Efisiensi : Kajian Ilmu Administrasi
Vol. 20 No. 2, Agustus 2023, Hal. 159-170

Available at https://journal.uny.ac.id/index.php/efisiensi
Published by Departemen Pendidikan Administrasi FEB UNY bekerjasama dengan ASPAPI Pusat

ISSN: 1412-1131
E-ISSN: 2528-5750

Analisis Butir Soal dengan *Classical Test Theory* untuk Mengukur Kemampuan Kognitif Siswa

Febrika Yogie Hermanto[1*], Agnes Dwi Anggraeni[2], Sutirman[3]
[1,2]Fakultas Ekonomika dan Bisnis, Universitas Negeri Surabaya, Indonesia
[3]Fakultas Ekonomi dan Bisnis, Universitas Negeri Yogyakarta, Indonesia
febrikahermanto@unesa.ac.id, agnesdwi.20006@mhs.unesa.ac.id, sutirman@uny.ac.id

Abstrak: Analisis Butir Soal dengan *Classical Test Theory* untuk Mengukur Kemampuan Kognitif Siswa
Penelitian ini bertujuan untuk menganalisis kelayakan pada butir soal yang digunakan untuk mengukur kemampuan kognitif siswa SMK Manajemen Perkantoran dan Layanan Bisnis (MPLB) pada pembelajaran Praktik Perkantoran. Penelitian ini merupakan penelitian deskriptif dengan menganalisis 50 butir instrumen tes berbentuk pilihan ganda. Analisis data penelitian ini menggunakan classical test theory yang terdiri dari uji validitas, reliabilitas, daya pembeda soal, tingkat kesukaran soal, dan efektivitas pengecoh di mana dianalisis dengan menggunakan SPSS dan Anbuso. Hasil penelitian ini menunjukkan bahwa terdapat 31 soal yang dapat digunakan dalam mengkur kemampuan kognitif siswa dengan tingkat ketepatan dan keajekan yang baik – valid dan reliabel. Daya pembeda menunjukkan bahwa seluruh soal dapat digunakan dengan kategori daya pembeda baik dan cukup baik. Tingkat kesukaran soal menunjukkan bahwa 6,5% soal dinyatakan sukar, 51% soal dinyatakan mudah, dan 41,9% soal dinyatakan sedang. Terakhir, efektivitas pengecoh menunjukkan bahwa 16 butir soal perlu direvisi pengecohnya apabila digunakan dalam paket soal yang digunakan karena menunjukkan kurang efektifnya alternatif jawaban untuk mengecoh peserta ujian.
Kata kunci: Analisis Butir Soal; *Classical Test Theory*; Memampuan Kognitif; SMK; anajemen Perkantoran dan Layanan Bisnis

Abstract: Analysis of test item with Classical Test Theory to Measure Students' Cognitive Ability
This study aims to analyze the feasibility of the items used to measure the cognitive abilities of vocational students of Office Management and Business Services (OMBS) in Office Practice learning. This study is a descriptive study by analyzing 50 items of multiple choice test instruments. The data analysis of this study uses classical test theory which consists of testing the validity, reliability, discriminat power, difficulty level, and distractor effectiveness which are analyzed using SPSS and Anbuso. This study indicate that there are 31 questions that can be used in measuring students' cognitive abilities with a good level of accuracy and reliability – valid and reliable. Discriminat power of test item shows that all questions can be used with good and fairly good discriminant power categories. The difficulty level of the questions showed that 6.5% of the questions were declared difficult, 51% of the questions were declared easy, and 41.9% of the questions were declared moderate. Then, the distractor effectiveness shows that 16 items need to be revised if they are used in the question package used because they show the ineffectiveness of alternative answers to trick participants.
Keyword: Analysis of Test Item; Classical Test Theory, Cognitive Ability, Vocational High School, Office Management and Business Services

History & License of Article Publication:

Received:	15/05/2023	***Revision:***	08/08/2023	***Published:***	31/08/2023

DOI: https://doi.org/10.21831/efisiensi.v20i2.80225

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

PENDAHULUAN

Penilaian adalah proses mencari dan mengolah informasi untuk mengetahui tingkat ketercapaian hasil belajar peserta didik. Penilaian bertujuan untuk mengawasi dan menilai proses, perkembangan, serta perbaikan hasil belajar peserta didik secara berkelanjutan (Sutami, 2020). Menurut Permendikbud nomor 23 tahun 2016 tentang Standar Penilaian, penilaian hasil belajar siswa terdiri dari aspek sikap, pengetahuan dan keterampilan. Dalam pelaksanaan penilaian dibutuhkan alat yang disebut instrumen penilaian. (Djaali, 2000) menambahkan bahwa instrumen penilaian harus memenuhi kriteria akademis sehingga dapat digunakan sebagai alat untuk mengukur sebuah obyek ukur atau mengumpulkan data mengenai suatu variabel atau obyek yang dinilai.

Lebih lanjut, instrumen penilaian dapat dibedakan menjadi dua, yaitu instrumen tes dan non-tes. (Matondang, 2009) menyatakan bahwa yang termasuk dalam instrumen tes adalah tes prestasi belajar, tes intelegensi, dan tes bakat, sedangkan yang termasuk instrumen non-tes adalah pedoman wawancara, angket atau kuisioner. Kedua jenis instrumen penilaian tersebut digunakan untuk mengukur obyek yang dinilai sesuai dengan kebutuhannya. Intsrumen tes dapat digunakan untuk mengukur kemampuan kognitif siswa, dan instrumen non tes dapat digunakan untuk mengukur kemampuan afektif, dan psikomotor siswa.

Fakta di lapangan menunjukkan bahwa Guru di Sekolah Menengah Kejuruan (SMK) umumnya menggunakan instrumen tes yang belum terjamin validitasnya (Sugiyono et al., 2018). Hal tersebut terjadi karena dalam pembuatan instrumen tes yang terukur membutuhkan waktu, tenaga dan biaya yang lebih. Penggunaan instrumen tes yang belum terukur tidak bisa menjadi acuan pencapaian hasil belajar siswa karena tidak adanya kepastian terhadap konsistensi instrumen tes yang digunakan. Dengan demikian, untuk mengukur obyek yang dinilai dengan baik, instrumen tes harus memenuhi uji validitas dan reliabilitas, sehingga instrumen tersebut dapat mengukur pencapaian hasil belajar secara tepat (Muhson et al., 2015). Lebih lanjut, instrumen tes perlu diuji atau dianalisis kelayakannya pada tiap butir soal, di mana instrumen tersebut harus lolos uji validitas, reliabilitas, daya pembeda, tingkat kesukaran dan efektivitas pengecoh (Muhson, 2017).

Pembelajaran di Sekolah Menengah Kejuruan (SMK) dirancang untuk menyiapkan siswa atau calon lulusan dalam memasuki dunia kerja (Hermanto et al., 2019, 2021, 2022; Widianingrum et al., 2020). Seluruh program yang dirancang harus lebih sejalan dengan persiapan lulusan yang berkualitas dan memenuhi kebutuhan industri, baik dari kognitif, afektif, dan psikomotor. Instrumen tes yang digunakan seharusnya dapat mengukur kemampuan yang dimiliki siswa secara tepat agar dapat memberikan informasi terkait perkembangan siswa serta

sebagai dasar dalam peningkatan program sekolah (Firdausa & Istiyono, 2019). Dengan demikian, penelitian ini bertujuan untuk menganalisis kelayakan pada butir soal yang digunakan untuk mengukur kemampuan kognitif siswa SMK Manajemen Perkantoran dan Layanan Bisnis (MPLB) pada pembelajaran Praktik Perkantoran.

METODE

Penelitian ini merupakan penelitian deskriptif (Sugiyono, 2017) dengan menganalisis 50 butir instrumen tes berbentuk pilihan ganda pada enam indikator utama, yakni 1) Administrasi Umum, 2) Kearsipan, 3) Komunikasi Kantor, 4) Teknologi Perkantoran, 5) Otomatisasi Perkantoran: Microsoft Office, dan 6) Teknologi Dunia Maya, di mana instrumen ini kemudian digunakan untuk penilaian Pretes, dan Postes pada Modul Praktik Perkantoran dengan kompetensi dasar terintegrasi (Hermanto et al., 2022). Instrumen tes yang diuji dikembangkan dari kompetensi dasar pada Kurikulum 2013, dan dianalisis dengan menggunakan *Classical Theory Test* (CTT). Instrumen tes diujicobakan pada 65 siswa SMK Negeri 1 dan 7 Yogyakarta, kemudian dianalisis menggunakan aplikasi SPSS untuk menganalisis validitas dan reliabilitas, serta aplikasi Anbuso untuk menganalisis daya pembeda, tingkat kesukaran dan efektivitas pengecoh (Muhson, 2017; Muhson et al., 2015).

Untuk melakukan uji validitas dapat mengacu pada teknik korelasi biserial dengan rumus sebagai berikut (Susanto et al., 2015) :

$$r_{pbi} = \frac{M_{p-} M_t}{SD_t} \sqrt{\frac{p}{q}}$$

Keterangan :

r_{pbi} = koefisien korelasi biserial

M_p = rata-rata skor dari subjek yang menjawab benar yang dicari validitynya

M_t = rata-rata skor total

SD_t = standar deviasi dari skor total proporsi

p = proporsi siswa yang menjawab benar dengan jumlah toral siswa

q = proposi siswa yang menjawab salah (q = 1-p)

Selanjutnya, nilai r_{pbi} akan dibandingkan dengan koefisien korelasi tabel nilai “r” *product moment* pada taraf signifikan 0,05. Apabila r_{pbi} hasil koefisien korelasi lebih besar (>) dari nilai r_{tabel}, maka hasil yang diperoleh adalah signifikan, artinya butir instrument tes dinyatakan valid. Pada aplikasi SPSS, corrected item-total correlation coefficient t ≥ r_{tabel}, maka instrument valid.

https://doi.org/10.21831/efisiensi.v%vi%i.80225

Untuk melakukan analisis reliabilitas, dapat menggunakan rumus (Sudijono, 2013) sebagai berikut :

$$r_{11} = \left(\frac{n}{n-1}\right)\left(\frac{S2-\sum pq}{S2}\right)$$

Keterangan :

r_{11} = reliability tes secara keseluruhan

n = banyaknya instrument tes

p = proporsi siswa menjawab instrument tes dengan benar

q = proporsi siswa menjawab instrument tes dengan salah

$\sum pq$ = jumlah perkalian p dan q

S = standar deviasi

Selanjutnya, nilai koefisien (r_i) akan dibandingkan dengan koefisien korelasi tabel r_{tabel}=$r_{(\alpha,n-2)}$. Jika $r_i > r_{tabel}$, maka instrument reliabel. Pada aplikasi SPSS, jika Cronbach's Alpha > r_{tabel}, maka instrument reliabel.

Analisis daya pembeda dilakukan dengan menggunakan rumus (Susanto et al., 2015) sebagai berikut :

D = P_A - P_B

Keterangan :

D = indeks discriminating power

P_A = jumlah skor pada kelompok atas

P_B = jumlah skor pada kelompok bawah

Apabila jumlah *testee* sebanyak 100 orang atau lebih maka hanya menggunakan 27% *testee* kelompok atas dan 27% *testee* kelompok bawah (Sudijono, 2013). Berikut interpretasi hasil *discriminating power* menurut (Saccuzzo, 2009):

Tabel 1. Analisis Daya Pembeda Soal

Kategori	Kriteria
Baik	> 0,3
Cukup Baik	0,2 – 0,29
Tidak Baik	< 0,2

Pada aplikasi Anbuso, hasil analisis daya pembeda dapat dilihat pada kolom Daya Pmbeda yang akan menunjukkan nilai koefisien disertai keterangan.

Analisis tingkat kesukaran dilakukan dengan menggunakan rumus sebagai berikut (Susanto et al., 2015) :

$$I = \frac{B}{J}$$

https://doi.org/10.21831/efisiensi.v%vi%i.80225

Keterangan :

I = indeks kesukaran butir

B = proporsi siswa yang menjawab dengan benar

J = jumlah siswa yang mengikuti tes

Berikut interpretasi tingkat kesukaran menurut (Saccuzzo, 2009) :

Tabel 2. Interpretasi Level of Tingkat Kesukaran Soal

Interval	Interpretasi
0,00 – 0,30	Sukar
0,31 – 0,70	Sedang
0,71 – 1,00	Sukar

Pada aplikasi Anbuso, hasil analisis tingkat kesukaran ditunjukkan pada kolom Tingkat Kesukaran yang akan menampilkan nilai koefisien disertai keterangan.

Lebih lanjut, menurut (Arikunto, 2013) pengecoh dapat dikatakan efektif apabila dipilih oleh siswa paling sedikit 5%. Pada aplikasi Anbuso, analisis efektivitas pengecoh ditunjukkan pada kolom alternatif jawaban tidak efektif dan keterangan.

HASIL DAN PEMBAHASAN

Uji Validitas

Dari hasil uji validitas terhadap 50 butir soal menggunakan aplikasi SPSS didapatkan hasil sebagai berikut:

Tabel 3. Hasil Uji Validitas

No	r-hitung	>/<	r-tabel	Ket	No	r-hitung	>/<	r-tabel	Ket
1	0,006	<	0,2441	Tidak Valid	26	0,045	<	0,2441	Tidak Valid
2	0,449	>	0,2441	Valid	27	0,327	>	0,2441	Valid
3	0,096	<	0,2441	Tidak Valid	28	0,374	>	0,2441	Valid
4	0,335	>	0,2441	Valid	29	0,412	>	0,2441	Valid
5	0,434	>	0,2441	Valid	30	0,245	>	0,2441	Valid
6	0,349	>	0,2441	Valid	31	0,294	>	0,2441	Valid
7	0,193	<	0,2441	Tidak Valid	32	0,176	<	0,2441	Tidak Valid
8	0,220	<	0,2441	Tidak Valid	33	0,516	>	0,2441	Valid
9	0,386	>	0,2441	Valid	34	0,431	>	0,2441	Valid
10	0,562	>	0,2441	Valid	35	0,422	>	0,2441	Valid
11	0,545	>	0,2441	Valid	36	-0,007	<	0,2441	Tidak Valid
12	0,368	>	0,2441	Valid	37	0,015	<	0,2441	Tidak Valid
13	0,062	<	0,2441	Tidak Valid	38	-0,212	<	0,2441	Tidak Valid
14	0,533	>	0,2441	Valid	39	-0,001	<	0,2441	Tidak Valid
15	0,101	<	0,2441	Tidak Valid	40	0,340	>	0,2441	Valid
16	0,296	>	0,2441	Valid	41	0,247	>	0,2441	Valid
17	0,113	<	0,2441	Tidak Valid	42	0,464	>	0,2441	Valid
18	-0,106	<	0,2441	Tidak Valid	43	0,451	>	0,2441	Valid
19	0,192	<	0,2441	Tidak Valid	44	0,370	>	0,2441	Valid
20	0,503	>	0,2441	Valid	45	0,055	<	0,2441	Tidak Valid

https://doi.org/10.21831/efisiensi.v%vi%i.80225

No	r-hitung	>/<	r-tabel	Ket	No	r-hitung	>/<	r-tabel	Ket
21	0,180	<	0,2441	Tidak Valid	46	0,414	>	0,2441	Valid
22	0,262	>	0,2441	Valid	47	0,368	>	0,2441	Valid
23	0,310	>	0,2441	Valid	48	0,406	>	0,2441	Valid
24	0,307	>	0,2441	Valid	49	0,271	>	0,2441	Valid
25	0,170	<	0,2441	Tidak Valid	50	0,205	<	0,2441	Tidak Valid

Dari hasil analisis uji validitas tersebut dapat diketahui bahwa dari 50 butir soal terdapat 31 soal yang dinyatakan valid yaitu butir soal nomor 2, 4, 5, 6, 9, 10 , 11, 12, 14, 16, 20, 22, 23, 24, 27, 28, 29, 30, 31, 33, 34, 35, 40, 41, 42, 43, 44, 46, 47, 48, 49. Serta terdapat 19 butir soal yang dinyatakan tidak valid yaitu 1, 3, 7, 8, 13, 15, 17, 18, 19, 21, 25, 26, 32, 36, 37, 38, 39, 45, 50. Untuk mendapatkan kualitas soal yang baik, maka seluruh soal yang dinyatakan tidak valid tidak akan digunakan dalam uji Reliabilitas, daya pembeda, Tingkat kesukaran soal, dan Pengecoh.

Uji Reliabilitas

Analisi reliabilitas pada penelitian ini menggunakan aplikasi SPSS yang dilakukan pada 31 butir soal yang telah dinyatakan valid dari hasil uji validitas. Dari hasil uji reliabilitas didapatkan nilai Cornbach's Alpha sebesar 0,822 sehingga instrumen dinyatakan reliabel.

Uji Daya Pembeda

Analisis daya pembeda soal dilakukan dengan berbantuan Aplikasi Anbuso terhadap 31 butir soal yang telah dinyatakan valid. Berikut hasil uji daya pembeda :

Tabel 4. Hasil Uji Daya Pembeda

Nomor Soal	Koefisien	Keterangan	Nomor Soal	Koefisien	Keterangan
2	0.417	Baik	29	0.378	Baik
4	0.298	Cukup Baik	30	0.203	Cukup Baik
5	0.411	Baik	31	0.258	Cukup Baik
6	0.309	Baik	33	0.494	Baik
9	0.354	Baik	34	0.396	Baik
10	0.532	Baik	35	0.409	Baik
11	0.518	Baik	40	0.313	Baik
12	0.329	Baik	41	0.219	Cukup Baik
14	0.515	Baik	42	0.432	Baik
16	0.272	Cukup Baik	43	0.414	Baik
20	0.471	Baik	44	0.331	Baik
22	0.220	Cukup Baik	46	0.378	Baik
23	0.268	Cukup Baik	47	0.329	Baik
24	0.275	Cukup Baik	48	0.375	Baik
27	0.286	Cukup Baik	49	0.237	Cukup Baik
28	0.339	Baik			

https://doi.org/10.21831/efisiensi.v%vi%i.80225

Dari hasil analisis daya pembeda dapat diketahui bahwa terdapat 21 butir soal dengan kriteria "Baik" yaitu soal nomor 2, 5, 6, 9, 10 , 11, 12, 14, 20, 28, 29, 33, 34, 35, 40, 42, 43, 44, 46, 47, 48. Serta terdapat 10 butir soal dengan kriteria "Cukup Baik" yaitu soal nomor 4, 16, 22, 23, 24, 27, 30, 31, 41, 49. Soal yang memiliki daya pembeda baik merupakan soal yang dapat membedakan siswa yang telah menguasai materi dengan siswa yang belum menguasai materi, sehingga dari 31 butir soal yang telah dilakukan uji daya pembeda seluruhnya layak untuk digunakan sebagai instrumen penilaian.

Tingkat Kesukaran

Analisis tingkat kesukarang dilakukan dengan bantuan aplikasi Anbuso terhadap 31 butir soal dengan hasil sebagai berikut :

Tabel 5. Tingkat Kesukaran Soal

Nomor Soal	Koefisien	Keterangan	Nomor Soal	Koefisien	Keterangan
2	0.738	Mudah	29	0.738	Mudah
4	0.292	Sulit	30	0.569	Sedang
5	0.892	Mudah	31	0.754	Mudah
6	0.585	Sedang	33	0.877	Mudah
9	0.785	Mudah	34	0.662	Sedang
10	0.631	Sedang	35	0.969	Mudah
11	0.769	Mudah	40	0.877	Mudah
12	0.631	Sedang	41	0.877	Mudah
14	0.923	Mudah	42	0.723	Mudah
16	0.908	Mudah	43	0.508	Sedang
20	0.677	Sedang	44	0.554	Sedang
22	0.569	Sedang	46	0.662	Sedang
23	0.477	Sedang	47	0.631	Sedang
24	0.815	Mudah	48	0.785	Mudah
27	0.585	Sedang	49	0.200	Sulit
28	0.738	Mudah			

Dari analisis tingkat kesukaran dapat diketahui terdapat 16 soal dengan kriteria "Mudah" yaitu soal nomor 2, 5, 9, 11, 14, 16, 24, 28, 29, 31, 33, 35, 40, 41, 42, 48. Terdapat 13 butir soal dengan kriteria "Sedang" yaitu soal nomor 6, 10, 12, 20, 22, 23, 27, 30, 34, 43, 44, 46, 47. Serta terdapat 2 butir soal dengan kriteria "Sulit" yaitu soal nomor 4 dan 49.

Analisis Efektivitas Pengecoh

Analisis efektivitas pengecoh dalam penelitian ini mengguanakan aplikasi Anbuso terhadap 31 butir soal dengan hasil yang tertera pada Tabel 6. Dari hasil analisis efektifitas pengecoh dapat diketahui bahwa terdapat 11 butir soal dengan kriteria "Baik" yaitu soal nomor 10, 12, 22, 23, 27, 30, 34, 43, 44, 46, 47. Terdapat 4 butir soal dengan kriteria "Cukup Baik" yaitu soal nomor

https://doi.org/10.21831/efisiensi.v%vi%i.80225

2, 5, 29, 49. Serta terdapat 16 butir soal dengan kriteria "Revisi Pengecoh" yaitu soal nomor 4, 6, 9, 11, 14, 16, 20, 24, 28, 31, 33, 35, 40, 41, 42, 48.

Tabel 6. Analisis Efektivitas Pengecoh

Nomor Soal	Alternatif Jawaban Tidak Efektif	Keterangan	Nomor Soal	Alternatif Jawaban Tidak Efektif	Keterangan
2	-	Cukup Baik	29	-	Cukup Baik
4	C	Revisi Pengecoh	30	-	Baik
5	-	Cukup Baik	31	E	Revisi Pengecoh
6	E	Revisi Pengecoh	33	E	Revisi Pengecoh
9	BE	Revisi Pengecoh	34	-	Baik
10	-	Baik	35	AB	Revisi Pengecoh
11	E	Revisi Pengecoh	40	DE	Revisi Pengecoh
12	-	Baik	41	D	Revisi Pengecoh
14	C	Revisi Pengecoh	42	C	Revisi Pengecoh
16	A	Revisi Pengecoh	43	-	Baik
20	A	Revisi Pengecoh	44	-	Baik
22	-	Baik	46	-	Baik
23	-	Baik	47	-	Baik
24	C	Revisi Pengecoh	48	E	Revisi Pengecoh
27	-	Baik	49	-	Cukup Baik
28	E	Revisi Pengecoh			

Implilkasi Praktik

Validitas soal menunjukkan kualitas instrumen tes secara faktual (Dachliyani, 2019), di mana soal yang dinyatakan valid dapat diartikan bahwa soal tersebut tepat digunakan untuk mengkur kemampuan kognitif siswa. Pertama, jumlah soal yang dinyataan valid sebanyak 31 soal dengan. Implikasi hasil analisis tersebut menunjukkan bahwa jumlah soal dan nomor soal tersebut dapat digunakan untuk mengukur kemampuan kognitif siswa (Handayani & Iba, 2020; Susanto et al., 2015) dengan tingkat ketepatan yang baik.

Kedua, hasil uji reliablitas menunjukkan bahwa paket soal dengan jumlah 31 soal tersebut dinyatakan reliabel, di mana paket instrumen tes tersebut apabila digunakan berkali-kali pada subjek yang sama – siswa SMK MPLB – dalam waktu yang berbeda akan menunjukkan hasil tes – kemampuan kognitif yang relatif sama (Dachliyani, 2019). Pada hasil uji reliabiltas dapat diketahui bahwa instrumen tes yang dikembangkan dinyatakan reliabel sehingga layak digunakan (Sutami, 2020). Berdasarkan analisis *classical theory test* – uji validitas dan reliabilitas pada instrumen – menunjukkan bahwa terdapat 31 soal yang dapat digunakan untuk mengukur kemampuan kognitif siswa.

Ketiga, hasil analisis butir soal, analisis daya pembeda soal menunjukkan bahwa 21 butir soal dengan kriteria "Baik" dan 10 butir soal dengan kriteria "Cukup Baik". Hasil ini

https://doi.org/10.21831/efisiensi.v%vi%i.80225

mengimplikasikan bahwa instrumen tes tersebut dapat membedakan siswa yang telah menguasai materi yang diajarkan oleh guru (Sutami, 2020), pada kemampuan kognitifnya. Selain itu, hasil analisis tersebut menunjukkan bahwa instrument tes yang dikembangkan mampu membedakan peeserta didik yang memiliki kemampuan tinggi dan rendah.

Keempat, analisis tingkat kesukaran menunjukkan bahwa 6,5% soal dinyatakan sukar, 51,6% soal dinyatakan mudah, dan 41,9% soal dinyatakan sedang. Hasil tersebut mengimplikasikan bahwa soal yang telah dianalisis tingkat kesukarannya dapat memudahkan guru dalam menentukan paket soal yang akan digunakan untuk mengkur kemampuan siswa (Sutami, 2020) dengan mempertimbangkan kebutuhan tes dengan tingkat kesukaran mudah, sedang dan sulit.

Kelima, hasil analisis efektivitas pengecoh menunjukkan bahwa 11 butir soal dengan kriteria "Baik", 4 butir soal dengan kriteria "Cukup Baik", dan 16 butir soal dengan kriteria "Revisi Pengecoh". Hasil analisis menunjukkan bahwa semua butir instrumen tes memiliki alternatif jawaban yang berfungsi, meskipun beberapa butir perlu direvisi untuk meningkatkan efektivitasnya, dan (Sutami, 2020) menyatakan bahwa efektivitas pilihan jawaban pada instrumen tes dapat dijadikan sebagai dasar untuk diperbaiki agar setiap pilihan jawaban secara efektif mengecoh peserta dalam memilih jawaban yang benar. Implikasinya adalah perlunya evaluasi dan analisis terhadap setiap item tes untuk menghasilkan soal yang berkualitas, melalui revisi atau penghapusan soal yang tidak efektif. Analisis butir soal juga memberikan informasi diagnostik terkait pencapaian materi sesuai dengan kompetensi dasar dan capaian pembelajaran yang telah diterapkan (Susanto et al., 2015). Lebih lanjut, (Fitrianawati, 2015) menambahkan bahwa identifikasi item tes bertujuan untuk mendapatkan umpan balik yang berguna untuk perbaikan dan penyempurnaan item tes, sehingga dapat mengukur apa yang ingin diukur. Zuriyanti (Fitrianawati, 2015) menjelaskan manfaat analisis butir instrumen tes sebagai berikut: 1) menentukan item tes yang tidak berfungsi dengan baik, 2) meningkatkan kualitas item tes melalui daya pembeda, tingkat kesukaran soal, dan efektivitas pengecoh, 3) meningkatkan validitas dan reliabilitas, dan 4) memperbaiki soal yang tidak relevan dengan materi yang diajarkan atau diujikan.

KESIMPULAN

Hasil penelitian ini menunjukkan bahwa terdapat 31 soal yang dapat digunakan dalam mengkur kemampuan kognitif siswa pada materi 1) Administrasi Umum, 2) Kearsipan, 3) Komunikasi Kantor, 4) Teknologi Perkantoran, 5) Otomatisasi Perkantoran: Microsoft Office, dan 6) Teknologi Dunia Maya. Butir soal yang dapat digunakan untuk mengukur kemampuan

https://doi.org/10.21831/efisiensi.v%vi%i.80225

kognitif siswa adalah butir soal nomor 2, 4, 5, 6, 9, 10 , 11, 12, 14, 16, 20, 22, 23, 24, 27, 28, 29, 30, 31, 33, 34, 35, 40, 41, 42, 43, 44, 46, 47, 48, 49 dengan tingkat ketepatan dan keajekan yang baik – valid dan reliabel. Kemudian, berdasarkan analisis butir soal pada uji daya pembeda menunjukkan bahwa seluruh soal dapat digunakan dengan kategori daya pembeda baik dan cukup baik, di mana soal tersebut dapat membedakan siswa yang telah menguasai materi yang dipelajari atau diujikan dan yang belum. Lebih lanjut, analisis butir soal pada tingkat kesukaran soal menunjukkan bahwa 6,5% soal dinyatakan sukar, 51% soal dinyatakan mudah, dan 41,9% soal dinyatakan sedang, di mana hasil tersebut dapat memudahkan guru dalam menentukan paket soal yang akan digunakan. Terakhir, efektivitas pengecoh menunjukkan bahwa 16 butir soal perlu direvisi pengecohnya apabila digunakan dalam paket soal yang digunakan karena menunjukkan kurang efektifnya alternatif jawaban untuk mengecoh peserta ujian.

Penelitian ini mengimplikasikan bahwa paket soal yang terdiri dari 31 soal dapat digunakan dalam mengkur kemampuan kognitif siswa pada Modul Praktik Perkantoran – pretes dan postes – sehingga siswa yang telah mempelajari modul tersebut dapat diukur kemampuan kognitifnya setelah menjawab soal yang telah ditentukan. Lebih lanjut, penelitian ini terbatas pada pengukuran instrumen tes yang dikembangkan dari Kurikulum 2013 dengan Kompetensi Dasar yang tertera pada Modul Praktik Perkantoran sehingga soal yang telah lolos uji hanya dapat digunakan untuk pengujian yang sesuai dengan kompetensi dasar atau capaian pembelajaran tersebut. Penelitian ini dapat dilanjutkan dengan mengembangkan instrumen tes atau soal dengan capaian pembelajaran yang terdapat pada Kurikulum Merdeka, dan diuji dengan menggunakan *classsical test theory* atau *item respon theory*, sehingga Guru SMK MPLB dapat terbantu dalam menyediakan soal yang baik untuk mengukur kemampuan kognitif siswa.

DAFTAR PUSTAKA

Arikunto. (2013). Dasar-dasar Evaluasi Pendidikan. Jakarta: *Bumi Aksara*.

Dachliyani, L. (2019). Instrumen yang Sahih: Sebagai Alat Ukur Keberhasilan Suatu Evaluasi Program Diklat (evaluasi pembelajaran). *MEDIKA: Media Informasi Dan Komunikasi Diklat Kepustakawanan*, *5*(1), 57–65.

Djaali. (2000). *Pengukuran Dalam Pendidikan*. Jakarta: Program Pascasarjana.

Firdausa, A. R., & Istiyono, E. (2019). Developing instrument for assessing student collaboration in vocational high schools. *Journal of Physics: Conference Series*, *1273*(1). https://doi.org/10.1088/1742-6596/1273/1/012043

Fitrianawati, M. (2015). Peran Analisis Butir Soal Guna Meningkatkan Kualitas Butir Soal, Kompetensi Guru Dan Hasil Belajar Peserta Didik. *Prosiding Seminar Nasional Pendidikan PGSD UMS & HDPGSDI Wilayah Jawa*, *5*(3), 282–295.

Handayani, S. L., & Iba, K. (2020). Karakteristik Tes Keterampilan Proses Sains: Validitas, Reliabilitas, Tingkat Kesukaran dan Daya Pembeda Soal. *Jurnal Publikasi Pendidikan*, *10*(2), 100–106. https://doi.org/10.26858/publikan.v10i2.13051

Hermanto, F. Y., Sutirman, Hidayati, B., & Sholikah, M. (2019). The Need of Practical Teaching in Vocational High School of Automation and Office Management Program in Yogyakarta City. *Jurnal Pendidikan Vokasi*, *9*(3), 238–248. https://doi.org/10.21831/jpv.v9i3.26734

Hermanto, F. Y., Sutirman, Mar'atus, S., & Ranu, M. E. (2021). The effectiveness of distance practice learning for facing covid-19 pandemic in Indonesia. *Journal of Theoretical and Applied Information Technology*, *99*(12), 2925–2936.

Hermanto, F. Y., Wulandari, S. S., Muyassaroh, J., & Hidayati, B. (2022). Validitas Modul Praktik Perkantoran Menggunakan Metode Pembelajaran Role-Playing berbasis Kebutuhan Industri. *Jurnal PROFIT: Kajian Pendidikan Ekonomi Dan Ilmu Ekonomi*, *9*(2), 103–111. https://doi.org/10.36706/jp.v9i2.19014

Matondang. (2009). Validitas dan Reabilitas Suatu Instrumen Penelitian. *Jurnal Tabularasa PPS Unimed*, *496–500*. https://doi.org/10.4028/www.scientific.net/AMM.496-500.1510

Muhson, A. (2017). The Development of Practical Item Analysis Program for Indonesian Teachers. *International Journal of Intrsuction*, *10*(2), 199–210.

Muhson, A., Lestari, B., Supriyanto, & Baroroh, K. (2015). Kelayakan Anbuso Sebagai Software Analisis Butir Soal Bagi Guru. *Jurnal Kependidikan*, *45*(2), 198–210. https://doi.org/doi.org/10.21831/jk.v45i2.7499

Saccuzzo, K. &. (2009). *Psychological testing: Principles, applications, and issues* (7th ed.). wadsworth.

Sudijono, A. (2013). *Pengantar Evaluasi Pendidikan*. Jakarta: PT. Raja Grafndo Persada.

Sugiyono. (2017). *Metode Penelitian Kebijakan: Pendekatan Kuantitatif, Kualitatif, Kombinasi, R&D dan Penelitian Evaluasi*. Alfabeta.

Sugiyono, S., Lastariwati, B., Budiastuti, E., & Yudianto, A. (2018). Development of Authentic Assessment Instruments for Saintifical Learning in Tourism Vocational High Schools. *Jurnal Pendidikan Teknologi Dan Kejuruan*, *24*(1), 52–61. https://doi.org/10.21831/jptk.v24i1.16670

Susanto, H., Rinaldi, A., & Novalia. (2015). Analisis Validitas Reabilitas Tingkat Kesukaran dan Daya Beda pada Butir Soal Ujian Akhir Semester Ganjil Mata Pelajaran Matematika. *Al-Jabar: Jurnal Pendidikan Matematika*, *6*(2), 203–217.

Sutami. (2020). Pengembangan Instrumen Asesmen Higher Order Thinking Skills (HOTS) pada Mata Pelajaran Bahasa Indonesia SMA dan SMK. *Diglosia: Jurnal Kajian Bahasa, Sastra, Dan Pengajarannya*, *3*(1), 102–113. https://doi.org/10.30872/diglosia.v3i1.24

Widianingrum, R., Suranto, Hermanto, F. Y., & Sholikah, M. (2020). Office Communication Competencies for Vocational High School in Industri 4.0. *Dinamika Pendidikan*, *15*(1), 77–86. https://doi.org/10.15294/dp.v15i1.24348

UCAPAN TERIMA KASIH

Terimakasih kepada seluruh pihak yang telah membantu dalam menyelesaikan penelitian ini, dan semoga penelitian ini dapat menyediakan soal atau instrumen tes yang berkualitas

sehingga dapat digunakan oleh guru dalam mengukur kemampuan siswa SMK MPLB dengan tepat.