

## **ANALISIS BUTIR SOAL UJIAN AKHIR SEMESTER GENAP MATA DIKLAT DASAR-DASAR MESIN**

### **TEST ITEM ANALYSIS OF EVEN SEMESTER FINAL EXAM IN THE BASIC OF ENGINE SUBJECT**

Oleh:

Farisal Windarto dan Martubi

Pendidikan Teknik Otomotif, Fakultas Teknik Universitas Negeri Yogyakarta

Email: farisalnews@gmail.com

#### **ABSTRAK**

Penelitian ini bertujuan untuk mengetahui (1) kualitas butir soal ujian akhir semester genap mata diklat Dasar-Dasar Mesin kelas X SMK Muhammadiyah Gamping tahun ajaran 2015/2016 ditinjau dari analisis kualitatif berupa aspek materi, konstruksi, dan bahasa/budaya dan (2) kualitas butir soal ujian akhir semester genap mata diklat Dasar-Dasar Mesin kelas X SMK Muhammadiyah Gamping tahun ajaran 2015/2016 ditinjau dari analisis kuantitatif berupa validitas, reliabilitas, daya pembeda, tingkat kesukaran, dan efektivitas pengecoh. Penelitian ini merupakan penelitian deskriptif. Sumber data adalah lembar soal, kunci jawaban, ksis-kisi soal, silabus, dan respon jawaban siswa. Teknik pengumpulan data dengan dokumentasi. Analisis data dilakukan dengan analisis kualitatif dan kuantitatif. Hasil penelitian ini menunjukkan bahwa berdasarkan analisis kualitatif ditinjau dari aspek materi sebanyak 2 soal tidak memenuhi, aspek konstruksi sebanyak 9 soal tidak memenuhi, aspek bahasa/budaya 24 soal tidak memenuhi, dan 6 soal sudah memenuhi ketiga aspek. Berdasarkan analisis kuantitatif ditinjau dari (1) validitas soal menunjukkan 70% soal termasuk kategori valid, 30% soal termasuk kategori tidak valid. (2) reliabilitas soal sebesar 0,637 sehingga soal termasuk belum memiliki reliabilitas yang tinggi. (3) daya pembeda soal menunjukkan 3,3% soal termasuk kategori baik sekali, 40% soal termasuk kategori baik, 33,3% soal termasuk kategori cukup, 16,7% soal termasuk kategori jelek, dan 6,7% soal termasuk kategori jelek sekali. (4) tingkat kesukaran soal menunjukkan 3,3% soal termasuk kategori mudah, 57,7% soal termasuk kategori sedang, dan 40% soal termasuk kategori sukar. (5) efektivitas pengecoh menunjukkan 33,3% termasuk kategori efektif dan 66,7% termasuk kategori tidak efektif. Dari 30 soal, 2 soal (6,7%) dipertahankan tanpa revisi, 22 soal (73,3%) dipertahankan dengan revisi, 5 soal (16,7%) harus diganti, dan 1 soal (3,3%) harus dibuang.

Kata kunci: analisis butir soal, Dasar-Dasar Mesin

#### **ABSTRACT**

*This study is aimed to identify (1) the quality of test item on even semester final exam in the basic of engine subject on 1<sup>ST</sup> grade at SMK Muhammadiyah Gamping in academic year 2015/2016 in terms of quality analysis in the form of material aspect, construction, and language/culture and (2) the quality of test item on even semester final exam in the basic of engine subject on 1<sup>ST</sup> grade at SMK Muhammadiyah Gamping in academic year 2015/2016 in terms of quantitative analysis in the form of validity, reliability, distinguishing feature, level of difficulty, and effectiveness of distractors. This study is a descriptive research. The data source is a booklet, key to the answer, grid of test item, syllabus, and the response of the students answer. Data collection technique is documentation. Data analysis is used qualitative and quantitative analysis. The result of this study indicate that based on qualitative analysis in term of material aspect as much as 2 test items not corresponden, construction aspect as much as 9 test item not correspondent, language/culture aspect as much as 24 test items not correspondent and as much as 6 test items is corresponden. Based on qualitative analysis in terms of (1) the validity test showed 70% items include a valid category, 30% items include invalid category (2) the reability of test 0,637 not have high reliability yet (3) distinguishing feature of test showed 3,3% include excellent category, 40% include good caetegory, 33,3% include enough category, 16,7% include bad category, and 6,7% include very bad category (4) the level of difficulty showed 3,3% include easy category, 57,7% include medium category, and 40% include difficult category (5) the effectiveness of destractor showed 33,3% include effective category and 66,7% include not effective. Of the 30 test items, 2 test items (6,7%) is retained without revision, 22 test items (73,3%) is retained with revision, 5 test items (16,7%) could be replaced, and 1 test item (3,3%) should be discarded.*

*Keywords: test item analysis, the basic of engine*

## PENDAHULUAN

Pendidikan mempunyai peranan penting sebagai ujung tombak dalam menentukan masa depan bangsa, melalui pendidikan ini cita-cita luhur untuk mencapai kesejahteraan dan kemajuan bangsa akan selalu tertanam dalam diri penerusnya. Pendidikan sebagai media pembangunan sumber daya manusia harus dapat berperan dalam pembentukan peserta didik agar mampu menjadi lulusan yang produktif, mampu menciptakan lulusan sesuai standar industri dan mampu menghadapi persaingan pada pasar global. Sudah selayaknya pendidikan tidak hanya memperhatikan proses berlangsungnya pembelajaran, namun juga perlu adanya evaluasi dalam pendidikan tersebut.

Evaluasi pendidikan yang dilakukan nantinya dapat digunakan untuk mengetahui seberapa besar ketercapaian proses pembelajaran itu sendiri. Proses belajar mengajar yang dilakukan oleh pendidik atau guru harus selalu diperbaiki agar nantinya dicapai hasil yang lebih baik. Salah satu upaya dalam meningkatkan kualitas proses dan hasil belajar sebagai bagian dari peningkatan kualitas pendidikan melalui sistem penilaian. Salah satu bentuk penilaian pembelajaran dalam pendidikan adalah ujian akhir semester.

Soal ujian akhir semester genap mata diklat Dasar-Dasar Mesin kelas X di SMK Muhammadiyah Gamping adalah salah satu contoh soal yang dibuat oleh guru atau pendidik mata diklat. Berdasarkan dari hasil ujian akhir semester genap mata diklat Dasar-Dasar Mesin tahun 2015/2016 diperoleh data nilai rata-rata dari kelas TKR A sebesar 60,2, untuk kelas TKR

B sebesar 67,56, dan untuk kelas TKR C sebesar 66,85. Hal ini membuktikan bahwa nilai hasil ujian akhir semester untuk mata diklat Dasar-Dasar Mesin masih belum mencapai KKM mata diklat tersebut yaitu sebesar 75. Salah satu indikasi dari belum terpenuhinya KKM adalah instrumen penilaian yang digunakan belum cukup memadai. Dalam hal ini instrumen yang dimaksud adalah soal tes.

Selama ini penilaian kurang mendapat perhatian dari pendidik. Indikasinya adalah pembuatan soal yang dinilai seadanya. Pada dasarnya soal yang dibuat oleh pendidik kebanyakan belum memperhatikan kaidah-kaidah dalam penulisan soal. Hal ini didasarkan pada wawancara dengan guru mata diklat Dasar-Dasar Mesin, jika penyusunan soal didasarkan pada kisi-kisi soal dan belum memperhatikan kaidah penulisan soal seperti aspek kognitif, daya pembeda, maupun pola sebaran jawaban. Soal ujian tersebut juga belum diketahui kualitasnya. Kualitas soal yang baik sangat dibutuhkan untuk mengetahui kelayakan soal ujian tersebut. Selain itu soal ujian tersebut juga belum pernah dilakukan analisis secara kualitatif maupun kuantitatif sehingga belum diketahui apakah sudah tepat untuk mengukur siswa atau belum.

Analisis butir soal dapat digunakan untuk memperoleh informasi tentang layak atau tidaknya soal ini digunakan. Hal ini dikarenakan untuk mengetahui kualitas dari soal tersebut agar mampu memperoleh hasil yang optimal dalam pembelajaran dan juga mampu meningkatkan kualitas pendidikan. Tujuan penelitian ini adalah untuk mengetahui kualitas butir soal ujian akhir semester genap mata diklat Dasar-Dasar Mesin tahun ajaran 2015/2016 ditinjau dari analisis

kualitatif berupa aspek materi, konstruksi dan bahasa, serta ditinjau dari analisis kuantitatif berupa validitas, reliabilitas, daya pembeda, tingkat kesukaran, dan efektivitas pengecoh.

## METODE PENELITIAN

### Jenis Penelitian

Penelitian ini merupakan penelitian deskriptif. Data yang dianalisis tidak untuk menerima atau menolak hipotesis, melainkan hasil analisis berupa deskriptif dari gejala yang diamati, berupa angka-angka atau koefisien antar variabel. Pendekatan yang digunakan adalah analisis kuantitatif karena data yang diperoleh dalam bentuk angka dan dianalisis melalui statistik menggunakan bantuan komputer dengan program *Microsoft Excel*. Namun dalam penelitian ini juga menggunakan analisis kualitatif yaitu menggunakan format penelaahan soal pilihan ganda yang dilakukan oleh peneliti.

### Waktu dan Tempat Penelitian

Penelitian ini dilaksanakan pada bulan Desember 2016. Adapun penelitian ini dilaksanakan di SMK Muhammadiyah Gamping yang beralamat di Jln. Wates Km.6 Depok, Ambarketawang, Gamping, Yogyakarta.

### Sumber Data/Subjek Penelitian

Subjek dalam penelitian ini adalah seluruh siswa kelas X Teknik Kendaraan Ringan SMK Muhammadiyah Gamping tahun ajaran 2015/2016 yang berjumlah 83 siswa dari kelas X TKR A sampai X TKR C. Objek penelitian ini adalah soal ujian beserta kunci jawaban dari seluruh jawaban siswa dan kisi-kisi soal.

### Teknik Pengumpulan Data

Teknik pengumpulan data pada penelitian ini adalah menggunakan metode dokumentasi. Metode dokumentasi digunakan untuk mendapatkan seperangkat soal beserta kunci jawaban siswa, silabus, kisi-kisi soal, hasil ujian siswa, serta respon jawaban dari siswa.

### Teknik Analisis Data

Teknik analisis data yang digunakan dalam penelitian ini menggunakan analisis kuantitatif dan kualitatif.

#### 1. Analisis Kuantitatif

Analisis soal obyektif secara kuantitatif adalah sebagai berikut:

##### a. Validitas

Analisis validitas bertujuan untuk mengetahui apakah suatu tes yang digunakan sudah tepat sebagai alat ukur atau belum. Validitas item tes dihitung dengan menggunakan rumus point biserial sebagai berikut:

$$Y_{pbi} = \frac{M_p - M_t}{S_t} \sqrt{\frac{p}{q}}$$

Keterangan:

$Y_{pbi}$  = koefisien korelasi biserial

$M_p$  = rerata skor dari subjek yang menjawab betul bagi yang dicari validasinya.

$M_t$  = rerata skor total

$S_t$  = standar deviasi dari skor total

$p$  = proporsi siswa yang menjawab benar

$$(p = \frac{\text{banyak siswa yang benar}}{\text{jumlah seluruh siswa}})$$

$q$  = proporsi siswa yang menjawab salah

$$(q = 1 - p)$$

(Suharsimi, 2009:79)

Berdasarkan patokan apabila  $Y_{pbi} > 0,195$  berarti valid dan apabila  $Y_{pbi} < 0,195$  maka

soal tidak valid. Indeks korelasi point biserial yang diperoleh dari hasil perhitungan dikonsultasikan dengan r tabel dengan taraf signifikansi 5% sesuai jumlah siswa yang diteliti.

#### b. Reliabilitas

Reliabilitas untuk soal pilihan ganda dapat dihitung dengan rumus KR-20 sebagai berikut:

$$r_{11} = \left( \frac{n}{n-1} \right) \left( \frac{S^2 - \sum pq}{S^2} \right)$$

Keterangan:

$r_{11}$  = reliabilitas tes secara keseluruhan

p = proporsi subjek yang menjawab item dengan benar

q = proporsi subjek yang menjawab item dengan salah

$\sum pq$  = jumlah hasil perkalian antara p dan q

$$(q=1-p)$$

n = banyaknya item

S = standar deviasi dari tes (standar deviasi adalah akar varians)

(Suharsimi, 2009:100)

Selanjutnya dalam pemberian interpretasi terhadap koefisien reliabilitas tes pada umumnya digunakan patokan sebagai berikut:

1) Apabila  $r_{11}$  sama dengan atau lebih besar daripada 0,70 berarti tes hasil belajar yang sedang diuji reliabilitasnya dinyatakan telah memiliki reliabilitas yang tinggi (*a reliable*)

2) Apabila  $r_{11}$  lebih kecil dari pada 0,70 berarti bahwa tes hasil belajar yang sedang diuji reliabilitasnya dinyatakan belum memiliki reliabilitas yang tinggi (*un reliable*)

#### c. Daya Pembeda

Daya pembeda soal dapat dihitung dengan menggunakan rumus:

$$DP = \frac{B_u - B_a}{\frac{1}{2} (N_u + N_a)}$$

Keterangan:

DP = indeks daya pembeda

$B_u$  = banyaknya peserta kelompok unggul/atas yang menjawab benar

$B_a$  = banyaknya peserta kelompok asor/bawah yang menjawab benar

N = jumlah peserta tes

$N_u$  = jumlah peserta dari kelompok atas

$N_a$  = jumlah peserta dari kelompok bawah

(Martubi, 2004:44)

Selanjutnya daya pembeda diklasifikasikan sesuai dengan pengklasifikasiannya untuk mengetahui klasifikasi butir soal tersebut.

Klasifikasi Daya Pembeda:

DP: 0,00-0,19 : jelek

DP: 0,20-0,39 : cukup

DP: 0,40-0,69 : baik

DP: 0,70-1,00 : baik sekali

DP:negatif, semua tidak baik (harus dibuang)

(Suharsimi, 2009:218)

#### d. Tingkat Kesukaran

Rumus yang digunakan untuk mencari tingkat kesukaran sebagai berikut:

$$TK = \frac{B_u + B_a}{N_u + N_a}$$

Keterangan:

TK = indeks kesukaran/tingkat kesukaran

$B_u$  = banyaknya peserta kelompok unggul/atas yang menjawab benar

$B_a$  = banyaknya peserta kelompok

asor/bawah yang menjawab benar

$N$  = jumlah peserta tes

$N_u$  = jumlah peserta tes pada kelompok unggul/atas

$N_a$  = jumlah peserta tes pada kelompok asor/bawah

Biasanya diambil :  $N_u = N_a = 27\% \times N$

( $N$  = jumlah seluruh testi)

(Martubi, 2004:41)

Menurut ketentuan yang sering diikuti, indeks kesukaran diklasifikasikan sebagai berikut:

TK : 0,00-0,30 = sukar

TK : 0,31-0,70 = sedang

TK : 0,71-1,00 = mudah

(Nana Sudjana, 2013:137)

e. Efektivitas Pengecoh

Efektivitas pengecoh dapat diketahui dengan melihat pola sebaran jawaban dari siswa. Sebaran jawaban dapat diperoleh dengan menghitung banyaknya testee yang memilih pilihan jawaban atau yang tidak memilih apapun. Efektivitas pengecoh adalah jika pengecoh telah dapat menjalankan fungsinya dengan baik apabila pengecoh tersebut telah dipilih sekurang-kurangnya 5% dari seluruh peserta tes (Anas Sudijono, 2012:411).

Semakin banyak peserta tes yang memilih pengecoh tersebut, maka pengecoh tersebut dapat menjalankan fungsinya dengan baik. Apabila peserta tes mengabaikan semua option (tidak memilih) disebut omit. Dilihat dari segi omit sebuah tes dikatakan baik

apabila omitnya tidak lebih dari 10% peserta didik.

f. Pemilihan kualitas butir soal

Kriteria pemilihan diadaptasi dari skala likert disesuaikan dengan kebutuhan kualitas butir soal sebagai berikut:

- 1) Butir soal yang memiliki kualitas sangat baik jika memenuhi empat kriteria, yaitu validitas, daya pembeda, tingkat kesukaran, dan efektivitas pengecoh.
- 2) Butir soal yang memiliki kualitas baik jika memenuhi tiga dari empat kriteria. Butir soal ini perlu dilakukan perbaikan.
- 3) Butir soal yang memiliki kualitas cukup jika memenuhi dua dari empat kriteria. Butir soal ini perlu dilakukan perbaikan yang lebih mendalam.
- 4) Butir soal yang memiliki kualitas jelek jika hanya memenuhi satu dari empat kriteria. Butir soal ini harus diganti dengan soal baru.
- 5) Butir soal yang memiliki kualitas jelek sekali jika tidak memenuhi satupun kriteria. Butir soal ini harus dibuang dan tidak digunakan.

2. Analisis Kualitatif

Untuk membantu analisis soal dapat dibuat kartu analisis soal obyektif mengacu pada kartu telaah soal yang ditulis oleh Pusat Pengembangan Sistem Pengujian (Pusbangsisjian) Jakarta.

Analisis soal dilakukan dengan mencocokkan rumusan soal dengan kriteria yang ada. Apabila rumusan soal sesuai dengan kriteria yang ada maka diberi tanda vek ( $\surd$ ) pada kolom soal. Sedangkan jika rumusan soal tidak sesuai dengan kriteria

yang ada diberi tanda silang (X) pada kolom soal.

## HASIL PENELITIAN DAN PEMBAHASAN

Penelitian ini dilakukan untuk mengetahui kualitas butir-butir soal Ujian Akhir Semester Genap mata diklat Dasar-Dasar Mesin kelas X SMK Muhammadiyah Gamping tahun ajaran 2015/2016 dilihat dari hasil analisis kuantitatif meliputi validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektifitas pengecoh, serta dari hasil analisis kualitatif meliputi aspek materi, konstruksi, dan bahasa. Instrumen yang digunakan berupa seperangkat soal ujian, respon jawaban siswa, beserta kunci jawaban soal ulangan akhir semester genap Dasar-Dasar Mesin kelas X SMK Muhammadiyah Gamping tahun ajaran 2015/2016 yang terdiri dari 30 soal obyektif dengan alternatif jawaban 5 (a, b, c, d, e) yang diujikan kepada 83 siswa kelas X Teknik Kendaraan Ringan. Instrumen diperoleh dengan menggunakan metode dokumentasi yaitu untuk memperoleh seperangkat soal beserta kunci jawaban, respon jawaban siswa, silabus, dan kisi-kisi pembuatan soal.

Selanjutnya instrumen tersebut dianalisis secara kualitatif menggunakan kartu telaah untuk mengetahui kualitas butir soal berdasarkan aspek materi, konstruksi dan bahasa dan secara kuantitatif menggunakan program *Microsoft Excel* untuk mengetahui kualitas butir soal berdasarkan validitas, reliabilitas, daya pembeda, tingkat kesukaran, dan efektivitas pengecoh. Selanjutnya disajikan hasil analisis soal secara kualitatif dan kuantitatif sebagai berikut:

### 1. Analisis Kualitatif

Tabel 1. Hasil Analisis Kualitatif Aspek Materi

Aspek yang Ditelaah	Nomor Soal yang Tidak Sesuai Kriteria
Soal sesuai dengan indikator	-
Pengecoh berfungsi	9, 16
Mempunyai satu jawaban yang benar atau yang paling benar	-

Ditinjau dari aspek materi ada 2 butir soal yang tidak memenuhi kriteria yaitu butir soal nomor 9 dan 16. Butir soal tersebut tidak memenuhi aspek pengecoh berfungsi. Pilihan jawaban yang baik hendaknya mengacu juga pada pengecoh soal. Pengecoh yang baik tidak menyimpang jauh dari materi atau bahan ajar. Pilihan jawaban juga tidak boleh sama antara pilihan satu dengan yang lainnya atau *double* dalam tiap satu soal. Selain itu pilihan jawaban jangan mudah ditebak.

Tabel 2. Hasil Analisis Kualitatif Aspek Konstruksi

Aspek yang Ditelaah	Nomor Soal yang Tidak Sesuai Kriteria
Pokok soal dirumuskan dengan jelas dan tegas	4, 6, 9, 13, 18, 19, 20
Rumusan soal dan rumusan jawaban hanya merupakan pernyataan yang diperlukan saja	-
Pokok soal tidak menunjukkan ke arah jawaban yang benar	-
Pokok soal tidak mengandung pernyataan negatif ganda	-
Pilihan jawaban homogen dan logis	30
Panjang rumusan jawaban relatif sama	-
Pilihan jawaban tidak menggunakan pernyataan "semua jawaban di atas benar atau semua jawaban di atas salah"	11
Pilihan jawaban yang berbentuk angka disusun berdasarkan urutan, sedangkan pilihan jawaban yang berbentuk waktu kejadian disusun secara kronologi	-
Gambar, grafik, tabel, dan diagram yang terdapat pada soal jelas dan berfungsi	-
Butir soal tidak bergantung jawaban sebelumnya	13, 18, 19

Hasil peninjauan butir soal dari aspek konstruksi ada 9 soal yang belum memenuhi kriteria. Soal yang belum memenuhi kriteria tersebut adalah nomor 4, 6, 9, 11, 13, 18, 19, dan

20. Beberapa soal yang belum memenuhi pokok soal dirumuskan dengan jelas dan tegas antara lain nomor 4, 6, 9, 13, 18, 19, dan 20. Misalkan pada soal nomor 6, soal masih terlalu panjang sedangkan kenyataannya soal dapat lebih singkat, jelas, dan tegas untuk dapat lebih mempermudah pemahaman dari siswa. Masih ada soal yang belum memenuhi pilihan jawaban tidak menggunakan pernyataan “semua jawaban diatas salah/benar” dan sejenisnya yaitu soal nomor 11. Seharusnya pilihan jawaban tidak menggunakan a, b, c, d tidak tepat, melainkan menggunakan pilihan yang lain yang lebih logis.

Tabel 3. Hasil Analisis Kualitatif Aspek Bahasa

Aspek yang Ditelaah	Nomor Soal yang Tidak Sesuai Kriteria
Soal menggunakan bahasa yang sesuai dengan kaidah bahasa Indonesia	1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 24, 28, 29, 30
Bahasa yang digunakan komunikatif	1, 2, 28
Tidak menggunakan bahasa yang berlaku setempat	-
Pilihan jawaban tidak mengulang kata/frasa yang bukan merupakan satu kesatuan pengertian	1, 2, 17, 18, 19, 28

Hasil peninjauan butir soal dari aspek bahasa/budaya ditemukan sebanyak 24 soal yang belum memenuhi kriteria tersebut. Kesalahan terjadi pada penulisan yang tidak sesuai dengan kaidah penulisan Bahasa Indonesia. Kesalahan lainnya pada penggunaan bahasa yang komunikatif serta terjadi pengulangan kata/sekelompok kata yang sama.

Penulisan soal yang tidak sesuai dengan kaidah penulisan Bahasa Indonesia terdapat pada penulisan pilihan jawaban tidak perlu diawali dengan huruf kapital kecuali stem diakhiri dengan tanda tanya (?). Rata-rata penulisan pilihan jawaban soal Ujian Akhir Semester Genap DDM tersebut diawali dengan huruf kapital. Pada kata-kata asing juga seharusnya dicetak miring dalam

penulisannya. Kesalahan dari aspek bahasa selanjutnya adalah dari segi penggunaan bahasa yang komunikatif. Namun dari hasil telaah pada soal Ujian Akhir Semester Genap ada beberapa butir soal yang penggunaan bahasanya kurang komunikatif. Pengulangan kata atau kelompok kata yang sama pada pilihan jawaban juga tidak dianjurkan pada penulisan soal pilihan ganda.

Salah satu kelemahan tes pilihan ganda adalah tes pilihan ganda memungkinkan peserta tes untuk menebak jawaban. Sehingga siswa yang tidak menguasai materi mempunyai kemungkinan untuk menjawab jawaban dengan benar. Untuk meminimalisir hal tersebut dapat diterapkan sistem denda. Maksud sistem denda ini misalnya untuk menjawab benar diberi nilai 4 poin, jawaban salah diberi -1 dan apabila tidak menjawab diberi nilai 0. Dengan menerapkan hal tersebut, siswa akan lebih berhati-hati dalam menjawab. Apabila benar-benar tidak tahu maka siswa akan memilih untuk tidak menjawab daripada hanya menebak jawaban karena takut nilainya dikurangi apabila salah dalam menjawab.

## 2. Analisis Kuantitatif

Analisis soal secara kuantitatif pada soal pilihan ganda dilakukan dengan program *Microsoft Excel*, yang meliputi validitas, reliabilitas, daya pembeda, tingkat kesukaran, dan efektivitas pengecoh.

### a. Validitas

Analisis validitas soal dengan menggunakan *Microsoft Excel* diperoleh hasil seperti berikut:

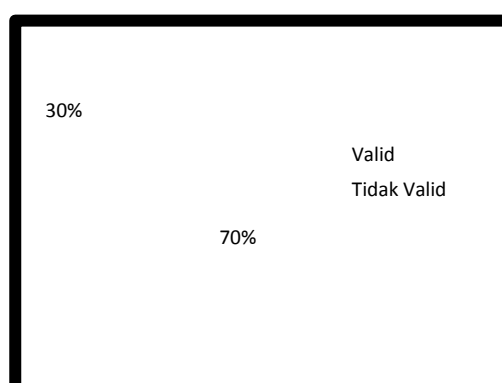
Tabel 4. Hasil Perhitungan Validitas Soal

Nomor Soal	Validitas	Kategori
1	0,521	Valid
2	0,485	Valid
3	0,504	Valid
4	0,558	Valid
5	0,367	Valid
6	0,247	Valid
7	0,159	Tidak Valid
8	0,355	Valid
9	0,265	Valid
10	0,445	Valid
11	0,346	Valid
12	0,156	Tidak Valid
13	0,140	Tidak Valid
14	0,575	Valid
15	0,223	Valid
16	0,616	Valid
17	-0,041	Tidak Valid
18	-0,041	Tidak Valid
19	0,478	Valid
20	0,314	Valid
21	-0,060	Tidak Valid
22	0,333	Valid
23	-0,22	Tidak Valid
24	0,262	Valid
25	0,440	Valid
26	0,247	Valid
27	0,069	Tidak Valid
28	0,297	Valid
29	0,164	Tidak Valid
30	0,381	Valid

Tabel 5. Distribusi Butir Soal Berdasarkan Validitas

Indeks Validitas	Nomor Butir Soal	Jumlah	Presentase
$Y_{pbi}$ 1. Valid $0,220 \leq Y_{pbi} \leq 1$	1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 14, 15, 16, 19, 20, 22, 24, 25, 26, 28, 30	21	70%
2. Tidak Valid $Y_{pbi} \leq 0,220$	7, 12, 13, 17, 18, 21, 23, 27, 29	9	30%

Berikut ini adalah diagram pie analisis validitas soal ujian akhir semester genap DDM kelas X SMK Muhammadiyah Gamping:



Gambar 1. Diagram Pie Analisis Validitas

Validitas dihitung dengan menggunakan rumus korelasi point biserial. Jumlah seluruh siswa kelas X Teknik Kendaraan Ringan adalah 83 siswa, sehingga diketahui  $n=83$ , nilai  $r$  tabel menunjukkan angka 0,220. Berdasarkan hasil analisis ke-30 butir soal ujian akhir semester genap Mata Diklat Dasar-Dasar Mesin kelas X SMK Muhammadiyah Gamping, dapat diketahui bahwa soal yang masuk kategori valid berjumlah 21 butir soal (70%). Sedangkan soal yang masuk kategori tidak valid berjumlah 9 butir soal (30%). Butir soal yang tidak valid tersebut sebaiknya direvisi dan butir soal yang valid dapat digunakan kembali untuk tes.

Berdasarkan uraian di atas dapat disimpulkan bahwa sebagian besar soal Ujian Akhir Semester Genap mata diklat Dasar-Dasar Mesin kelas X SMK Muhammadiyah Gamping tahun ajaran 2015/2016 termasuk soal yang cukup baik berdasarkan validitasnya. Untuk butir soal yang tidak valid sebaiknya diadakan revisi dengan cara meningkatkan penguasaan teknis tentang cara-cara pembuatan soal tes. Soal dapat menjadi valid karena konstruksinya baik dan mencakup materi yang benar-benar mencakup keseluruhan yang akan diukur.

#### b. Reliabilitas

Hasil analisis soal dengan menggunakan *Microsoft Excel* diketahui bahwa reliabilitas soal sebesar 0,637. Dengan koefisien reliabilitas sebesar 0,637 berarti tes yang digunakan belum memiliki reliabilitas yang tinggi (*un reliable*).

#### c. Daya Pembeda

Analisis daya pembeda soal dengan menggunakan *Microsoft Excel* diperoleh hasil seperti berikut:



Tabel 6. Hasil Perhitungan Daya Pembeda Soal

Nomor Soal	Daya Pembeda	Klasifikasi Daya Pembeda
1	0,54	Baik
2	0,62	Baik
3	0,45	Baik
4	0,62	Baik
5	0,27	Cukup
6	0,22	Cukup
7	0,27	Cukup
8	0,36	Cukup
9	0,45	Baik
10	0,54	Baik
11	0,40	Baik
12	0,13	Jelek
13	0,22	Cukup
14	0,62	Baik
15	0,22	Cukup
16	0,76	Baik Sekali
17	0,00	Jelek
18	0,00	Jelek
19	0,40	Baik
20	0,45	Baik
21	-0,09	Jelek Sekali
22	0,31	Cukup
23	-0,27	Jelek Sekali
24	0,22	Cukup
25	0,54	Baik
26	0,31	Cukup
27	0,13	Jelek
28	0,36	Cukup
29	0,13	Jelek
30	0,40	Baik

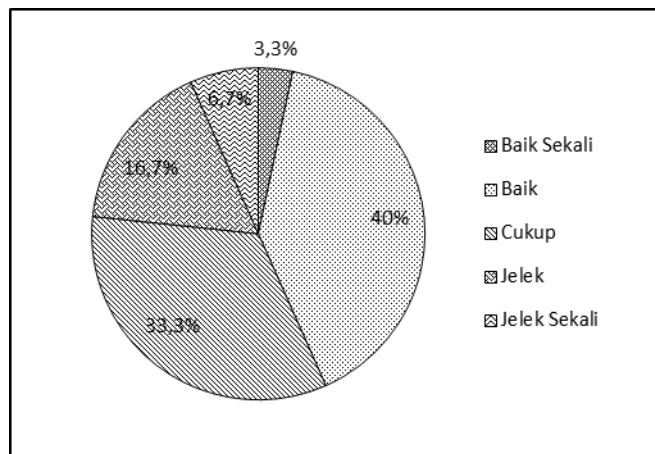
beberapa soal yang memiliki daya beda jelek sekali atau yang bernilai negatif. Dari 30 soal yang diujikan, soal dengan daya beda baik sekali sebanyak 1 butir soal, soal dengan daya beda baik sebanyak 12 butir soal, soal dengan daya beda cukup sebanyak 10 butir soal, soal dengan daya beda jelek sebanyak 5 butir soal, dan soal dengan daya beda jelek sekali sebanyak 2 butir soal.

Berdasarkan hasil penelitian tersebut dapat disimpulkan bahwa sebanyak 76,8% dari soal yang diujikan memiliki daya beda yang memadai sedangkan 23,2% dari soal memiliki daya beda yang lemah. Soal pada kriteria diterima atau baik berarti soal tersebut mempunyai kemampuan untuk membedakan antara siswa yang sudah memahami materi dan siswa yang belum memahami materi. Soal pada kriteria tidak diterima/tidak dapat dipakai maka harus dibuang karena tidak dapat membedakan antara siswa yang sudah memahami materi dan siswa yang belum memahami materi. Suatu soal tidak dapat membedakan antara siswa yang memahami materi dan siswa yang belum memahami materi dimungkinkan karena kunci jawaban soal tidak tepat, butir soal mempunyai dua atau lebih kunci jawaban, kompetensi yang diukur tidak jelas, pengecoh tidak berfungsi, dan materi yang ditanyakan terlalu sulit

Tabel 7. Distribusi Butir Soal Berdasarkan Daya Pembeda

Klasifikasi	Nomor Soal	Jumlah	Presentase (%)
Baik Sekali	16	1	3,3%
Baik	1, 2, 3, 4, 9, 10, 11, 14, 19, 20, 25, 30	12	40%
Cukup	5, 6, 7, 8, 13, 15, 22, 24, 26, 28	10	33,3%
Jelek	12, 17, 18, 27, 29	5	16,7%
Jelek Sekali	21, 23	2	6,7%

Berikut ini adalah diagram pie analisis daya pembeda soal ujian akhir semester genap DDM kelas X SMK Muhammadiyah Gamping:



Gambar 2. Diagram Pie Analisis Daya Pembeda

Pada analisis yang dilakukan secara keseluruhan daya beda soal diketahui ada

d. Tingkat Kesukaran

Analisis tingkat kesukaran soal dengan menggunakan *Microsoft Excel* diperoleh hasil seperti berikut:

Tabel 8. Hasil Perhitungan Tingkat Kesukaran Soal

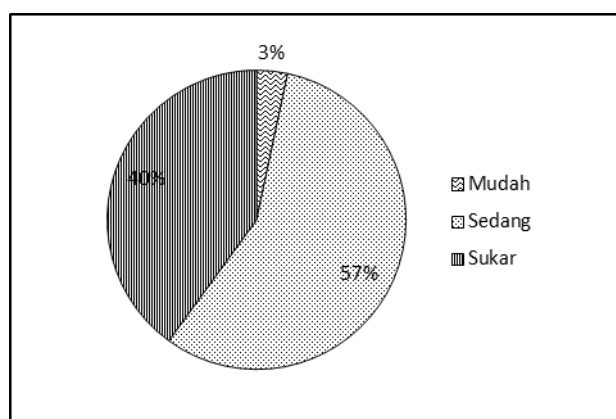
Nomor Soal	Tingkat Kesukaran	Klasifikasi
1	0,45	Sedang
2	0,40	Sedang
3	0,22	Sukar
4	0,54	Sedang
5	0,18	Sukar
6	0,29	Sukar

Nomor Soal	Tingkat Kesukaran	Klasifikasi
7	0,18	Sukar
8	0,22	Sukar
9	0,49	Sedang
10	0,45	Sedang
11	0,20	Sukar
12	0,42	Sedang
13	0,42	Sedang
14	0,62	Sedang
15	0,16	Sukar
16	0,56	Sedang
17	0,40	Sedang
18	0,27	Sukar
19	0,74	Mudah
20	0,31	Sedang
21	0,27	Sukar
22	0,20	Sukar
23	0,18	Sukar
24	0,33	Sedang
25	0,27	Sukar
26	0,56	Sedang
27	0,38	Sedang
28	0,31	Sedang
29	0,38	Sedang
30	0,38	Sedang

Tabel 9. Distribusi Soal Berdasarkan Tingkat Kesukaran

Klasifikasi	Nomor Soal	Jumlah	Presentase (%)
Mudah	19	1	3,3%
Sedang	1, 2, 4, 9, 10, 12, 13, 14, 16, 17, 20, 24, 26, 27, 28, 29, 30	17	57,7%
Sukar	3, 5, 6, 7, 8, 11, 15, 18, 21, 22, 23, 25	12	40%

Berikut ini adalah diagram pie analisis tingkat kesukaran soal ujian akhir semester genap DDM kelas X SMK Muhammadiyah Gamping:



Gambar 3. Diagram Pie Analisis Tingkat Kesukaran

Hasil analisis tingkat kesukaran soal yang dilakukan terhadap 30 butir soal pilihan ganda tersebut dapat diketahui bahwa 1 butir soal termasuk kategori mudah (3,3%), 17 butir soal termasuk kategori sedang (57,7%), dan 12 butir soal termasuk kategori sukar (40%). Pada soal

Ujian Akhir Semester Genap mata diklat Dasar-Dasar Mesin kelas X SMK Muhammadiyah Gamping jumlah tingkat kesukaran kategori sukar masih terlalu banyak yaitu sebesar 40% untuk kategori sukar atau 12 soal. Jumlah tersebut jauh lebih banyak dari jumlah ideal yang ditetapkan.

Perbandingan tingkat kesukaran soal sebaiknya dibuat proporsional yaitu 3:5:2. Artinya 30% soal kategori mudah, 50% kategori sedang, dan 20% kategori sukar. Namun dari hasil perhitungan diperoleh tingkat kesukaran soal kategori sukar mencapai 40%. Berdasarkan jumlah soal dengan kategori sukar tersebut dapat diartikan bahwa sebagian siswa tidak mampu mengerjakan soal tersebut. Hal ini sesuai dengan hasil tes yang diperoleh oleh siswa dimana sebagian besar siswa mendapatkan hasil yang kurang memuaskan.

#### e. Efektivitas Pengecoh

Analisis efektivitas pengecoh soal dengan menggunakan *Microsoft Excel* diperoleh hasil seperti berikut:

Tabel 10. Hasil Perhitungan Efektivitas Pengecoh Soal

No	Distribusi Jawaban					Kunci Jawaban	Efektivitas Pengecoh
	A	B	C	D	E		
1	0,446	0,349	0,157	0,036	0,012	A	E Tidak Efektif
2	0,169	0,373	0,325	0,096	0,024	C	E Tidak Efektif
3	0,313	0,337	0,193	0,012	0,145	C	D Tidak Efektif
4	0,048	0,084	0,277	0,000	0,590	E	A dan D Tidak Efektif
5	0,241	0,410	0,181	0,133	0,012	D	E Tidak Efektif
6	0,012	0,145	0,325	0,217	0,277	C	A Tidak Efektif
7	0,289	0,157	0,036	0,241	0,253	E	C Tidak Efektif
8	0,217	0,253	0,193	0,084	0,205	A	Efektif
9	0,036	0,253	0,060	0,482	0,169	D	A Tidak Efektif
10	0,410	0,313	0,205	0,048	0,024	A	D dan E Tidak Efektif
11	0,096	0,217	0,241	0,036	0,410	C	D Tidak Efektif
12	0,108	0,422	0,048	0,277	0,145	B	C Tidak Efektif
13	0,217	0,108	0,060	0,410	0,205	D	Efektif
14	0,024	0,133	0,735	0,072	0,036	C	A dan E Tidak Efektif
15	0,265	0,120	0,096	0,253	0,253	A	Efektif

No	Distribusi Jawaban					Kunci Jawaban	Efektivitas Pengecoh
	A	B	C	D	E		
16	0,735	0,096	0,096	0,000	0,048	A	D dan E Tidak Efektif
17	0,434	0,289	0,181	0,048	0,024	A	D dan E Tidak Efektif
18	0,193	0,313	0,289	0,036	0,169	B	D Tidak Efektif
19	0,048	0,000	0,084	0,855	0,012	D	A, B dan E Tidak Efektif
20	0,373	0,012	0,181	0,024	0,398	A	B dan D Tidak Efektif
21	0,277	0,120	0,193	0,301	0,108	C	Efektif
22	0,217	0,169	0,398	0,169	0,036	A	E Tidak Efektif
23	0,084	0,361	0,169	0,265	0,108	C	Efektif
24	0,217	0,361	0,108	0,229	0,072	B	Efektif
25	0,169	0,181	0,253	0,108	0,277	E	Efektif
26	0,133	0,145	0,494	0,181	0,048	C	E Tidak Efektif
27	0,096	0,349	0,253	0,205	0,096	B	Efektif
28	0,313	0,301	0,205	0,133	0,048	A	E Tidak Efektif
29	0,349	0,084	0,157	0,265	0,145	A	Efektif
30	0,060	0,108	0,325	0,145	0,361	E	Efektif

Pada analisis efektivitas pengecoh juga harus ada perbedaan frekuensi jawaban antara siswa kelompok atas dan siswa kelompok bawah. Berdasarkan analisis efektivitas pengecoh yang telah dilakukan, diketahui bahwa kebanyakan pengecoh pada soal Ujian Akhir Semester Genap mata diklat Dasar-Dasar Mesin kelas X SMK Muhammadiyah Gamping tidak berfungsi dengan baik. Sebanyak 10 butir soal (33,3%) mempunyai pengecoh yang baik dan 20 butir soal (66,7%) mempunyai pengecoh yang tidak baik.

Pengecoh dikatakan berfungsi bila dipilih oleh minimal 5% dari jumlah siswa yang mengikuti tes serta dipilih oleh siswa yang kurang menguasai materi ujian. Jika pengecoh justru dipilih oleh siswa yang termasuk kategori pandai maka dapat dikatakan pengecoh tersebut menyesatkan. Dari 30 butir soal yang diujikan, ada 2 soal yang pengecohnya menyesatkan yaitu nomor 21 dan 23. Pengecoh pada soal tersebut justru lebih banyak dipilih oleh siswa yang tergolong pandai dari pada dipilih oleh siswa yang tergolong kurang pandai.

## SIMPULAN DAN SARAN

### Simpulan

Simpulan dari penelitian ini adalah sebagai berikut:

Kualitas butir soal Ujian Akhir Semester Genap Mata Diklat Dasar-Dasar Mesin kelas X SMK Muhammadiyah Gamping tahun ajar 2015/2016 ditinjau dari aspek materi, konstruksi, dan bahasa/budaya yaitu 2 soal tidak memenuhi aspek materi, 9 soal tidak memenuhi aspek konstruksi, dan 24 soal tidak memenuhi aspek bahasa/budaya. Hanya 6 soal yang sudah memenuhi ketiga aspek tersebut.

Berdasarkan aspek validitas soal, 70% butir soal termasuk kategori valid dan 30% butir soal termasuk kategori tidak valid, yang berarti soal tersebut cukup valid untuk mengukur kemampuan siswa. Berdasarkan aspek reliabilitas diperoleh angka 0,637 yang berarti soal tes belum memiliki reliabilitas yang tinggi (*un reliable*). Berdasarkan aspek daya pembeda soal, 3,3% soal termasuk kategori baik sekali, 40% termasuk kategori baik, 33,3% termasuk kategori cukup, 16,7% termasuk kategori jelek, dan 6,7% termasuk kategori jelek sekali yang berarti soal termasuk kategori cukup. Berdasarkan aspek tingkat kesukaran soal, 3,3% termasuk kategori mudah, 57,7% termasuk kategori sedang, dan 40% termasuk kategori sukar yang berarti soal termasuk kategori sukar. Berdasarkan efektivitas pengecoh, 33,3% termasuk kategori efektif dan 66,7% termasuk kategori tidak efektif yang berarti pengecoh termasuk kategori tidak efektif. Berdasarkan analisis kuantitatif tersebut 2 soal (6,7%) dipertahankan tanpa perbaikan/layak digunakan, 22 soal (73,3%) dipertahankan dengan

perbaikan, 5 soal harus diganti (16,7%) dan 1 soal (3,3%) harus dibuang.

### **Saran**

Saran yang dapat disampaikan adalah sebagai berikut:

Melakukan upaya untuk meningkatkan keterampilan penyusunan instrumen evaluasi berdasarkan pada kaidah penulisan soal dari segi materi, konstruksi, bahasa, validitas, reliabilitas, tingkat kesukaran, dan efektivitas pengecoh soal terutama soal pilihan ganda.

Melakukan koordinasi antar sesama pendidik dalam penyusunan kisi-kisi, pembuatan soal, dan analisis soal untuk memperoleh soal berkualitas.

Mengikuti pelatihan meningkatkan keterampilan dalam penyusunan soal yang baik dan cara penganalisisannya serta perlu untuk mengadakan berbagai metode pembelajaran sehingga dapat memperoleh nilai maksimal.

### **DAFTAR PUSTAKA**

- Anas Sudijono. (2012). *Pengantar Evaluasi Pendidikan*. Jakarta: PT. Raja Grafindo Persada.
- Martubi. (2004). *Evaluasi Pembelajaran Teori (Cognitif)*.
- Nana Sudjana. (2013). *Penilaian Hasil Proses Belajar Mengajar*. Bandung: PT. Remaja Rosdakarya
- Suharsimi Arikunto. (2009). *Dasar-Dasar Evaluasi*. Jakarta: PT. Bumi Aksara.