

ANALISIS SOAL TES HASIL BELAJAR HIGH ORDER THINKING SKILLS (HOTS) MATEMATIKA MATERI PECAHAN UNTUK KELAS 5 SEKOLAH DASAR

Maria Agustina Amelia

Dosen Program Studi Pendidikan Guru Sekolah Dasar, FKIP, Universitas Sanata Dharma
Alamat korespondensi: Jl. Affandi Mrican Tromol Pos 29 Yogyakarta 55022
Email: amelia0284@gmail.com

ABSTRACT

This study conducted to analyze the quality of teacher's made mathematics achievement test. Research conducted using survey research methods and implemented at five elementary schools located in Bandung on 357 learners. The instrument used is a mathematics achievement test, subject of learning fraction. The test is multiple choice with 4 option. Based on the results the quality of the teacher's made test are: (1) 100% item test are valid (20 items), (2) The reliability of the test is high, (3) 3 items must be revised because it do not have good discriminations index, 17 items have good discrimination index, (4) 1 item (5%) categorized as easy, 15 items (75%) categorized as moderate, and 4 items (20%) categorized as difficult, (5) There 11 options that have to revised.

Keywords : test quality, fractions, reliability, discriminations indexes, item difficulty.

1. PENDAHULUAN

Kemampuan peserta didik dapat diketahui dari hasil pengujian. Pengujian dilakukan menggunakan alat ukur/instrumen berupa tes maupun non-tes. Alat ukur yang baik akan menghasilkan data yang baik. Guru dapat mengetahui kemampuan siswa dengan tepat jika alat ukur yang digunakan merupakan alat ukur yang baik. Arikunto (2008: 57) menyatakan bahwa suatu tes dapat dikatakan baik apabila memenuhi lima persyaratan, yaitu: validitas, reliabilitas, objektivitas, praktikabilitas dan ekonomis. Berdasar pendapat Arikunto di atas, kriteria minimal suatu alat ukur yang baik adalah alat ukur tersebut harus valid dan reliabel. Selain valid dan reliabel, tes dikatakan baik jika daya pembeda, tingkat kesulitan dan analisis pengecoh (soal pilihan ganda) juga baik.

Kualitas tes hasil belajar yang baik: (1) Validitas. Azwar (2009: 5) memaparkan bahwa validitas berasal dari kata *validity* yang mempunyai arti sejauh mana ketepatan dan kecermatan suatu alat ukur dalam melakukan fungsi ukurnya. Suatu tes atau instrumen pengukur dapat dikatakan mempunyai validitas yang tinggi apabila alat tersebut menjalankan fungsi ukurnya sesuai dengan maksud

dilakukannya pengukuran tersebut. (2) Reliabilitas. Masidjo (1995: 208) memaparkan bahwa reliabilitas adalah taraf kemampuan tes dalam menunjukkan konsistensi hasil pengukurannya yang diperlihatkan dalam taraf ketepatan dan ketelitian hasil. (3) Daya Pembeda. Masidjo (1995: 196) menyatakan bahwa daya pembeda adalah taraf jumlah jawaban benar siswa yang tergolong kelompok (pandai = *upper group*) berbeda dari siswa yang tergolong kelompok bawah (kurang pandai = *lower group*) untuk suatu item. (4) Tingkat kesukaran. Sulistyorini (2009: 176) menjelaskan bahwa tingkat kesulitan merupakan kemampuan siswa untuk menjawab soal dengan kriteria soal mudah, sedang, dan sukar. Widoyoko (2014: 165) mengungkapkan bahwa tingkat kesukaran yang baik pada suatu tes adalah 25% mudah, 50% sedang, dan 25% sukar. (5) Analisis Pengecoh. Purwanto (2009: 75) memaparkan bahwa pengecoh (*distractor*) adalah pilihan yang bukan merupakan kunci jawaban. Arikunto (2012: 234) memaparkan bahwa pengecoh dapat berfungsi dengan baik apabila pengecoh tersebut mempunyai daya tarik bagi peserta tes yang kurang memahami materi.

Alat ukur dalam pembelajaran dapat dilakukan dengan menggunakan instrumen tes maupun non-tes. Dalam ranah kognitif, dalam mengukur

kemampuan peserta didik digunakan instrumen tes. Mardapi (2008: 67) mengemukakan bahwa tes adalah beberapa pertanyaan yang membutuhkan jawaban atau beberapa pernyataan yang membutuhkan tanggapan untuk mengukur tingkat kemampuan suatu individu yang diberikan tes tersebut melalui jawaban terhadap beberapa pertanyaan atau tanggapan dari beberapa pernyataan. Widoyoko (2016: 57) mengemukakan bahwa bentuk tes dikategorikan menjadi dua, yaitu tes objektif dan tes subjektif. Tes objektif dalam hal ini memiliki pengertian yaitu bentuk tes yang pemeriksaan atau penskoran jawaban/respon peserta tes sepenuhnya dapat dilakukan secara objektif oleh korektor. Karena sifatnya yang objektif ini maka tidak perlu harus dilakukan oleh manusia. Pekerjaan tersebut dapat dilakukan oleh mesin, misalnya mesin *scanner*. Dengan demikian skor hasil tes dapat dilakukan secara objektif.

Salah satu kompetensi yang perlu dimiliki oleh guru adalah kemampuan menyusun soal sebagai instrumen tes peserta didik dengan baik dan analisis hasil tes tersebut. Namun dari hasil wawancara yang dilakukan peneliti pada beberapa guru didapat fakta bahwa masih ada guru yang melakukan konstruksi tes tanpa memperhatikan validitas, reliabilitas dan karakteristik butir soal. Menurut Mardapi (2008: 71) Tes bentuk pilihan ganda adalah tes yang jawabannya dapat diperoleh dengan memilih alternatif jawaban yang telah disediakan. Dalam tes pilihan ganda ini, bentuk tes terdiri atas: pernyataan (pokok soal), alternatif jawaban yang mencakup kunci jawaban dan pengecoh. Bentuk tes pilihan ganda ini banyak digunakan dalam ujian tengah semester, ujian akhir semester, ujian sekolah maupun ujian Nasional. Tes yang dikonstruksi tanpa memperhatikan kualitasnya dimungkinkan tidak dapat mengungkapkan kemampuan peserta didik dengan tepat. Tujuan penelitian ini adalah menganalisis kualitas tes hasil belajar matematika buatan guru mengenai materi pecahan untuk siswa kelas 5 SD?

2. METODE PENELITIAN

2.1 Variabel Penelitian

Variabel dalam penelitian ini adalah soal matematika pada materi pecahan yang diujikan pada siswa SD kelas IV

2.2 Populasi Penelitian

Populasi penelitian adalah seluruh siswa Sekolah Dasar Negeri yang bersekolah di 575 Sekolah Dasar dikota Bandung.

2.3 Sampel

Sampel dipilih dari siswa kelas IV yang bersekolah di 575 Sekolah Dasar Negeri di kota Bandung.

2.4 Teknik Sampling

Pengambilan sampel menggunakan teknik *random sampling*. Menurut Taniredja dan Mustafidah (2012: 35) "Teknik *random sampling* disebut juga acak, serampangan, tidak pandang bulu/tidak pilih kasih, objektif, sehingga seluruh elemen populasi mempunyai kesempatan untuk jadi sampel penelitian". Jadi teknik *randomsampling* dilakukan agar semua populasi subjek memiliki peluang yang sama untuk dipilih menjadi sampelsubjek. Alasan dipilih teknik pengambilan sampel *randomsampling* karena pada tingkat sekolah dasar, penerimaan peserta didik tidak didasarkan pada nilai tertentu.

2.5 Teknik Pengumpulan Data

Pengumpulan data mengenai kemampuan berpikir kritis materi pecahan pada siswa kelas IV menggunakan tes pilihan ganda dengan 4 pilihan jawab.

2.6 Instrumen Penelitian

Instrumen penelitian yang digunakan adalah 20 soal tes pilihan ganda mengenai materi pecahan. Materi pecahan pada tes hasil belajar terdiri dari beberapa indikator yaitu: 2.6.1. Mengenal arti pecahan sebagai perbandingan sebagian dengan keseluruhan, 2.6.2. Memahami berbagai bentuk pecahan, 2.6.3. Operasi penjumlahan dan pengurangan, 2.6.4. Menjumlah dan mengurangi berbagai bentuk pecahan, 2.6.5. Pemecahan masalah sehari-hari yang melibatkan penjumlahan dan pengurangan pecahan, 2.6.6. operasi perkalian dan pembagian.

2.7 Teknik Analisis Data

Teknik analisis data yang digunakan adalah analisis data secara deskriptif. Data tes hasil belajar akan diuji kualitasnya menggunakan *software* TAP (*Test Analysis Program*) versi 14.7.4. *Software* TAP

ini dapat diunduh secara cuma-cuma dan memiliki hak cipta atas nama Gordon P. Brooks. Software TAP dipilih untuk analisis soal tes karena penggunaannya relatif mudah, dan dalam satu kali input data dapat diperoleh hasil mengenai analisis validitas, reliabilitas, daya pembeda, tingkat kesukaran, dan pengecoh.

2.7.1 Analisis Validitas

Suatu tes dikatakan valid apabila tes tersebut benar-benar mengukur apa yang dimaksudkan untuk diukur. Pengertian validitas ini dapat dilihat dari dua segi, yaitu (1) bila dalam penyusunan suatu tes, penyusun berusaha memilih soal-soal yang secara logis diperkirakan mengukur apa yang mau diukur baik menurut pertimbangan sendiri maupun setelah bertukar pikiran dengan orang-orang lain atau bahkan ahli-ahli di bidang pengetahuan yang bersangkutan, (2) bila suatu tes dipergunakan, maka validitasnya bisa diukur dengan memperbandingkan hasil-hasil pengukurannya dengan hasil pengukuran-pengukuran lainnya. (Joni, 1984: 35). Teknik yang digunakan dalam penelitian ini adalah teknik korelasi biserial. Korelasi biserial digunakan untuk menghitung validitas setiap item. (Arikunto, 1986: 70). Rumus mencari korelasi biserial adalah sebagai berikut:

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

Hasil analisis validitas pada penelitian ini dapat dilihat melalui hasil *point biserial* pada TAP. Hasil *point biserial* dibandingkan dengan r_{tabel} dengan taraf signifikan 5% (Sugiyono, 2010: 258). Jika *point biserial* lebih besar dari r_{tabel} maka butir soal tersebut valid. Besar r_{tabel} untuk jumlah siswa sebanyak 357 siswa yaitu $\approx 0,1048$. Jika *point biserial* lebih besar dari 0,1048 maka butir soal valid.

2.7.2 Analisis Reliabilitas

Reliabilitas adalah salah satu hal yang penting dalam menganalisis setiap butir. Reliabilitas setiap butir suatu model tes adalah derajat tingkat kemantapan dan keterandalan tes itu secara keseluruhan. Tes yang reliabel selalu memberikan hasil yang sama bila dicobakan kepada kelompok yang sama dalam waktu yang berbeda. (Kartawidjaja, 1987: 125). Dalam penelitian ini, peneliti menggunakan metode belah dua atau *split-half method*. Pembelahan dilakukan dengan cara membagi dua sama banyak butir soal berdasar nomor soal genap dan ganjil yang selanjutnya disebut belahan ganjil-genap. Hasil reliabilitas yang dihitung menggunakan TAP

Tabel 1: Kriteria Reliabilitas

Koefisien Korelasi	Kualifikasi
0,91 – 1,00	Sangat Tinggi
0,71 – 0,90	Tinggi
0,41 – 0,70	Cukup
0,21 – 0,40	Rendah
Negatif – 0,20	Sangat Rendah

Rumus 1: Rumus korelasi biserial

$$r_{pbi} = \frac{Mp - Mt}{St} \sqrt{\frac{p}{q}}$$

Keterangan:

- r_{pbi} = koefisien korelasi biserial
 Mp = rerata skor dari subjek yang menjawab betul bagi item yang dicari validitasnya.
 Mt = rerata skor total
 St = standar deviasi dari skor total
 p = proporsi siswa yang menjawab benar

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

kemudian dianalisis menggunakan tabel kriteria reliabilitas menurut Masidjo (1995: 209).

2.7.3 Analisis Daya Pembeda

Daya pembeda soal adalah kemampuan suatu butir soal dapat membedakan antara siswa yang telah menguasai materi yang ditanyakan dan siswa yang belum menguasai materi yang diujikan. (Kusaeri dan Suprananto, 2012: 175). Daya pembeda dalam suatu tes bertujuan untuk membedakan siswa yang pandai dengan siswa yang kurang pandai. Dalam penelitian ini, peneliti menggunakan indeks daya pembeda menurut Kusaeri dan Suprananto (2012: 176) dengan perhitungan sebagai berikut:

Rumus 2: Indeks daya pembeda

$$DP = \frac{BA - B}{\frac{1}{2}N}$$

Keterangan:

D = indeks daya pembeda soal (Indeks Diskriminasi)
 BA = jumlah jawaban benar pada kelompok atas
 BB = jumlah jawaban benar pada kelompok bawah
 N = jumlah peserta tes

Kriteria daya pembeda atau indeks diskriminatif menurut Cracker & Algina (dalam Kusaeri dan Surapranata, 2012: 177) yang digunakan untuk menganalisis daya pembeda dalam penelitian ini adalah sebagai berikut:

Rumus 3: Indeks tingkat kesukaran

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

Keterangan:

P = indeks kesukaran
 B = banyaknya siswa yang menjawab soal itu dengan betul
 JS = jumlah seluruh siswa peserta tes

Indeks kesukaran menurut Arikunto (2012: 225) dapat dilihat pada Tabel 3. Distribusi tingkat kesukaran, digunakan pendapat Widoyoko (2014: 165) yaitu: 25% mudah, 50% sedang, dan 25% sukar

2.7.5 Analisis Pengecoh

Pengecoh (*distractor*) yang juga dikenal dengan istilah penyesat atau penggoda adalah pilihan

Tabel 2: Kriteria Daya Pembeda

No	Range Daya Pembeda	Kategori	Keputusan
1.	0,40-1,00	Sangat memuaskan	Diterima
2.	0,30-0,39	Memuaskan	Diterima
3.	0,20-0,29	Tidak memuaskan	Ditolak/direvisi
4.	0,00-0,19	Sangat tidak memuaskan	Direvisi total

2.7.4 Analisis Tingkat Kesukaran

Soal yang baik adalah soal yang tidak terlalu mudah atau tidak terlalu sukar. Besarnya indeks kesukaran antara 0,00 sampai dengan 1,0. Indeks kesukaran ini menunjukkan taraf kesukaran soal. Soal dengan indeks kesukaran 0,0 menunjukkan bahwa soal itu terlalu sukar, sebaliknya indeks 1,0 menunjukkan bahwa soalnya terlalu mudah. (Arikunto, 2012: 223)

jawaban yang bukan merupakan kunci jawaban. Pengecoh diadakan untuk menyesatkan siswa agar tidak memilih kunci jawaban. Pengecoh dikatakan berfungsi efektif apabila paling tidak ada siswayang terkecoh memilih. Pengecoh yang berdasarkan hasil uji coba tidak efektif direkomendasikan untuk diganti dengan pengecoh yang lebih menarik. (Purwanto, 2009: 108). Menurut Sudijono (2011:

Tabel 3: Indeks Kesukaran

No	Indeks kesukaran	Kategori
1.	0,00-0,30	Sukar
2.	0,31-0,70	Sedang
3.	0,71-1,00	Mudah

Didalam istilah evaluasi, indeks kesukaran ini diberi simbol P, dengan singkatan dari kata “proporsi”. Dengan demikian maka soal dengan P = 0,70 lebih mudah jika dibandingkan dengan P = 0,20. Sebaliknya soal dengan P = 0,30 lebih sukar daripada soal P = 0,80.

411) pengecoh dinyatakan telah dapat menjalankan fungsinya dengan baik apabila distraktor/pengecoh tersebut sekurang-kurangnya sudah dipilih oleh 5% dari seluruh peserta tes. Arikunto (2012: 234) mengatakan bahwa suatu distraktor dapat dikatakan berfungsi baik jika paling sedikit dipilih oleh 5% peserta tes.

3. PEMBAHASAN

Berdasar hasil tes matematika materi pecahan yang telah diberikan pada peserta didik tersebut, diperoleh hasil analisis mengenai kualitas tes yaitu: Analisis validitas, reliabilitas, daya pembeda, tingkat kesulitan, dan pengecoh.

3.1 Analisis Validitas

Analisis validitas dengan menggunakan TAP (*Test Analysis Program*) digunakan untuk mengetahui soal valid atau tidak validnya suatu butir soal. Hasil analisis dapat dilihat dengan menggunakan TAP (*Test Analysis Program*) pada bagian *point biserial*. *Point biserial* dibandingkan dengan r_{tabel} dengan taraf signifikansi 5%. Koefisien validitas pada r_{tabel} berdasarkan taraf signifikansi 5% untuk 357 siswa yaitu $\approx 0,1048$. Jika hasil *point biserial* lebih besar dari 0,1048 maka soal tersebut dinyatakan valid. Hasil analisis validitas soal dapat dilihat pada Tabel 4.

$\approx 0,1048$. Maka butir soal dikatakan sudah mengukur kemampuan peserta didik mengenai materi pecahan dengan tepat.

3.2 Analisis Reliabilitas

Hasil uji analisis reliabilitas soal tipe A menggunakan TAP (*Test Analysis Program*) dapat dilihat dari *Split-Half (odd/even) reliability* yaitu 0,711. Hasil uji reliabilitas pada soal menurut Masidjo (1995: 209) tergolong dalam kriteria “tinggi”. Jadi butir soal memiliki konsistensi yang tinggi dalam mengukur kemampuan peserta didik mengenai materi pecahan.

3.3 Analisis Daya Pembeda

Daya pembeda atau Indeks Diskriminasi (ID) soal menurut Cracker & Algina dalam Kusaeri dan Surapranata (2012: 177) dapat diterima jika daya pembeda 0,30-1,00, ditolak/direvisi jika daya pembeda 0,20-0,29 dan daya pembeda ditolak jika

Tabel 4. Hasil Analisis Validitas

No Item	Poin Biserial	r tabel	keterangan
Item 01	0,20	0,1048	Valid
Item 02	0,32	0,1048	Valid
Item 03	0,38	0,1048	Valid
Item 04	0,57	0,1048	Valid
Item 05	0,48	0,1048	Valid
Item 06	0,56	0,1048	Valid
Item 07	0,56	0,1048	Valid
Item 08	0,31	0,1048	Valid
Item 09	0,50	0,1048	Valid
Item 10	0,46	0,1048	Valid
Item 11	0,59	0,1048	Valid
Item 12	0,25	0,1048	Valid
Item 13	0,57	0,1048	Valid
Item 14	0,70	0,1048	Valid
Item 15	0,57	0,1048	Valid
Item 16	0,56	0,1048	Valid
Item 17	0,65	0,1048	Valid
Item 18	0,49	0,1048	Valid
Item 19	0,57	0,1048	Valid
Item 20	0,32	0,1048	Valid

Berdasar Tabel 4, dapat dilihat bahwa seluruh butir soal dinyatakan valid karena koefisien point biserial yang diperoleh lebih besar dari t tabel

0,00-0,19. Hasil analisis daya pembeda dapat dilihat pada Tabel 5.

Tabel 5: Hasil Analisis Daya Pembeda

No Item	Dicrimination Index	kriteria
Item 01	0,20	ditolak/direvisi
Item 02	0,32	diterima
Item 03	0,44	diterima
Item 04	0,68	diterima
Item 05	0,50	diterima
Item 06	0,70	diterima
Item 07	0,69	diterima
Item 08	0,24	ditolak/direvisi
Item 09	0,52	diterima
Item 10	0,52	diterima
Item 11	0,72	diterima
Item 12	0,28	ditolak/direvisi
Item 13	0,72	diterima
Item 14	0,86	diterima
Item 15	0,71	diterima
Item 16	0,63	diterima
Item 17	0,76	diterima
Item 18	0,50	diterima
Item 19	0,66	diterima
Item 20	0,36	diterima

Berdasar tabel 5 diperoleh hasil, dari 20 butir soal tes, terdapat 3 soal yang perlu direvisi yaitu butir soal nomor 1, 8, dan 12 dan 17 soal dapat diterima. Soal-soal yang perlu direvisi disebabkan karena memiliki indeks daya pembeda yang belum baik sehingga belum dapat membedakan peserta didik dengan kemampuan tinggi dengan peserta didik dengan kemampuan rendah. Soal-soal yang sudah dapat diterima memiliki indeks daya pembeda yang baik sehingga sudah dapat membedakan peserta didik dengan kemampuan tinggi dengan peserta didik dengan kemampuan rendah

3.4 Analisis Tingkat Kesukaran

Menurut Arikunto (2012: 225), secara umum tingkat kesukaran diklasifikasikan kedalam 3 kategori yaitu sukar, sedang, dan mudah. Kategori sukar berada pada rentang nilai 0,00-0,30, kategori sedang berada pada rentang nilai 0,32-0,75 dan kategori mudah berada pada rentang 0,71-1,00. Distribusi tingkat kesukaran, dari 20 yang digunakan adalah: 25% mudah (5 soal), 50% sedang (10 soal), dan 25% sukar (5 soal). Hasil analisis tingkat kesukaran dapat dilihat pada Tabel 6.

Tabel 6: Analisis Tingkat Kesukaran

No Item	Item Difficulty	kriteria
Item 01	0,77	sedang
Item 02	0,81	mudah
Item 03	0,73	sedang
Item 04	0,60	sedang
Item 05	0,27	sukar
Item 06	0,53	sedang
Item 07	0,59	sedang
Item 08	0,18	sukar
Item 09	0,31	sukar
Item 10	0,66	sedang

Tabel 6: Lanjutan

No Item	Item Difficulty	kriteria
Item 11	0,51	sedang
Item 12	0,20	sukar
Item 13	0,54	sedang
Item 14	0,59	sedang
Item 15	0,47	sedang
Item 16	0,59	sedang
Item 17	0,35	sedang
Item 18	0,40	sedang
Item 19	0,46	sedang
Item 20	0,50	sedang

Berdasar tabel 6. Didapatkan hasil 1 soal (5%) memiliki tingkat kesukaran kategori mudah, 15 soal (75%) memiliki tingkat kesukaran kategori sedang dan 4 soal (20%) yang memiliki tingkat kesukaran kategori sukar. Dapat dilihat bahwa distribusi soal belum memenuhi kriteria sebagai soal yang baik. Untuk mendapatkan distribusi soal yang baik, maka soal-soal dalam kategori mudah perlu ditambahkan 4 soal, dalam kategori sedang perlu dikurangi 5 soal, dan dalam kategori sukar perlu ditambah 1 soal.

3.5 Analisis Pengecoh

Sudijono (2011: 410) mengatakan bahwa pengecoh adalah alternatif yang bukan merupakan

jawaban yang digunakan agar peserta tes dapat tertarik dengan pengecoh jawaban tersebut. Semakin banyak peserta tes yang memilih pengecoh, maka pengecoh tersebut sudah menjalankan fungsinya. Sebaliknya apabila pengecoh yang dipasang tidak ada yang memilih maka pengecoh tersebut tidak berfungsi Arikunto (2012: 234) memaparkan sebuah distraktor dapat dikatakan berfungsi dengan baik jika paling sedikit dipilih oleh 5 % (0,05) peserta tes. Hasil analisis pengecoh dapat dilihat pada Tabel 7.

Berdasar Tabel 7, dapat dilihat bahwa secara keseluruhan ada 11 pengecoh tidak berfungsi. Pengecoh disebut tidak berfungsi jika dipilih kurang

Tabel 7: Hasil Analisis Pengecoh

No Item	Pilihan Jawaban				Keterangan
	A	B	C	D	
Item 01	0.157	0.020	Kunci Jawab	0.056	Pengecoh B tidak berfungsi
Item 02	0,31	0,067	0,053	Kunci Jawab	-
Item 03	0,160	Kunci Jawab	0,031	0,028	Pengecoh C dan D tidak berfungsi
Item 04	0,020	Kunci Jawab	0,048	0,294	Pengecoh A dan C tidak berfungsi
Item 05	0,042	Kunci Jawab	0,415	0,204	Pengecoh A tidak berfungsi
Item 06	0,034	0,134	0,165	Kunci Jawab	Pengecoh A tidak berfungsi
Item 07	0,174	Kunci Jawab	0,059	0,078	-
Item 08	0,244	0,084	0,370	Kunci Jawab	-
Item 09	0,098	Kunci Jawab	0,275	0,160	-
Item 10	0,081	0,034	Kunci Jawab	0,067	Pengecoh B tidak berfungsi
Item 11	0,286	0,034	Kunci Jawab	0,087	Pengecoh B tidak berfungsi
Item 12	Kunci Jawab	0,303	0,076	0,126	-
Item 13	0,232	0,042	Kunci Jawab	0,067	-
Item 14	0,174	0,064	Kunci Jawab	0,034	Pengecoh D tidak berfungsi
Item 15	0,059	0,199	Kunci Jawab	0,070	-

Tabel 7: Hasil Analisis Pengecoh

No Item	Pilihan Jawaban				Keterangan
	A	B	C	D	
Item 16	0,081	Kunci Jawab	0,067	0,090	-
Item 17	0,221	0,087	0,106	Kunci Jawab	-
Item 18	Kunci Jawab	0,076	0,143	0,106	-
Item 19	0,118	0,151	Kunci Jawab	0,028	Pengecoh D tidak berfungsi
Item 20	Kunci Jawab	0,185	0,076	0,056	-

dari 5% keseluruhan peserta tes. Pengecoh yang tidak berfungsi perlu direvisi kembali.

4. PENUTUP

4.1 Kesimpulan

Kualitas produk tes hasil belajar matematika materi pecahan dengan indikator-indikator yaitu: 2.6.1. Mengenal arti pecahan sebagai perbandingan sebagian dengan keseluruhan, 2.6.2. Memahami berbagai bentuk pecahan, 2.6.3. Operasi penjumlahan dan pengurangan, 2.6.4. Menjumlah dan mengurangi berbagai bentuk pecahan, 2.6.5. Pemecahan masalah sehari-hari yang melibatkan penjumlahan dan pengurangan pecahan, 2.6.6. operasi perkalian dan pembagian dapat disimpulkan sebagai berikut:

- 1) Hasil analisis validitas soal dengan taraf signifikan 5% untuk siswa berjumlah 357 diperoleh 20 soal (100%) valid.
- 2) Hasil uji reliabilitas soal diperoleh indeks reliabilitas dalam kriteria “tinggi”. Jadi butir soal memiliki konsistensi yang tinggi dalam mengukur kemampuan peserta didik mengenai materi pecahan.
- 3) Hasil uji daya pembeda pada soal terdapat terdapat 3 soal yang perlu direvisi karena belum dapat membedakan peserta didik berkemampuan tinggi dengan peserta didik berkemampuan rendah. 17 soal dapat diterimakarena sudah dapat membedakan

peserta didik berkemampuan tinggi dengan peserta didik berkemampuan rendah.

- 4) Hasil uji analisis tingkat kesukaran soal yaitu: 1 soal (5%) memiliki tingkat kesukaran kategori mudah, 15 soal (75%) memiliki tingkat kesukaran kategori sedang dan 4 soal (20%) yang memiliki tingkat kesukaran kategori sukar.
- 5) Hasil uji pengecoh pada soal secara keseluruhan ada 11 pengecoh tidak berfungsi. Pengecoh disebut tidak berfungsi jika dipilih kurang dari 5% keseluruhan peserta tes. Pengecoh yang tidak berfungsi perlu direvisi kembali.

4.2 Saran

Saran untuk peneliti selanjutnya yang akan mengembangkan produk tes hasil belajar matematika adalah sebagai berikut: Sebaiknya soal yang akan diberikan untuk mengukur kemampuan peserta didik perlu diuji kualitasnya terlebih dahulu. Dengan melakukan analisis butir soal, guru dapat mengetahui kualitas soal yang dibuat. Soal yang berkualitas baik akan dapat mengukur kemampuan peserta didik secara tepat. Namun jika kualitas butir soal belum baik, dimungkinkan kemampuan peserta didik tidak diukur secara tepat dan soal tersebut perlu diperbaiki. Untuk menganalisis soal, dapat digunakan software untuk memudahkan kerja guru, saat ini sudah banyak software yang dapat digunakan untuk melakukan analisis butir soal yang mudah digunakan dan dapat diperoleh secara Cuma-Cuma.

DAFTAR PUSTAKA

- Arikunto, Suharsimi. 2008. *Dasar-Dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.
- Balitbang. 2007. *Panduan Penulisan Soal Pilihan Ganda*. Depdiknas: Pusat Penilaian Pendidikan.

Kartawidjaja, Eddy Soewardi. 1987. *Pengukuran Dan Hasil Evaluasi Belajar*. Bandung: C.V. Sinar Baru.

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

- Masidjo, Ign. 1995. *Penilaian Pencapaian Hasil Belajar Siswa di Sekolah*. Yogyakarta: Kanisius.
- Purwanto. 2009. *Evaluasi Hasil Belajar*. Yogyakarta: Pustaka Pelajar.
- Putri, Ida Ayu Putu Giri, dkk. 2013. *Pengembangan Tes Matematika Berbasis SK/KD dengan Teknik Concurent pada Siswa Kelas VI di SD Negeri Se-Kecamatan Gianyar*. Jurnal Penelitian Pasca Sarjana Undiksha Vol. 3 Tahun 2013.
- Sulistiyorini. 2009. *Evaluasi Pendidikan*. Yogyakarta: Teras.
- Widoyoko, S.E. 2014. *Penilaian Hasil Pembelajaran di Sekolah*. Yogyakarta: Pustaka Pelajar.
- Widoyoko, Eko Putro. 2015. *Teknik Penyusunan Instrumen Penelitian*. Yogyakarta: Pustaka Pelajar