

## PENYUSUNAN DAN ANALISIS BUTIR SOAL MATA PELAJARAN TEKNOLOGI DASAR OTOMOTIF DI SMK

Bayu N. Witarsa<sup>1</sup>, Wahid Munawar<sup>2</sup>, Ega T. Berman<sup>3</sup>

Universitas Pendidikan Indonesia  
Jl. Dr. Setiabudhi No. 229 Bandung 40154  
bayunoorwitarsa@gmail.com

### ABSTRAK

Tujuan penelitian ini yaitu untuk menghasilkan soal-soal pilihan ganda mata pelajaran TDO yang telah terkalibrasi dengan cara mengetahui validitas isi, validitas konstruk, dan reliabilitas soal. Bentuk tes pilihan ganda dipilih karena mampu mewakili banyak indikator dengan waktu pengerjaan yang singkat. Alat tes yang dibuat tidak diperuntukkan menjadi alat tes utama, akan tetapi sebagai rekomendasi untuk guru. Penelitian menggunakan metode deskriptif. Subjek penelitian adalah soal-soal pilihan ganda yang telah dirancang oleh peneliti. Data penelitian diperoleh melalui 2 subjek uji coba yaitu (1) data validitas isi diperoleh melalui penilaian ahli, (2) data validitas konstruk, validitas soal, dan reliabilitas diperoleh melalui hasil uji coba tes pada 34 siswa. Pengolahan data validitas isi menggunakan rumus CVR. Pengolahan data validitas konstruk, validitas soal, dan reliabilitas menggunakan program *Iteman*. Hasil penelitian menunjukkan bahwa dari 95 butir soal yang dibuat, 42 butir soal dinyatakan valid yang memenuhi kriteria soal. 42 butir soal yang valid terdiri dari 1 butir soal C1, 23 butir soal C2, 17 butir soal C3, dan 1 butir soal C4. Nilai reliabilitas untuk 42 butir soal yang valid menunjukkan kategori sangat tinggi.

Kata kunci: analisis butir soal, otomotif, validitas, reliabilitas, kalibrasi

### PENDAHULUAN

Penilaian merupakan sebuah rangkaian kegiatan pengumpulan informasi tentang proses dan hasil belajar untuk mengukur tingkat penguasaan peserta didik terhadap kompetensi yang telah diberikan. Penilaian dapat mengetahui perbedaan individual antara peserta didik yang satu dengan peserta didik lainnya. Perbedaan ini akan menentukan berhasil atau tidaknya peserta didik dalam menjalankan kewajiban yang nantinya akan menentukan prestasi yang berbeda setiap peserta didik (Sudjana, 2011). Penilaian hasil belajar tercantum pada Peraturan Menteri Pendidikan dan Kebudayaan (Permendikbud) No. 53 tahun 2015 Pasal 1 ayat 1 menyatakan bahwa: Penilaian hasil belajar oleh pendidik adalah proses pengumpulan informasi/data tentang capaian pembelajaran peserta didik dalam aspek sikap, aspek pengetahuan, dan aspek keterampilan yang dilakukan secara terencana dan sistematis yang dilakukan untuk memantau proses, kemajuan belajar, dan perbaikan hasil belajar melalui penugasan dan evaluasi hasil belajar.

---

<sup>1</sup> Mahasiswa Departemen Pendidikan Teknik Mesin FPTK, UPI

<sup>2</sup> Dosen Departemen Pendidikan Teknik Mesin FPTK, UPI

<sup>3</sup> Dosen Departemen Pendidikan Teknik Mesin FPTK, UPI

Mekanisme penilaian hasil belajar oleh pendidik dalam Permendikbud No. 53 tahun 2015 Pasal 8 ayat 5 yaitu, penilaian aspek pengetahuan dilakukan melalui tes tertulis, tes lisan, dan penugasan sesuai dengan kompetensi yang dinilai. Penilaian hasil belajar yang dilakukan oleh pendidik dapat berupa tes tertulis, tes lisan, dan penugasan. Penilaian tersebut harus memenuhi prinsip-prinsip penilaian hasil belajar yang baik (Uno dan Koni, 2014). Penilaian hasil belajar peserta didik diharapkan memenuhi beberapa prinsip, seperti pada Permendikbud No.53 tahun 2015 Pasal 4 ayat: 1) Sahih, berarti penilaian didasarkan pada data yang mencerminkan kemampuan yang diukur, 2) Objektif, berarti penilaian didasarkan pada prosedur dan kriteria yang jelas, tidak dipengaruhi subjektivitas penilai, 3) Akuntabel, berarti penilaian dapat dipertanggungjawabkan, baik dari segi teknik, prosedur, maupun hasilnya (Lestariningsih, 2014).

Instrumen penilaian sebagai alat untuk mengetahui pencapaian pembelajaran peserta didik diharapkan memenuhi persyaratan kelayakan instrumen penilaian (Tayibnapis, 2008). Permendikbud No.53 tahun 2015 Pasal 7 ayat 2 mengemukakan bahwa: Instrumen penilaian yang digunakan oleh Satuan Pendidikan dalam bentuk Penilaian Akhir dan/atau Ujian Sekolah/Madrasah memenuhi persyaratan substansi, konstruksi, dan bahasa serta memiliki bukti validitas empirik. Seiring dengan persyaratan instrumen penilaian yang harus terpenuhi, maka Permendikbud No. 23 tahun 2016 tentang Standar Penilaian Pendidikan Pasal 13 ayat 2 poin 4 menyatakan bahwa prosedur penilaian hasil belajar oleh satuan pendidikan dilakukan dengan mengkoordinasikan kegiatan, salah satunya dengan melakukan analisis kualitas instrumen. Alat tes yang baik harus dilakukan analisis alat tes dengan cara uji validitas isi, validitas konstruk, dan reliabilitas soal (Hendryadi, 2014; Djanuarsih, 2010).

Analisis validitas isi mencakup tentang isi dan bentuk alat tes, termasuk penggunaan bahasa yang baik dan benar. Validitas isi ialah derajat di mana sebuah tes evaluasi mengukur cakupan substansi yang ingin diukur (Sukardi, 2010). Validitas isi lebih cocok digunakan dalam bidang pendidikan atau pembelajaran di sekolah. Suatu tes dikatakan memiliki validitas isi, apabila butir-butir yang disusun sesuai dengan materi-materi pelajaran dan indikator yang telah ditetapkan. Validitas isi adalah validitas yang akan mengecek kecocokan di antara butir-butir tes yang dibuat dengan indikator, materi atau tujuan pembelajaran yang telah ditetapkan (Susetyo, 2015).

Analisis validitas konstruk dan reliabilitas soal mencakup tentang pengukuran tingkat kesukaran butir soal, daya beda, indeks pengecoh, dan reliabilitas butir soal. Validitas konstruk adalah ide teoritis yang dikembangkan untuk menjelaskan dan mengatur beberapa aspek pengetahuan yang ada (Taufiq, 2015). Reliabilitas juga dapat diartikan sama dengan konsistensi atau keajegan. Suatu instrumen evaluasi, dikatakan mempunyai nilai reliabilitas tinggi, apabila tes yang dibuat mempunyai hasil yang konsisten dalam mengukur yang hendak diukur (Sukardi, 2010). Hal ini untuk mengetahui tingkat kesukaran butir-butir soal, untuk menyeleksi siswa yang memiliki prestasi tinggi atau rendah, untuk mengetahui lemah atau kuat pengecoh soal, dan untuk mempertimbangkan kelayakan butir-butir soal.

Kenyataan dilapangan/sekolah setelah melakukan observasi dan wawancara guru, instrumen penilaian/alat tes yang ada di sekolah belum memenuhi prinsip sah, objektif, dan akuntabel (Kustiawan, Munawar, dan Sriyono, 2016). Alat tes yang ada di sekolah belum dilakukan validasi/analisis alat tes. Indikasinya adalah pembuatan soal yang seadanya. Buktinya, Alat tes tidak dilakukan analisis validitas isi, validitas konstruk dan reliabilitas soal, dimana mutu dan kelayakan alat tes belum terukur. Jika ditinjau dari validitas isi, yaitu bahasa atau tingkat keterbacaan, maka dapat mengakibatkan butir soal sulit dimengerti kemana arah dan tujuan soal. Sedangkan jika ditinjau dari segi validitas konstruk dan reliabilitas soal yang belum diketahui, penelaahan atau pengembangan butir-butir soal tidak bisa dilakukan. Pertimbangan antara layak atau tidak layaknya butir-butir soal tidak diketahui. Mutu alat tes dapat dikatakan statis, yaitu tidak mengarah ke arah yang lebih baik (Purnomo & Munadi, 2005).

Penelitian ini berfokus pada penyusunan dan analisis butir soal. Penelitian ini dibatasi hanya pada aspek kognitif dengan membuat alat tes pilihan ganda. Aspek kognitif dipilih karena mata pelajaran Teknologi Dasar Otomotif (TDO) merupakan mata pelajaran kelas X di SMK yang berisi tentang teori-teori dasar otomotif (Harriyanto, 2013). Pilihan ganda dipilih karena dapat mewakili banyak indikator dan waktu pengerjaan alat tes yang singkat.

## **METODE PENELITIAN**

Metode penelitian adalah metode deskriptif. Metode deskriptif adalah suatu metode yang digunakan untuk menggambarkan atau menganalisis suatu hasil penelitian tetapi tidak digunakan untuk membuat kesimpulan yang lebih luas. Survei dilakukan secara langsung ke

sekolah sebagai tahap pendefinisian. Mengumpulkan informasi-informasi yang terkait dengan keadaan alat tes yang ada di sekolah, dan mengkaji kompetensi inti, kompetensi dasar, indikator pencapaian, dan materi pembelajaran. Informasi-informasi yang sudah dikumpulkan menjadi bahan untuk penyusunan alat tes. Kemudian menyusun kisi-kisi alat tes. Penyusunan kisi-kisi alat tes ini merupakan penentuan indikator yang ingin dicapai dan tingkat kognitif soal. Selanjutnya penulisan soal pilihan ganda sesuai dengan kisi-kisi yang telah dibuat. Alat tes yang telah disusun dilakukan penelaahan soal untuk memperbaiki kesalahan keterbacaan dan tujuan soal. Penelaahan soal dapat dilakukan oleh ahli, guru, dan siswa, dan alat tes diperbaiki sesuai dengan hasil telaah. Selanjutnya alat tes dilakukan uji coba yang digunakan sebagai sarana untuk memperoleh data-data terkait alat tes yang telah disusun.

## HASIL PENELITIAN

Hasil penyusunan butir-butir soal berdasarkan tabel kisi-kisi soal berjumlah 95 butir soal, sebagai berikut:

Tabel 1. Jumlah Penyusunan Butir Soal.

No	Kompetensi/ Materi	Nomor Soal	Jumlah
1.	Dasar-dasar mesin	1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50,51.	51 butir
2.	Pembentukan logam	52,53,54,55,56,57,58,59,60,61,62,63,64,65,66.	15 butir
3.	Mesin konversi energi	67,68,69,70,71,72,73,74,75,76,77,78,79,80,81,82,83,84,85,86,87,88,89,90,91,92,93,94,95.	29 butir
Total			95 butir

Pengolahan data validitas isi (Tabel 2) mendasarkan pada rasio kecocokan para validator. Penilaian pada uji keterbacaan didasarkan pada dipahami atau tidak dipahami, dan untuk *judgement* guru didasarkan pada penting atau tidak penting pada setiap butir tes. Pengolahan data validitas isi menggunakan rumus CVR dari Lawshe (Susetyo, 2015), sebagai berikut:

$$CVR = \frac{Mp - \frac{M}{2}}{\frac{M}{2}} = \frac{2Mp}{M} - 1$$

Keterangan:

Mp = Jumlah ahli yang menyatakan penting

M = Jumlah ahli yang memvalidasi

Tabel 2. Hasil Uji Validitas Isi

Kriteria	Nomor Soal	Jumlah	Presentase
Penting	1,2,3,4,5,6,7,8,9,10,12,14,15,16,17,18,19,20, 21,22,24,25,26,27,28,29,31,32,33,34,35,36,38, 39,40,42,43,44,45,46,47,48,49,50,51,55,56,57, 59,60,61,62,63,64,65,66,67,68,69,70,71,72,73, 74,75,77,78,79,80,81,82,83,84,85,86,87,88,89, 91,92,93,95.	82 butir	86%
Tidak Penting	11,13,23,30,37,41,52,53,54,58,76,90,94	13 butir	14%

Validitas Konstruk terdiri atas tingkat kesukaran, daya pembeda, indeks pengecoh dan validitas soal. Analisis validitas konstruk dilakukan dengan menggunakan program *Iteman*. Nilai tingkat kesukaran (Tabel 3) ditunjukkan oleh nilai *Prop.Corect*. Tingkat kesukaran diklasifikasikan pada kriteria tingkat kesukaran, sebagai berikut:

Tabel 3. Tingkat Kesukaran

Kriteria	Nomor Soal	Jumlah	Presentase
Mudah (0,71-1,00)	2,3,27,40,42,47,66,77.	8 butir	10%
Sedang (0,31-0,70)	1,5,6,7,8,9,10,14,15,18,19,20,21,22,24,25, 26,28,29,31,32,33,34,35,36,38,39,43,44,45, 46,49,59,61,62,63,64,65,67,68,69,70,72,74, 75,78,79,81,82,86,89,91,92,95.	55 butir	67%
Sukar 0,00-0,30)	4,12,16,17,48,50,51,55,56,57,60,71,73,80, 83,84,85,87,93.	19 butir	23%

Analisis daya pembeda (Tabel 4) dilakukan menggunakan program *Iteman*. Nilai daya pembeda ditunjukkan oleh nilai *Biser*. Daya pembeda dapat diklasifikasikan pada kriteria daya pembeda, sebagai berikut:

Tabel 4. Daya Pembeda

Kriteria	Nomor Soal	Jumlah	Presentase
Baik Sekali (0,71 – 1,00)	4,32,40,45,56,57,68,73,80,87	10 butir	12%
Baik (0,41 – 0,70)	2,5,6,7,8,9,10,12,14,15,16,19,20,21,22,25, 27,28,31,33,34,35,36,38,39,42,43,46,49, 59,60,61,62,63,64,65,66,69,70,71,77,78, 79,84.	44 butir	54%
Cukup (0,21 – 0,40)	1,3,17,18,24,29,44,50,67,72,75,81,82,85, 86,89,92.	17 butir	21%
Jelek (0,00 – 0,20)	26,51,55,74,83,88,91,95.	8 butir	9%
Tidak Baik (DP < 0)	47,48,93.	3 butir	4%

Analisis indeks pengecoh (Tabel 5) dilakukan menggunakan program *Iteman*. Nilai indeks pengecoh ditunjukkan oleh nilai *Prop.Endorsing* pada masing-masing alternatif jawaban. Indeks pengecoh soal dikatakan berfungsi apabila lebih dari 5% siswa menjawab/memilih pada setiap alternatif jawaban, dan kurang dari 10% siswa tidak menjawab/omit. Indeks pengecoh soal dapat diklasifikasikan pada kriteria berfungsi dan tidak berfungsi (Arikunto, 2012), sebagai berikut:

Tabel 5. Indeks Pengecoh

Pengecoh Berfungsi	Nomor Soal	Jumlah	Presentase
3 (Baik)	1,3,4,7,8,10,15,16,17,20,21,22,25,28,32, 33,34,35,38,42,43,45,46,47,48,49,50,51, 55,56,59,60,62,64,65,72,73,75,78,79,80, 81,82,83,84,87,88,92,93.	49 butir	60%
2 (Cukup)	2,5,6,12,14,18,19,24,27,31,36,39,44,57, 61,63,67,68,69,70,74,85,86,95.	24 butir	29%
1 (Kurang)	9,26,29,40,66,71,77,89,91.	9 butir	11%
0 (Tidak Baik)	-	-	-

Validitas soal (Tabel 6) akan dilihat dari kolerasi *Point Biseral* hasil analisis program *Iteman*. Hasil tersebut akan di interpretasikan dengan nilai  $r_{\text{tabel}}$  *product moment* pada taraf signifikan 5% sesuai dengan jumlah peserta tes yaitu 34 siswa ( $n=34$ ). Nilai  $r_{\text{tabel}}$  dengan  $n=34$  adalah 0,339. Soal dikatakan valid apabila  $\gamma_{\text{pBis}} \geq r_{\text{tabel}}$  dan tidak valid apabila  $\gamma_{\text{pBis}} \leq r_{\text{tabel}}$ .

Tabel 6. Validitas Soal

Korelasi signifikan r hitung	Nomor soal	Jumlah	Persentase
$\gamma_{\text{pBis}} \geq 0,339$ (valid)	4,5,6,7,8,9,10,14,16,19,20,21,22,27,28, 31,32,34,35,38,39,42,43,45,46,49,56,57, 59,60,61,62,63,64,65,68,69,70,73,77,79, 80,87.	44 butir	54%
$\gamma_{\text{pBis}} \leq 0,339$ (tidak valid)	1,2,3,12,15,17,18,24,25,26,29,33,36,40, 44,47,48,50,51,55,66,67,71,72,74,75, 81,82,83,84,85,86,88,89,91,92,93,95.	38 butir	46%

Analisis reliabilitas soal dilakukan untuk memberikan konsistensi bahwa suatu instrumen telah valid. Semakin besar nilai koefisien reliabilitas, maka reliabilitas instrumen pun tinggi/baik. Temuan pada hasil analisis butir soal menggunakan program *Iteman* dapat diketahui bahwa reliabilitas soal sebesar 0,909 untuk 82 butir soal yang dianalisis. Hasil ini dilihat dari besarnya korelasi Alpha.

## PEMBAHASAN

Pengujian validitas isi diuji melalui dua cara yaitu uji keterbacaan dan *judgement* guru. Uji keterbacaan mendapatkan 11 butir soal yang tidak dipahami, yakni soal nomor 3, 26, 33, 35, 39, 40, 41, 46, 52, 59, dan 74. Butir-butir soal yang tidak dipahami kemudian dilakukan revisi bahasa atau keterbacaan. Kemudian dilakukan *judgement* guru yang mendapatkan 13 butir soal yang tidak penting, yakni soal nomor 11, 13, 23, 30, 37, 41, 52, 53, 54, 58, 76, 90, dan 94. Butir-butir soal yang tidak penting merupakan soal yang tidak rasional dengan materi pengajaran guru, tidak sesuai dengan indikator, dan memiliki jawaban ganda. Tindak lanjut untuk butir-butir yang tidak penting adalah dengan mengeliminasi/membuang butir-butir soal tersebut.

Analisis tingkat kesukaran dilakukan menggunakan program *Iteman*. Nilai tingkat kesukaran ditunjukkan oleh nilai *Prop.Corect*. Hasil analisis tingkat kesukaran butir soal yang telah dilakukan yaitu terdapat 8 (10%) butir soal dengan kriteria mudah, 55 (67%) butir soal

dengan kriteria sedang, 19 (23%) butir soal dengan kriteria sukar. Tingkat kesukaran soal-soal berdasarkan presentase kriteria dapat dikatakan cukup. Adanya variasi kriteria tingkat kesukaran dapat mengukur hasil belajar siswa dengan baik. Butir-butir soal dengan kriteria mudah dapat menjadi pemicu kepercayaan diri siswa dalam mengerjakan tes, karena mayoritas siswa dapat menjawab soal dengan benar. Butir-butir soal dengan kriteria sedang menjadi komposisi terbesar dalam alat tes, karena banyaknya variasi jawaban siswa berdasarkan sejauh mana kemampuannya. Butir-butir soal dengan kriteria sukar dapat membedakan antara siswa dengan prestasi atas dengan prestasi bawah (Arif, 2014).

Analisis daya pembeda dilakukan menggunakan program *Iteman*. Nilai daya pembeda ditunjukkan oleh nilai *Biser*. Hasil analisis daya pembeda butir soal yang telah dilakukan yaitu terdapat 10 (12%) butir soal dengan kriteria baik sekali, 44 (54%) butir soal dengan kriteria baik, 17 (21%) butir soal dengan kriteria cukup, 8 (9%) butir soal dengan kriteria jelek, dan 3 (4%) butir soal dengan kriteria tidak baik. Daya pembeda soal-soal berdasarkan presentase kriteria dapat dikatakan baik. Artinya, kemampuan soal dalam membedakan siswa yang sudah menguasai materi dan siswa yang belum menguasai materi tergolong baik.

Butir-butir soal yang termasuk pada kriteria jelek dan tidak baik, yakni soal nomor 26, 47, 48, 51, 55, 74, 83, 88, 91, 93, dan 95 merupakan soal yang tidak layak digunakan. Nilai daya pembeda yang kecil seperti ini diakibatkan banyaknya siswa prestasi atas yang memilih jawaban salah. Tindak lanjut untuk butir-butir soal dengan kriteria jelek dan tidak baik adalah dengan mengeliminasi/membuang butir-butir soal tersebut.

Analisis indeks pengecoh dilakukan menggunakan program *Iteman*. Nilai indeks pengecoh ditunjukkan oleh nilai *Prop.Endorsing* pada masing-masing alternatif jawaban. Hasil analisis indeks pengecoh yang telah dilakukan yaitu terdapat 49 (60%) butir soal yang semua pengecohnya berfungsi dengan baik, terdapat 24 (29%) butir soal yang pengecohnya cukup berfungsi, dan 9 (11%) butir soal yang pengecohnya kurang berfungsi. Butir-butir soal yang semua pengecohnya berfungsi dengan baik dan cukup dapat digunakan. Sedangkan, untuk butir-butir soal yang pengecohnya tidak berfungsi perlu dilakukan revisi terlebih dahulu jika ingin digunakan.

Butir soal dikatakan memiliki validitas soal apabila tes tersebut dapat mengukur apa yang seharusnya diukur. Nilai validitas soal dilihat dari kolerasi *Point Biserial* hasil analisis program *Iteman*. Hasil tersebut akan diinterpretasikan dengan nilai  $r_{\text{tabel}}$  *product moment* pada



taraf signifikan 5% sesuai dengan jumlah peserta tes yaitu 34 siswa ( $n=34$ ). Nilai  $r_{\text{tabel}}$  dengan  $n=34$  adalah 0,339. Soal dikatakan valid apabila  $\gamma_{\text{pBis}} \geq r_{\text{tabel}}$  dan tidak valid apabila  $\gamma_{\text{pBis}} \leq r_{\text{tabel}}$ . Hasil analisis validitas butir soal yang telah dilakukan terdapat 44 (54%) butir soal dinyatakan valid, dan 38 (46%) butir soal dinyatakan tidak valid. Artinya, dari jumlah sebanyak 82 butir soal yang valid sesuai dengan indikator dapat disimpulkan bahwa 44 (54%) butir soal dapat menjalankan fungsi ukurnya dengan baik.

Butir soal valid dengan kualitas yang baik (Tabel 7) dapat diketahui dengan cara membuat keputusan berdasarkan pemenuhan kriteria dari daya pembeda, indeks pengecoh, dan validitas soal. Apabila dua dari tiga kriteria butir tes yang baik dapat terpenuhi atau konsisten, maka butir tes tersebut dapat digunakan. Sebaliknya apabila dua dari tiga butir kriteria butir tidak dapat memenuhi kualitas butir yang baik maka butir tes perlu diganti atau direvisi. Kemudian dapat disimpulkan pula bahwa berdasarkan dengan hasil analisis menggunakan pertimbangan tersebut, maka sebanyak 42 butir soal memiliki kualitas yang baik.

Tabel 7. Butir-butir Soal Valid yang Memenuhi Kriteria

Kognitif	Nomor Soal	Jumlah butir	Presentase
C1	70	1 butir	2%
C2	4,5,6,14,19,27,28,31,32,34,35, 42,45,59,60,61,62,63,64,65,68, 69,92.	23 butir	55%
C3	7,8,10,16,20,38,39,43,46,49,50, 56,57,78,79,80,87.	17 butir	41%
C4	73	1 butir	2%

## KESIMPULAN

Penelitian ini dapat disimpulkan bahwa alat tes dapat dikatakan belum sempurna karena butir-butir soal tidak dapat digunakan seluruhnya. Penelitian ini menghasilkan 42 butir soal yang dinyatakan valid dengan kualitas baik setelah melalui pengujian validitas isi, validitas konstruk, validitas soal, dan reliabilitas soal. 42 butir soal yang valid terdiri dari 1 butir soal C1, 23 butir soal C2, 17 butir soal C3, dan 1 butir soal C4. Nilai reliabilitas untuk 42 butir soal yang valid dengan kualitas baik adalah 0,924. Berarti alat tes memiliki keajegan sangat tinggi.

**REFERENSI**

- Arif, M. (2014). Penerapan Aplikasi Anates Bentuk Soal Pilihan Ganda. *Jurnal Ilmiah Edutic*. 1. hlm 1-9.
- Arikunto, S. (2012). *Dasar-dasar Evaluasi Pendidikan*. Jakarta: PT Bumi Aksara.
- Djanuarsih, E. (2010). Validitas dan Reliabilitas Butir Soal. *Jurnal Dinas Pendidikan Kota Surabaya*. 1. hlm. 1-12.
- Harriyanto. (2013). *Teknologi Dasar Otomotif*. Jakarta: Kementerian Pendidikan dan Kebudayaan.
- Hendryadi. (2014). *Validitas Isi (Content Validity)*. *Jurnal Online*. 1. hlm. 1-5.
- Lestariningsih. (2014). Evaluasi Hasil Belajar Siswa yang Diberi Umpan Balik Positif dan Negatif pada Pokok Bahasan Pecahan. *Jurnal Pendidikan Sidoarjo*. 2. hlm. 65-74.
- Kustiawan, I., Munawar, W., dan Sriyono. (2016). Evaluasi Impementasi Pembelajaran Keterampilan Tune Up Sepeda Motor pada Siswa Sekolah Menengah Kejuruan. *Journal of Mecahanical Engineering Education*. 3 (1). hlm. 83-89.
- Purnomo & Munadi. (2005). Evaluasi Hasil Belajar dalam Implementasi Kurikulum Berbasis Kompetensi di Sekolah Menengah Kejuruan. *Jurnal Cakrawala Pendidikan*. 2. hlm. 259-272.
- Sudjana, N. (2011). *Penilaian Hasil Proses Belajar Mengajar*. Bandung: PT Remaja Rosdakarya.
- Sukardi, H.M. (2010). *Evaluasi Pendidikan Prinsip dan Operasionalnya*. Jakarta: Bumi Aksara.
- Susetyo, B. (2015). *Prosedur Penyusunan & Analisis tes*. Bandung: PT Refika Aditama.
- Taufiq, A.U. (2015). Pengembangan Tes Kognitif Berbasis Revisi Taksonomi Bloom Pada Materi Sistem Reproduksi Untuk Siswa SMA. *Jurnal Biotek*. 3. hlm. 1-17.
- Tayibnapis, F.Y. (2008). *Evaluasi Program dan Instrumen Evaluasi*. Jakarta: Rineka Cipta.
- Uno, H.B. & Koni, S. (2014). *Assesment Pembelajaran*. Jakarta: Bumi Aksara.