

Desain dan Validasi Aplikasi Tes Literasi Matematika Berbasis Komputer dengan Pendekatan *Item Response Theory*

Design and Validation of Computer-Based Test of Mathematical Literacy with Item Response Theory Approach

Ilham Falani

Program Studi Pendidikan Matematika Universitas Jambi

*corresponding author: ilhamfalani@unja.ac.id

Abstract

This study aims to develop a computer-based mathematical literacy test application with an item response theory approach that has been tested for its feasibility and advantages. This type of research is Research and Development (R&D) with the Bord and Gall model. Data were collected through observation, questionnaires, and mathematical literacy tests at the junior high school level. This research begins with a need's analysis related to the essence of the application being developed. Then proceed with a literature study to find a suitable development model for developing this mathematical literacy test application. The development of this application is integrated with the Item Response Theory (IRT) approach to improve accuracy in measuring mathematical literacy skills. The application design that is made is then validated by experts who have expertise in accordance with this research variable. This computer application design uses Visual Basic for Application (VBA) in Microsoft Excel. The use of VBA can increase the effectiveness and efficiency of the application design process. The resulting display also becomes more interactive and easier to use. The application that has been designed is then tested on 250 junior high school students. Based on the results of the analysis of the test data, it can be concluded that the application of the mathematical literacy test that has been designed can provide the results of measuring students' mathematical literacy levels with a high level of accuracy.

Keywords: *Development, application, mathematical literacy, computer-based test*

Abstrak

Penelitian ini bertujuan mengembangkan aplikasi tes literasi matematika berbasis komputer dengan pendekatan item response theory yang teruji kelayakan dan keunggulannya. Jenis penelitian ini adalah *Research and Development* (R&D) dengan Model Bord and Gall. Pengumpulan data dilakukan melalui observasi, angket, dan tes literasi matematika pada level sekolah menengah pertama. Penelitian ini diawali dengan analisis kebutuhan terkait esensi aplikasi yang dikembangkan. Kemudian dilanjutkan dengan studi literatur untuk menemukan model pengembangan yang cocok untuk pengembangan aplikasi tes literasi matematika ini. Pengembangan aplikasi ini diintegrasikan dengan pendekatan *Item Response Theory* (IRT) untuk meningkatkan akurasi pada pengukuran kemampuan literasi matematika. Rancangan aplikasi yang dibuat kemudian divalidasi oleh para pakar yang memiliki keahlian yang sesuai. Rancangan aplikasi komputer ini menggunakan *Visual Basic for Application* (VBA) pada *Microsoft Excel*. Hal ini dapat meningkatkan efektifitas dan efisiensi proses perancangan aplikasi. Tampilan yang dihasilkan juga menjadi lebih interaktif dan mudah digunakan. Aplikasi yang telah dirancang selanjutnya diujicobakan kepada 250 siswa sekolah menengah pertama. Berdasarkan hasil analisis terhadap data hasil uji coba, dapat disimpulkan bahwa aplikasi tes literasi matematika yang telah dirancang dapat memberikan hasil pengukuran tingkat literasi matematika siswa dengan tingkat akurasi yang tinggi.

Kata Kunci: Pengembangan, aplikasi, literasi matematika, tes berbasis komputer

PENDAHULUAN

Revolusi industri menjadi isu yang populer dan banyak dibahas saat ini. Era relovusi industry memunculkan tantangan di berbagai bidang kehidupan, perkembangan suatu negara tergantung untuk dapat bersaing di era ini sangat ditentukan oleh tingkat keterampilan yang dimiliki oleh sumber daya manusianya. Untuk bisa bersaing di Era Industri 4.0, masyarakat harus mampu responsif terhadap perubahan dan mampu mengembangkan berbagai keterampilan sesuai perkembangan teknologi informasi dan komunikasi. Perkembangan teknologi informasi dan komunikasi di era Revolusi Industri mengubah budaya literasi individu menjadi literasi digital. Literasi digital diarahkan pada tujuan peningkatan kemampuan membaca, menganalisis, dan menggunakan informasi di dunia digital (*Big Data*). Hal ini sangat berdampak pada dunia pendidikan yang dituntut untuk dapat mencetak sumber daya manusia yang memiliki kemampuan literasi digital yang tinggi.

Literasi matematika merupakan salah satu bagian penting dari literasi digital. Literasi matematika merupakan kemampuan individu untuk merumuskan, menerapkan, dan menafsirkan matematika dalam berbagai konteks, termasuk kemampuan melakukan penalaran secara matematis dan menggunakan konsep, prosedur, fakta, dan alat matematika, untuk mendeskripsikan, menjelaskan, dan memprediksi suatu fenomena atau kejadian (Niss & Jablonka, 2020).

Peningkatan kemampuan literasi matematika perlu dilakukan sejak dini melalui pendidikan. Teknologi dan pembelajaran matematika memiliki kaitan erat, terlebih di era Revolusi Industri 4.0 (Rafiqoh, 2020). Menurut Lase (2019), pendidikan matematika harus berusaha

mengakomodasi tuntutan kemampuan yang dibutuhkan di era Revolusi Industri 4.0 yang dinamis agar sesuai dengan perkembangan teknologi. Pemanfaatan teknologi yang tepat dalam pembelajaran akan memberi penguatan terhadap pola perubahan paradigma pembelajaran. Penggunaan teknologi adalah salah satu cara yang efektif dan efisien dalam menyampaikan informasi kepada peserta didik. Komputer merupakan salah satu teknologi yang berpotensi besar meningkatkan kualitas pembelajaran matematika. Sifat abstrak objek kajian matematika yang sulit dipikirkan oleh peserta didik dapat dipresentasikan melalui simulasi komputer agar lebih konkrit.

Ujian berbasis komputer atau biasa disebut CBT (*Computer-Based Test*) merupakan model percontohan digitalisasi bagi sekolah. Digitalisasi merupakan salah satu dampak langsung dari kemajuan teknologi yang tak dapat dihindari (Hariri et al., 2018; Khoshshima, Hosseini, & Toroujeni, 2017). Kertas dan alat tulis tak dibutuhkan lagi, sementara pendidikan tetap berjalan baik menggunakan metode digital. Ujian berbasis komputer merupakan salah satu bentuk usaha untuk beradaptasi menyesuaikan dengan perkembangan teknologi. Tidak dapat dipungkiri bahwa CBT memiliki banyak keunggulan, diantaranya adalah CBT dapat mengefisiens penggunaan anggaran pada ujian yang berskala besar, pelaksanaan CBT dapat megefektifkan waktu dalam pelaksanaan proses serta administrasi ujian, CBT dapat memudahkan siswa saat mengerjakan ujian, CBT dapat menampilkan tes dalam bentuk yang lebih interaktif disertai dengan stimulus bagi siswa, selain itu administrative hasil ujian dapat menjadi lebih cepat dan akurat, melalui CBT, kita berpartisipasi aktif menjaga kelestarian bumi dengan mengurangi penggunaan kertas (Jan dkk.,

2016; Clarke, 2018; Rezaie & Golshan, 2015; Luo dkk., 2016).

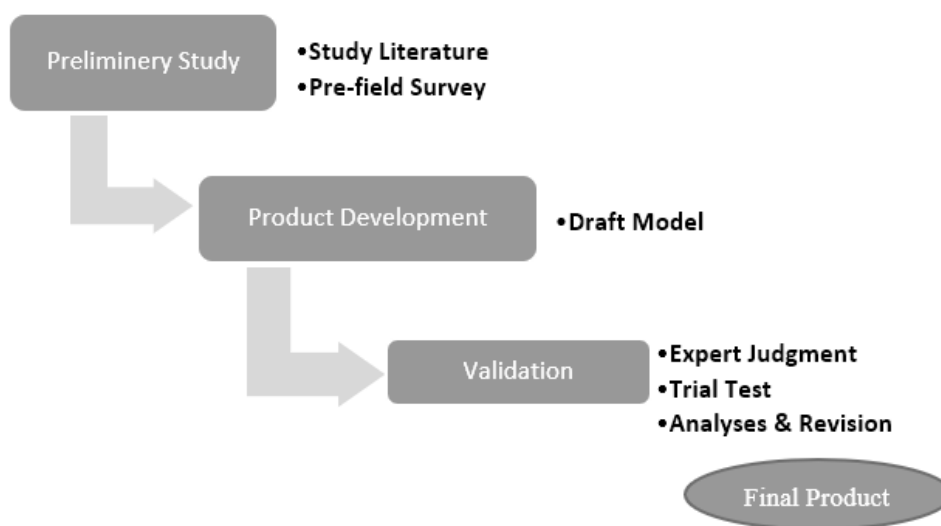
Analisis data respon siswa merupakan hal yang esensi dalam pengukuran pendidikan. Oleh karena itu, proses analisis ini perlu menjadi fokus penting dalam sebuah pengembangan instrument. Prosedur analisis dimulai dari proses mendapatkan informasi mengenai siswa melalui skor hasil ujiannya (Silvia dkk, 2015; Marwiyah dkk, 2015; Frianto dkk, 2019).

Item Response Theory (IRT) merupakan sebuah pendekatan modern dalam menganalisis hasil ujian, dimana pendekatan ini diusulkan untuk memperbaiki kelemahan pada teori tes klasik yang banyak digunakan (Ayala, 2013; Baker, 2012; Bastari, 2000). Pendekatan IRT diharapkan pengukuran yang dilakukan menghasilkan informasi yang lebih akurat (Naga, 2012; Falani dkk, 2020; Falani & Kumala, 2017; Supriyati dkk, 2021). Berdasarkan paparan yang telah dijelaskan tersebut, perlu adanya sebuah penelitian yang berfokus pada pengembangan instrument tes literasi matematika berbasis komputer

dengan pendekatan IRT. Hal ini bertujuan untuk mengoptimalkan pengukuran literasi siswa dengan memanfaatkan kemajuan teknologi.

METODE PENELITIAN

Jenis penelitian ini adalah *Research and Development* (R&D) dengan Model pengembangan yang digunakan adalah Model Bord and Gall (1983). R&D bertujuan untuk mengembangkan atau memvalidasi produk yang digunakan dalam pendidikan. Pengembangan ini dilakukan melalui beberapa tahapan yang sistematis. Hasil pengembangan ini bertujuan untuk menemukan pengetahuan yang secara praktis dapat diaplikasikan. Penelitian ini dilakukan pada bulan Januari-Agustus tahun 2020, di departemen program studi penelitian dan evaluasi pendidikan, Universitas Negeri Jakarta. Berikut adalah tahapan-tahapan penelitian pengembangan yang dilakukan,



Gambar 1. Alur penelitian dan pengembangan

Studi Pendahuluan (*Preliminary Study*)

Penelitian pengembangan ini diawali dengan sebuah studi pendahuluan yang bertujuan untuk menetapkan spesifikasi dan rancangan produk yang akan dikembangkan. Studi pendahuluan ini terdiri dari dua kegiatan utama, yaitu studi literatur dan studi lapangan. Studi pendahuluan ini menghasilkan gambaran awal terkait aplikasi tes literasi matematika berbasis komputer yang akan dikembangkan. Studi literatur dilakukan melalui pengkajian teori-teori yang berkaitan dengan produk yang akan dirancang. Studi literatur ini juga mengkaji penelitian-penelitian terdahulu yang relevan dengan produk yang akan dikembangkan. Studi lapangan pada penelitian pendahuluan dilakukan melalui observasi di sekolah untuk mengetahui kondisi proses pembelajaran matematika di lapangan, khususnya terkait dengan pelaksanaan ujian matematika.

Development Planning Model

Pada tahapan ini dilakukan penyusunan rancangan produk yakni aplikasi tes literasi matematika berbasis komputer. Perancangan ini dilakukan berdasarkan studi pendahuluan yang telah dilakukan. Perancangan ini mencakup beberapa kegiatan, yakni mendefinisikan standar kompetensi tes yang akan diujikan, mendefinisikan tujuan pengembangan produk, mengidentifikasi pihak-pihak terkait yang terlibat, menyusun prosedur kerja serta uji kelayakan yang akan dilakukan. Perencanaan ini merupakan tahapan yang penting dalam pengembangan yang akan dilakukan. Perencanaan yang tepat menentukan keberhasilan pengembangan produk yang akan dilakukan. Hasil dari perencanaan pengembangan ini adalah draft model aplikasi yang siap untuk diujicobakan.

Validasi Rancangan Model Aplikasi

Tahapan ini mencakup beberapa kegiatan, yakni validasi pakar, uji coba, analisis, dan revisi. Validasi pakar merupakan proses penilaian kelayakan rancangan aplikasi yang telah dibuat. Kelayakan aplikasi mencakup: kesesuaian antara butir-butir tes dengan indikator standar kompetensi literasi digital, tingkat kesulitan soal dengan level kemampuan peserta tes, dasar-dasar konsep maupun teori yang digunakan, bahasa, tampilan dan desain grafis aplikasi, dan kepraktisan aplikasi untuk digunakan. Pakar yang dilibatkan dalam penilaian kelayakan rancangan produk ini adalah pakar yang memiliki keahlian terkait dengan produk yang dikembangkan. Para pakar yang dilibatkan dalam penelitian ini adalah dosen dan praktisi yang dalam dunia pendidikan khususnya pendidikan matematika. Penilaian kelayakan rancangan produk yang dilakukan oleh para ahli ini dilakukan secara kualitatif dan kuantitatif.

Perhitungan kuantitatif indeks penilaian pakar dilakukan dengan menggunakan Metode Aiken. Setelah dinyatakan layak oleh para ahli, selanjutnya rancangan aplikasi tersebut diujicobakan. Uji coba rancangan aplikasi dilakukan terhadap 250 siswa sekolah menengah pertama. Uji coba ini ditujukan untuk menilai atau menjustifikasi kepraktisan penggunaan aplikasi yang telah dirancang. Selanjutnya, data yang diperoleh dari hasil uji coba dianalisis secara statistik, terutama untuk menjamin validitas dan reliabilitas. Kemudian tahap berikutnya, adalah revisi, penyempurnaan, dan penyesuaian aplikasi. Hasil akhir dari proses ini adalah sebuah produk akhir berupa sebuah aplikasi tes literasi matematika berbasis komputer yang teruji kehandalannya.

HASIL DAN PEMBAHASAN

Hasil dari penelitian pengembangan yang telah dilakukan adalah berupa aplikasi tes literasi matematika berbasis komputer dengan pendekatan IRT, berikut beberapa hasil dari analisis yang telah dilakukan,

Berdasarkan hasil studi literature yang telah dilakukan, butir-butir yang telah disusun dalam instrument tes literasi matematika memiliki tiga domain utama, yakni.

Tabel 1. Struktur butir tes

Format Butir	Jumlah soal	Penyekoran	Model Analisis IRT
Pilihan Ganda	30	Dikotomi 0 (Salah) 1 (benar)	3 Parameter Logistic Model

Tabel 2. Indikator domain proses literasi matematika

No	Kategori Domain	Indikator	Jumlah Soal
1	Merumuskan situasi matematis (<i>formulate</i>)	<ul style="list-style-type: none"> Mengidentifikasi aspek-aspek matematika dalam permasalahan yang berkaitan dengan konteks nyata serta mengidentifikasi variable-variabel yang signifikan Menyederhanakan situasi atau masalah untuk menjadikannya mudah diterima dengan analisis matematika Memahami dan menjelaskan hubungan Antara Bahasa, symbol, dan konteks sehingga dapat disajikan secara matematika Mengubah permasalahan ke dalam Bahasa matematika atau model matematika Merancang dan mengimplementasikan strategi untuk menemukan solusi matematika 	10
2	Menggunakan konsep matematika, prosedur, fakta, dan penalaran (<i>employ</i>)	<ul style="list-style-type: none"> Menerapkan fakta, aturan, algoritma, dan struktur matematika dalam menemukan solusi Memanipulasi bilangan, grafik, data statistic, bentuk aljabar, informasi, persamaan, dan bentuk geometri dan menggantikan berbagai macam situasi dalam proses menemukan solusi Merefleksikan pendapat matematika dan menjelaskan serta memberikan penguatan hasil matematika. 	10
3	Menafsirkan, menerapkan, dan mengevaluasi hasil dari suatu proses matematika (<i>interpret</i>)	<ul style="list-style-type: none"> Menginterpretasikan kembali hasil matematika ke konteks nyata Mengevaluasi alasan-alasan yang masuk akal dan solusi matematika ke dalam masalah nyata Menjelaskan mengapa hasil matematika sesuai atau tidak sesuai dengan permasalahan konteks yang diberikan. 	10

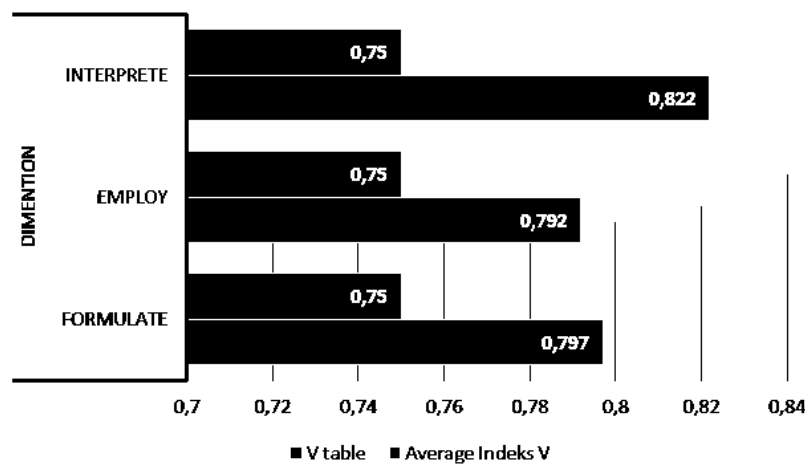
(Nirmala, Isnarto, & Mulyono; 2019).

Hasil Validasi Pakar

Validasi tes oleh para ahli dilakukan secara kuantitatif dan kualitatif. Pada tahap ini, peneliti memberikan instrumen tes matematika dengan format butir Pilihan Ganda atau *Multiple Choice* (MC) yang telah disusun kepada para ahli di bidang Matematika, ahli matematika yang dilibatkan adalah sebanyak tujuh orang yang merupakan ahli dan praktisi dibidang matematika. Para ahli memberikan masukan dan saran untuk perbaikan tes.

Pada pengajuan awal jumlah butir yang diajukan kepada para ahli adalah sebanyak 32 butir, namun setelah penilaian dilakukan terdapat dua butir yang didrop sehingga tertinggal 30 butir.

Berikut adalah indeks kesepakatan hasil penilaian para ahli terhadap kelayakan butir-butir tes literasi matematika untuk tiap dimensi.



Gambar 2. Indeks Aiken dimensi tes literasi matematika

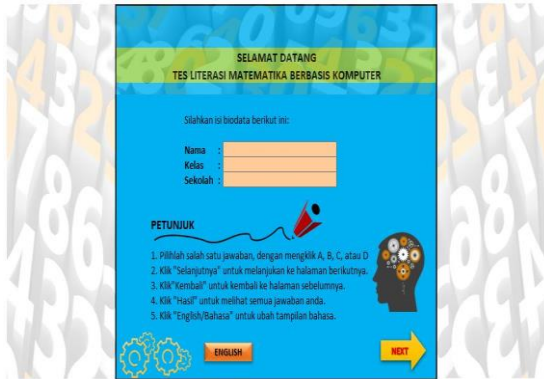
Rangkuman hasil penilaian dari para ahli untuk tiap butir tes dapat dilihat pada gambar di atas. Perhitungan validasi isi oleh para ahli dilakukan dengan menggunakan metode Aiken. Perhitungan nilai V dengan menggunakan persamaan,

$$V = \frac{\sum s}{n(c-1)}, \text{ dengan } s = r - I_0.$$

V adalah indeks kesepakatan ahli atau pakar mengenai validitas butir, s adalah skor yang ditetapkan setiap penilai @ dikurangi skor terendah dalam kategori penilaian I_0 yang digunakan. n adalah banyaknya penilai, dan c adalah banyaknya kategori penilaian. Hasil perhitungan nilai V untuk tiap butir lebih besar dibandingkan dengan kriteria minimal sesuai nilai V tabel yakni sebesar 0.75. Hal ini menunjukkan bahwa butir-butir pada tiap dimensi yang disusun telah memenuhi kriteria validitas yang layak.

Struktur Aplikasi Tes Literasi Matematika

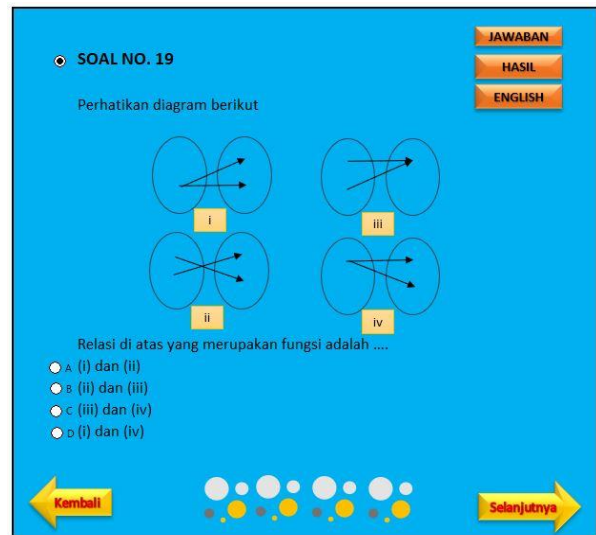
Aplikasi tes literasi matematika berbasis komputer yang dikembangkan dalam penelitian ini memiliki beberapa komponen, sebagai berikut:



Gambar 3. *Homepage*

Halaman ini merupakan tampilan awal pada aplikasi tes literasi matematika yang telah dibuat. *Homepage* ini juga dikenal dengan istilah “cover”. Pada halaman ini terdapat beberapa komponen tampilan meliputi, kotak biodata peserta tes (nama, kelas, asal sekolah), petunjuk penggunaan, petunjuk penggunaan aplikasi tes. Selain itu terdapat button “ENGLISH” untuk merubah tampilan menjadi Bahasa Inggris dari Bahasa Indonesia yang digunakan sebagai Default Bahasa. Tes literasi matematika yang dirancang terdiri dari dua Bahasa yakni Bahasa Indonesia (default) dan Bahasa Inggris. Selanjutnya ada tombol “NEXT” untuk melanjutkan ke halaman berikutnya atau berfungsi juga untuk memulai tes.

Halaman Soal



Gambar 4. Halaman Soal (Contoh Soal Nomor 19)

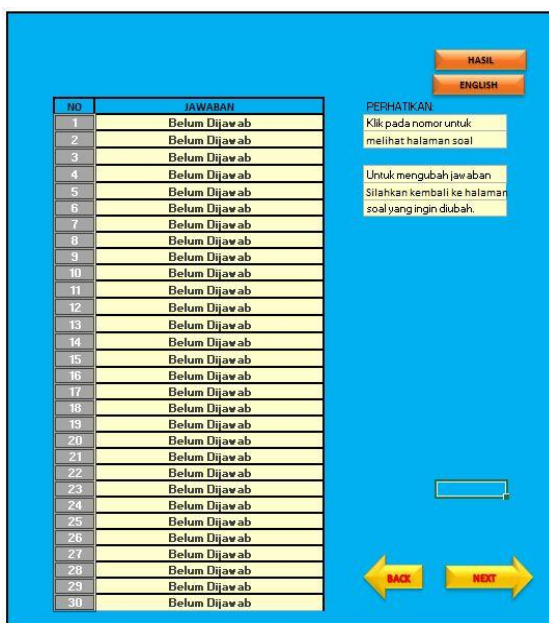


Gambar 5. Halaman Soal (Contoh Soal Nomor 28)

Halaman ini merupakan bagian inti dari aplikasi tes yang telah disusun, pada bagian ini ditampilkan butir soal yang harus di jawab oleh peserta tes dengan mengklik salah satu pilihan jawaban yang disediakan (A, B, C, atau D). pada bagian ini terdapat beberapa komponen/tombol bantu, yakni tombol “JAWABAN” tombol ini merupakan tompol yang berfungsi menampilkan rangkuman jawaban dari seluruh soal yang telah dijawab oleh peserta tes tersebut. Tombol “Hasil”

berfungsi menampilkan hasil akhir dari tes yang telah dikerjakan oleh peserta tersebut. Tombol “Kembali” berfungsi untuk mengembalikan halaman menuju halaman sebelumnya, sedangkan tombol “selanjutnya” berfungsi mengarahkan peserta ke halaman berikutnya. Dengan menekan tombol “Selanjutnya” jawaban peserta tes akan tersimpan secara otomatis. Tombol “ENGLISH” berfungsi untuk alih Bahasa menjadi Bahasa Inggris.

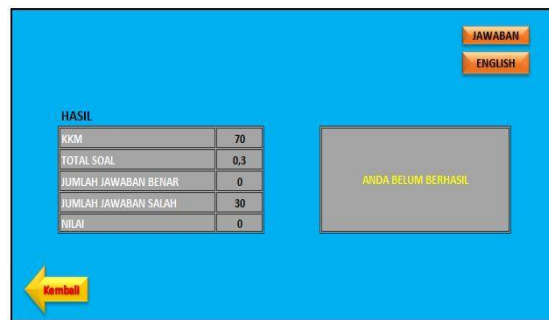
Halaman Jawaban



Gambar 6. Halaman Jawaban

Halaman ini menampilkan secara keseluruhan jawaban yang telah diberikan oleh peserta tes untuk masing-masing butir. Halaman ini juga disertai dengan petunjuk bagaimana untuk mengubah jawaban, jika peserta tes ingin melakukan revisi jawaban. Tombol ini juga disertai tomo “ENGLISH” untuk alih Bahasa ke Bahasa Inggris, “Result” untuk menampilkan hasil akhir penyekoran, tombol *BACK* untuk Kembali ke halaman sebelumnya, tombol *NEXT* untuk melangkah ke halaman berikutnya.

Halaman Hasil



Gambar 7. Halaman hasil

Halaman ini menghasilkan hasil dari penyekoran dari jawaban peserta tes. Pada halaman ini ditampilkan juga Standar cut off score yang digunakan, bobot tiap butir, jumlah jawaban benar, jumlah jawaban salah, total skor akhir, serta keterangan yang menyatakan “selamat anda berhasil” saat skor yang diperoleh oleh peserta tes telah melewati batas minimal kriteria kelulusan yang telah ditentukan.

Hasil Uji coba Lapangan

Aplikasi tes literasi matematika yang telah dibuat diujicobakan kepada siswa yang berjumlah 250 siswa sekolah menengah pertama di kota Depok, Indonesia. Berikut adalah analisis berdasarkan hasil ujicoba yang telah dilakukan,

Uji reliabilitas dilakukan terhadap data response siswa menggunakan Metode Cronbach's Alpha. Berikut adalah hasil uji reliabilitas berdasarkan output SPSS 21.

Tabel 3. Uji Reliabilitas

<i>Case Processing Summary</i>		N	%
<i>Cases</i>	<i>Valid</i>	35	100.0
	<i>Excluded^a</i>	0	.0
	<i>Total</i>	35	100.0

a. Listwise deletion based on all variabels in the procedure.

Case Processing Summary

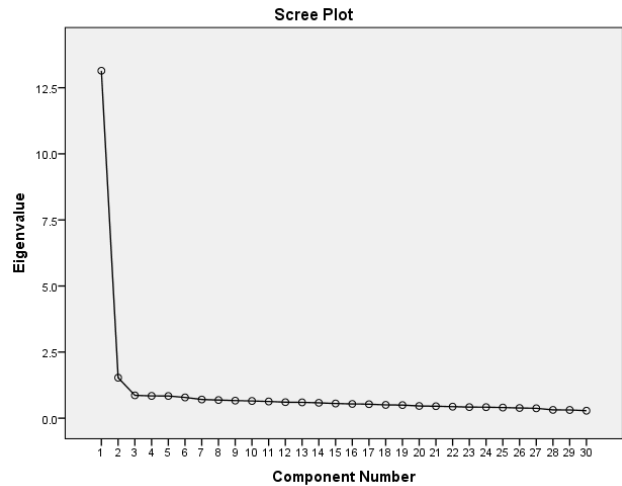
		N	%
Cases	Valid	35	100.0
	Excluded ^a	0	.0
	Total	35	100.0

a. Listwise deletion based on all variabls in the procedure.

Berdasarkan tabel di atas diperoleh Nilai Cronbach’s Alpha sebesar 0,949, lebih dari kriteria nilai minimal Cronbach’s Alpha yakni 0,6. Dapat disimpulkan tes yang digunakan dalam pengumpulan dalam penelitian ini handal atau reliabel. Berdasarkan hasil uji-uji statistik pada data respon siswa yang telah dilakukan, dapat disimpulkan bahwa instrumen tes pada penelitian ini telah siap digunakan pengumpulan data penelitian karena telah memenuhi kriteria validitas dan reliabilitas.

Hasil Uji Model Item Response Theory

Model yang digunakan dalam analisis data respon siswa hasil ujicoba adalah tiga parameter logistic model (3PLM). Sebelum dilakukan analisis data lebih lanjut menggunakan PARSCALE 4.1. Perlu adanya pengujian asumsi prasyarat IRT. Menurut Alagoz (2000), asumsi unidimensi dapat diuji menggunakan analisis faktor, dengan bantuan *software Statistical Package for Social Science (SPSS) 21* (Almquist, Ashir, & Brännström, 2020). Uji unidimensi dilakukan terhadap data-data sebelum digunakan untuk estimasi parameter kemampuan peserta tes. Berikut adalah grafik *scree plot* untuk data hasil uji coba.



Gambar 8. *Scree Plot* untuk Data dengan 3PLM

Hasil perhitungan analisis faktor terhadap data-data yang diuji serta *Scree Plot* pada Gambar di atas menunjukkan bahwa faktor utama tiap data dapat menjelaskan sebagian besar dari total varians. Sehingga dapat disimpulkan bahwa butir-butir tes yang digunakan unidimensi. Sebagian besar butir-butir soal membentuk sebuah faktor yang dapat disebut literasi matematika.

Setelah dilakukan uji prasyarat unidimensi, selanjutnya dilakukan uji kecocokan model. Berdasarkan hasil uji kecocokan pada butir tes dengan bantuan PARSCALE 4.1 diperoleh item fit statistics nilai chi kuadrat 3PLM sebesar 434.71 (p-value= 0.101), nilai yang dihasilkan kurang dari nilai kritis distribusi χ^2 nya, hal ini menunjukkan bahwa model yang digunakan memiliki kecocokan yang tinggi dengan data empiris. (Naga, 2012; Retnawati, 2014; Yılmaz, 2019). Selanjutnya model-model tersebut dapat digunakan untuk mengestimasi parameter kemampuan literasi matematika peserta tes. Berikut adalah deskripsi statistik hasil estimasi parameter kemampuan literasi matematika yang telah dilakukan.

Tabel 4. Distribusi frekuensi hasil estimasi parameter kemampuan literasi matematika dengan 3PLM

No	Kelompok	Median	Frek.	Prob.	Frek. Kum	Prob. Kum
1	-3,45 – 2,87	-3,16	1	0%	1	0%
2	-2,86 – -2,29	-2,58	0	0%	1	0%
3	-2,28 – -1,71	-2,00	0	0%	1	0%
4	-1,70 – -1,13	-1,42	194	13%	195	13%
5	-1,12 – -0,55	-0,84	324	21%	519	34%
6	-0,54 – 0,03	-0,25	351	23%	870	57%
7	0,04 – 0,61	0,33	254	17%	1124	73%
8	0,62 – 1,19	0,91	199	13%	1323	86%
9	1,20 – 1,77	1,49	107	7%	1430	93%
10	1,78 – 2,36	2,07	73	5%	1503	98%
11	2,37 – 2,94	2,65	26	2%	1529	100%
12	2,95 – 3,52	3,23	2	0%	1531	100%

Penelitian ini menghasilkan aplikasi tes literasi matematika berbasis komputer dengan pendekatan item response theory. Aplikasi yang dirancang telah divalidasi oleh sejumlah pakar dan juga telah diujicobakan kepada siswa sekolah menengah pertama. Berdasarkan hasil analisis yang telah dilakukan dapat dilihat bahwa aplikasi yang dikembangkan telah memenuhi kriteria validitas dan validitas yang tinggi. Hal ini dapat dilihat dari hasil uji validitas oleh para ahli yang menunjukkan bahwa indeks Aiken untuk dimensi interprete, employ, dan formulate berturut-turut adalah 0.822, 0.792, dan 0.797. Ketiga nilai ini telah melebihi batas minimal kriteria Indeks Tabel Aiken yakni 0.75. Lebih lanjut, hasil uji reliabilitas yang telah dilakukan menunjukkan bahwa indeks reliabilitas yang dihasilkan adalah 0.94 (lebih dari standar 0.6). Sehingga instrument yang disusun dapat dinyatakan layak untuk mengukur variable laten pada penelitian ini kemampuan literasi matematika.

3PLM memiliki prinsip memprediksi probabilitas respons tertentu berdasarkan tingkat kemampuan/sifat peserta tes (Hambleton & Swaminathan, 1985; Hambleton, Swaminathan, & Rogers,

1991; Stenbeck et al., 1992). Pemilihan model 3PLM dalam penelitian ini memiliki hubungan erat dengan format tes yang digunakan, yakni format multiple-choice. Model IRT 3PLM lebih sering digunakan untuk tes format MC. Model ini memiliki tiga parameter yakni, a (daya beda), b (tingkat kesulitan), dan c (tebakan) (Naga, 2012). Menurut Naga (2012), Retnawati (2014), dan Samejima (1979), butir format MC sangat relevan untuk butir MC, karena model ini dapat probabilitas peserta menjawab secara tebakan. Sehingga model ini sangat cocok untuk digunakan pada tes penilaian kemampuan atau pengetahuan dibandingkan model 1PLM atau 2PLM (Żóltak, & Golonka, 2015).

Hasil analisis data uji coba menunjukkan bahwa 3PLM yang merupakan model yang digunakan dalam analisis response siswa terhadap aplikasi ini menunjukkan kecocokan dengan data lapangan. Selain itu diperoleh hasil nilai fungsi informasi dalam estimasi kemampuan literasi yang tinggi. Hal ini menunjukkan bahwa estimasi yang dilakukan akurat. Nilai fungsi informasi memiliki kaitan erat dengan akurasi estimasi, semakin tinggi nilai fungsi informasi maka akan semakin tinggi hasil estimasi yang dihasilkan oleh

sebuah tes, dan juga sebaliknya (Baker 2001; Ercikan, 1998; Naga, 2012).

Desain aplikasi tes literasi matematika yang dikembangkan pada penelitian ini menggunakan fitur *Visual Basic for Application* (VBA) pada Microsoft Excel. Hal ini membuat perancangan yang dilakukan lebih mudah dan sederhana (Bernard, 2018; Alexander & Kusleika, 2019)., sehingga mudah untuk direplikasi dan dicontoh oleh penelitian selanjutnya. Pemilihan fitur ini juga ditujukan untuk menghemat pengeluaran biaya yang berlebih, namun dengan kualitas yang baik dan praktis (mudah untuk diaplikasikan) (Roth, Goldwasser, & Parker, 2016).

Penelitian ini memiliki keterbatasan yang bisa menjadi peluang untuk penelitian lebih lanjut. Model yang digunakan dalam penelitian ini terbatas hanya pada model 3PLM, penelitian yang akan datang dapat meneliti lebih lanjut penggunaan model-model IRT lainnya sebagai upaya meningkatkan akurasi pengukuran dalam pendidikan. Selain itu, tes yang digunakan pada aplikasi ini hanya dibatasi pada tes format multiple choice. Tes format campuran dapat menjadi salah satu alternatif lain dalam format tes yang dapat diteliti lebih lanjut (Ercikan, 1998).

KESIMPULAN

Penelitian pengembangan yang dilakukan pada penelitian ini bertujuan menghasilkan aplikasi tes literasi matematika berbasis komputer yang mengintegrasikan *item response theory* dalam analisisnya. Berdasarkan hasil analisis dan pembahasan yang telah dilakukan dalam penelitian ini, menunjukkan bahwa aplikasi tes literasi matematika berbasis komputer yang dihasilkan telah memenuhi kriteria validitas dan reliabilitas yang terbukti secara ilmiah kehandalannya. Oleh karena

itu aplikasi tes ini dapat digunakan untuk mengukur literasi matematika. Selain itu, penelitian ini juga dapat digunakan sebagai referensi penelitian pengembangan lebih lanjut terkait aplikasi tes berbasis komputer untuk bidang kajian lainnya, terutama yang memiliki karakteristik yang serupa dengan matematika.

DAFTAR PUSTAKA

- Alagoz, C. (2000b). *Scoring Tests with Dichotomous and Polytomous Items*. Gazy Univeristy.
- Almquist, B., Ashir, S., & Brännström, L. (2020). *A Guide to Quantitative Methods*. Sweden: Stockholm University.
- Alexander, M., & Kusleika, D. (2019). *Excel 2019 Power Programming with VBA*. John Wiley & Sons.
- Ayala, R. J. de. (2013). *The theory and practice of item response theory*. (D. A. Kenny & Todd D. Little, Eds.). New York: Guilford Publications.
- Baker, F. B., & Kim, S. (2012). *Item Response Theory Parameter Estimation Techniques*. New York: Taylor & Francis.
- Baker, F. B., & Kim, S. (2017). *The Basics of Item Response Theory Using R*. New York: Springer International Publishing.
- Bastari. (2000a). An Investigation of Linear and Non-linear Estimates for Multidimensional Graded Response Model. *Unpublished*.
- Bernard, M., Minarti, E. D., & Hutajulu, M. (2018). Constructing Student's Mathematical Understanding Skills and Self Confidence: Math Game with Visual Basic Application for Microsoft Excel in Learning Phytagogoras at Junior High School. *International Journal of Engineering & Technology*, 7(3.2), 732-736.

- Clarke, A. (2018). *Designing computer-based learning materials*. Routledge.
- Dan Gall, B. (1983). *Educational research: An introduction*. New York Longman
- Ercikan, K., Schwarz, R. D., Julian, M. W., Burket, G. R., Weber, M. M., Link, V., ... Link, V. (1998). Calibration and Scoring of Tests With Multiple-Choice and Constructed-Response Item Types. *Journal of Educational Measurement*, 35(2), 137–154.
- Falani, I., Akbar, M., & Naga, D. S. (2020). The Precision of Students' Ability Estimation on Combinations of Item Response Theory Models. *International Journal of Instruction*, 13(4), 545-558.
- Falani, I., & Kumala, S. A. (2017). Kestabilan Estimasi Parameter Kemampuan Pada Model Logistik Item Response Theory Ditinjau dari Panjang Tes. *SAP (Susunan Artikel Pendidikan)*, 2(2).
- Frank B, B. (2001). *The Basics of Item Response Theory*. USA: ERIC Clearinghouse on Assessment and Evaluation.
- Frianto, A., Risnita, R., & Maison, M. (2019). Pengembangan Alat Evaluasi Berbasis Guided Discovery Learning untuk Mengukur Kemampuan Berpikir Kritis Siswa SMA Kelas X: Development of Evaluation Tools Based on Guided Discovery Learning to Measure the Critical Thinking Skill of High School Students in Class X. *Edu-Sains: Jurnal Pendidikan Matematika dan Ilmu Pengetahuan Alam*, 8(1), 17-22.
- Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: Principles and applications* (8th ed.). New York: Springer Science+Business media, LLC. <https://doi.org/10.1007/978-94-017-1988-9>
- Hambleton, R. K., Swaminathan, H., & Rogers, D. J. (1991). *Fundamentals of Item Response Theory*. Sage Publications.
- Hariri-Akbari, M., Shokrvash, B., Mahmoodi, F., Jahanjoo-Aminabad, F., Yousefi, B., & Azabdaftari, F. (2018). Conversion of extrinsic into intrinsic motivation and computer-based testing (CBT). *BMC medical education*, 18(1), 1-8.
- Jan, S. R., Shah, S. T. U., Johar, Z. U., Shah, Y., & Khan, F. (2016). An innovative approach to investigate various software testing techniques and strategies. *International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET)*, Print ISSN, 2395-1990.
- Khoshsima, H., Hosseini, M., & Toroujeni, S. M. H. (2017). Cross-Mode Comparability of Computer-Based Testing (CBT) versus Paper-Pencil Based Testing (PPT): An Investigation of Testing Administration Mode among Iranian Intermediate EFL Learners. *English Language Teaching*, 10(2), 23-32.
- Lase, D. (2019). Pendidikan di era revolusi industri 4.0. *SUNDERMANN: Jurnal Ilmiah Teologi, Pendidikan, Sains, Humaniora dan Kebudayaan*, 12(2), 28-43.
- Luo, W., Pelletier, J., Duffin, K., Ormand, C., Hung, W. C., Shernoff, D. J., ... & Furness, W. (2016). Advantages of computer simulation in enhancing students' learning about landform evolution: A case study using the Grand Canyon. *Journal of Geoscience Education*, 64(1), 60-73.
- Marwiyah, S., Kamid, K., & Risnita, R. (2015). Pengembangan instrumen penilaian keterampilan berpikir kreatif pada mata pelajaran IPA terpadu materi atom, ion, dan molekul SMP Islam Al Falah. *Edu-Sains: Jurnal Pendidikan Matematika Dan Ilmu Pengetahuan Alam*, 4(1).
- Naga, D. S. (2012). *Teori Sekor pada Pengukuran Mental*. Jakarta: PT Nagarani Citrayasa.

- Niss, M., & Jablonka, E. (2020). Mathematical literacy. *Encyclopedia of mathematics education*, 548-553.
- Nirmala, M. D., Isnarto, I., & Mulyono, M. (2019). Kemampuan Literasi dalam Menyelesaikan Soal Ujian Nasional Berbasis Komputer yang Mempunyai Daya Serap Rendah Siswa Kelas XII. In *Prosiding Seminar Nasional Pascasarjana (PROSNAMPAS)* (Vol. 2, No. 1, pp. 1126-1135).
- Rafiqoh, S. (2020). Arah Kecenderungan dan Isu Dalam Pembelajaran Matematika Sesuai Pembelajaran Abad 21 Untuk Menghadapi Revolusi Industri 4.0. *Jurnal MathEducation Nusantara*, 3(1), 58-73.
- Rezaie, M., & Golshan, M. (2015). Computer adaptive test (CAT): Advantages and limitations. *International Journal of Educational Investigations*, 2(5), 128-137.
- Retnawati, H. (2014b). *Teori Respons Butir dan Penerapannya*. Yogyakarta: Nuha Medika.
- Roth, A., Goldwasser, D., & Parker, A. (2016). There's a measure for that!. *Energy and Buildings*, 117, 321-331.
- Samejima, F. (1979). *A New Family of Models for the multiple Choice Item*. London.
- Silvia, F., Risnita, R., & Syaiful, S. (2015). Pengembangan rubrik keterampilan berpikir kreatif dalam memecahkan masalah matematika siswa kelas VIII SMP Attaufiq Jambi. *Edu-Sains: Jurnal Pendidikan Matematika dan Ilmu Pengetahuan Alam*, 4(1).
- Stenbeck, M., Hambleton, R. K., Swaminathan, H., Rogers, H. J., Rogers, D. J., Stenbeck, M., ... Rogers, D. J. (1992). Fundamentals of Item Response Theory. *Contemporary Sociology*, 21(2), 289.
- Supriyati, Y., Iriyadi, D., & Falani, I. (2021). The development of equating application for computer based in physics HOTS category. *JOTSE*, 11(1), 117-128.
- Tout, D., & Spithill, J. (2015). The challenges and complexities of writing items to test mathematical literacy. In *Assessing Mathematical Literacy* (pp. 145-171). Springer, Cham.
- Yılmaz, H. B. (2019). A Comparison of IRT Model Combinations for Assessing Fit in a Mixed Format Elementary School Science Test. *International Electronic Journal of Elementary Education*, 11(5), 539-545.
- Żóltak, T., & Golonka, G. (2015). Does guessing matter? Differences between ability estimates from 2PL and 3PL IRT models in case of guessing. *Edukacja*, 3, 63-76.