

Analisis Pemetaan Kemampuan Numerik Siswa SMK Model Teori Respon Butir

¹Panyahuti, ²Krismadinata, ¹Nizwardi Jalinus, ¹Rusnardi Rahmat, ³Ambiyar

¹Jurusan Pendidikan Tekonologi & Kejuruan, Fakultas Teknik, Universitas Negeri Padang

²Program Studi Teknik Elektro , Fakultas Teknik, Universitas Negeri Padang

³Program Studi Teknik Mesin, Fakultas Teknik, Universitas Negeri Padang

*Corresponding author, e-mail: p.matondang @gmail.com

Abstrak– Pada Sekolah Menengah Kejuruan, matematika memainkan peran penting untuk menguasai mata pelajaran yang berhubungan dengan ilmu-ilmu keteknikan dan sains. Salah satu kemampuan terpenting dalam matematika adalah kemampuan berhitung (numerik). Kemampuan numerik yang baik akan mempermudah siswa untuk mempelajari mata pelajaran lainnya disekolah. Penelitian ini dilakukan berdasarkan rendahnya kemampuan siswa SMK khususnya kemampuan numerik siswa SMK yang mendukung pelajaran keteknikan. Tujuan penelitian ini adalah untuk mengetahui kesahihan alat ukur yang digunakan sekolah untuk menguji kemampuan numerik siswa. Analisis dilakukan terhadap parameter item dari tes kemampuan numerik (kesulitan & daya beda) nilai yang dapat diterima, butir-butir soal-soal numerik yang efektif untuk digunakan untuk pengukuran, tingkat kesesuaian butir soal (item fit), dan indikator validitas uji kemampuan numerik. Penelitian ini adalah penelitian *ex post facto*, Teknik analisis data yang digunakan adalah analisis data secara kuantitatif pendekatan teori respon butir model logistik dua parameter dengan menggunakan bahasa pemrograman R, eirt dan SPSS. Hasil analisis akan membantu sekolah dalam membuat program-program yang lebih efektif untuk meningkatkan hasil belajar siswa. Butir-butir soal yang ada, efektif untuk digunakan dalam pengukuran, dibuktikan dengan hasil korelasi Item-Total signifikan untuk semua butir soal. Penelitian di masa depan diharapkan sebuah tes dapat memetakan kemampuan siswa secara baik pada dimensi pengetahuan *taxonomy bloom* dan variable-variabel kemampuan numerik lainnya.

Kata Kunci : analisis, hasil belajar, item response theory

Abstract– Mathematics has an essential function as a basis for learning other subjects such as engineering and science in vocational schools. The arithmetic (numerical) ability must be mastered to learn mathematics. Great arithmetic ability will help students learn other lessons in school. This research was conducted based on the low ability of vocational students, especially the numerical ability of vocational students who support engineering lessons. The main purpose of this study is to determine the validity of measuring instruments used by schools to test the numerical ability of students. Analyzes are performed on item parameters from numerical ability tests (difficulty & different power) acceptable values, effective numerical items to use for measurement, item fit level, and indicators of the validity of numerical ability tests. This research is an *ex post facto* research, the data analysis technique used is quantitative data analysis approach to the theory of response items in the two-parameter logistic model using the programming languages R, eirt and SPSS. The goal is certainly to improve student learning outcomes. The types of questions existed were considered effective in measurement. This was evidenced by the results of the Total Item-Total correlation at the 0.01 level for all items. The research study in the future is expected to reveal that a test can map students' abilities thoroughly in the dimensions of taxonomy bloom knowledge and other numerical ability variables.

Keywords : analysis, learning outcomes, item response theory



This is an open access article distributed under the Creative Commons 4.0 Attribution License

I. PENDAHULUAN

Salah satu cara untuk mengetahui sejauh mana pencapaian peserta didik adalah dengan penilaian hasil belajar. Penilaian hasil belajar dilakukan dengan memberikan tes yang akan menilai kemampuan peserta didik dan mengetahui ketuntasan dan pencapaian pada bidang studi tertentu [1]. Tes adalah pertanyaan yang diberikan kepada *teste* untuk mendapatkan jawaban dari *teste* dalam bentuk lisan atau tes *oral*, atau dalam bentuk tes perbuatan atau tes tindakan [2]. Tes adalah bagian dari sebuah pengukuran hasil belajar. Pengertian dari pengukuran adalah kegiatan untuk membedakan karakteristik atau atribut seseorang [3]. Beberapa pendapat lain tentang pengukuran menyatakan bahwa pengukuran dinyatakan sebagai sebuah kegiatan untuk menetapkan angka pada seseorang individu atau karakteristik individu menurut aturan tertentu [4,5].

Apakah tes yang telah dibuat dan diujikan tersebut sudah sesuai dengan kemampuan siswa yang menjawabnya, maka perlu adanya validitas untuk mengetahui apakah tes tersebut sudah mampu mengukur aspek yang akan diukur. Pengertian pengukuran bisa disimpulkan bahwa tujuan pengukuran adalah untuk mengetahui karakteristik objek tersebut. Untuk validitas dan realibilitas item yang akan dipilih sebagai test dapat memanfaatkan fungsi indeks realibilitas dan indeks validitas item. Pendekatan yang lebih modern antara lain menggunakan pendekatan logistik yang dikenal dengan *Item Response Theory* (IRT). Dengan analisis IRT, kelemahan penerapan teori tes klasik dapat diselesaikan [6], yaitu: (1) estimasi kemampuan pengambil tes tidak tergantung pada karakteristik dari tes yang digunakan; (2) estimasi parameter item yang tidak tergantung pada kemampuan *teste*; dan (3) kesalahan pengukuran dapat dicari untuk setiap individu.

IRT diklaim dapat menyediakan metode yang baik untuk menjelaskan, memilih item tes yang baik dan membandingkan beberapa macam tes. Ciri-ciri teori respon butir [7] antara lain sebuah karakteristik item butir tidak akan tergantung kepada *teste*, skor hasil dari *teste* tidak tergantung pada tes. Model respon butir akan menekankan pada tingkat butir, bukan dari hasil tes. Model ini tidak mensyaratkan secara ketat tes paralel untuk mengukur nilai reliabilitas sebuah soal, dan merupakan model yang menguraikan sebuah ukuran keputusan untuk tiap skor kemampuan yakni hubungan fungsional antara *teste* dengan tingkat kemampuan yang dimiliki [8].

IRT model logistik satu parameter dapat digunakan untuk melihat tingkat kesukaran butir tes. IRT model logistik dua parameter memungkinkan untuk selain melihat tingkat kesukaran juga untuk melihat tingkat daya beda butir tes. IRT model logistik tiga parameter memungkinkan selain mengetahui tingkat kesukaran dan daya beda butir tes, model ini memungkinkan untuk melihat *pseudo-chance level*, yaitu kemungkinan menjawab benar secara kebetulan. Selain ketiga model logistik tersebut, IRT juga menyediakan beberapa fungsi informasi lainnya. Fungsi informasi butir yang tersedia dapat menggambarkan kekuatan butir tes serta untuk membandingkan beberapa butir tes dengan indikator yang sama. Fungsi informasi butir test ini sekaligus juga dapat menyediakan fungsi informasi test secara keseluruhan [9].

Beberapa asumsi pendukung pada IRT yang secara tidak langsung dapat diukur dan dibuktikan antara lain, *monotonicity* (*Shape of curve*), *unidimensionality*, *item invariance* dan *local independence*. *Monotonicity* paling baik ditampilkan pada kurva yang disebut kurva karakteristik butir (ICC), yang dianggap mencerminkan hubungan sebenarnya antara sifat dan respons terhadap butir. Ini dapat menunjukkan ketika tingkat kemampuan meningkat, probabilitas untuk mendapatkan butir yang benar meningkat secara monoton. Kemampuan peserta yang diukur dalam sebuah soal harus tunggal (asumsi *unidimensi*). Setiap butir soal yang digunakan untuk mengukur harus mengukur satu kemampuan. Tidak boleh mengukur dua atau lebih kemampuan *teste* [8]. Pada penelitian ini, dilakukan analisis karakteristik soal matematika khusus kemampuan numerik. Analisis item ini bertujuan untuk mengetahui karakteristik atau kualitas butir yang baik secara empiris dibatasi untuk uji asumsi *unidimensi* saja. Penelitian ini Model logistik yang digunakan pada analisis penelitian ini adalah menggunakan model 2PL (logistik 2 Parameter) yaitu terdiri dari kesulitan butir soal dan daya beda butir butir soal. Secara matematis, model logistik ini dapat dinyatakan sebagai berikut [7] :

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

θ = kemampuan *teste*, $P_i(\theta)$ = jawaban benar *teste* butir ke i benar, a_i = indeks daya beda, b_i = indeks kesukaran butir ke- i , c_i = indeks tebakan ke- i , e = bilangan natural $\sim 2,719$, D = faktor penskalaan 1,7, dan n = banyaknya butir dalam tes, $i = 1 \dots n$.

Besarnya kemampuan *teste* dilambangkan dengan θ , terletak -3 dan +3, sesuai dengan grafik arsipiran distribusi normal. Nilai akan menjadi dasar asumsi besarnya nilai b_i yaitu terletak diantara $-\infty$ dan (+tidak terhingga sampai +tidak terhingga). Sebuah butir soal akan dikatakan baik jika nilainya berkisar antara 2 dan +2 [7]. Apabila nilai b_i lebih cenderung ke -2, maka indeks kesukaran butir soal tersebut bisa disimpulkan sangat rendah, apabila nilai b_i lebih cenderung ke +2 maka disimpulkan indeks kesukaran butir soal tersebut tinggi dan sangat tinggi. Apabila menggunakan 3 parameter maka suatu butir tes, nilai c_i harus berkisar antara 0 dan 1. Berdasarkan teori modern, kriteria kualitas model logistik tiga parameter menunjukkan bahwa soal yang baik bila - nilai p lebih kecil dari 0.05, nilai a berada antara 0.5 dan 2, nilai b antara -2 dan +2 serta nilai c kurang atau sama dengan 0.2.

Penelitian ini juga akan melihat fungsi informasi dan kemampuan *teste*. Fungsi informasi tes (FIT), digunakan untuk mengetahui atau melihat *validan/keakuratan* atau kepresisian hasil pengukuran suatu tes. Persamaan atau rumus matematika untuk fungsi informasi butir soal menurut Birnbaum dalam Hambleton [7] adalah :

$$I_i(\theta) = \frac{2.89(1-c_i)a_i^2}{\left[c_i + e^{1.7a_i(\theta-b_i)} \right] \left[1 + e^{-1.7a_i(\theta-b_i)} \right]^2} \quad i=1,2,3..n \quad \dots\dots (2)$$

Fungsi informasi tes adalah gabungan dari beberapa informasi item soal untuk tes tersebut. Fungsi perangkat informasi tes dari $I(\theta)$ adalah penjumlahan dari seluruh fungsi informasi item soal. Nilai fungsi ini akan ditentukan dari fungsi informasi item soal yang akan membentuk tes tersebut. Secara matematika rumus fungsi tersebut adalah sebagai berikut:

$$I(\theta) = \sum_{i=1}^n I_i(\theta) \dots\dots\dots(3)$$

$I(\theta)$ = fungsi informasi butir soal pada θ , fungsi $I_i(\theta)$ adalah fungsi informasi butir mulai dari i sampai pada θ .

Fungsi informasi sendiri berbentuk suatu kurva yang disebut FIB (Fungsi Informasi Butir) digunakan untuk masing-masing butir. Kurva yang terbentuk bisa ditambahkan untuk menentukan bentuk dari sebut kurva FIT.

Berdasarkan teori-teori analisis soal yang sudah dijelaskan, maka pada penelitian ini analisis IRT digunakan untuk melakukan analisis pada soal-soal

kemampuan numerik siswa kelas X SMK. Pertimbangan untuk memutuskan menggunakan IRT adalah untuk memperoleh informasi yang lebih banyak dibandingkan menggunakan teori analisis yang lain. Model analisis IRT yang digunakan adalah 2PL.

Matematika adalah ilmu yang penting dan perlu bagi setiap individu dalam kehidupan dimasa sekarang dan masa yang akan datang. Matematika berperan penting untuk memecahkan masalah yang dihadapi individu dalam banyak situasi kehidupan. Pada lingkungan sekolah SD, SMP, SMA/SMK matematika memainkan peran penting dalam untuk menguasai mata pelajaran-pelajaran lain, seperti Kimia, Fisika, Biologi, Akuntansi, Ilmu-ilmu Keteknikan. Mata pelajaran teknik misalnya Teknik Produksi harus didukung oleh logika matematis, visual-spasial, kinestik jasmani dan juga bidang administrasi didukung oleh logika matematis, *linguistik, interpersonal* [10].

Namun sayangnya kemampuan siswa-siswa Indonesia dibidang matematika dan sains masih tergolong sangat rendah. Salah satu indikator kemampuan siswa Indonesia dibidang numerik dan matematika adalah skor PISA. Indonesia untuk matematika dan sains yang juga menjadi pekerjaan rumah besar. Rata-rata skor PISA untuk negara-negara yang tergabung keanggotaan OECD matematika dan sains 489. Sedangkan skor PISA siswa-siswa Indonesia pada tahun 2018 skor PISA Indonesia untuk matematika berkisar di angka 379 dan sains di skor 396. Skor tersebut jauh berada dibawah China dan Singapura menempati peringkat tinggi untuk skor matematika dengan skor 591 dan 569 (Harususilo, 2019).

Hasil studi kemampuan numerik matematika menemukan, bahwa kemampuan numerik siswa adalah salah satu prediktor dalam pembelajaran kimia [11,14]. Untuk mengikuti sebuah tes bakat dengan hasil yang baik siswa juga harus menguasai keterampilan verbal dan numerik [12]. Pemerintah Indonesia telah menetapkan mata pelajaran Matematika menjadi salah satu pelajaran utama disemua tingkatan sekolah.

Pengertian kemampuan numerik dalam penelitian ini adalah kemampuan untuk berurusan dengan angka, dan melakukan pemecahan masalah sederhana [15]. Kemampuan numerik yang dimaksud adalah kemampuan siswa untuk mengerjakan operasi numerik (+, -, ×, ÷) secara akurat dan cepat. Kemampuan numerik dalam penelitian ini meliputi: bilangan real, perhitungan numerik, perbandingan, aljabar, dan persamaan pertidaksamaan. Kemampuan numerik ini diyakini menjadi dasar untuk menguasai kemampuan

matematika yang akan mendukung perhitungan-perhitungan pada mata pelajaran lain.

Dari penjelasan pentingnya matematika, penelitian ini akan memberikan panduan praktis yang akan mendeteksi tingkat kemampuan numerik siswa SMK. Penelitian ini juga akan membantu untuk memahami matematika secara mendalam, memetakan kemampuan para siswa SMK dibidang numerik yang erat hubungannya dengan mata pelajaran lain. Hasil Analisa akan membantu sekolah dalam membuat program-program yang lebih efektif seperti persiapan guru, merencanakan pelatihan, mengarahkan, dan membangun model-model yang sesuai untuk meningkatkan hasil belajar siswa. Hasil penelitian ini secara tidak langsung akan menguntungkan dalam pengembangan kurikulum, buku teks matematika, dan termasuk panduan guru dengan kegiatan yang sesuai untuk berbagai tingkat kemampuan numerik. Tujuan penelitian yang dilakukan ini adalah untuk mengetahui: 1) Parameter item dari tes kemampuan numerik (kesulitan & daya beda) nilai yang dapat diterima. 2) Butir-butir soal-soal numerik yang efektif untuk digunakan untuk pengukuran kemampuan numerik. 3) Tingkat kesesuaian butir soal (item fit). 4) Indikator validitas uji kemampuan numerik.

II. METODE PENELITIAN

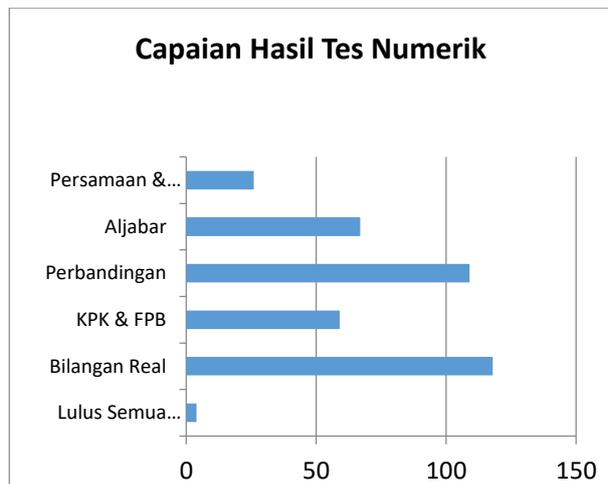
Penelitian ini dilakukan untuk mengetahui soal-soal yang digunakan dalam memetakan kemampuan numerik siswa SMK. Objek penelitian ini adalah butir soal matematika kemampuan numerik sederhana yang disusun oleh tim pembuat soal SMK dengan tipe pilihan ganda. Jumlah soal yang dianalisis sebanyak 50 butir terdiri dari bilangan real 10 butir, perhitungan numerik 10 butir, perbandingan 10 butir, aljabar 10 butir, dan persamaan pertidaksamaan 10 butir. Soal-soal yang dianalisis tersebut digunakan oleh sekolah untuk memetakan kemampuan numerik siswa. Siswa yang kemampuan rendah akan diberikan matrikulasi pembelajaran matematika dasar, khususnya kemampuan numerik.

Populasi penelitian adalah semua siswa kelas X SMK Negeri 1 Lahat, SMK Negeri 3 Lahat bidang Teknologi dan Rekayasa dengan rincian siswa SMK Negeri 1 Lahat sebanyak 315 dan siswa SMK Negeri 3 Lahat sebanyak 102. Pengumpulan data dilakukan dengan menggunakan teknik dokumentasi yaitu dengan mengutip respons siswa *teste* pemetaan. Penelitian ini adalah penelitian *ex post facto* [16]. Teknik analisis data yang digunakan adalah analisis data secara kuantitatif

dengan menggunakan pendekatan teori respon butir model logistik dua parameter dengan menggunakan bahasa pemrograman R dan *eirt* sebagai *add-ins* pada Microsoft Excel [17]. Pada R menggunakan pustaka ltm, psych, eRm dan difR [18]. Langkah penelitian meliputi: (1) Mengumpulkan jawaban siswa dari database program aplikasi pemeriksa jawaban siswa; (2) Melakukan konversi data dari database aplikasi pemeriksa jawaban menjadi menjadi tipe data .txt (3) Mempersiapkan *library* aplikasi pemrograman R yaitu *library* “ltm version 3.6.1”, “eRm version 3.6.1”, “psych” versi 3.6, dan “difR version 3.6.1; (4) Melakukan analisis data dengan R, eirt dan menyajikan hasilnya; (5) Menginterpretasikan data hasil analisis.

III. HASIL DAN PEMBAHASAN

Hasil tes numerik kemampuan siswa yang dilakukan pada 2 sekolah, ditemukan lulus semua materi dengan nilai lebih besar dari 60 adalah 4 (empat) orang. Hasil ujian kemampuan siswa pada gambar 1.



Gambar 1. Hasil ujian kemampuan siswa

Sebelum melakukan analisis tingkat kesulitan dan daya beda, dilakukan analisis *unidimensi*. Tujuan analisis *unidimensi* adalah untuk mempertahankan invariansi pada IRT [19]. Apabila satu butir tes dipakai mengukur lebih dari satu dimensi, maka jawaban *teste* pada butir tersebut adalah campuran dari beberapa macam kemampuan *teste*. Misalkan kita membuat sebuah tes matematika yang bertujuan untuk mengukur logika matematika siswa. Apabila kita tidak berhati-hati dalam membuat butir soal tersebut, selain butir soal tersebut akan mengukur

kemampuan logika matematika juga dapat mengukur kemampuan berbahasa. Apabila salah seorang *teste* tidak bisa menjawab soal ujian dengan benar, maka kemungkinan besar pembuat soal tersebut tidak bisa melihat dimana kekurangan kemampuan tersebut, bisa saja disebabkan oleh kemampuan peserta di bidang logika matematika atau kemampuan dibidang bahasa. Sangat penting untuk mengetahui sebuah perangkat tes bersifat *unidimensi* atau tidak. Untuk mengetahuinya kita bisa menggunakan uji analisis faktor [18]. Pada waktu praktik pengukurannya untuk mengetahuinya cukup menemukan 1 (satu) dimensi yang dominan. Dimensi ini yang akan menjadi *unidimensi* pada respon atau karakteristik butir soal. Berdasarkan hasil analisis *factor* menggunakan bahasa pemrograman R telah diperoleh hasil sebagai berikut:

Second eigenvalue in the observed data: 1.9709
Average of second eigenvalues in Monte Carlo samples: 1.5542

Monte Carlo samples: 100

p-value: 0.0495

Dari hasil analisis tersebut diperoleh nilai *p-value*=0.0495, artinya *p-value* >0.01 memenuhi syarat untuk melanjutkan analisis IRT dengan model 2PL [18].

A. Tingkat Kesulitan dan Daya Beda

Hasil analisis tingkat kesulitan soal dan daya bedanya di sajikan pada table 1. Analisis ini dilakukan dengan aplikasi Add-ins “eirt” pada Microsoft Excel.

Tabel 1. Tingkat Kesulitan dan Daya Beda Soal

#Item	Kesulitan	Daya Beda	#Item	Kesulitan	Daya Beda
S001	-1.319	2.882	S026	5.658	0.286
S002	0.977	0.609	S027	0.060	1.619
S003	-0.853	2.268	S028	-0.270	1.441
S004	-0.440	1.464	S029	0.418	0.951
S005	0.686	0.924	S030	0.644	0.910
S006	0.112	1.570	S031	0.000	1.702
S007	-0.469	1.697	S032	-0.328	1.651
S008	0.309	0.827	S033	-0.358	1.387
S009	0.907	0.631	S034	-0.019	1.509
S010	-0.503	1.470	S035	-0.518	2.063
S011	-0.663	2.052	S036	0.279	1.492
S012	-0.720	1.902	S037	-0.332	1.901
S013	0.000	1.702	S038	9.263	0.546
S014	-0.078	1.334	S039	-0.322	1.927
S015	0.298	0.852	S040	-0.286	1.963
S016	0.546	0.545	S041	0.000	1.702
S017	5.881	0.097	S042	-0.187	1.509
S018	2.348	0.257	S043	0.186	1.261
S019	1.376	0.352	S044	0.180	1.691
S020	0.860	0.813	S045	0.447	1.218
S021	-0.961	0.471	S046	1.137	0.447
S022	-1.027	1.999	S047	0.778	0.983
S023	-0.663	1.930	S048	0.065	0.992
S024	-0.648	1.268	S049	5.173	0.214
S025	-0.115	0.563	S050	9.263	0.546

Hambleton dan kawan-kawan [7] menyebutkan, bahwa nilai parameter tingkat kesulitan bervariasi antara -2.0 hingga +2.0.

Nilai yang mendekati -2,0 adalah butir soal yang sangat mudah, dan nilai yang mendekati angka +2,0 adalah butir soal sangat sulit. Dari

hasil analisis yang dilakukan pada tabel 1 mengungkapkan nilai yang dapat diterima untuk parameter kesulitan adalah berkisar - 1.319 sampai 1.376. Butir soal 17, 18, 26, 38, 49 dan 50 perlu diteliti ulang, soal tersebut tergolong sangat sulit melebihi ambang batas yang diberikan para ahli.

Untuk bersarnya indeks daya beda butir soal, Alagumalai [20] telah mengelompokkan indeks tersebut menjadi: sangat bagus > 0,40, bagus 0,30 – 0,39, cukup 0,20 - 0,29, tidak mampu membedakan 0,00-0,19, membutuhkan pemeriksaan terhadap butir < 0,00. Nilai yang

tinggi akan menghasilkan kurva karakteristik item yang sangat curam. Tabel 1 mengungkapkan nilai yang dapat diterima untuk parameter daya beda adalah 0.257 sampai 2.882. Butir soal ke 17 tidak memiliki indeks daya beda, sebaiknya di revisi.

B. Analisis Item-Dimension Butir Soal-Soal Numerik yang Efektif

Analisis Item-Dimension dilakukan dengan bantuan aplikasi SPSS, hasil analisis disajikan pada table 2. Item-dimension, Total Korelasi dan Cronbach's Alpha item deleted

Table 2. Item-dimension, Total Korelasi dan Cronbach's Alpha item deleted

# Item	Item-dimension Correlation	Item-Total correlation	Cronbach's Alpha if Item Was Deleted	# item	Item-dimension correlation	Item-Total Correlation	Cronbach's Alpha if Item Was Deleted
1	0.416**	0.275**	0.803	26	0.361**	0.431**	0.804
2	0.432**	0.373**	0.811	27	0.423**	0.342**	0.806
3	0.520**	0.412**	0.821	28	0.522**	0.322**	0.808
4	0.472**	0.309**	0.811	29	0.435**	0.411**	0.809
5	0.353**	0.219*	0.821	30	0.261**	0.344**	0.810
6	0.427**	0.224**	0.833	31	0.018**	0.323**	0.811
7	0.321**	0.361**	0.821	32	0.512**	0.432**	0.804
8	0.165**	0.412**	0.851	33	0.461**	0.332**	0.820
9	0.251**	0.362**	0.810	34	0.361**	0.321**	0.813
10	0.151*	0.415**	0.831	35	0.319**	0.261**	0.811
11	0.262**	0.318**	0.829	36	0.338**	0.331*	0.815
12	0.351*	0.331**	0.818	37	0.441**	0.222**	0.814
13	0.012*	0.225**	0.808	38	0.267**	0.253**	0.816
14	0.451**	0.423**	0.803	39	0.431**	0.322**	0.813
15	0.162**	0.313**	0.821	40	0.522**	0.315**	0.802
16	0.281**	0.281**	0.833	41	0.011**	0.232**	0.805
17	0.411**	0.312**	0.812	42	0.421**	0.421**	0.814
18	0.488**	0.321**	0.818	43	0.477**	0.442**	0.808
19	0.502**	0.415**	0.816	44	0.327**	0.372**	0.817
20	0.184*	0.316**	0.823	45	0.611**	0.343**	0.812
21	0.433**	0.452**	0.801	46	0.522**	0.333**	0.821
22	0.541**	0.422**	0.801	47	0.518**	0.361**	0.814
23	0.245**	0.356**	0.811	48	0.432**	0.436**	0.811
24	0.324**	0.310**	0.822	49	0.333**	0.322**	0.808
25	0.441**	0.453**	0.831	50	0.432**	0.411**	0.810

Correlation is significant at the 0.05. **Correlation is significant at the 0.01.

Tabel 2 mengungkapkan bahwa korelasi Item-dimensi signifikan pada level 0,01 untuk semua item tes. Butir 13 signifikan pada tingkat 0,05. Ini menunjukkan keefektifan item tes untuk mengukur dimensi yang akan diukur sudah baik. Korelasi Item-Total signifikan pada level 0,01 untuk semua item. Ini menunjukkan keefektifan item tes untuk mengukur tes sudah baik. Juga Tabel 2 mengungkapkan nilai-nilai *Cronbach's Alpha* jika item telah dihapus (0.801, 0.833), nilai-nilai konvergen ini memberikan lebih banyak bukti untuk efektivitas item.

C. Analisis Tingkat Kesesuaian Butir (Item Fit)

Analisis *Item Fit* digunakan untuk melihat butir-butir soal yang cocok (fit). Apabila hasil analisis menunjukkan sesuai dengan kriteria, item butir soal itu, akan menunjukkan perilaku yang tetap dengan apa yang diharapkan oleh model [21]. Apabila dari hasil analisis ditemukan soal tidak fit, hal ini menunjukkan telah terjadi perbedaan persepsi siswa terhadap butir soal tersebut. Analisis *item fit* dilakukan dengan aplikasi *Add-ins "eirt"* pada Microsoft Excel, hasil analisis di sajikan pada tabel 3.

Tabel 3. Analisis Item Fit

Item	Chi-square	Degrees of freedom	P-Value	Item	Chi-square	Degrees of freedom	P-Value
S001	6.099	8	0.636	S026	3.051	8	0.931
S002	1.873	8	0.985	S027	5.974	8	0.65
S003	4.072	8	0.851	S028	6.568	8	0.584
S004	2.108	8	0.978	S029	3.02	8	0.933
S005	5.961	8	0.652	S030	4.12	8	0.846
S006	2.469	8	0.963	S031	111.234	8	0
S007	2.769	8	0.948	S032	3.396	8	0.907
S008	1.497	8	0.993	S033	5.923	8	0.656
S009	15.139	8	0.057	S034	7.589	8	0.475
S010	12.657	8	0.124	S035	3.797	8	0.875
S011	4.055	8	0.852	S036	8.579	8	0.379
S012	2.117	8	0.977	S037	6.926	8	0.545
S013	146.241	8	0	S038	1.206	8	0.997
S014	3.149	8	0.925	S039	13.104	8	0.108
S015	2.6	8	0.957	S040	2.616	8	0.956
S016	11.418	8	0.179	S041	171.578	8	0
S017	9.773	8	0.281	S042	1.339	8	0.995
S018	15.423	8	0.051	S043	4.73	8	0.786
S019	4.972	8	0.761	S044	5.569	8	0.695
S020	2.983	8	0.935	S045	4.654	8	0.794
S021	4.139	8	0.844	S046	7.122	8	0.523
S022	2.403	8	0.966	S047	10.294	8	0.245
S023	11.841	8	0.158	S048	12.896	8	0.115
S024	0.879	8	0.999	S049	3.437	8	0.904
S025	9.019	8	0.341	S050	1.206	8	0.997

Berdasarkan hasil analisis pada table 3, terlihat 3 (tiga) butir soal nilai *p-value* dibawah 0.05, hal ini menunjukkan bahwa butir soal S013, S031, dan

S041 telah terjadi miskonsepsi siswa pada butir-butir soal tersebut.

Butir-butir soal yang tidak fit menurut hasil analisis tersebut adalah butir soal 13, 31, dan 41, semua nilai *p-value* dari butir soal tersebut adalah 0.000. Pada soal no 13 materinya adalah mencari faktor prima dari 140, butir ini jika dikonfirmasi pada table 1 diperoleh indeks kesukaran 0.00 dengan daya beda 1.702, total *teste* yang menjawab soal dengan benar sebesar 13%. Pada soal nomor 31 indeks kesukaran 0.00 dengan daya beda 1.702 materi yang diujikan mencari koefisien x dari bentuk aljabar, dengan jumlah peserta tes yang mampu menjawab dengan benar 24%. Soal butir ke 41 indeks kesukaran 0.00 dengan daya beda 1.702 materi yang diujikan Persamaan dan Pertidaksamaan, total *testet* yang menjawab dengan benar sama adalah 13%.

Menurut Smith [22], analisis statistik menggunakan *Item Fit* memiliki kekuatan untuk mendeteksi gangguan pengukuran dengan besaran yang masuk akal. Hasilnya menunjukkan bahwa, ketika data tersebut sesuai dengan model, sifat distribusi dari statistik *item fit*, dimungkinkan untuk membangun tingkat kesalahan yang masuk akal. Analisis ini dapat digunakan sebagai kerangka acuan ketika menyelidiki kecocokan data

aktual dengan model Rasch dan *Item Characteristic Curves*.

D. Indikator validitas uji kemampuan numerik

Berdasarkan hasil wawancara pada tim pembuat soal, uji validitas konstruk sudah dilakukan pada soal-soal yang dibuat. Validitas soal diuji oleh guru-guru senior sejumlah 5 orang dengan pengalaman kerja sudah lebih 20 tahun mengajar matematika di SMK. Beberapa catatan dan masukan dari guru-guru sudah dilakukan perbaikan pada butir-butir soal terkait.

Analisis faktor digunakan untuk melihat variabel-variabel yang paling dominan dalam butir-butir soal yang diujikan. Variabel-variabel kemampuan numerik siswa yang diukur adalah bilangan real, perhitungan numerik, perbandingan, aljabar, dan persamaan pertidaksamaan. Analisis faktor dilakukan dengan menggunakan aplikasi SPSS, menggunakan komponen utama dan rotasi *varimax*, digunakan untuk mengeksplorasi apakah tes mengukur lebih dari satu konstruk. Hasil analisis di sajikan pada tabel 4.

Tabel 4. Total Variance of Explained

Component	Initial Eigensvalue			Extraction sums of Squared Loadings		
	Total	% of Variance	Cummulative %	Total	% of Variance	Cummulative %
1	2.541	50.814	50.814	2.541	50.814	50.814
2	0.586	17.124	67.938			
3	0.648	12.955	80.893			
4	0.519	10.374	91.267			
5	0.437	8.733	100			

Faktor yang terbentuk adalah 1 faktor, ini ditunjukkan nilai yang terbentuk dari *Extraction sums of Squared Loadings* hanya 1 yaitu 2.541. Sebagai syarat untuk menjadi faktor adalah nilai *eigensvalue* > 1 [23]. Faktor 1 mampu menjelaskan sebanyak 50.814 % variasi. Berdasarkan analisis faktor dapat ditarik kesimpulan bahwa faktor yang menentukan dalam tes kemampuan numerik adalah komponent 1 yaitu Bilangan Real.

Untuk menguji kecocokan model pada penelitian ini digunakan pemrograman R. Secara umum kecocokan model sudah dijelaskan pada tabel 3 yaitu analisis *Item Fit*. Hal ini sesuai dengan penjelasan Hattie [24], bahwa ketika model sesuai dengan data maka model tersebut sudah menunjukkan kesesuaian. Berdasarkan pengujian analisis faktor telah ditemukan bahwa indikator yang muncul dalam pengujian kemampuan numerik adalah 1 (satu) indikator yaitu pada

variabel “Bilangan Real. Empat variabel lain yaitu perhitungan numerik, perbandingan, aljabar, dan persamaan pertidaksamaan tidak berperan banyak dalam menguji kemampuan numerik siswa SMK, hal ini bisa disebabkan kemampuan siswa dalam menyelesaikan soal-soal pada keempat variabel tersebut rendah.

IV. KESIMPULAN

Pelaksanaan pengujian kemampuan numerik untuk siswa SMK digunakan untuk memetakan kemampuan matematika dasar kelas X, data hasil pengujian akan digunakan oleh sekolah dalam menentukan program-program untuk meningkatkan kemampuan siswa dalam dalam berhitung yang akan mendukung mata pelajaran lain yang membutuhkan kemampuan matematika. Dari hasil analisis terhadap soal-soal yang dikembangkan oleh tim, tidak ditemukan soal-soal

yang terlalu mudah, dan ditemukan beberapa butir soal yang sangat sulit.

Penelitian ini menyelidiki sifat psikometrik dari uji kemampuan numerik melalui pengujian model *Item Response Theory*. Kecocokan model diuji dengan *unidimensi* dan uji *Item Fit*. Berdasarkan pengujian *unidimensi*, model analisis yang digunakan adalah model 2 parameter (2PL). Kecocokan butir-butir terhadap model pengujian 2PL diuji dengan pengujian *Item Fit*.

Tes ini dapat digunakan sebagai tindakan mengeksplorasi kekuatan dan kelemahan dalam kemampuan numerik untuk mendukung kekuatan dan mengevaluasi kelemahan dengan menggunakan strategi dan kegiatan yang sesuai untuk mengajarkan kemampuan numerik. Tes kemampuan numerik untuk siswa dapat digunakan untuk mengkategorikan siswa tergantung pada dimensi tes.

Penelitian di masa depan diharapkan sebuah tes dapat memetakan kemampuan siswa baik pada dimensi pengetahuan dan variable-variabel kemampuan numerik. Melalui analisis yang tepat dan data yang lengkap sekolah bisa menentukan langkah-langkah yang tepat untuk meningkatkan kemampuan siswa dibidang matematika numerik yang akan mendukung mata pelajaran lain di SMK.

DAFTAR PUSTAKA

- [1] Gronlund, N. E. (1993). *Assessment of student achievement*. Allyn & Bacon Publishing. Longwood Division, 160 Gould Street, Needham Heights, MA 02194-2310; tele.
- [2] Wihaskoro, Ahmad Maburi. (2013). *Modul Evaluasi Pembelajaran*.
- [3] Oriondo, L. L., & Antonio, E. M. D. (1998). *Evaluating Educational. Outcomes*, Manila: Rex Book Store.
- [4] Griffin, P., & Nix, P. (1991). *Assessment and reporting: A new approach*.
- [5] Ebel, R. L., & Frisbie, D. A. (1986) *Essentials of. Educational Measurement*.
- [6] Susongko, P. (2016). Validation of science achievement test with the rasch model. *Jurnal Pendidikan IPA Indonesia*, JPII 5(2), 268-277.
- [7] Hambleton, R.K., Swaminathan, H. & Rogers, H.J. (1991). *Fundamental of item response theory*. Newbury Park, CA : Sage Publication Inc.
- [8] Hambleton, R.K. & Swaminathan, H. (1985). *Item response theory: Principles and applications*. Boston. Kluwer. Nijhoff Publishing.
- [9] Lord, Frederick, M. (1990). *Applications of Item Response Theory to Practical Testing Problems*. New Jersey: Lawrence Erlbaum Associates, Publishers.
- [10] M. Candiasa, N. Natajaya, and K. Widiartini. (2018). *Vocational Aptitude Test*. SHS Web Conf., vol. 42, p. 00044.
- [11] G. Sugiharti and B. J. W. Habeahan. (2018). *Influence of Learning Model Using Laboratory and Numeric Ability to Student Learning Result on Thermochemical Material*. *Int. Educ. Stud.*, vol. 11, no. 5, p. 154.
- [12] Harususilo, YE. (2019). *Skor PISA Terbaru Indonesia, Ini 5 PR Besar Pendidikan pada Era Nadiem Makarim*. Retrived from: <http://> "<https://edukasi.kompas.com/read/2019/12/04/13002801/skor-pisa-terbaru-indonesia-ini-5-pr-besar-pendidikan-pada-era-nadiem-makarim?page=all>". Diakses : 24 Januari 2020
- [13] R. . Olatoye and A. . Aderogba. (2011). *Performance of Senior Secondary School Science Students in Aptitude Test: The Role of Student Verbal and Numerical Abilities*. *J. Emerg. Trends Educ. Res. Policy Stud.*, vol. 2, no. 6, pp. 431 – 435.
- [14] Ayodelle, O. D. (2014). *Teacher Instructional Time, Student–Engaged Time and Numerical Ability as Predictors of Student Achievement in Senior Secondary School Chemistry*. *Journal of Emerging Trends in Educational Research an Policy Studies (JETERAPS)*, 5(3), 377-380.
- [15] Othman, F. (2006). *The psychology of individual differences, and mental ability*. Al-Ameen press, Egypt, Cairo.
- [16] Simon, M. K., & Goes, J. (2013). *Ex post facto research*. Retrieved from. dissertationrecipes.com, 2013, Diakses 18 Nov. 2019.
- [17] Eirt, Retrived from <https://libirt.psychometricon.net/>, Diakses 18 Nov 2019.
- [18] Wood, J. *Logistic IRT Models*. Retrived from https://quantdev.ssri.psu.edu/sites/qdev/files/IRT_tutorial_FA17.html, Diakses 18 Nov 2019.
- [19] Djemari Mardapi. (2012). *Pengukuran, Penilaian, dan Evaluasi Pendidikan*. Yogyakarta: Nuha Litera.

- [20] Alagumalai, S., Curtis, D. D., & Hungi, N. (2005). *Applied Rasch Measurement: A Book of Exemplars*. Dordrecht: Springer.
- [21] Hanif, A., Retrived from <https://www.semestapsikometrika.com/2017/07/analisis-dan-seleksi-item-menggunakan.html> , Diakses 18 Nov 2019.
- [22] Smith, R. M. (1991). The distributional properties of Rasch item fit statistics. *Educational and psychological measurement*, 51(3), 541-565.
- [23] <https://www.spssindonesia.com/2018/12/analisis-faktor-dan-interpretasi-spss.html>, Diakses 27 Nov 2019.
- [24] Hattie, J. (1994). An empirical study of various indices for determining unidimensionality Multivariate. *Behavioral Research*, 19, 49-78, http://dx.doi.org/10.1207/s15327906mbr1901_3.

Biodata Penulis

Panyahuti, Sarjana Pendidikan di Jurusan Pendidikan Teknik Mesin Universitas Sriwijaya. Tahun 1995 memperoleh gelar Magister Pendidikan di jurusan Pendidikan Teknologi Kejuruan Program Pascasarjana UNP dengan bidang Komputer & Informatika. Mahasiswa S3 Pendidikan Teknologi dan Kejuruan UNP tahun 2018. Guru SMK.

Krismadinata, menyelesaikan studi S1 di Jurusan Teknik Elektro Fakultas Teknik UNAND. S2 di Institut Teknologi Bandung, dan S3 di Universitas Malaya. Dosen di Jurusan Teknik Elektro Fakultas

Teknik Universitas Negeri Padang sejak tahun 2000 – sekarang.

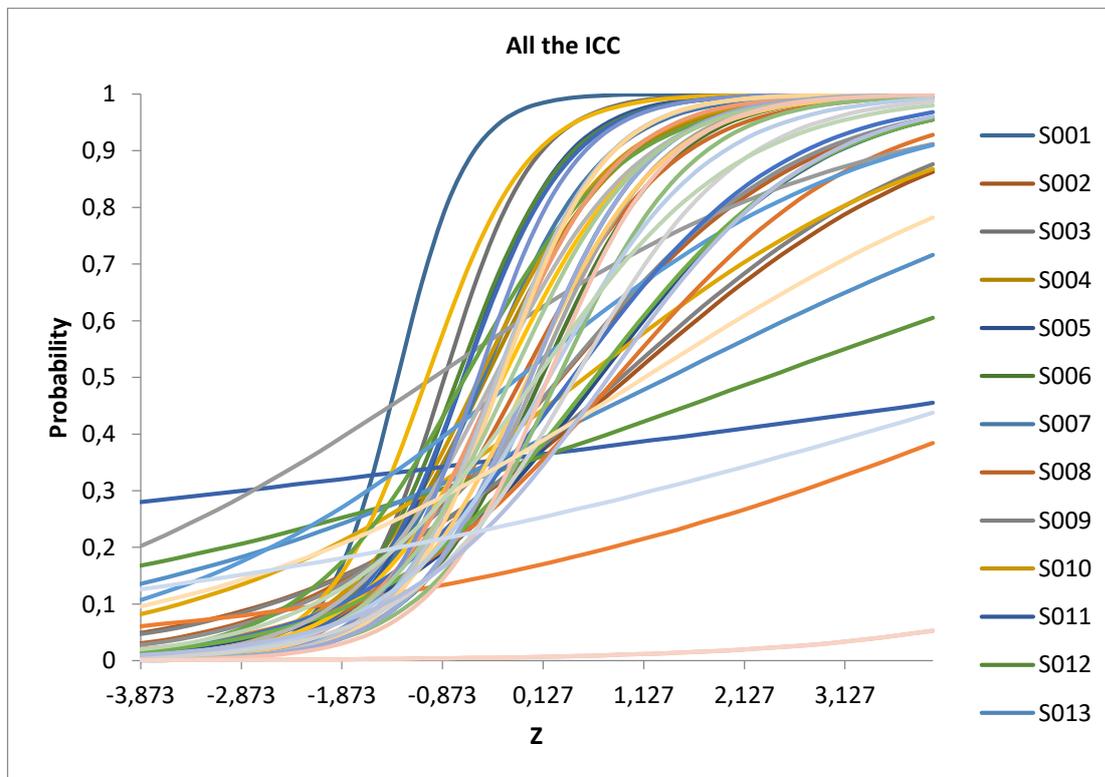
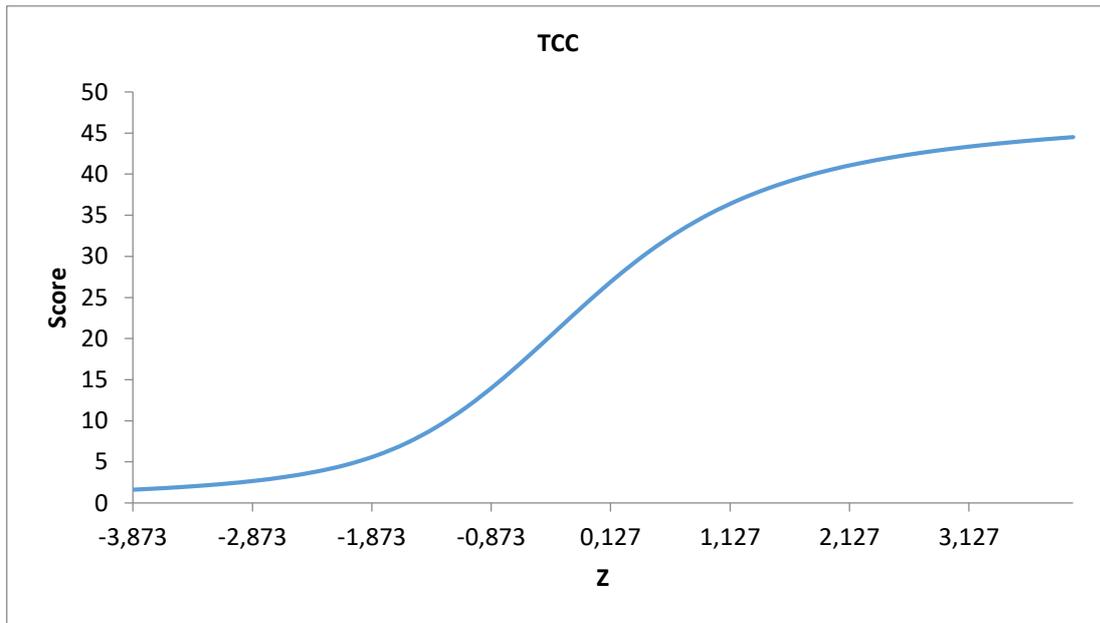
Nizwardi Jalinus, menyelesaikan Program Sarjana Pendidikan Teknik Mesin di IKIP Padang pada tahun 1978 dan Program Pendidikan Vokasi pada tahun 1998 di University of Tasmania, Australia. Saat ini menjadi Guru Besar di Fakultas Teknik Universitas Negeri Padang.

Rusnardi Rahmat, menyelesaikan studi S1 Civil Engineering di Universitas Bung Hatta pada tahun 2000, S2 Structure Engineering di Universitas Gadjah Mada pada tahun 2006, dan S3 Earthquake Engineering and Disaster Mitigation di Kyoto University, Japan Pada tahun 2012. Saat ini Dosen Jurusan Teknik Sipil Fakultas Teknik Universitas Negeri Padang

Ambiyar, menyelesaikan S1 di IKIP Padang Jurusan Teknik Mesin. Lulus sarjana muda pada tahun 1977 dan sarjana pada tahun 1979. Melanjutkan studi S2 Jurusan Pendidikan Teknologi dan Kejuruan di IKIP Jakarta dan lulus tahun 1986. Program studi S3 di UNJ Jakarta lulus tahun tahun 2005

APENDIKS

Characteristic Curves



Information Functions Test

