi 5,299, reliabilitas KR-20 menjadi 0,892, dan kesalahan baku pengukuran menjadi 1,553.

Dari matrik korelasi antar item, diperoleh grafik nilai eigen terhadap jumlah komponen seperti ditunjukkan pada GAMBAR 4. Terdapat lima eigenvalue yang nilainya lebih dari besar dari 1. Komponen ke satu dominan, dan mampu menjelaskan varian data respon para peserta tes sebesar 32,53%. Hal ini mengkonfirmasi unidimensi sebagai syarat model item respon teori (Chou & Wang 2012).

Teori Respon Item

Dari GAMBAR 3 tidak ada hubungan antara tingkat kesukaran dengan daya pembeda item. Hal ini berarti bahwa tingkat kesukaran dan daya pembeda item tes mempunyai peluang yang besar untuk dilibatkan dalam model yang dipilih. Dengan demikian, model 2PL atau 3PL adalah alternatif pilihan yang lebih rasional dibandingkan model 1PL. Dengan menggunakan software *eirt versi 1.3.0*, diperoleh parameter untuk setiap model logistik.

Informasi masing-masing model dijelaskan melalui fungsi informasi atau *information function* (IF), yang memberi petunjuk sejauh mana model yang dipilih (1PL, 2PL, atau 3PL) mampu memberikan informasi tentang estimasi traits-level sepanjang skala kemampuan. Semakin tinggi puncak IF, makin informatif pula model yang dipilih mampu menjelaskan kemampuan para peserta tes. Secara matematis, fungsi informasi item (IF) dinyatakan dengan persamaan:

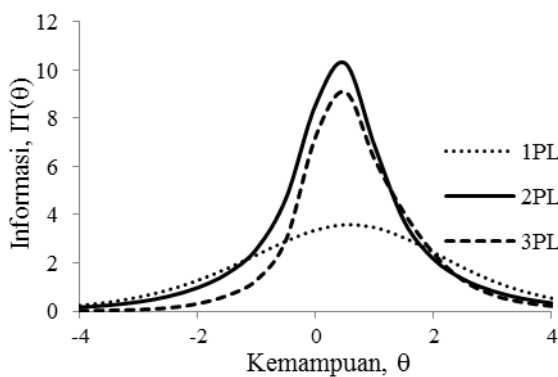
$$IF_i(\theta) = \frac{\{P'_i(\theta)\}^2}{P_i(\theta)\{1-P_i(\theta)\}} \quad (4)$$

Akumulasi keseluruhan fungsi informasi item, diperoleh fungsi informasi tes atau *test information* (TI), yang secara matematis formulanya adalah:

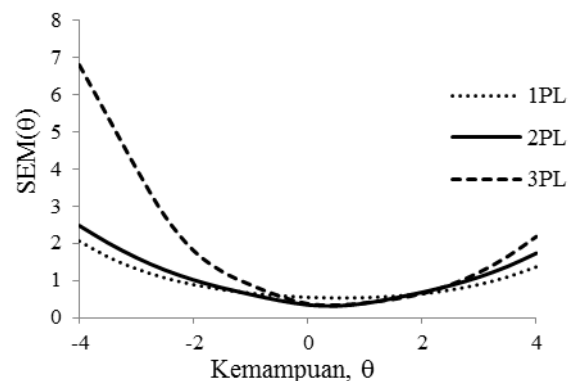
$$TI(\theta) = \sum IF_i(\theta) \quad (5)$$

Untuk masing-masing model IRT, fungsi informasinya tes KBK-BA ditunjukkan pada GAMBAR 5.

Dari gambar 5 tampak bahwa IF pada model 2PL mempunyai puncak paling tinggi. Hal ini berarti bahwa model 2PL dapat memberikan informasi yang lebih baik tentang hubungan antara pola respon para peserta tes dengan keseluruhan karakteristik masing-masing item. Hal ini juga berimplikasi pada kepresisian estimasi kemampuan para peserta tes, semakin tinggi IF maka makin presisi sebuah model dalam mengestimasi kemampuan para peserta.



GAMBAR 5. Fungsi informasi 1PL, 2PL, dan 3PL



GAMBAR 6. SEM 1PL, 2PL, dan 3PL

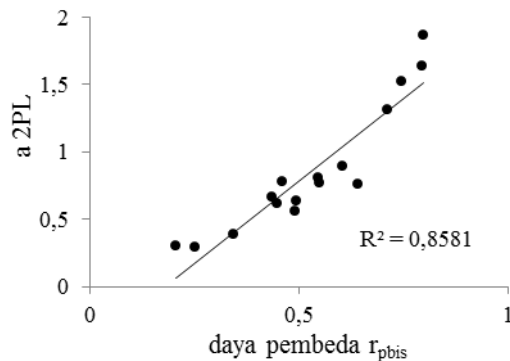
Tingkat presisi model dapat dilihat pula dari simpangan baku pengukuran atau *standard error of measurement* (SEM). SEM merupakan fungsi kebalikan dari TI. SEM tiap-tiap kemampuan θ , dinyatakan dengan formula:

$$SEM(\theta) = \frac{1}{\sqrt{TI(\theta)}} \quad \dots (6)$$

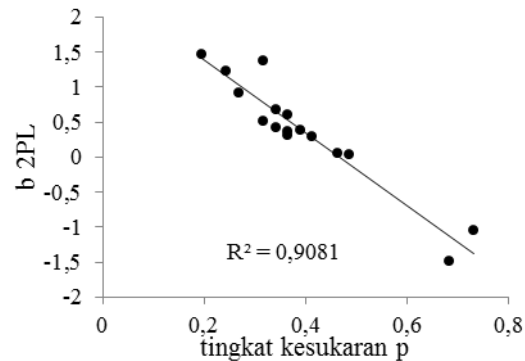
SEM untuk masing-masing model ditunjukkan pada GAMBAR 6.

Dari GAMBAR 6, fungsi SEM model 2PL memiliki lembah paling rendah dibanding dua model yang lain, jadi model 2PL adalah yang paling presisi dalam mengestimasi kemampuan para peserta tes.

Korelasi antara daya pembeda r_{pbis} dengan parameter a, serta hubungan antara tingkat kesukaran p dengan parameter b, ditunjukkan pada GAMBAR 7 dan GAMBAR 8.



GAMBAR 7. Hubungan r_{pbis} dengan a 2PL



GAMBAR 8. Hubungan p dengan b 2PL

KESIMPULAN

Dari analisis dengan teori item klasik menunjukkan bahwa tes KBK-BA memiliki reliabilitas internal sebesar 0,892. Tingkat kesukaran terentang dari 0,244 sampai dengan 0,732 dan daya beda terentang dari 0,269 sampai dengan 0,66. Tes dapat membedakan kelompok rendah dan kelompok tinggi, dan pengecoh berfungsi dengan baik.

Dari analisis dengan teori respon item diperoleh model parameter logistik yang paling rasional adalah model 2PL. Terdapat korelasi linier positif antara daya pembeda r_{pbis} dengan parameter a 2PL, dan korelasi linier negatif antara tingkat kesukaran p dengan parameter b 2PL.

UCAPAN TERIMAKASIH

Terima kepada DP2M Dikti Kemendiknas melalui LPPM UPI, atas dukungan biaya penelitian yang diberikan. Terima kasih juga disampaikan staf dosen departemen pendidikan Fisika FPMIPA UPI atas dukungannya selama pelaksanaan penelitian.

DAFTAR ACUAN

- Caruth, G. D 2013, 'Demystifying Mixed Methods Research Design', *Mevlana International Journal of Education*, vol. 3, pp. 112-122.
- Chou, Y & Wang, W 2012, 'Checking Dimensionality Item Response Models With PCA Standardized Residuals', *Journal Education and Psychological Meas.*, vol. 70, pp. 717-731.
- Facione, P A 2013, *Core Critical Thinking Skills*, Insight Assessment/California Academic Press, San Jose CA.
- Jonh, C & King, J 2010, 'An Analysis of Misconceptions in Science Textbooks: Earth science in England and Wales', *International Journal of Science Education*, vol. 32, no. 5, pp. 565-601.
- Jurecki, K & Wander, C. F 2012, 'Science Literacy, Critical Thinking, and Scientific Literature: Guidelines for Evaluating Scientific Literature in the Classroom', *Journal of Geoscience Education*, vol. 60, pp. 100-105

- Kohli, N, Koran, J & Henn, L 2015, 'Relationships Among CTT and IRT Frameworks via Factor Analytic Models', *Journal Educational and Psychological Measurement*, vo. 75, no. 3, pp. 389-405.
- LaDue, N & Clark S 2012,' Educator Perspectives on Earth System Science Literacy: Challenges and Priorities', *Journal of Geoscience Education*, vol. 60, pp. 372-383.
- Lelliott, A & Rollnick, M 2009, 'Big Ideas: A review of astronomy education research 1974-2008', *International Journal of Science Education*, vol. 32 no.13, pp. 1771-1799.
- Miller, B & Brewer, W 2010, 'Misconceptions of Astronomical Distances', *International Journal of Science Education*, vol. 32, no.12, pp. 1549-1560.