

## Analisis Butir Soal Tes Matematika Materi Trigonometri

M.M Endang Susetyawati<sup>1)</sup>, Bintang Wicaksono<sup>2)</sup>

<sup>1,2</sup> Program Studi Pendidikan Matematika, FKIP, Universitas PGRI Yogyakarta

\*Email: [magda@upy.ac.id](mailto:magda@upy.ac.id)

[bintang@upy.ac.id](mailto:bintang@upy.ac.id)

---

### Abstrak

Dalam kegiatan pengukuran kompetensi peserta didik, seorang guru berperan aktif dalam menyusun alat ukur kompetensi. Alat ukur kompetensi harus memiliki kriteria berkualitas yang layak digunakan dalam mengukur kompetensi peserta didik. Penelitian ini bertujuan untuk mengetahui karakteristik dari instrumen tes matematika yang dilihat dari validitas butir soal, tingkat kesulitan butir, daya pembeda butir soal dan keberfungsian pengecoh. Instrumen yang dikembangkan adalah instrumen tes kompetensi materi Trigonometri. Bentuk instrumen ini adalah multiple choice. Sampel dari penelitian ini adalah siswa SMA Negeri 5 Yogyakarta. Analisis butir soal tes kompetensi Trigonometri dilakukan dengan telaah kualitatif dengan hasil 85% dari segi materi memiliki kriteria baik, 82,5% dari segi konstruksi berkriteria baik dan 92,5% segi Bahasa berkriteria baik. Telaah kuantitatif tes kompetensi Trigonometri, diperoleh hasil 27 butir soal (67,5%) memiliki validitas yang baik, tes memiliki nilai reliabilitas yaitu 0,709 dengan kriteria baik. Diperoleh 26 butir soal (65%) dengan tingkat kesulitan butir soal sedang, 28 butir soal (70%) memiliki daya pembeda kategori baik, dan keberfungsian pengecoh 37 butir soal (92,5%) berkriteria baik.

**Kata kunci:** validitas, reliabilitas, tingkat kesukaran, daya pembeda

---

### Abstract

*In learning outcomes assessment activities, teachers play an active role in preparing measuring instruments. Competency measurement tools must have appropriate quality criteria used in measuring student competence. This study aims to determine the characteristics of the mathematical competency test instrument in terms of item validity, item difficulty level, item discriminating power and distractor function. The instrument developed is a competency test instrument for trigonometry material. The form of this instrument is multiple choice. The sample of this research is SMA Negeri 5 Yogyakarta. Analysis of the Trigonometry competency test items was carried out using qualitative analysis with the result that 85% in terms of material had good criteria, 82.5% in terms of construction had good criteria and 92.5% in terms of language had good criteria. Quantitative analysis of the Trigonometry competency test shows that 27 items (67.5%) have good validity, the test has a reliability value of 0.709 with good criteria. There were 26 questions (65%) with moderate difficulty level, 28 questions (70%) had good category discrimination, and 37 questions (92.5%) had good criteria for the distraction function.*

**Keywords:** validity, reliability, difficulty level, differentiating power

---

## 1. PENDAHULUAN

Meningkatkan sumber daya manusia yang memiliki keterampilan berpikir kritis, kreatif, kolaboratif, komunikatif, serta cakap dalam mengambil keputusan, sehingga dapat mengelola, mengolah dan merespon arus informasi dan pemutakhiran teknologi sangat penting dipersiapkan di era globalisasi ini (Lutfianto & Sari, 2017). Peraturan Menteri Pendidikan Nasional Nomor 23 Tahun 2016 “Tentang Standar Penilaian Pendidikan” menjelaskan bahwa: Standar Penilaian pendidikan merupakan kriteria mengenai lingkup, tujuan, manfaat, prinsip, mekanisme, prosedur, dan instrumen penilaian hasil belajar peserta didik. Standar Penilaian Pendidikan digunakan sebagai dasar dalam penilaian hasil

belajar peserta didik pada pendidikan dasar dan pendidikan menengah. (Kementerian Pendidikan dan Kebudayaan, 2013)

Kegiatan mengukur hasil belajar peserta didik bertujuan untuk mengetahui kemampuan berpikir peserta didik, merupakan kegiatan yang tidak lepas dari proses belajar peserta didik. (Mardapi, 2017). Pengukuran dilakukan untuk melihat pencapaian tujuan instruksional. Untuk mengetahui kebutuhan dan kemampuan peserta didik perlu dilakukan penilaian dan pengukuran. (Tshabalala & Ncube, 2014). Hasil pengukuran tersebut merupakan gambaran hasil belajar mengenai derajat pencapaian kompetensi peserta didik. Berdasarkan hasil pengukuran, penilaian dan evaluasi, guru dapat mendiagnosa kekuatan dan kelemahan peserta didik, serta mengambil tindakan untuk kemajuan dan peningkatan hasil belajar mereka. Pengukuran dan evaluasi yang efektif dapat digunakan untuk memperbaiki situasi belajar.

Instumen penilaian atau alat ukur yang berfungsi sebagai alat untuk mengukur pencapaian kompetensi peserta didik, harus baik. Instrumen dikatakan baik apabila memenuhi kaidah-kaidah alat ukur yang baik. Karakteristik alat ukur yang baik harus memenuhi syarat valid dan reliabel. Suatu alat ukur dapat dikatakan valid, jika dapat dengan tepat mengukur apa yang seharusnya diukur. Alat ukur dikatakan reliabel, jika digunakan akan memberikan hasil tes yang ajeg/tetap (*consistent*). Instrumen/alat ukur penilaian hasil belajar berupa tes, atau biasa disebut soal merupakan salah satu alat ukur yang digunakan untuk mendeteksi kemampuan peserta didik. Matematika dikenal sebagai ilmu dasar, pembelajaran matematika akan melatih kemampuan kritis, logis, analitis dan sistematis. Tetapi peran matematika tidak hanya sebatas hal tersebut, seperti bidang lain, seperti fisika, ekonomi, biologi tidak terlepas dari peran matematika. Dengan demikian dalam pembelajaran matematika perlu adanya pembelajaran yang mengarah pada kemampuan berpikir tinggi. (Moore & Stanley, 2009). Menurut De Lange (2004) dalam (Sari, 2015) menyatakan kemampuan yang harus dipelajari siswa dalam pembelajaran matematika adalah "*Mathematical argumentation. Knowing what proofs are; knowing how proofs differ from other forms of mathematical reasoning; following and assessing chains of arguments; having a feel for heuristics; creating and expressing mathematical arguments*".

Soal tes matematika merupakan salah satu instrumen yang penting dalam proses pembelajaran matematika, sehingga butir-butir soal matematika harus terstruktur dengan baik. Butir soal tes kompetensi yang baik akan memberikan hasil pengukuran yang baik (Mardapi, 2008). Soal tes pilihan ganda yang digunakan untuk menilai hasil belajar materi Trigonometri, Kompetensi Dasar Menyelesaikan masalah yang berkaitan dengan rumus jumlah dan selisih sinus dan cosinus, analisis butir soal yang dilakukan masih sangat sederhana. Guru Matematika SMA Negeri 5 Yogyakarta dalam menilai baik atau tidaknya suatu soal dilakukan dengan uji validitas logis, dan berdasarkan bisa atau tidaknya siswa menjawab soal tersebut.

Validitas logis dari suatu alat/instrumen evaluasi yang memenuhi persyaratan valid berdasarkan hasil penalaran. Uji validitas logis belum cukup untuk membangun sebuah tes yang berkualitas. Uji validitas untuk membangun suatu tes yang baik sebaiknya dilakukan dengan analisis butir soal secara empirik.

Analisis butir soal tes secara empirik adalah analisis mutu butir soal tes untuk mengetahui karakteristik dari soal tersebut. Analisis butir soal tes dilakukan berdasarkan data hasil respon peserta tes dalam menjawab butir soal. Analisis butir soal bertujuan untuk mengetahui tingkat kesukaran butir, daya pembeda butir, dan keberfungsian pengecoh. Analisis butir soal secara keseluruhan dilakukan dengan menghitung beberapa aspek yaitu validitas, reliabilitas, daya pembeda, dan tingkat kesukaran. Analisis butir soal dapat dihitung secara manual dan dapat menggunakan bantuan perangkat komputer. Dalam memudahkan menganalisis butir soal, peneliti menggunakan program *Iteman versi 3.00*.

Tujuan penelitian ini adalah 1) Mengetahui validitas butir soal 2) Mengetahui reliabilitas butir soal 3) Mengetahui tingkat kesukaran butir soal 4) Mengetahui daya pembeda butir soal. Penelitian ini diharapkan dapat menjadi masukan kepada guru untuk mengembangkan bank soal matematika yang berkualitas.

## 2. METODE PENELITIAN

### a. Jenis Penelitian

Penelitian ini merupakan penelitian deskriptif dengan menggunakan pendekatan kuantitatif. Dalam pelaksanaannya, penelitian ini bermaksud untuk mencari informasi dan data yang dapat digunakan untuk mendeskripsikan karakteristik butir soal ulangan tengah semester gasal kelas XI di SMK Muhammadiyah 1 Bambanglipuro Bantul. Pendekatan yang digunakan dalam penelitian ini adalah pendekatan kuantitatif karena semua data atau informasi yang diperoleh berupa data numerik yaitu data dalam bentuk angka-angka dan diolah dengan menggunakan program *Iteman versi 3.00*.

### b. Prosedur

Prosedur yang dilakukan untuk melakukan analisis butir soal, sebagai berikut:

1. Mengembangkan Kisi-kisi tes
2. Menulis butir soal tes sesuai kisi-kisi yang telah dibuat
3. Melakukan analisis
4. Melakukan uji coba tes
5. Melakukan analisis empirik dengan program *Iteman versi 3.00*
6. Melakukan penilaian.

### c. Teknik Analisis Data

#### 1. Analisis Validitas Soal

Validitas mencerminkan sejauh mana soal dapat dengan tepat mengukur apa yang seharusnya diukur. Perangkat soal bersifat valid apabila butir-butir soalnya valid. Menghitung validitas butir soal menggunakan koefisien korelasi point biserial ( $r_{pbis}$ ). Menurut Weiresma dan Jurs (1990), secara umum jika koefisien korelasi sudah lebih besar dari 0,3 maka butir instrumen tersebut sudah dikategorikan valid.

Kriteria validitas yang dapat digunakan untuk menginterpretasikan hasil analisis adalah sebagai berikut:

**Tabel 1.** Interpretasi Nilai  $r$

Besarnya nilai $r$	Interpretasi
$0,800 < r \leq 1,000$	Sangat tinggi
$0,600 < r \leq 0,800$	Tinggi
$0,400 < r \leq 0,600$	Sedang
$0,200 < r \leq 0,400$	Rendah
$0,000 < r \leq 0,200$	Sangat rendah

(Sukirman, 2012: 184)

#### 2. Analisis Reliabilitas Soal

Reliabilitas adalah kestabilan score yang diperoleh orang yang sama ketika diuji ulang dengan tes yang sama pada situasi yang berbeda. Jadi reliabilitas dapat dikatakan tingkat konsistensi hasil dua pengukuran terhadap hal yang sama. Tes yang terdiri dari banyak butir soal, tentu saja lebih valid dibandingkan dengan tes yang hanya terdiri dari beberapa butir soal. Tinggi rendahnya validitas menunjukkan tinggi rendahnya reliabilitas tes. Dengan demikian semakin panjang tes, maka reliabilitas semakin tinggi. Menurut Nunally (1972) dan Kaplan & Saccuzo, koefisien reliabilitas 0,7 sampai dengan 0,8 sudah termasuk cukup tinggi.

**Tabel 2.** Interpretasi Reliabilitas Soal

Indeks $\alpha$	Interpretasi
$0,90 < \alpha \leq 1,00$	Sangat tinggi
$0,70 < \alpha \leq 0,90$	Tinggi
$0,40 < \alpha \leq 0,70$	Cukup
$0,20 < \alpha \leq 0,40$	Rendah
$0,00 < \alpha \leq 0,20$	Sangat rendah

(Hariyanto & Basuki, 2014)

### 3. Analisis Tingkat Kesukaran Soal

Kriteria indeks kesulitan soal yaitu:

**Tabel 3.** Kriteria Indeks Kesukaran

Kriteria	Kategori	Keterangan
$> 0,90$	Sangat Mudah	Ditolak
$0,70 < IK \leq 0,90$	Mudah	Direvisi
$0,30 < IK \leq 0,70$	Sedang	Diterima
$0,10 \leq IK \leq 0,30$	Sukar	Direvisi
$< 0,10$	Sangat Sukar	Ditolak

(Hariyanto & Basuki, 2014)

### 4. Analisis Daya Pembeda Soal

Indeks yang dimiliki butir soal dalam membedakan kemampuan antara siswa yang memiliki kemampuan tinggi dan siswa yang memiliki kemampuan rendah disebut indeks daya pembeda (Azwar, 2012). Indeks Daya Beda mengkaji butir-butir soal dengan tujuan untuk mengetahui kesanggupan soal dengan membedakan antara siswa yang tergolong mampu dengan siswa yang tergolong kurang mampu. Untuk menentukan baik tidaknya butir soal dilihat dari indeks daya beda terhadap beberapa acuan yang dapat digunakan. Jika dibuat suatu rentangan, interpretasi indeks daya pembeda suatu butir tes adalah sebagai berikut

**Tabel 4.** Klasifikasi Daya Pembeda

Indeks Daya Pembeda	Interpretasi
$0,00 < DP \leq 0,20$	Kurang baik ( <i>poor</i> )
$0,20 < DP \leq 0,40$	Cukup ( <i>satisfactory</i> )
$0,40 < DP \leq 0,70$	Baik ( <i>good</i> )
$0,70 < DP \leq 1,00$	Baik sekali ( <i>excellent</i> )

(Prestasi & Akuntansi, 2014)

Pengertian distraktor yaitu “*Distractor are classified as the incorrect answer in a multiple-choice question.*”. Dalam setiap tes obyektif selalu digunakan alternatif jawaban yang mengandung 2 unsur sekaligus, yaitu jawaban tepat dan jawaban yang salah sebagai penyesat (distraktor). Untuk analisis pengecoh dikatakan baik jika memiliki nilai korelasi *point biserial* negatif. Nilai negatif ini menunjukkan bahwa siswa yang berkemampuan rendah cenderung memilih pengecoh tersebut sebagai jawaban dan sebaliknya peserta didik yang berkemampuan tinggi akan memilih kunci jawaban sebagai jawabannya. Menurut (Mardapi, 2017) pengecoh dikatakan berfungsi jika dipilih paling sedikit 3% siswa.

## 3. HASIL PENELITIAN dan PEMBAHASAN

### a. Hasil Penelitian

Analisis data dalam penelitian ini diawali dengan analisis kualitatif butir soal, yang meliputi review butir soal dari aspek materi, bahasa dan konstruksi oleh *expert judgment*. Untuk melihat produk awal maka instrumen tes ini di review oleh bidang ahli pendidikan matematika. Langkah ini dimaksudkan untuk mendapatkan masukan,

saran dan memberikan penilaian terhadap instrumen tes yang dikembangkan. Kegiatan validasi ahli dilakukan mulai dari melihat kisi-kisi soal dan instrumennya. Hasil telaah butir soal tes ditunjukkan pada tabel 5.

**Tabel 5.** Telaah kualitatif tes Kompetensi Matematika

Aspek	Kriteria (%)		Nomor Soal
	Baik	Kurang Baik	
<b>Materi</b>	85%	15%	8, 15, 16, 25, 29,38
<b>Konstruksi</b>	82,5%	17,5%	5, 8, 15, 16, 25, 29,38
<b>Bahasa</b>	92,5%	7,5%	15, 25, 38

Dari hasil analisis menunjukkan bahwa pada aspek materi terdapat 34 butir soal yang memenuhi kriteria materi baik. Berarti 85% dari 40 butir soal yang disusun memenuhi kriteria baik. Pada aspek konstruksi terdapat 33 butir soal pada kriteria baik, atau terdapat 82,5% butir soal memenuhi kriteria konstruksi. Pada aspek bahasa terdapat 37 butir soal dengan keterbacaan baik. Analisis kuantitatif menggunakan program ITEMAN versi 3,00 diperoleh statistik butir soal dan statistik skala. Karakteristik butir mencakup: tingkat kesukaran, daya beda dan keberfungsian pengecoh dapat dilihat dari statistik butir soal. Hasil Analisa validitas butir dipaparkan pada tabel 6.

**Tabel 6.** Tabel Hasil Analisis Validitas

Kategori	Nomor Butir	Jumlah
<b>Valid</b>	1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 17, 18, 19, 20, 22, 23, 24, 27, 28, 31, 32, 33, 34, 36, 37, 39, 40	27
<b>Tidak Valid</b>	8, 12, 13, 14, 15, 16, 21, 25, 26, 29, 30, 35, 38	13

Analisis butir soal berdasarkan tingkat kesukaran butir disajikan pada tabel berikut:

**Tabel 7.** Kategori Butir Soal berdasarkan Tingkat Kesukaran

Kategori	Nomor Butir	Jumlah	Ket.
<b>Mudah</b>	6, 7, 11, 13	4	revisi
<b>Sedang</b>	1, 2, 3, 4, 9, 10, 12, 16, 17, 18, 19, 20, 21, 22, 23, 24, 27, 31, 32, 33, 34, 35, 36, 37, 39, 40	26	terima
<b>Sukar</b>	5, 8, 14, 15, 25, 26, 28, 29, 30, 38	10	revisi

Berdasarkan tingkat kesukaran butir soal terdapat 4 butir soal dengan kategori mudah, ada tiga kemungkinan yang dilakukan yaitu butir soal itu dibuang, atau direvisi dan digunakan untuk tes dengan karakteristik kemampuan peserta didik yang lebih rendah. Butir soal yang termasuk dalam kategori sukar, butir soal tersebut direvisi sehingga dapat diujicobakan pada tes yang akan datang atau butir soal tersebut dibuang atau tidak digunakan lagi dalam tes berikutnya. Hasil analisis butir soal berdasarkan indeks daya pembeda, dapat ditunjukkan seperti pada tabel berikut:

**Tabel 8.** Kategori Butir Soal berdasarkan Daya Pembeda

Kategori	Nomor Butir Soal	Jumlah	Ket.
<b>Baik</b>	1, 2, 3, 4, 9, 10, 23, 31, 32, 33, 34	11	terima
<b>Cukup</b>	5, 6, 7, 11, 13, 14, 17, 18, 19, 20, 24, 27, 28, 36, 37, 39, 40	17	terima
<b>Kurang baik</b>	8, 12, 15, 16, 21, 22, 25, 26, 29, 30, 35, 38	12	tolak

Berdasarkan data tabel 8 dapat diketahui bahwa terdapat 17 butir soal dengan kategori cukup baik dan 11 butir soal dengan kategori baik, ini berarti terdapat 28 butir soal memiliki kemampuan membedakan peserta didik yang mampu dan yang tidak mampu. Distraktor digunakan untuk mengecohkan peserta didik yang kurang mampu atau tidakb tahu, untuk dapat dibedakan dengan yang mampu. Distraktor yang baik adalah yang tidak dipilih oleh peserta didik yang mampu dan terpilih oleh peserta didik

yang kurang mampu. Dan apabila terpilih minimal 5% dari jumlah peserta. Berdasarkan analisis program ITEMAN, nilai point Biser pada kolom Alternative Statistics, keberfungsian pengecoh butir soal dapat ditunjukkan pada tabel berikut.

**Tabel 9.** Tabel Keberfungsian Pengecoh

Keberfungsian	Jumlah	Keterangan
Sangat Baik	26	Diterima
Baik	9	Diterima
Cukup Baik	2	Diterima
Kurang baik	3	Ditolak

Dari tabel 9 dapat diketahui bahwa distractor/pengecoh dari 26 butir soal masuk pada kriteria sangat baik, 9 butir soal pada kriteria baik dan 2 butir soal pada kriteria cukup baik berfungsi. Artinya pilihan lima jawaban itu mempunyai keberfungsian pengecoh yang baik. Reliabilitas menyokong terbentuknya validitas. Sebuah tes dimungkinkan reliabel, akan tetapi tidak valid. Sebaliknya tes dinyatakan valid umumnya reliabel. Reliabilitas yang tinggi disebabkan banyaknya butir soal yang digunakan guru. Semakin banyak butir soal semakin tinggi reliabilitas soal.

**Tabel 10.** Karakteristik Tes Skala

Kriteria	Nilai
Jumlah item	40
Jumlah peserta tes	143
Skor rata-rata	17,503
Varians	27,257
Standar deviasi	5,221
Kemiringan distribusi skor -	0,084
Puncak distribusi skor	-0,589
Skor terendah	7,000
Skor tertinggi	28,000
Median	18,000
Koefisien reliabilitas (Alpha)	0,709
Kesalahan tingkat pengukuran	2,818
Rata-rata tingkat kesulitan	0,438
Rata-rata daya beda semua item	0,278
Rata-rata daya beda korelasi biserial	0,362

Dari hasil analisis tes skala diatas menunjukkan bahwa tes ini berisi 40 butir soal dengan peserta tes 143. Tes memiliki reliabilitas yang baik, ditunjukkan oleh nilai Alpha sebesar 0,709.

**b. Pembahasan**

Hasil analisis instrumen tes bentuk obyektif ini menunjukkan nilai reliabilitas 0,709 indeks ini tergolong tinggi (baik).

**Tabel 11.** Kategori Butir Soal berdasarkan Validitas, daya Pembeda dan Tingkat Kesukaran

Butir	r	DP	TK	Kategori	Butir	r	DP	TK	Kategori
1	√	√	√	Baik	21	-	-	√	buang
2	√	√	√	Baik	22	√	-	√	revisi
3	√	√	√	Baik	23	√	√	√	Baik
4	√	√	√	Baik	24	√	√	√	Baik
5	√	√	-	revisi	25	-	-	-	buang
6	√	√	-	revisi	26	-	-	-	buang
7	√	√	-	revisi	27	√	√	√	Baik
8	-	-	-	buang	28	√	√	-	revisi

Butir	r	DP	TK	Kategori	Butir	r	DP	TK	Kategori
9	√	-	√	revisi	29	-	-	-	buang
10	√	√	√	Baik	30	-	-	-	buang
11	√	√	-	revisi	31	√	√	√	Baik
12	-	-	√	buang	32	√	√	√	Baik
13	-	√	-	buang	33	√	√	√	Baik
14	-	√	-	buang	34	√	√	√	Baik
15	-	-	-	buang	35	-	-	√	buang
16	-	-	√	buang	36	√	√	√	Baik
17	√	√	√	Baik	37	√	√	√	Baik
18	√	√	√	Baik	38	-	-	-	buang
19	√	√	√	Baik	39	√	√	√	Baik
20	√	√	√	Baik	40	√	√	√	Baik

**Keterangan:** √ memenuhi kriteria baik

Tabel 11, menunjukkan pengambilan keputusan butir soal yang dikembangkan baik, direvisi atau pun diterima. Pertimbangan tersebut diambil berdasarkan kriteria validitas butir, daya pembeda butir dan tingkat kesukaran butir. Berdasarkan pertimbangan tersebut, terdapat 20 butir soal memiliki kriteria baik (diterima), 7 butir soal direvisi dan 13 butir soal dibuang. Untuk soal yang diterima atau layak digunakan 50% berkriteria baik untuk indeks kesulitan dan daya beda soal. Dari hasil analisis tes kompetensi ini, memiliki nilai reliabilitas dan keberfungsian pengecoh dengan kriteria tinggi. Pengambilan keputusan butir soal diterima, direvisi atau ditolak ditunjukkan dengan pertimbangan sebagai berikut:

**Butir Soal diterima: Nomor 3**

3	0-3	0.552	0.614	0.488	A	0.133	-0.390	-0.247
					B	0.552	0.614	0.488 *
					C	0.140	-0.157	-0.101
					D	0.084	-0.305	-0.169
					E	0.091	-0.471	-0.268
					Other	0.000	-9.000	-9.000

Butir soal ini memiliki tingkat kesukaran butir sebesar 0,552 (sedang), yang berarti sebanyak 55,2% peserta tes dapat menjawab soal dengan benar. Daya pembedanya,  $R_{bis} = 0,614$  menunjukkan bahwa peserta tes yang pintar (memperoleh skor total tinggi) cenderung menjawab benar soal ini dan peserta tes yang kurang pintar cenderung menjawab salah soal ini. Alternatif jawaban B merupakan kunci jawaban, maka tanda positif ini menunjukkan bahwa kunci jawaban sudah berfungsi sebagaimana mestinya.

Ditinjau dari distribusi jawaban yaitu persentase peserta tes merespons alternatif jawaban, semua pengecoh tampak berfungsi dengan baik. Dapat dilihat pada kolom Prop Endorsing, tampak bahwa sebanyak 13,3% peserta tes merespon jawaban A, 14,0% peserta tes merespon jawaban C, 8,4% merespon jawaban D, dan 9,1% peserta tes merespon jawaban E. Untuk alternatif jawaban A, C, D dan E seluruhnya bernilai negatif artinya peserta tes yang pintar cenderung tidak memilih alternatif jawaban itu dan peserta tes yang kurang pintar cenderung memilih alternatif jawaban selain kunci. Kesimpulannya, butir soal nomor 3 ini sudah baik dan dapat digunakan.

**Butir Soal direvisi: Nomor 28**

28	0-28	0.294	0.513	0.388	A	0.224	-0.117	-0.084
					B	0.294	0.513	0.388 *
					C	0.154	-0.289	-0.190
					D	0.189	-0.082	-0.057
					E	0.140	-0.229	-0.147
					Other	0.000	-9.000	-9.000

Tingkat kesukaran soal ini adalah 0,294 dengan kategori sukar. Soal ini hanya dijawab benar oleh 29,4% peserta didik. Dilihat dari daya pembedanya Rbis 0,513 dan Rpbis 0,388 dan keduanya bernilai positif maka peserta didik yang pintar akan cenderung menjawab soal ini dengan benar sedangkan peserta didik yang kurang pintar akan cenderung menjawab soal ini dengan salah. Kunci jawaban adalah alternatif B, tanda positif ini menunjukkan bahwa kunci jawaban sudah berfungsi sebagaimana mestinya. Dari distribusi jawaban yaitu persentase peserta didik merespon alternatif jawaban, semua pengecoh sudah berfungsi dengan baik. Dilihat pada kolom Prop Endorsing tampak bahwa 22,4% siswa merespons alternatif jawaban A. 15,4% siswa merespons alternatif jawaban C. 18,9% merespons alternatif jawaban D. dan 14% siswa merespons alternatif jawaban E. Dilihat dari distribusi jawaban, pengecoh sudah berfungsi dengan baik (> 5%). Dapat disimpulkan bahwa soal ini termasuk dalam kategori diterima namun perlu direvisi.

**Butir Soal dibuang: Nomor 15**

15	0-15	0.119	-0.011	-0.006	A	0.238	-0.100	-0.073
					B	0.182	0.035	0.024
					C	0.203	0.078	0.055 ?
					D	0.119	-0.011	-0.006 *
					E	0.259	0.006	0.004
					Other	0.000	-9.000	-9.000

CHECK THE KEY  
D was specified, C works better

Tingkat kesukaran soal ini adalah 0,119 termasuk dalam kategori sukar, yaitu sebanyak 11,9% siswa dapat menjawab soal dengan benar. Dilihat dari daya pembedanya Rbis = -0,011 dan Rpbis = -0,006 dan keduanya bernilai negatif maka siswa yang kurang pintar (memperoleh skor rendah) cenderung menjawab benar soal ini dan siswa yang pintar cenderung menjawab salah soal ini. Karena alternatif jawaban D merupakan kunci, maka tanda negatif ini menunjukkan bahwa kunci jawaban belum berfungsi sebagaimana mestinya.

Ditinjau dari besarnya nilai tingkat kesukaran maupun daya pembeda, soal ini termasuk dalam kategori gagal(ditolak). Hal ini didukung pula oleh belum berfungsinya seluruh pengecoh dengan baik yang tampak dari distribusi jawaban yaitu sebanyak 23,8% siswa merespons alternatif jawaban A, 18,2% merespons alternatif jawaban B, 20,3% merespons alternatif jawaban C, dan 25,9% merespons alternatif jawaban E. Selain itu, ditinjau dari daya pembedanya, alternatif jawaban C dapat dikatakan belum berfungsi baik karena rbis atau rpbis untuk alternatif jawaban C ternyata bernilai positif dan lebih besar nilainya daripada alternatif jawaban D sebagai kuncinya. Karena alternatif jawaban D merupakan kunci, maka tanda negatif ini menunjukkan bahwa kunci jawaban belum berfungsi sebagaimana mestinya. Ditinjau dari besarnya nilai tingkat kesukaran (masuk kategori sukar) dan daya pembeda (buruk/ditolak). Kesimpulannya bahwa soal ini termasuk dalam kategori gagal/ditolak

**4. KESIMPULAN**

Dari hasil analisis dalam penelitian ini dapat disimpulkan instrumen ini menunjukkan valid yang dinilai oleh validator ahli diperoleh 85% dari segi materi memiliki kriteria baik,

82,5% dari segi konstruksi berkriteria baik dan 92,5% segi Bahasa berkriteria baik. Analisis empirik butir tes kompetensi Trigonometri, diperoleh 27 butir soal (67,5%) memiliki validitas yang baik, tes memiliki nilai reliabilitas yaitu 0,709 dengan kriteria baik. Diperoleh 26 butir soal (65%) dengan tingkat kesulitan butir soal sedang, 28 butir soal (70%) memiliki daya pembeda kategori baik, dan keberfungsian pengecoh 37 butir soal (92,5%) berkriteria baik.

## **5. REFERENSI**

- Azwar, S. (2012). Reliabilitas dan validitas edisi 4. *Yogyakarta: Pustaka Pelajar*.
- Hariyanto, & Basuki, I. (2014). Asesmen Pembelajaran Bandung. In *Bandung: PT. Remaja Rosdakarya Offset*.
- Kementerian Pendidikan dan Kebudayaan. (2013). Peraturan Menteri Pendidikan dan Kebudayaan RI Nomor 23 Tahun 2013 tentang Perubahan Atas Peraturan Menteri Pendidikan Nasional Nomor 15 Tahun 2010 tentang Standar Pelayanan Pendidikan Dasar Di Kabupaten/Kota. *Menteri Kesehatan Republik Indonesia Peraturan Menteri Kesehatan Republik Indonesia, 69(127)*.
- Lutfianto, Moch., & Sari, A. F. (2017). Respon Siswa terhadap Soal Matematika Mirip PISA dengan Konteks Berintegrasi Nilai Islam. *Jurnal Elemen, 3(2)*. <https://doi.org/10.29408/jel.v3i2.342>
- Mardapi, D. (2008). Teknik Penyusunan Instrumen Tes dan NonTes. In *Indonesian Digital Journal of Mathematics and Education*.
- Mardapi, D. (2017). Pengukuran, Penilaian, dan Evaluasi Pendidikan. *Academia Edu, 7(2)*.
- Moore, B., & Stanley, T. (2009). Critical Thinking and Formative Assessments. *Critical Thinking and Formative Assessments: Increasing the Rigor in Your Classroom*.
- Prestasi, M., & Akuntansi, P. B. (2014). Arikunto, Suharsimi. 2012. Dasar-Dasar Evaluasi Pendidikan. Jakarta: Bumi Aksara. *Universitas, 137*.
- Sari, E. F. P. (2015). Pengembangan soal matematika model PISA untuk mengetahui argumentasi siswa di Sekolah Menengah Pertama. *Jurnal Pendidikan Matematika, 9(2)*.
- Tshabalala, T., & Ncube, A. C. (2014). Teachers' perceptions on the causes and effects of highrate teacher-turnover of mathematics and science teachers in rural secondary schools - a case study of Nkayi District. *International Journal of Innovation and Applied Studies, 8(1)*.
- Arifin, Zainal. (2016). *Evaluasi Pembelajaran: Prinsip, Teknik, Prosedur Cet. 8*. Bandung: PT Remaja Rosdakarya
- An, D., Instrument, A., Higher, O. F., & Thinking, O. (n.d.). Pengembangan Instrumen Asesmen Higher Order Thinking Skill (HOTS) ... (Agus Budiman, Jailani) - 139, *1(November 2014)*, 139–151.
- Brookhart, S. (2010). *How to assess Higher Order Thinking Skills in your classroom*. Alexandria: VA: ASCD.
- Fanani, M. Z. (2018). Strategi pengembangan soal hots pada kurikulum 2013. *Edudeena, 2(1)*, 57–76.
- Gais, Zakkina, dkk. (2017). Analisis kemampuan siswa dalam menyelesaikan soal high. *Mosharafa, 6(2)*, 255–266.

- Makur, A. P. dkk. (2018). Kemampuan berpikir tingkat tinggi. *Jurnal Pendidikan Matematika*, 12(2), 5–24.
- Munayati, Zulva, Zulkardi, dan S. (2015). Kajian soal buku teks matematika kelas X Kurikulum 2013 menggunakan framework PISA. *Jurnal Pendidikan Matematika*, 9(2), 188–206.
- Nitko, A.J. dan Brookhart, S. M. (2011). *Educational assessment of students: Xth edition*. New Jersey: Prentice Hall Eglewood Cliffs.
- OECD. (2016). PISA 2015 Result in Focus. (Online), ([www.oecd.org](http://www.oecd.org)), diakses 9 Januari 2019.
- Sudijono, Anas. (2015). *Pengantar Evaluasi Pendidikan Ed. 1, Cet. 14*. Jakarta: PT RajaGrafindo Persada
- Sugiyono. (2015). *Metode Penelitian Pendidikan (Pendekatan Kuantitatif, Kualitatif, dan R&D) Cet. 21*. Bandung: CV Alfabeta
- Sukirman. (2012). *Pengembangan Sistem Evaluasi*. Yogyakarta: Insan Madani
- Shafizan Sabri, *Item Analysis of student Comprehensive Test for Reasearch Teaching Begiliner String Ensemble Using Model Based Teaching Among Music Students In Public Universities*”, *International Journal of Rducation and Reasearch*, 1 (12), hlm. 1.
- Simalango, M. M. (2018). Kesulitan siswa dalam menyelesaikan soal-soal PISA pada konten change and relationship level 4, 5, dan 6 di SMP N 1 Indralaya. *I2(1)*, 43–58.