

**EVALUASI PEMBELAJARAN UJIAN AKHIR SEMESTER
MATA PELAJARAN BISNIS ONLINE KELAS XII SMK SUNAN DRAJAT
LAMONGAN**

Miftachul Ulum, Abdul Mun'im, Erly Juliyani, Pusvyta Sari

Institut Pesantren Sunan Drajat Lamongan, Indonesia

Abstract

At the end of each semester in an education unit, a learning evaluation is always carried out. One form of learning evaluation in schools is to carry out final semester exams. This study aims to determine the quality of online business lesson items for class XII SMK Sunan Drajat Lamongan in the odd semester 2017/2018 academic year in terms of validity, reliability, difficulty level and difference tests. This quantitative research is based on data from 21 student answer sheets from 25 multiple choice questions. The use of SPSS and Anates V4 software was used to measure the level of validity, reliability, difficulty level and item difference. The results showed that the validity value obtained was 15 items (60%) were declared invalid and 10 items (40%) were declared valid. Reliability of the items with a Cronbach Alpha value of $0.640 > 0.6$, which means that the instrument of the 25 items was declared reliable. The level of difficulty of the tested items obtained 5 items in the very easy category, 10 items in the easy category and 10 items in the moderate category. Of the 25 items, 8 items (32%) are good, 7 (28%) are sufficient and 10 (40%) are bad.

Keywords: *Validity, Reliability, Level of Difficulty, Distinction Power*

Abstrak

Setiap akhir semester pada satuan pendidikan selalu dilaksanakan evaluasi pembelajaran. Salah satu bentuk evaluasi pembelajaran di sekolah adalah dengan melaksanakan ujian akhir semester. Penelitian ini bertujuan untuk menentukan kualitas butir soal pelajaran bisnis online kelas XII SMK Sunan Drajat Lamongan pada tahun pelajaran 2017/2018 semester ganjil dalam hal tingkat validitas, reliabilitas, tingkat kesukaran dan uji beda. Penelitian kuantitatif ini bersumber pada data 21 lembar jawaban siswa dari 25 butir soal pilihan ganda.

Corresponding author: drajatulum@insud.ac.id

EVALUASI: Jurnal Manajemen Pendidikan Islam is licensed under

The CC BY License (<https://creativecommons.org/licenses/by-sa/4.0/>)

Penggunaan software SPSS dan Anates V4 digunakan untuk mengukur tingkat validitas, reabilitas, tingkat kesukaran dan daya beda butir soal. Hasil penelitian diperoleh nilai validitas diperoleh 15 butir soal (60%) dinyatakan tidak valid dan 10 butir soal (40%) dinyatakan valid. Reliabilitas butir soal dengan nilai Cronbach Alpha sebesar 0,640 > 0,6 yang berarti instrumen butir soal dari 25 butir soal dinyatakan reliabel. Tingkat kesukaran dari butir soal yang diujikan diperoleh 5 butir soal dikategorikan sangat mudah, 10 butir soal dikategorikan mudah dan 10 butir soal dikategorikan sedang. Dari 25 butir soal dapat dikelompokkan 8 butir soal (32%) baik , 7 butir soal (28%) cukup dan 10 (40%) butir soal jelek.

Kata kunci : Validitas, Reliabilitas, Tingkat Kesukaran , Daya Beda

A. PENDAHULUAN

Proses kegiatan dalam pembelajaran sangat erat sekali dalam mewujudkan kualitas pendidikan disekolah. Kegiatan pembelajaran harus ditunjang dengan fasilitas dan perencanaan yang baik sehingga hasil yang diharapkan dapat tercapai. Dalam proses pembelajaran perlu ditetapkan tujuan dari pengajaran, pengalaman belajar dan hasil belajar yang akan dicapai. Ketiga komponen tersebut merupakan tiga unsur dalam pembelajaran. Dalam proses belajar mengajar (pengalaman belajar) kegiatan penilaian merupakan bagian dalam proses belajar mengajar. Kegiatan penilaian adalah kegiatan evaluasi dari proses belajar mengajar yang dilakukan oleh guru untuk memantau proses kegiatan belajar, kemajuan pembelajaran dan perbaikan hasil belajar peserta didik secara keseluruhan ¹. Abdul Razaq mengatakan bahwa keberhasilan siswa dapat diukur dalam proses pembelajaran dikelas dikarenakan proses belajar mengajar yang dapat meningkatkan minat dan motivasi dalam menjalani proses belajar ².

¹ Sukardi, *Evaluasi Pendidikan Prinsip dan Operasionalnya* (Jakarta: Bumi Aksara, 2011).

² Abdul Razaq Ahmad Irwan Fariza, Mohd Mahzan Awang, "The Relationship between Students' Involvement in 21st Century Classroom Learning Activities and Higher Order Thinking Skills (Hubungan antara Keterlibatan Pelajar dalam Aktiviti Pembelajaran Abad ke-21 dan Kemahiran Berfikir Aras Tinggi)," *Jurnal Pendidikan Malaysia* 44, no. 1 (2019): 59–64.

Kegiatan evaluasi pembelajaran merupakan tolak ukur dari tercapainya suatu tujuan. Suharsini Arikunto mengatakan evaluasi adalah kegiatan mengukur dan mengumpulkan data untuk mengetahui sejauh mana tujuan telah tercapai³. Kegiatan evaluasi dilakukan seorang guru bertujuan untuk mengetahui tingkat pencapaian prestasi siswa dalam mengelola dan menjalankan tugasnya. Evaluasi berarti mampu mengetahui kelemahan siswa dalam proses pembelajaran. Evaluasi yaitu mengukur dengan membandingkan sesuai dengan standar yang telah ditetapkan yang bersifat kuantitatif. Evaluasi juga berarti mengambil keputusan dengan standar baik atau buruk yang bersifat kualitatif. Wringstone mengatakan bahwa evaluasi pendidikan adalah penaksiran terhadap kemajuan dan pertumbuhan siswa menuju arah dan tujuan-tujuan yang telah ditetapkan didalam kurikulum sebelumnya⁴. Sedangkan pelaksanaan evaluasi pembelajaran dapat dilakukan saat pelaksanaan pembelajaran sedang berlangsung, tengah semester atau akhir semester.

Bentuk evaluasi dalam pembelajaran dengan pemberian tugas, tes tulis atau praktek. Pemberian tes dalam setiap akhir pembelajaran merupakan alat ukur atas keberhasilan dalam pembelajaran. Tes merupakan bagian yang terkecil atau tersempit dari suatu penilaian. Suharsini Arikunto menyatakan bahwa tes adalah prosedur atau alat yang digunakan untuk mengetahui dan mengukur dengan cara dan aturan yang telah ditentukan sebelumnya⁵. Djemari Mardapi mengatakan tes adalah salah satu alat untuk mengukur tingkat kemampuan siswa secara tidak langsung dengan melalui respon pertanyaan⁶. Dari beberapa pendapat diatas dapat disimpulkan bahwa evaluasi adalah proses mengukur hasil belajar yang meliputi prestasi peserta didik dalam proses pembelajaran yang meliputi awal hingga akhir dari proses pembelajaran.

Pelaksanaan evaluasi yang dilaksanakan dalam pembelajaran ingin menjamin apakah kegiatan yang dilaksanakan dalam pembelajaran telah memenuhi standar yang telah ditetapkan. Zainul Arifin mengapresiasi dari

³ Suharsimi Arikunto, *Dasar-dasar Evaluasi Pendidikan*, 2 ed. (Jakarta: Bumi Aksara, 2013), 40.

⁴ Ngalim Purwanto, *Prinsip-prinsip dan Teknik Evaluasi Pengajaran* (Bandung: PT Remaja Rosdakarya, 2013), 3.

⁵ Suharsimi Arikunto, *Dasar-dasar Evaluasi Pendidikan*, 40.

⁶ Djemari Mardapi, *Teknik Penyusunan Instrumen Tes dan Nontes* (Yogyakarta: Mitra Cendikia, 2008), 67.

pelaksanaan evaluasi pembelajaran bertujuan untuk mengetahui efisiensi dan efektifitas dalam sistem pembelajaran yang meliputi materi, metode, media, sumber belajar, sistem penilaian dan lingkungan belajar⁷. Sistem pembelajaran yang dikembangkan dalam satuan lembaga pendidikan berbeda-beda. Perbedaan sistem pembelajaran yang berbeda sangat bergantung dari visi dan misi setiap satuan lembaga pendidikan. Suharsini Arikunto memaparkan beberapa fungsi dari evaluasi belajar. 1. *Fungsi selektif* yaitu digunakan untuk menyeleksi siswa yang naik kelas, siswa yang lulus dan siswa yang berhak mendapatkan beasiswa prestasi. 2. *Fungsi diagnostik* yaitu mengetahui kelemahan dan kelebihan peserta didik, sehingga dengan mengetahui kelemahan maka akan memudahkan mencari solusi cara mengatasinya. 3. *Fungsi penempatan* yaitu menentukan kelompok mana siswa tersebut dapat dikelompokkan dalam katagori yang sama atas kemampuan yang dimilikinya. 4. *Fungsi keberhasilan* yaitu mengetahui sejauh mana program yang telah direncanakan dapat tercapai⁸.

Pelaksanaan evaluasi yang dilaksanakan oleh guru terhadap peserta didik haruslah dapat menjamin perubahan dari setiap evaluasi yang dilakukannya. Pelaksanaan dalam setiap evaluasi menunjukkan tingkat antisipasi atas kekurangan dan kelemahan selama pelaksanaan pembelajaran. Anas Sudijono menyampaikan pelaksanaan evaluasi belajar dapat dikatakan baik bila pelaksanaan evaluasi menerapkan 3 prinsip dasar yaitu (1). *Prinsip komprehensif* yaitu evaluasi harus dilaksanakan secara keseluruhan yang harus mencakup aspek perkembangan atau perubahan perilaku yang terjadi pada peserta didik, (2). *Prinsip kesinambungan* yaitu evaluasi harus dilaksanakan secara teratur dari waktu ke waktu, (3). *Prinsip objektif* yaitu evaluasi harus dijauhkan dari faktor subjektif dari guru dan memperlakukan peserta didik secara wajar⁹.

Dariyanto mengatakan beberapa prinsip yang harus dimiliki oleh seorang pendidik dalam memerankan sebagai seorang evaluator yaitu (1). Keterpaduan antara perencanaan evaluasi dalam penyusunan satuan pengajaran dengan tujuan pembelajaran dan materi pengajaran, (2).

⁷ Zainal Arifin, *Evaluasi Pembelajaran* (Bandung: PT Remaja Rosdakarya, 2013), 14.

⁸ Suharsimi Arikunto, *Dasar-dasar Evaluasi Pendidikan*, 10.

⁹ Anas Sudijono, *Pengantar Evaluasi Pendidikan* (Jakarta: PT Raja Grafindo Persada., 2011), 209.

Keterlibatan siswa untuk mengetahui keberhasilan menerima materi pembelajaran, (3). Koherensi bahwa materi yang disampaikan dalam pembelajaran harus sesuai dengan ranah kemampuan siswa, (4). Pedagogis yang berarti evaluasi yang dilakukan dapat merubah sikap dan perilaku siswa lebih baik, (5). Akuntabilitas yang berarti evaluasi yang dilakukan harus dapat dipertanggungjawabkan kepada semua pihak yang berkepentingan¹⁰. Salah satu bentuk evaluasi pembelajaran dengan melaksanakan ujian atau tes. Tes juga dapat diartikan sebagai sejumlah pertanyaan yang harus diberikan tanggapan dengan tujuan untuk mengukur tingkat kemampuan seseorang atau mengungkapkan aspek tertentu dari orang yang dikenai tes. Respon peserta tes terhadap sejumlah pertanyaan maupun pernyataan adalah menggambarkan kemampuan peserta tes dalam bidang tertentu. Pemberian butir soal dalam pelaksanaan ujian akhir semester dapat memberikan gambaran tentang evaluasi pembelajaran selama pelaksanaan kegiatan belajar mengajar. Keberhasilan peserta didik dalam mengerjakan soal sangat dipengaruhi kualitas soal yang diujikan. Dari gambaran uraian di atas maka peneliti ingin mengetahui kualitas butir soal yang diujikan dilihat dari validitas, reliabilitas, tingkat kesukaran dan uji beda pada pelajaran bisnis online kelas XII SMK Sunan Drajat Lamongan tahun pelajaran 2017/2018 semester ganjil.

B. METODE PENELITIAN

Penelitian dengan metode kuantitatif dilaksanakan di kelas XII SMK Sunan Drajat Lamongan pada mata pelajaran bisnis online dalam akhir semester ganjil tahun pelajaran 2017/2018. Penilaian dilakukan terhadap lembar jawaban 21 siswa yang telah melaksanakan tes tulis soal pilihan ganda atau multiple choice sebanyak 25 butir soal. Variabel penelitian berupa butir soal pilihan ganda atau multiple choice yang ditinjau dari aspek validitas, reliabilitas, tingkat kesukaran dan daya pembeda dengan menggunakan software SPSS versi 16 dan software Anates V4. Validitas merupakan suatu alat ukur yang digunakan untuk mengetahui sejauhmana alat ukur itu mampu mengukur apa yang seharusnya diukur¹¹. Analisa uji

¹⁰ Daryanto, *Evaluasi Pendidikan* (Jakarta: PT Rineka Cipta, 2008), 21.

¹¹ Miftachul Ulum, *Mahir Analisa Data SPSS Statistical Product, Service Solution* (Yogyakarta: Ghaneswara, 2013).

validitas butir soal dapat dilakukan dengan menggunakan teknik poin biserial yaitu :

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

Analisa uji validitas butir soal dapat juga dengan menggunakan korelasi product moment¹² yaitu :

$$r = \frac{n\sum(xy) - (\sum x)(\sum y)}{\sqrt{(n(\sum x^2) - (\sum x)^2)(n(\sum y^2) - (\sum y)^2)}}$$

Penerapan Uji validitas dapat dilakukan dengan menggunakan kriteria pengujian yaitu nilai r_{hitung} akan dibandingkan dengan koefisien korelasi r_{tabel} , jika $r_{hitung} > r_{tabel}$, maka instrumen butir soal valid. Jika $r_{hitung} < r_{tabel}$ maka instrumen butir soal tidak valid. Uji validitas juga dapat dilakukan dengan membandingkan hasil output pada program SPSS dengan kriteria jika hasil *corrected item-total correlation coefficient* $r > r_{tabel}$, maka instrumen valid. Jika *corrected item-total correlation coefficient* $r < r_{tabel}$, maka instrumen valid.

Reliabilitas tes berkenaan dengan pertanyaan apakah suatu tes yang diteliti sesuai dengan kriteria yang telah ditetapkan. Reliabilitas adalah tingkat atau derajat konsistensi dari sebuah instrumen¹³. Suatu instrumen pengukuran dikatakan reliabel jika pengukurannya konsisten, cermat dan akurat. Rumus pengujian uji reliabilitas yang dapat digunakan adalah Alfa Cronbach.

$$r_i = \frac{k}{(k-1)} \left\{ 1 - \frac{\sum si^2}{st^2} \right\}$$

¹² Miftachul Ulum, *BASIC STATISTIC With Statistical Package for Social Sciences (SPSS)*, 1 ed. (Lamongan: CV. Pustaka Ilalang, 2020).

¹³ Zainal Arifin, *Evaluasi Pembelajaran*.

Kreteria pengujian nilai *koefesien* (r) akan dibandingkan dengan koefesien korelasi tabel r_{tabel} . Jika $r_{hitung} > r_{tabel}$, maka instrumen reliabel. Pada Output SPSS, jika *Cronbach's Alpha* $> r_{tabel}$, maka instrumen reliabel. Tingkat reliabilitas soal pada rentangan koefesien korelasi dapat di inteprestasikan apabila *koefesien* (r) lebih besar daripada 0,60 berarti tes hasil belajar yang sedang diuji reliabilitasnya dinyatakan telah memiliki reliabilitas yang tinggi (*reliable*). Apabila *koefesien* (r) lebih kecil daripada 0,60 berarti tes hasil belajar yang sedang diuji reliabilitasnya dinyatakan belum memiliki reliabilitas yang tinggi (*un-reliable*)¹⁴.

Analisis tingkat kesukaran soal artinya mengkaji soal-soal tes dari segi kesulitannya sehingga dapat diperoleh soal-soal mana yang termasuk mudah, sedang, dan sukar¹⁵. Angka indeks kesukaran item dapat diperoleh dengan menggunakan rumus sebagai berikut:

$$P = \frac{B}{JS}$$

Witherington menafsirkan tingkat kesukaran butir tes yang digunakan dalam kriteria¹⁶ sebagai berikut :

Tabel 1. Interpretasi tingkat kesukaran butir soal

Interval	Interprestasi
0,00 – 0,30	Soal Sukar
0,32 – 0,70	Soal Sedang
0,71 – 1,00	Soal Mudah

Penafsiran analisis tingkat kesukaran butir soal dalam Anates sebagai berikut:

Tabel 2. Interpretasi Tingkat Kesukaran Anates

Interval	Interprestasi
0% - 15%	Sangat Sukar

¹⁴ Anas Sudijono, *Pengantar Evaluasi Pendidikan*, 209.

¹⁵ Muslikah Purwanti, "Analysis Of Final Exam Questions In Financial Accounting Using Microsoft Excel 2010," *Jurnal Pendidikan Akuntansi Indonesia* XII, no. 1 (2014): 81–94.

¹⁶ Anas Sudijono, *Pengantar Evaluasi Pendidikan*.

0% - 15%	Sukar
31% - 70%	Sedang
71% - 85%	Mudah
86% - 100%	Sangat Mudah

Daya beda butir soal ialah indeks yang menunjukkan tingkat kemampuan butir soal membedakan kelompok yang berprestasi tinggi (kelompok atas) dari kelompok yang berprestasi rendah (kelompok bawah) diantara para peserta tes¹⁷. Tujuan pokok mencari daya beda adalah untuk menentukan apakah butir soal tersebut memiliki kemampuan membedakan kelompok dalam aspek yang diukur, sesuai dengan perbedaan yang ada pada kelompok itu¹⁸. Suharsini mengintegrasikan daya beda dapat diklasifikasikan sebagai berikut¹⁹ :

Tabel 3. Interpretasi Daya Beda

Interval	Interprestasi
0,00 – 0,19	Jelek
0,20 – 0,39	Cukup
0,40 – 0,69	Baik
0,70 – 1,00	Baik sekali

Daya Pembeda juga dapat diklasifikasikan²⁰ sebagai berikut :

Tabel 4. Klasifikasi Daya Beda

Interval	Interprestasi
Negatif – 9%	Sangat Buruk , harus diulang
10% - 19%	Buruk , sebaiknya dibuang
20% - 29%	Cukup baik , perlu direvisi

¹⁷ Asmawi Zainul, *Penilaian Hasil Belajar* (Jakarta: Pusat Antar Universitas, Direktorat Jenderal Pendidikan Tinggi: Departemen Pendidikan Dan kebudayaan, 1997).

¹⁸ Suryabrata, S., *Pengembangan Alat Ukur Psikologis* (Jakarta: Direktorat Jenderal Pendidikan Tinggi: Departemen Pendidikan dan Kebudayaan, 1999).

¹⁹ Suharsimi Arikunto, *Dasar-dasar Evaluasi Pendidikan*, 218.

²⁰ Karno To, *Mengenal Analisis Tes Pengantar ke Program Komputer Anates* (Bandung: FIP UPI, 2003), 14.

C. HASIL PENELITIAN DAN PEMBAHASAN

Data hasil penelitian pada ujian akhir semester mata pelajaran pemasaran online diperoleh , siswa mengerjakan soal dengan benar 13 soal sebanyak 1 orang, 14 soal sebanyak 3 orang, 16 soal sebanyak 1 orang, 17 soal sebanyak 4 orang, 18 soal sebanyak 5 orang, 19 soal sebanyak 1 orang, 20 soal sebanyak 2 orang, 22 soal sebanyak 2 orang, 23 soal sebanyak 1 orang dan 24 soal sebanyak 1 orang. Adapun secara rinci dapat dilihat pada tabel berikut :

Tabel 5. Data Nilai Mata Pelajaran Pemasaran Online Kelas XII

Jumlah Jawaban Benar	Jumlah Siswa	Nilai
13	1	52
14	3	56
16	1	64
17	4	68
18	5	72
19	1	76
20	2	80
22	2	88
23	1	92
24	1	96

Nilai terendah yang diperoleh siswa dalam mata pelajaran Pemasaran Online sebesar 44 dan nilai tertinggi 96. Dengan menggunakan program SPSS diperoleh data statistik sebagai berikut :

Tabel 6. Descriptive Statistics

	N	Range	Min	Max	Sum	Mean	Std. Deviation	Variance
Jawaban Benar	21	11	13	24	379	18.05	3.025	9.148
Nilai	21	44	52	96	1516	72.19	12.098	146.362
Valid N (listwise)	21							

Dari data jumlah jawaban benar dan salah dapat dijabarkan soal-soal yang dapat dikerjakan oleh siswa dan soal yang tidak dapat dikerjakan sebagaimana terangkum tabel berikut ini :

Tabel 7. Data Jumlah Jawaban benar dan Salah

Nomer Soal	Jumlah Jawaban Benar	Jumlah Jawaban Salah
1	16	5
2	18	3
3	18	3
4	18	3
5	13	8
6	8	13
7	13	8
8	11	10
9	15	6
10	15	6
11	17	4
12	14	7
13	13	8
14	17	4
15	15	6
16	14	7
17	16	5
18	13	8
19	18	3
20	14	7
21	15	6
22	15	6
23	17	4

24	17	4
25	18	3

Nilai validitas dari butir soal dinyatakan valid dengan kriteria jika $r_{hitung} > r_{tabel}$ dan dinyatakan tidak valid dengan kriteria jika $r_{hitung} < r_{tabel}$. Dengan tingkat taraf signifikan sebesar 0,05 diperoleh nilai r_{tabel} sebesar 0,381. Dengan menggunakan Anates di peroleh nilai validitas dari 25 butir soal dengan diperoleh 10 (40 %) soal dinyatakan valid dan 15 (60%) dinyatakan tidak valid yang tercermin nilai koefisien korelasi tiap butir soal sebagaimana dalam tabel berikut :

Tabel 8. Nilai Korelasi Butir Soal

Nomer Soal	Korelasi	Signifikan
1	0,123	
2	0,145	
3	0,422	Signifikan
4	0,007	
5	0,411	Signifikan
6	0,253	
7	0,411	Signifikan
8	0,405	Signifikan
9	0,439	Signifikan
10	0,422	Signifikan
11	0,254	
12	0,491	Signifikan
13	0,577	Sangat Signifikan
14	0,053	
15	0,08	
16	0,01	
17	0,539	Sangat Signifikan
18	0,079	
19	0,337	
20	0,08	
21	0,439	Signifikan

22	0,312
23	0,213
24	-0,033
25	0,191

Nilai validitas dari butir soal dapat juga dinyatakan dari hasil output program SPSS. Butir soal dinyatakan valid jika nilai *corrected item-total correlation coefficient* $r > r_{\text{tabel}}$, dan dinyatakan tidak valid jika *corrected item-total correlation coefficient* $r < r_{\text{tabel}}$. Dengan tingkat taraf signifikan sebesar 0,1 diperoleh nilai r_{tabel} sebesar 0,3233 diperoleh nilai validitas dari 25 butir soal dengan diperoleh 10 (40%) soal dinyatakan valid dan 15(60%) dinyatakan tidak valid yang tercermin dari nilai *corrected item-total correlation coefficient* tiap butir soal dan hasil pengolahan dengan pengolahan SPSS sebagaimana tabel berikut :

Tabel 9 Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Soal_1	35.33	36.133	.051	.642
Soal_2	35.24	36.090	.086	.639
Soal_3	35.24	34.890	.371	.626
Soal_4	35.24	36.690	-.053	.646
Soal_5	35.48	34.362	.340	.623
Soal_6	35.71	35.314	.174	.634
Soal_7	35.48	34.362	.340	.623
Soal_8	35.57	33.557	.467	.613
Soal_9	35.38	34.348	.374	.622
Soal_10	35.38	35.148	.224	.631
Soal_11	35.29	35.514	.191	.634
Soal_12	35.43	33.957	.426	.617
Soal_13	35.48	33.362	.519	.610
Soal_14	35.29	35.714	.149	.636
Soal_15	35.38	37.148	-.137	.653

Soal_16	35.43	35.957	.069	.641
Soal_17	35.33	33.933	.485	.616
Soal_18	35.48	36.362	-.003	.646
Soal_19	35.24	35.890	.133	.637
Soal_20	35.43	36.357	.000	.645
Soal_21	35.38	34.348	.374	.622
Soal_22	35.33	35.133	.245	.630
Soal_23	35.29	35.714	.149	.636
Soal_24	35.29	36.914	-.099	.649
Soal_25	35.24	35.890	.133	.637
Total_skor	18.05	9.148	1.000	.482

Reliabilitas tes menunjukkan sejauh mana alat ukur yang digunakan dapat dipercaya. Reliabilitas tes merupakan ketepatan atau keajegan alat dalam menilai apa yang dinilai. Reliabilitas menunjukkan kestabilan skor yang diperoleh peserta tes yang sama ketika diuji ulang dengan tes yang sama pada situasi yang berbeda atau dari suatu pengukuran ke pengukuran lainnya. Suatu instrumen dikatakan reliabel jika Cronbach's Alpha > 0,6 Berdasarkan hasil pengolahan dengan program SPSS diperoleh nilai Cronbach Alpha sebesar 0,640 yang berarti diatas nilai ketentuan 0,6 yang berarti Ujian Akhir Semester untuk mata pelajaran bisnins online untuk soal pilihan ganda dari 25 soal dinyatakan instrumen soal reliabel.

Analisa tingkat kesukaran soal merupakan proporsi dari jumlah soal yang mampu di jawab oleh siswa dengan benar. Analisa ini mengkaji soal-soal yang termasuk katagori mudah, sedang dan sulit dijawab. Tingkat kesukaran dipandang dari segi siswa dalam kemampuannya menjawab soal bukan guru dalam kesulitan membuat soal. Berdasarkan hasil analisis tingkat kesukaran dengan *software* Anates diperoleh tingkat kesukaran dari 25 soal yang tergolong soal sangat mudah sebanyak 5 soal yaitu soal nomer 2, 3,4,14 dan 25 . Soal yang tergolong mudah sebanyak 10 soal yaitu soal nomer 1, 9, 11, 16, 17, 19, 21, 22, 23, dan 24. Soal yang tergolong sedang 10 soal yaitu soal nomer 5,6,7,8,10, 12, 13, 15, 18 dan 20. Secara rinci hasil pengolahan dengan menggunakan *software* Anates diperoleh tingkatan kesukaran sebagaimana tabel 10.

Tabel 10 . Tingkat Kesukaran

Nomer Soal	Jumlah Jawaban Benar	Tingkat Kesukaran	Tafsiran
1	16	76,19	mudah
2	18	85,71	sangat mudah
3	18	85,71	sangat mudah
4	18	85,71	sangat mudah
5	13	61,9	sedang
6	8	38,1	sedang
7	13	61,9	sedang
8	11	57,14	sedang
9	15	71,43	mudah
10	15	66,67	sedang
11	17	80,95	mudah
12	14	66,67	sedang
13	13	61,9	sedang
14	17	85,71	sangat mudah
15	15	66,67	sedang
16	14	71,43	mudah
17	16	76,19	mudah
18	13	61,9	sedang
19	18	80,95	mudah
20	14	66,67	sedang
21	15	71,43	mudah
22	15	76,19	mudah
23	17	80,95	mudah
24	17	80,95	mudah
25	18	85,71	sangat mudah

Witherington menafsirkan dalam tingkat kesukaran dari 25 soal yang diujikan dapat disimpulkan bahwa tingkat kesukaran sedang sebanyak 10 soal yaitu soal nomer 5, 6, 7, 8, 10, 12, 13, 15, 18, dan 20. Sedangkan kelompok soal dengan tingkat kesukaran mudah sebanyak 15 soal yaitu soal nomer 1, 2, 3, 4, 9, 11, 14, 16, 17, 19, 21, 22, 23, 24, dan 25.

Perbedaan kemampuan dalam menjawab soal sangat bervariasi. Adanya analisa uji beda berfungsi untuk mengkaji soal-soal dilihat dari kesanggupan kemampuan siswa. Uji beda bertujuan mengelompokkan

kemampuan siswa yang dapat dikelompokkan dalam kategori rendah dan kategori tinggi. Analisis daya beda artinya mengkaji soal-soal tes dari segi kesanggupan tes tersebut dalam membedakan siswa yang termasuk kedalam kategori rendah dan kategori tinggi. Daya pembeda butir adalah kemampuan suatu butir tes untuk dapat membedakan antara *testee* yang berkemampuan tinggi dan berkemampuan rendah²¹. Analisis daya beda pada penelitian ini menggunakan rumus dan *software* Anates diperoleh hasil uji beda sebagai berikut :

Tabel 10 . Inteprestasi Uji beda

Nomer Soal	Indeks	Inteprestasi
1	0,00	Jelek
2	0,00	Jelek
3	33,33	Cukup
4	-16,67	Jelek
5	50,00	Baik
6	33,33	Cukup
7	50,00	Baik
8	66,67	Baik
9	33,33	Cukup
10	50,00	Baik
11	33,33	Cukup
12	50,00	Baik
13	66,67	Baik
14	0,00	Jelek
15	16,67	Jelek
16	0,00	Jelek
17	50,00	Baik
18	16,67	Jelek
19	16,67	Jelek
20	16,67	Jelek
21	50,00	Baik

²¹ Nuraini Soleiman Dewi Juliah Ratnaningsih, Isfarudi, "Analisis Butir Soal Pilihan Ganda Ujian Akhir Semester Mahasiswa Di Universitas Terbuka Dengan Pendekatan Teori Tes Klasik," *Jurnal Pendidikan Terbuka dan Jarak Jauh* 12, no. 2 (2011): 93.

22	33,33	Cukup
23	33,33	Cukup
24	0,00	Jelek
25	33,33	Cukup

Dari dua puluh lima soal yang diujikan dapat dikategorikan soal baik 8 butir atau 32 % yaitu butir soal nomer 5, 7, 8, 10,12, 13, dan 21. Soal dalam katagori cukup 7 atau 28 % yaitu butir soal nomer 3, 6, 9, 11, 22, 23, dan 25 . Soal dalam katagori jelek 10 butir atau 40 % yaitu butir soal nomer 1, 2, 4, 14, 15, 16, 18, 19, 20, dan 24. Fernandes mengatakan dengan Indeks daya beda butir soal dapat digunakan sebagai bahan pertimbangan sebuah butir baik atau tidak baik. Butir soal yang baik adalah butir soal yang mempunyai indeks daya beda lebih dari 0,2.²² Sementara Ebel menjelaskan suatu butir soal dikatakan berkualitas apabila indeks diskriminasi atau daya pembedanya paling sedikit 0,41²³. Dari uji beda dalam butir soal dapat ditindak lanjuti terhadap soal yang termasuk katogari jelek untuk yaitu diperbaiki kualitas daya beda soal atau tidak digunakan lagi untuk diujikan. Sedangkan soal yang termasuk katagori daya beda yang baik untuk dapat dipertahankan. Butir soal nomer 4 yang angka indeks diskriminanya bertanda negatif tidak perlu digunakan pada tes hasil belajar atau direvisi kembali

D. KESIMPULAN

Hasil penelitian dan pembahasan dapat disimpulkan 15 butir soal (60%) dinyatakan tidak valid dan 10 butir soal (40%) dinyatakan valid. Reliabilitas soal dinyatakan reliabel sesuai ketentuan nilai Cronbach Alpha sebesar 0,640 > 0,6 dari 25 butir soal. Tingkat kesukaran dari 25 butir soal dapat dikelompokkan 5 butir soal (20%) sangat mudah, 10 butir soal (40%) mudah dan 10 butir soal (40%) sedang. Untuk jenis soal dapat dikategorikan baik 8 butir soal (32%), katagori cukup 7 butir soal (28%) dan katagori jelek 10 butir soal (40%).

²² Fernandes, H.J. X., *Evaluation of educational program* (Jakarta: National Education Planning , Evaluating and Curriculum Development, 1984).

²³ Ebel, R.L., *Essentials of educational measurement*, 3rd ed. (Englewood Cliffts,NJ: Prentice Hall Inc., 1972).

EVALUASI, 5 (1), Maret 2021, ISSN 2580-3387 (print) |
ISSN 2615-2886 (online)

Homepage : <http://e-journal.staima-alhikam.ac.id/index.php/evaluasi>

DOI : <http://doi.org/10.32478/evaluasi.v5i1.487>

Article type : Original Research Article

REFERENSI

Anas Sudijono. *Pengantar Evaluasi Pendidikan*. Jakarta: PT Raja Grafindo Persada., 2011.

Asmawi Zainul. *Penilaian Hasil Belajar*. Jakarta: Pusat Antar Universitas, Direktorat Jenderal Pendidikan Tinggi: Departemen Pendidikan Dan kebudayaan, 1997.

Daryanto. *Evaluasi Pendidikan*. Jakarta: PT Rineka Cipta, 2008.

Dewi Juliah Ratnaningsih, Isfarudi, Nuraini Soleiman. "Analisis Butir Soal Pilihan Ganda Ujian Akhir Semester Mahasiswa Di Universitas Terbuka Dengan Pendekatan Teori Tes Klasik." *Jurnal Pendidikan Terbuka dan Jarak Jauh* 12, no. 2 (2011): 92–99.

Djemari Mardapi. *Teknik Penyusunan Instrumen Tes dan Nontes*. Yogyakarta: Mitra Cendikia, 2008.

Ebel, R.L. *Essentials of educational measurement*. 3rd ed. Englewood Cliffts,NJ: Prentice Hall Inc., 1972.

Fernandes, H.J. X. *Evaluation of educational program*. Jakarta: National Education Planning , Evaluating and Curriculum Development, 1984.

Irwan Fariza, Mohd Mahzan Awang, Abdul Razaq Ahmad. "The Relationship between Students' Involvement in 21st Century Classroom Learning Activities and Higher Order Thinking Skills (Hubungan antara Keterlibatan Pelajar dalam Aktiviti Pembelajaran Abad ke-21 dan Kemahiran Berfikir Aras Tinggi)." *Jurnal Pendidikan Malaysia* 44, no. 1 (2019): 59–64.

Karno To. *Mengenal Analisis Tes Pengantar ke Program Komputer Anates*. Bandung: FIP UPI, 2003.

Miftachul Ulum. *BASIC STATISTIC With Statistical Package for Social Sciences (SPSS)*. 1 ed. Lamongan: CV. Pustaka Ilalang, 2020.

EVALUASI, 5 (1), Maret 2021, ISSN 2580-3387 (print) |

ISSN 2615-2886 (online)

Homepage : <http://e-journal.staima-alhikam.ac.id/index.php/evaluasi>

DOI : <http://doi.org/10.32478/evaluasi.v5i1.487>

Article type : Original Research Article

———. *Mahir Analisa Data SPSS Statistical Product, Service Solution*. Yogyakarta: Ghaneswara, 2013.

Muslikah Purwanti. "Analysis Of Final Exam Questions In Financial Accounting Using Microsoft Excel 2010." *Jurnal Pendidikan Akuntansi Indonesia* XII, no. 1 (2014): 81–94.

Ngalim Purwanto. *Prinsip-prinsip dan Teknik Evaluasi Pengajaran*. Bandung: PT Remaja Rosdakarya, 2013.

Suharsimi Arikunto. *Dasar-dasar Evaluasi Pendidikan*. 2 ed. Jakarta: Bumi Aksara, 2013.

Sukardi. *Evaluasi Pendidikan Prinsip dan Operasionalnya*. Jakarta: Bumi Aksara, 2011.

Suryabrata, S. *Pengembangan Alat Ukur Psikologis*. Jakarta: Direktorat jenderal Pendidikan Tinggi: Departemen Pendidikan dan Kebudayaan, 1999.

Zainal Arifin. *Evaluasi Pembelajaran*. Bandung: PT Remaja Rosdakarya, 2013.