

PENGEMBANGAN TES ILMU PENGETAHUAN ALAM TERKOMPUTERISASI

Suwarto

Universitas Veteran Bangun Nusantara Sukoharjo
Gadingan, Jombor, Bendosari, Sukoharjo, Jawa Tengah 57521, Indonesia
Email: suwartowarto@yahoo.com

Abstrak

Penelitian ini bertujuan untuk: (1). Memaparkan karakteristik tes IPA yang dikembangkan, dan (2) memaparkan program komputer yang telah dikembangkan dalam menentukan skor tes siswa dengan teori tes modern. Tes IPA terdiri dari 600 butir soal dan telah diujicobakan di SMP se-Solo Raya dengan responden 1.800 siswa. Teknik analisis data yang digunakan adalah program R. Program komputer yang telah dikembangkan diujicobakan di SMP Negeri 4 Polokarto. Hasil analisis tes IPA dengan program R menunjukkan: (1) tingkat kesulitan butir dalam kisaran -2,001 sampai 1,713, (2) daya beda butir dalam kisaran 0,553 sampai 2,394, dan (3) terkaan butir dalam kisaran 0,000 sampai 0,379. Program komputer yang dikembangkan dapat menentukan skor tes siswa menurut teori tes modern dengan baik.

Kata kunci: *pengembangan tes IPA*

DEVELOPING OF COMPUTERIZED NATURAL SCIENCE TEST

Abstract

This study is aimed at: (1) describing the characteristics of Natural Science test developed, and (2) describing a computer program which had been developed to determine the test scores of students with modern test theory. Natural Science test consists of 600 items and had been tested in junior high school throughout Solo Raya with 1,800 respondents. Data analysis technique which was used is R programming. The computer program had been developed and tested in SMP Negeri 4 Polokarto (Polokarto 4 Junior High School). The results of the analysis of Natural Science program test with R programming show that: (1) item difficulty is in the range of -2.001-1.713, (2) discriminatory power of items is in the range of 0.553-2.394, and (3) the supposition of items is in the range of 0.000-0.379. The computer program developed can determine students' test scores according to the modern test theory well.

Keywords: *natural science test development*

Permalink/DOI: <http://dx.doi.org/10.21831/pep.v21i2.13144>

Pendahuluan

Setiap sekolah selalu melaksanakan tes untuk keperluan evaluasi di dalam proses pembelajaran. Untuk mengetahui seberapa jauh siswa memahami materi yang dipelajari, maka perlu diberikan tes (Surjono, 1996). Perkembangan ilmu pengetahuan dan teknologi dapat mendukung tercapainya belajar secara efektif dan efisien (Subardjono, 1988).

Tes digunakan untuk ujian kompetensi dasar, ujian tengah semester atau ujian semester. Tes yang baik akan memberikan hasil ukur yang baik. Menurut Gronlund & Linn (1990), *Test an instrument or systematic procedure for measuring a sample of behavior. (Answers the question "How well does the individual perform--either in comparison with others or in comparison with a domain of performance tasks?")*. Cronbach (1970) mendefinisikan sebuah tes: *a systematic procedure for observing a person's behavior and describing it with the aid of a numerical scale or a category-system*. Widoyoko (2012) mendefinisikan bahwa tes merupakan salah satu alat untuk melakukan pengukuran, yaitu alat untuk mengumpulkan informasi karakteristik suatu objek. Cronbach, (1970), Gronlund & Linn (1990), dan Widoyoko (2012) sejalan dalam mendefinisikan tes. Mereka mendefinisikan tes yang dipandang dari sudut prosedur pengukuran dalam rangka mengumpulkan informasi. Prosedur pengukuran dilakukan secara sistematis untuk mendapatkan informasi tingkah laku siswa (Suwanto, 2016).

Tes yang ada di sekolah baru berupa *item pool* (tes yang belum *standard* dan tes tersebut belum diketahui karakteristiknya, baik daya beda butir, tingkat kesulitan butir maupun tebaan butir) belum berupa *item bank* (tes yang *standard* dan tes tersebut sudah diketahui karakteristiknya, baik daya beda butir, tingkat kesulitan butir maupun tebakan butir). Tes yang belum *standard* akan memberikan informasi tentang kemampuan siswa yang bias, tidak akurat sehingga data atau informasi yang diperoleh masih diragukan kebenarannya (Suwanto, 2009). Dari kenyataan ini diperlukan ketersediaan bank soal dari berbagai bidang studi. Pada peneli-

tian ini difokuskan pada pengembangan bank soal dibidang Ilmu Pengetahuan Alam (IPA) yang biasa digunakan untuk Sekolah Menengah Pertama (SMP). Standar Kompetensi (SK) dan Kompetensi Dasar (KD) IPA di SMP/MTs merupakan standar minimum yang secara nasional harus dicapai oleh peserta didik dan menjadi acuan dalam pengembangan kurikulum di setiap satuan pendidikan. Pendidikan IPA diharapkan mampu membekali peserta didik agar memiliki landasan epistemologis dan aksiologis yang mapan agar dapat berkompetisi di dunia yang semakin kompleks. IPA merupakan bidang ilmu yang senantiasa berdekatan dengan realitas alam yang menjadi tempat hidup peserta didik (Nurohman, 2008). IPA merupakan pengetahuan yang dibangun berdasarkan pengamatan dan klasifikasi data yang melibatkan aplikasi penalaran matematis dan analisis data terhadap gejala-gejala alam. Pembelajaran IPA terpadu pada hakekatnya merupakan suatu pendekatan pembelajaran yang dapat membiasakan peserta didik secara individual ataupun kelompok dengan aktif mengeksplorasi, mengelaborasi, mengkonfirmasi, dan mengkomunikasikan hasilnya (Lukum, 2015). Selain proses pembelajaran, faktor lain yang perlu mendapatkan perhatian adalah evaluasi. Evaluasi pembelajaran dapat menentukan apakah tujuan pembelajaran sudah tercapai atau belum (Prilantari, 2007). Oleh sebab itu bank soal IPA yang dikembangkan harus mengungkap tentang SK dan KD yang sudah ditentukan.

Ketersediaan dan kecepatan memperoleh butir soal yang berkualitas baik selalu diharapkan baik oleh guru maupun pembuat tes. Banyaknya koleksi soal yang berkualitas akan membantu guru untuk lebih berkonsentrasi pada proses pembelajaran tanpa menghabiskan waktu untuk menyusun butir soal. Hal ini juga dapat menjamin bahwa hanya soal-soal yang berkualitas yang digunakan. Salah satu ciri dari bank soal adalah butir-butir harus mudah diakses (Suwanto, 2009). Ini berarti perlunya pemanfaatan program komputer di dalam pendidikan (Umar, 1999). Program komputer dapat bekerja secara sistematis, cepat, tepat, dan akurat

(Suwanto, 2009, 2010). Pengaruh pada fleksibilitas dan kualitas pengukuran tes sangat besar. Diharapkan bahwa meningkatnya fleksibilitas merupakan dorongan utama terhadap kesatuan tes bidang pendidikan dan pengajaran (van der Linden, 1999). Untuk menerapkan butir-butir ini pada siswa, dibutuhkan program komputer. *Software* yang dikembangkan secara lokal lebih mudah untuk memenuhi kebutuhan lokal dan sering lebih mudah dari pada sistem yang dikembangkan secara profesional (Linacre, 1995). Berdasarkan latar belakang tersebut, rumusan penelitian ini adalah: (1) bagaimana karakteristik tes IPA yang telah dikembangkan?; (2) bagaimana program komputer yang telah dikembangkan dapat menentukan skor tes pada siswanya?

Teori tes modern dikembangkan oleh para ahli pengukuran bidang psikologi dan pendidikan sebagai upaya meminimalkan kekurangan-kekurangan yang ada dalam teori tes klasik. Perhitungan dalam analisis butir berdasarkan teori tes modern dilakukan dengan menggunakan bantuan program komputer. Ada dua postulat dasar dari teori tes modern (Hambleton, Swaminathan, & Rogers, 1991, p. 7), yaitu: (1) hasil kerja seorang peserta tes pada suatu butir soal dapat diprediksikan dari suatu beberapa jenis faktor yang disebut sifat-sifat, sifat-sifat laten, atau kemampuan; (2) hubungan antara hasil kerja peserta tes pada suatu butir tes dengan sifat-sifat yang mendasarinya dapat dideskripsikan oleh fungsi yang meningkat yang bersifat *monotonic* yang disebut dengan fungsi karakteristik butir (*item characteristic function* atau *item characteristic curve-ICC*). Fungsi ini menjelaskan, jika taraf sifat-sifat (kemampuan) meningkat, maka probabilitas menjawab benar pada suatu butir tes juga meningkat.

Asumsi-asumsi yang mendasari teori tes modern. Asumsi pertama; dalam setiap perangkat tes kesesuaian model terhadap data dapat dinilai, hal ini disebut dengan *overall fit*. Asumsi kedua; dalam setiap perangkat tes, butir-butir tes hanya mengukur satu kemampuan (Suryabrata, 2000, p. 28), dengan kata lain apabila seseorang mampu menjawab dengan benar butir yang sulit,

maka butir yang mudah tentu akan dijawab dengan benar (Mardapi, 2000), asumsi ini disebut dengan unidimensi (*unidimensionality*). Asumsi ketiga; fungsi karakteristik dari suatu butir tertentu merefleksikan adanya hubungan yang sebenarnya dengan kemampuan merespon butir dengan benar. Karakteristik ini sebagai syarat apabila pengembang tes menggunakan *bank* soal. Parameter utama yang menjadi dasar perhitungan dalam teori tes modern adalah kemampuan peserta. Parameter dengan kemampuan peserta ini disebut dengan θ (=teta). Oleh Hambleton batasan nilai θ tidak terbatas, sehingga dapat terbentang dari $-\infty$ sampai $+\infty$ (Linn, 1989, p. 155). Meskipun demikian nilai teta dapat ditentukan pada suatu batasan baku. Dapat ditentukan dari -4 sampai +4 (Hambleton et al., 1991, p. 65; Naga, 1992, p. 207). Batasan nilai teta ini sangat wajar karena peluang nilai $\theta < -4$ dan $\theta > +4$ sangat kecil. Dalam model *logistik* (teori tes modern) yang menggunakan 1 parameter, tingkat kesukaran butir didefinisikan sebagai nilai skala kemampuan peserta tes yang memiliki probabilitas 0,50 untuk menjawab dengan benar pada butir tertentu (Hambleton, 1989, p. 154). Jadi pada dasarnya $b = \theta$ sehingga dalam kurva karakteristik (Lord, 1980, p. 14): $b_i = \theta$, terletak pada $P_i(\theta) = 0,5 + (1 + c_i)$. Jika kemampuan peserta tes dalam suatu kelompok ditransformasi dengan nilai rata-rata 0 dan standar deviasi 1 (kurva normal), maka nilai parameter tingkat kesukaran butir rentangan dari -2,0 sampai +2,0, dengan pengertian nilai b yang mendekati -2,0 dikatakan soal sangat mudah, dan nilai b yang mendekati +2,0 dikatakan soal sangat sukar. Dalam model 1 Parameter Logistik (PL) yang terkenal dengan model Rasch, probabilitas kemampuan peserta tes dapat dibuat dalam model matematika seperti berikut:

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

Model 2 PL menggunakan parameter taraf kesukaran butir dan taraf daya beda butir. Daya beda butir berfungsi untuk me-

menentukan dapat tidaknya suatu butir membedakan kelompok dalam aspek yang diukur sesuai dengan perbedaan yang ada dalam kelompok itu. Dengan menambahkan konstanta diskriminansi (a_i) dari daya beda soal bertindak sebagai arah pada lengkungan ojaif normal (kurva karakteristik). Nilai a_i yang positif menghasilkan lengkungan ke arah kanan, yakni ke arah peningkatan nilai θ . Dan nilai $a_i = 0$ menunjukkan pada kurva yang berbeda. Sehingga lengkungan kurva yang positif memiliki nilai a_i positif, atau berada pada interval: $a_i > 0$ sedangkan a_i bernilai negatif menunjukkan arah lengkungan ke kiri yaitu arah penurunan nilai θ . Bentuk persamaan matematika model 2 PL, probabilitas kemampuan peserta tes dapat dibuat dalam bentuk matematika:

$$P_i(\theta) = \frac{e^{D_i a_i (\theta - b_i)}}{1 + e^{D_i a_i (\theta - b_i)}} \quad (2)$$

Model 3 PL menggunakan Parameter-parameter tingkat kesukaran butir - b_i -, daya pembeda - a_i -, dan faktor terkaan - c_i -. Faktor terkaan adalah kebetulan menjawab dengan benar. Parameter ini mirip dengan probabilitas menjawab benar. Sehingga jika terkaannya benar maka jawabannya benar, dan jika terkaannya salah maka jawabannya juga salah. Apabila menggunakan nilai 0 untuk jawaban salah dan nilai 1 untuk jawaban benar, maka probabilitas jawaban benar, berada dalam rentang 0 hingga 1. Model persamaan matematika untuk 3PL ini adalah sebagai berikut:

$$P_i(\theta) = c_i + \left\{ (1 - c_i) \frac{e^{D_i a_i (\theta - b_i)}}{1 + e^{D_i a_i (\theta - b_i)}} \right\} \quad (3)$$

Metode Penelitian

Penelitian ini merupakan penelitian pengembangan dengan mengadaptasi model yang dikembangkan oleh Hopkins & Clark yaitu model *Research Development and Diffusion* atau “*The R, D & D model*” (Havelock, 1976).

Pengembangan di Tahun 1 (*Research*)

Tahap pertama, pengembangan soal IPA. Soal IPA terdiri dari soal IPA untuk SLTP kelas 7, 8, dan 9. Adapun kegiatan yang akan dilakukan berupa *workshop* yang melibatkan para guru-guru IPA di daerah Solo raya. *Workshop-workshop* yang dilakukan meliputi: *Workshop* 1 mengembangkan *bank* soal IPA untuk siswa kelas 7 SLTP, *workshop* 2 mengembangkan *bank* soal IPA untuk siswa kelas 8 SLTP, dan *workshop* 3 mengembangkan *bank* soal IPA untuk siswa kelas 9 SLTP. Semua soal diujicobakan di sekolah-sekolah se Solo Raya dan kemudian dilakukan analisis untuk mengetahui karakteristik dari masing-masing butir tes. Dalam keperluan tersebut digunakan R Programming untuk mengetahui masing-masing karakteristik dari masing-masing butir soal (tingkat kesukaran, daya beda, terkaan, maupun fungsi informasinya). *Bank* soal dan karakteristiknya disimpan di dalam *database*.

Tahapan kedua, membuat sistem informasi dengan menggunakan Delphi yang digunakan untuk membuat program komputer. Program komputer tersebut dapat menampilkan butir-butir soal yang telah standar dari *bank* soal yang tersimpan di dalam *database*. Butir-butir soal yang keluar dari *database* tentunya disesuaikan dengan kehendak operator dengan menginput atau memilih pilihan (ruang lingkup materi dan jumlah butir soal) yang telah disediakan oleh program komputer. Program komputer juga dapat melakukan penskoran menurut teori tes modern dengan baik.

Tahap ketiga, ujicoba program komputer (produk) secara terbatas dengan melibatkan para guru dan para pendidik (sebagai pengguna program). Ujicoba ini untuk mengetahui bagaimana program yang sudah dibuat, adakah yang perlu diperbaiki terkait dengan pengalaman mereka di lapangan, adakah yang perlu direvisi, saran apa yang dapat diperoleh untuk perbaikan program. Semuanya ini diperlukan untuk sempurnanya program yakni prototipe 1 program komputer.

Pengembangan di Tahun 2 (*Development and Diffussion*)

Buku panduan 1 dibuat, buku ini yang dapat memberikan informasi tentang gambaran umum, bagaimana melakukan instalasi program ke dalam *Personal Computer* (PC), cara menjalankan program, melakukan penskoran menurut teori tes modern, mencetak berbagaimacam hasil penskoran), dan memaknai hasil berbagai macam *printout*. Program komputer prototipe 1 dan buku panduan 1 yang dibuat kemudian di validasi oleh seorang pakar dan direvisi sesuai dengan masukan seorang pakar. Pada tahap berikutnya ujicoba, yaitu diujicobakan di SMP Negeri 4 Polokarto. Ujicoba ini melibatkan para guru maupun para siswa SLTP di sekolah tersebut. Ujicoba ini untuk mengetahui bagaimana program yang sudah dibuat, adakah yang perlu diperbaiki terkait dengan pengalaman mereka di lapangan, adakah yang perlu direvisi, saran apa yang dapat diperoleh untuk perbaikan program. Semuanya ini diperlukan untuk sempurnanya program maupun buku panduan. Ujicoba berlanjut sehingga terbentuk program komputer prototipe 2 dan buku panduan 2. Prototipe 2 Program Komputer dan Buku Panduan 2 yang telah terbentuk diujicobakan lagi sehingga menjadi program komputer 3 dan buku panduan 3.

Penentuan skor siswa dapat dijelaskan sebagai berikut. Jika jawaban siswa sesuai dengan kunci jawaban, maka dikode 1, sedang jawaban siswa yang tidak sesuai dengan kunci jawaban dikode 0. Dengan demikian maka dari kode-kode dikotomus ini dapat diketahui pola respon siswa dalam menjawab tes. Sebagai ilustrasi akan diuraikan bahwa suatu tes terdiri dari 3 butir soal. Masing-masing butir soal mempunyai karakteristik seperti yang terlihat dalam Tabel 1.

Tabel 1. Karakteristik 3 Butir dan Pola Respon Jawaban Siswa

Butir ke	Parameter Butir			Jawaban Peserta
	a	b	c	
1	0,750	-2,000	0,100	1
2	1,250	0,000	0,180	1
3	1,000	1,750	0,160	0

Seorang siswa dengan pola respon 110, maka untuk menentukan *ability* (θ) siswa dengan mempertimbangkan tingkat kesulitan butir, daya beda butir, dan terkaan jawaban, maka dapat menggunakan formula berikut.

$$P_j(\theta) = c_j + (1 - c_j) \frac{1}{1 + e^{-D a_j(\theta - b_j)}} \quad (4)$$

Dimana $D=1,7$

$$L(\underline{U}|\theta) = \prod_{i=1}^n P^{u_i}_i Q^{1-u_i}_i, \underline{U} = u_1, u_2, \dots, u_n \quad (5)$$

L maksimum dicapai pada teta sebesar 0,810 (dengan metode *Maximum Likelihood Estimation = MLE*). *Ability* siswa 0,810 belum dapat dipahami oleh banyak orang, kemudian perlu dilakukan transformasi. Dalam penelitian ini peneliti melakukan transformasi secara linier.

$$\theta^* = l\theta + k$$

$$\theta^* = 10.0,810 + 200$$

$$\theta^* = 8,10 + 200$$

$$\theta^* = 208,10$$

Jadi skor siswa = 208,10

Hasil dan Pembahasan

Hasil penelitian ditahun pertama, pengembangan soal IPA SLTP telah dicapai dengan melalui workshop 1, 2, dan 3 yang dilakukan oleh para guru IPA se Solo Raya. Soal IPA yang berfokus pada biologi telah terbentuk enam paket, yaitu: BIO7-SMT1, BIO7-SMT2, BIO8-SMT1, BIO8-SMT2, BIO9-SMT1, dan BIO9-SMT2. Masing-masing paket ada 100 butir, sehingga seluruhnya ada 600 butir. BIO7-SMT1 dan BIO7-SMT2 diujicobakan pada 600 siswa, BIO8-SMT1 dan BIO8-SMT2 diujicobakan pada 600 siswa, BIO9-SMT1 dan BIO9-SMT2 diujicobakan pada 600 siswa. Semua siswa yang terlibat dalam ujicoba paket-paket tes tersebut adalah 1.800 siswa SMP se Solo raya. Semua paket tes telah dianalisis dengan R Programing, sehingga diketahui karakter-

istik masing-masing butir. Tingkat kesulitan butir dalam kisaran -2,001 sampai 1,713, daya beda butir dalam kisaran 0,553 sampai 2,394, terkaan butir dalam kisaran 0,000 sampai 0,379.

Selanjutnya semua butir dan karakteristik butir disatukan dalam program dalam bentuk *database*. Program komputer telah diujicobakan secara terbatas dengan melibatkan para guru dan pendidik sehingga terbentuk prototipe 1 program komputer yang berisi bank soal dalam *database*, program untuk membuat tes, dan program untuk melakukan penskoran menurut teori tes modern secara terintegrasi.

Hasil penelitian ditahun yang kedua, tersusunlah Buku Panduan 1 yang memberikan informasi tentang gambaran umum, bagaimana melakukan instalasi program ke dalam *Personal Computer* (PC), membuka program, menu utama yang ada dalam program (siswa, soal, transaksi, SQL, buka *file* QRP, dan exit), cara menjalankan program (*Entry Data* Trkerja, menghapus Data Trkerja Per Kelas, cetak soal, cetak lembar jawab, mengisi jawaban soal per siswa, analisis skor menurut teori tes modern, mencetak berbagai macam hasil penskoran), dan memaknai hasil berbagai macam *printout*. Buku panduan 1 dan prototipe 1 program komputer divalidasi oleh seorang pakar dan mendapatkan masukan tentang gambar yang perlu diperbesar agar dapat dilihat dengan jelas pada buku panduan 1, kemudian direvisi sesuai dengan masukan validator, maka prototipe 1 program komputer dan buku panduan 1 diujicobakan di SMP Negeri 4 Polokarto. Ujicoba dilakukan tiga kali sehingga diperoleh program komputer dan buku panduan yang baik. Masukan dari para guru selama ujicoba dapat dirangkum sebagai berikut: (1). *Printout* lembar soal pada petunjuk umum tidak perlu ada kata-kata tulislah dahulu nama, kelas, dan nama sekolah, dikarenakan pada *printout* pada lembar jawab soal sudah tertulis nama, kelas, dan nama sekolah tersebut. (2). Tahapan untuk mencari NIS dan nama siswa tidak diperlukan, karena NIS dan nama siswa juga sudah ada pada lembar jawab. (3). Pada *printout* hasil nilai

skor ujian siswa, garis-garis mendatar belum seragam pada semua siswa, untuk itu perlu diseragamkan agar kelihatan lebih baik. (4). Perlu ditambahkan program dan pedoman untuk menghapus program dari komputer atau laptop. Dari ujicoba yang ketiga tersebut semua sudah dapat berjalan dengan baik, saran guru juga telah memberikan komentar bahwa program maupun buku panduan telah dapat berjalan dan dapat dipahami. Dengan demikian telah terbentuk Program Tes IPA Biologi (PTIBP) dan Buku Panduan PTIBP yang baik. Adapun salah satu hasil skor ujian siswa dapat disajikan pada Gambar 1.

Hasil ini adalah skor semua siswa dalam satu kelas dan dapat dicetak dengan cara tekan gambar printer. Dari hasil *printout* di atas dapat diketahui bahwa siswa yang bernama Nanang Pamungkas memperoleh skor 210,00 (penskoran dengan memperhatikan tingkat kesulitan butir yang diselesaikan), 210,00 (penskoran dengan memperhatikan tingkat kesulitan butir dan daya beda butir yang diselesaikan), dan 208,00 (penskoran dengan memperhatikan tingkat kesulitan butir, daya beda butir, dan terkaan jawaban yang diselesaikan). Siswa yang bernama Aliya Rohmaningsih memperoleh skor 210,00 (penskoran dengan memperhatikan tingkat kesulitan butir yang diselesaikan), 208,00 (penskoran dengan memperhatikan tingkat kesulitan butir dan daya beda butir yang diselesaikan), dan 178,00 (penskoran dengan memperhatikan tingkat kesulitan butir, daya beda butir, dan terkaan jawaban yang diselesaikan). Demikian juga pemaknaan perolehan skor pada siswa-siswa yang lain.

Nanang Pamungkas, Aliya Rohmaningsih, Dimas Wahyu Saputro, Eka Nur Fahma, Handika Restu Candra, Intan Ayu Listiana, dan Nurul Ilmi sama-sama menyelesaikan 6 butir soal dari 10 butir yang disediakan oleh program komputer, tetapi skor perolehan mereka menurut teori tes modern berbeda-beda. Hal ini dikarenakan dalam penskoran siswa menurut teori tes modern telah menyertakan tingkat kesulitan butir, daya beda butir, dan terkaan jawaban dalam menentukan besarnya skor siswa.

HASIL NILAI SKOR UJIAN SISWA						
NAMA SEKOLAH : SMPN 4 POLOKARTO						
KELAS : IXA			TANGGAL UJIAN : 13/10/2016			
NO IND SIS	NAMA SISWA	SKOR 1PL	SKOR 2PL	SKOR 3PL	JML BENAR	NILAI
1527	Nanang Pamungkas	210,00	210,00	208,00	6	60,00
1602	Aliya Rohmaningsih	210,00	208,00	178,00	6	60,00
1603	Ardian Nur Fausan	214,00	214,00	212,00	7	70,00
1604	Avian Miffakhul Nur Huda	206,00	210,00	188,00	5	50,00
1605	Baridho Rama Dhoni	218,00	220,00	218,00	8	80,00
1606	Beni Prasetyo	214,00	216,00	216,00	7	70,00
1607	Desy Putri Nurhalisah	214,00	210,00	180,00	7	70,00
1608	Dila Amanda Syahri	214,00	212,00	210,00	7	70,00
1609	Dimas Wahyu Saputro	210,00	212,00	208,00	6	60,00
1610	Eka Nur Fahma	210,00	208,00	208,00	6	60,00
1611	Evin Nur Adila	214,00	214,00	176,00	7	70,00
1612	Handika restu Candra	210,00	210,00	180,00	6	60,00
1613	Indra Rusianto	214,00	216,00	216,00	7	70,00
1615	Intan Ayu Listiana	210,00	214,00	214,00	6	60,00
1616	Ira Istikhawati	214,00	212,00	210,00	7	70,00
1617	Mardani Ika Rahayu	214,00	214,00	214,00	7	70,00
1618	Muh. Wadyananto Nugroho	214,00	216,00	214,00	7	70,00
1619	Nurul Ilmi	210,00	212,00	206,00	6	60,00
1620	Oktaviana Kusula Dewi M	218,00	218,00	218,00	8	80,00
1621	Putri Wulandari	206,00	206,00	204,00	5	50,00
1622	Regi Pradita	214,00	214,00	180,00	7	70,00
1623	Shalwati Amelia Putri	210,00	214,00	212,00	6	60,00

Gambar 1. Hasil Nilai Skor Ujian Siswa

Kalibrasi 10 butir tersebut bila menggunakan model 1PL dengan daya beda butir dibuat konstan yaitu 1,06 dan terkaan jawaban adalah 0, maka diperoleh sebagai berikut (Tabel 2).

Tabel 2. Karakteristik 10 butir 1PL

Butir	b
1	-1,342
2	-0,334
3	-0,860
4	-0,982
5	-1,548
6	0,023
7	-0,396
8	-0,838
9	-0,871
10	0,503

b= Tingkat kesulitan butir

Kalibrasi 10 butir dengan menggunakan model 2PL dimana terkaan jawaban adalah 0, maka diperoleh sebagai Tabel 3.

Tabel 3. Karakteristik 10 Butir 2PL

Butir	a	b
1	0,633	-2,001
2	1,547	-0,268
3	1,253	-0,773
4	0,984	-0,977
5	1,641	-1,194
6	1,470	0,018
7	0,563	-0,651
8	2,288	-0,568
9	1,800	-0,648
10	1,114	0,485

a= Daya beda,

b= Tingkat kesulitan

Kalibrasi 10 butir dengan menggunakan model 3PL, maka diperoleh sebagai Tabel 4.

Tabel 4. Karakteristik 10 Butir 3PL

Butir	a	b	c
1	0,644	-1,919	0,022
2	1,735	-0,142	0,061
3	1,253	-0,766	0,000
4	0,981	-0,974	0,000
5	1,625	-1,194	0,000
6	1,463	0,028	0,000
7	0,553	-0,659	0,000
8	2,394	-0,507	0,031
9	1,791	-0,637	0,000
10	1,645	0,683	0,109

a = Daya beda,
b = Tingkat kesulitan, dan
c = Terkaan

Simpulan

Tes IPA yang telah dikembangkan mempunyai karakteristik: (1). tingkat kesulitan butir dalam kisaran -2,001 sampai 1,713, (2). daya beda butir dalam kisaran 0,553 sampai 2,394, dan (3). terkaan butir dalam kisaran 0,000 sampai 0,379. Program komputer yang dikembangkan dapat menentukan skor tes siswa dengan menggunakan teori tes modern dengan baik.

Daftar Pustaka

- Cronbach, L. J. (1970). *Essentials of psychological testing* (6th ed.). New York: Harper & Row.
- Gronlund, N. E., & Linn, R. L. (1990). *Measurement and evaluation in teaching* (6th ed.). New York: Collier Macmillan Publishers.
- Hambleton, R. K. (1989). Principles and selected applications of item response theory. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 147–200). New York: Macmillan.
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage Publications.
- Havelock, R. G. (1976). *Planning for Innovation through Dissemination and Utilization of Knowledge*. Michigan: Institute for Social Research The University of Michigan.
- Linacre, J. M. (1995). Individualized testing in the classroom. In *International Encyclopedia of Teaching and Teacher Education* (pp. 295–299). Oxford, New York, Tokyo: Elsevier Science.
- Linn, R. L. (1989). *Educational measurement* (3rd ed.). New York: Macmillan Publishing Company.
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. New Jersey: Lawrence Erlbaum Associates.
- Lukum, A. (2015). Evaluasi program pembelajaran IPA SMP menggunakan model countenance stake. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 19(1), 25–37. Retrieved from <https://journal.uny.ac.id/index.php/jpep/article/view/4552>
- Mardapi, D. (2000). Peran statistika pada bidang pengukuran pendidikan. Makalah disampaikan pada ., In *Seminar Peran Statistik Dalam Bidang ilmu Pengetahuan*. FMIPA UGM.
- Naga, D. S. (1992). *Pengantar teori sekor pada pengukuran pendidikan*. Jakarta: Besbats.
- Nurohman, S. (2008). Peningkatan thinking skills melalui pembelajaran ipa berbasis konstruktivisme di Sekolah Alam. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 11(1). Retrieved from <https://journal.uny.ac.id/index.php/jpep/article/view/1423>
- Prilantari, H. (2007). Pengembangan bank soal keterampilan proses IPA (Biologi) kelas X tingkat Madrasah Aliyah. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 10(2). Retrieved from <https://journal.uny.ac.id/index.php/jpep/article/view/1984>

- Subardjono, S. (1988). Pengembangan pengajaran berbantuan komputer suatu strategi untuk mencapai tujuan instruksional. *Cakrawala Pendidikan*, *VII*(2). Retrieved from <https://journal.uny.ac.id/index.php/article/view/7573>
- Surjono, H. D. (1996). Pengembangan program pengajaran berbantuan komputer (CAI) dengan sistem authoring. *Cakrawala Pendidikan*, *XV*(2). Retrieved from <https://journal.uny.ac.id/index.php/article/view/9224>
- Suryabrata, S. (2000). *Pengembangan alat ukur psikologis*. Yogyakarta: Andi.
- Suwarto, S. (2009). Pengembangan tes dan analisis hasil tes yang terintegrasi dalam program komputer. *Jurnal Penelitian Dan Evaluasi Pendidikan*, *13*(1). Retrieved from <https://journal.uny.ac.id/index.php/jpep/article/view/1401>
- Suwarto, S. (2010). Pengembangan the two-tier diagnostic tes pada bidang biologi secara terkomputersisasi. *Jurnal Penelitian Dan Evaluasi Pendidikan*, *14*(2). Retrieved from <https://journal.uny.ac.id/index.php/jpep/article/view/1079>
- Suwarto, S. (2016). Karakteristik tes Biologi kelas 7 semester gasal. *Jurnal Penelitian Humaniora*, *17*(1), 1. <https://doi.org/10.23917/humaniora.v17i1.2346>
- Umar, J. (1999). Item banking. In G. N. Master & J. P. Keeves (Eds.), *Advances in measurement in educational research and assessment* (pp. 207–219). Amsterdam-Lausanne-New York-Oxford-Shannon-Singapore-Tokyo: Pergamon.
- van der Linden, W. J. (1999). Computerized educational testing. In *Advances in measurement in educational research and assessment* (pp. 138–150). Oxford: Elsevier Science.
- Widoyoko, S. E. P. (2012). *Evaluasi program pembelajaran*. Yogyakarta: Pustaka Pelajar.