

**PENGEMBANGAN INSTRUMEN PENILAIAN UNTUK MENGUKUR
CRITICAL THINKING SKILL MATERI RELASI DAN FUNGSI**

**THE DEVELOPMENT OF ASSESSMENT INSTRUMENTS TO
MEASURE CRITICAL THINKING SKILL IN RELATION AND
FUNCTIONS MATERIALS**

Erni Ekafitria Bahar¹, Rosmiyati²
erniekafitria@unismuh.ac.id

Pendidikan Matematika, FKIP Universitas Muhammadiyah Makassar

Abstrak

Pengembangan instrumen penilaian untuk mengetahui *critical thinking skill* siswa kelas VIII MTs pada materi relasi dan fungsi serta memberikan gambaran tentang kualitas instrumen (validitas, reliabilitas, tingkat kesukaran, dan daya pembeda) menjadi tujuan utama penelitian ini. Model pengembangan mengacu pada *formative research Tessmer* dengan tahapan yaitu: (1) *preliminary*, (2) *self evaluation*, (3) *prototyping (expert reviews one-to-one, dan small group)*, dan (4) *field test*. Kisi-kisi (peta distribusi soal), soal tes performa (unjuk kerja) dan esai (uraian) *critical thinking skill*, serta rubrik penilaian merupakan instrument penilaian yang dikembangkan dengan karakteristik yang mengacu kepada indikator *critical thinking skill* menurut Facione (*interpretation, analysis, evaluation dan inference*). Hasil penelitian menunjukkan: (a) tes unjuk kerja dan tes uraian memiliki validitas tinggi; (b) kategori reliabilitas adalah baik; (c) Soal nomor 1 tes unjuk kerja dan uraian, tingkat kesulitannya mudah dan nomor 2 tingkat kesulitannya sedang; (d) Daya pembeda soal nomor 1 pada tes unjuk kerja dan uraian berada pada kategori memuaskan dan soal nomor 2 kategorinya sangat memuaskan. Hasil penelitian menunjukkan bahwa kategori instrumen penelitian yang dikembangkan layak digunakan

Kata kunci: instrumen penilaian, *critical thinking skill*

Abstract

The development of critical thinking skills assessment instruments for students of class VIII MTs on the subject of relations and functions as well as describing the quality of the instrument (validity, reliability, level of difficulty, discriminatory power) are the main objectives of this research. The development model refers to Tessmer's formative research with stages: (1) preliminary, (2) self-evaluation, (3) prototyping (expert reviews one-to-one, and small group), and (4) field test. The grid, performance test questions and critical thinking skills descriptions, as well as assessment rubrics are assessment instruments developed with characteristics that refer to the critical thinking skill indicators according to Facione (interpretation, analysis, evaluation and inference). The results showed: (a) the performance test and the description test had high validity; (b) the reliability category of the performance test and description test items is good; (c) Question number 1 on the performance test and description has an easy level of difficulty and item number 2 has a moderate level of difficulty; (d) The distinguishing power of item number 1 on the performance test and description is in the satisfactory category and item number 2 is in the very satisfactory category. The results of the study indicate that the category of research instruments developed for relation and function material has a quality that is feasible to use.

Keywords: *assessment instruments, critical thinking skill*

PENDAHULUAN

Kurikulum yang berubah menjadi Kurikulum 2013 merupakan suatu cara untuk mengembangkan taraf Pendidikan yang ada di Indonesia yang mempunyai ciri 4C yaitu *Communication* (komunikasi), *Collaboration* (kolaborasi), *Creativity* (Kreativitas) dan *Innovation* (inovasi), *Critical Thinking* (berpikir kritis) dan *Problem Solving* (pemecahan masalah). Ashim et al., (2019) menjabarkan keempatnya sebagai berikut: (1) Komunikasi merupakan aktivitas menyampaikan keterangan baik secara tertulis maupun melalui ucapan. (2) Kolaborasi merupakan sikap mampu untuk bekerja bersama, menyesuaikan diri dalam keadaan apapun dan bertanggung jawab; menghasilkan sesuatu ketika bertindak; bersikap tenggang rasa sesuai porsinya; menghargai beragam opini. (3) Kreativitas dan inovasi adalah kesanggupan untuk menciptakan, melakukan, memberikan beragam gagasan yang inovatif, dan terbuka dalam menerima hal-hal yang baru. (4) Berpikir kritis dan pemecahan masalah merupakan kesanggupan dalam mengerti permasalahan sulit, mengaitkannya dengan pengetahuan lain sehingga diperoleh beragam sudut pandang, serta memperoleh penyelesaian dari masalah tersebut.

Fokus utama penilaian yang dilakukan selama ini di MTs Izzatul Ma'arif Tappina hanya berupa pengulangan materi dengan soal-soal yang masih merupakan kategori soal rutin serta belum mendorong siswa untuk mampu berpikir secara analitis dan kritis. Latihan tugas yang diberikan di sekolah dan di rumah, khususnya untuk mata pelajaran Matematika belum mampu untuk mengukur kecakapan siswa dalam berpikir secara kritis karena tugas tersebut hampir sama dengan contoh yang telah dijelaskan serta tidak mempertimbangkan unsur berpikir kritis. Informasi ini diperoleh dari hasil pengamatan serta penjelasan dari guru matematika yang dilakukan pada bulan Mei tahun 2020.

Berdasarkan pemaparan tersebut, maka pada mata pelajaran matematika perlu dikembangkan suatu penilaian yang berpusat menggali keterampilan siswa untuk berpikir kritis. Sehingga penelitian ini memiliki tujuan utama yaitu mengajukan alternatif penyelesaian bagi guru mata pelajaran di sekolah tersebut untuk menilai keterampilan siswa dalam berpikir kritis. Penilaian kemampuan berpikir kritis ini perlu dikembangkan untuk melihat sejauh mana siswa menggunakan penalaran kognitifnya dan mengembangkan idenya untuk memperoleh penyelesaian atau solusi terbaik dalam menyelesaikan permasalahan yang terdapat di dalam pelajarannya.

Siswono (2018) mengemukakan bahwa berpikir kritis merupakan suatu mekanisme untuk memakai kecakapan berpikir secara efektif yang dapat menolong seseorang menciptakan sesuatu, menilai, dan menerapkan keputusan berdasarkan apa yang telah dilakukan. Dewey (Nur Kholidah,

2019) mengemukakan berpikir kritis merupakan peninjauan secara sungguh-sungguh, terus menerus, dan cermat serta tidak langsung menerima begitu saja ketika memperoleh suatu pengetahuan melainkan menganalisa terlebih dahulu dengan memberikan pendapat-pendapat yang menyokong. Terdapat beberapa aspek keterampilan berpikir kritis berdasarkan pendapat Facione, yaitu: (a) Interpretasi adalah saat siswa memiliki kemampuan pemahaman untuk mengungkapkan arti dari bermacam pengetahuan, suasana, informasi, kejadian, analisa, kebiasaan, norma, metode, atau kategori; (b) Analisis merupakan kondisi ketika siswa memiliki kemampuan untuk mengidentifikasi korelasi antara yang ditanyakan, pernyataan, konsep yang dimaksudkan untuk menyimpulkan penilaian, alasan, atau penyelesaian soal; (c) Evaluasi merupakan kecakapan untuk meninjau kredibilitas data-data yang diperoleh; dan (d) Inferensi merupakan kecakapan seseorang untuk membuat konklusi dari permasalahan yang telah diselesaikan berdasarkan informasi yang diperoleh.

Langrehr (2006) mengemukakan bahwa hal-hal yang terdapat dalam tes yang digunakan untuk melatih berpikir kritis adalah: (1) mengetahui konsekuensi dari keputusan yang diambil; (2) menentukan opini terhadap satu *statement*; (3) membuat berbagai pertanyaan masalah; (4) mengemukakan sudut pandang lain; (5) mengemukakan alasan atau sebab dari suatu permasalahan; (6) menetapkan alasan-alasan atau hal-hal pendukung untuk pengambilan keputusan. Berbagai macam jenis soal dapat digunakan untuk mengukur *critical thinking skill*, diantaranya pilihan ganda disertai alasan, soal uraian (esai), serta soal performa (unjuk kerja). Soal pilihan ganda beralasan menuntut siswa untuk memberikan alasan yang menjadi dasar untuk memilih jawaban. Soal esai mengarahkan siswa untuk menggabungkan semua gagasan atau ide yang dimiliki sehingga mampu memberikan jawaban sesuai dengan kalimatnya sendiri. Soal unjuk kerja digunakan untuk mengukur sejauh mana kompetensi yang diajarkan telah tercapai.

Instrumen penilaian yang dikembangkan memuat materi pada semester gasal kelas VIII SMP/ sederajat yaitu relasi dan fungsi. Materi relasi dan fungsi adalah materi yang dianggap sangat penting pada matematika karena merupakan salah satu materi fundamental dan menjadi prasyarat untuk mempelajari materi selanjutnya sebagaimana dikemukakan oleh Sumarsih (Muthmainnah et al., 2019). Urgennya materi ini belum berbanding lurus dengan tingkat penguasaan siswa. Beberapa siswa masih kurang memahami konsep pada materi tersebut. Hal ini karena masalah yang diberikan untuk melatih dan menilai siswa berpikir kritis terkait materi ini masih terbilang kurang yang berakibat pada kesulitan siswa memahami

konsep relasi dan fungsi. Sehingga pengembangan instrumen penilaian berkualitas perlu dilakukan.

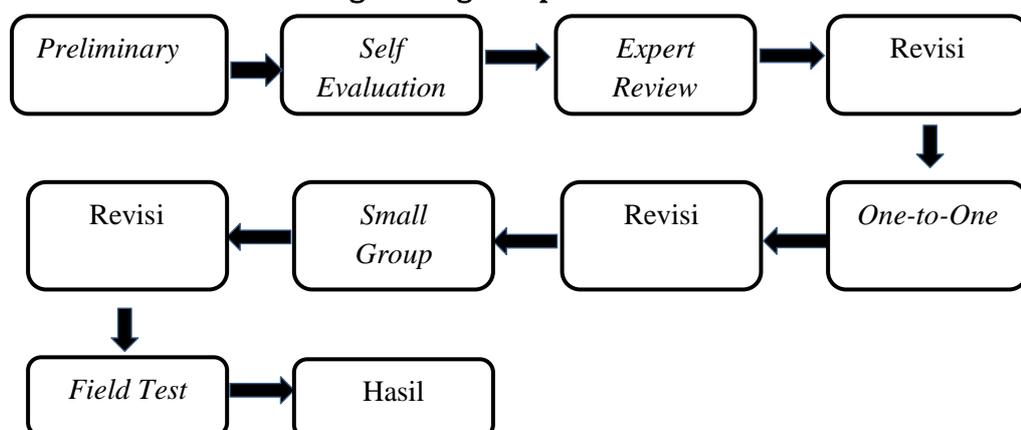
Permendikbud No 104 Tahun 2014, menyebutkan bahwa ketercapaian siswa selama proses pembelajaran berlangsung diukur dengan menggunakan instrument penilaian. Instrumen yang terlebih dahulu diujicobakan dan mempunyai kualitas yang baik merupakan instrument yang layak untuk mengukur ketercapaian siswa dalam pembelajaran. Instrumen dikatakan layak apabila mampu memberikan data yang tepat yaitu yang syarat validitasnya terpenuhi, reliabilitas, tingkat kesulitan serta daya pembeda yang baik.

Berdasarkan hal tersebut, maka penelitian ini akan memberikan gambaran mengenai ciri-ciri instrument penilaian untuk menilai keterampilan berpikir kritis siswa pada materi relasi dan fungsi dan menggambarkan validitas, konsistensi (reliabilitas), tingkat kesulitan, dan daya pembeda. Pengembangan alat penilaian menggunakan indikator menurut Facione yaitu interpretasi, analisis, evaluasi, serta inferensi dalam bentuk soal uraian dan soal unjuk kerja. Hal ini karena kecakapan siswa berpikir kritis dalam menentukan penyelesaian masalah, dapat diukur dengan tepat melalui kedua bentuk soal itu

METODE

Penelitian ini adalah penelitian pengembangan dengan model *formative research* Tessmer yang meliputi beberapa tahap, yaitu tahap *preliminary*, tahap *self evaluation* (analisis kurikulum, analisis materi, analisis siswa, dan desain), tahap *prototyping* (*expert review*, *one - to - one*, dan *small group*) dan tahap *field test*. Subjek dalam penelitian ini adalah siswa kelas VIII MTs Izzatul Ma'arif Tappina. Prosedur penelitian secara sistematis digambarkan sebagai berikut.

Gambar 1 Alur Model Pengembangan Tipe *Formative Research* Tessmer



Data penelitian dikumpulkan dengan menggunakan instrumen berupa catatan lapangan dan lembar validasi instrumen penilaian *critical thinking*

skill. Catatan lapangan memuat setiap langkah-langkah pengembangan produk. Informasi kevalidan produk yang dikembangkan diperoleh dari hasil lembar validasi. Proses validasi dilaksanakan oleh pakar dalam bidang pengembangan instrumen penilaian *critical thinking skill*.

Arikunto (2012) mengemukakan suatu tes disebut baik jika memenuhi kelima syarat berikut:

- a. Validitas yaitu jika alat ukur tersebut dapat menilai apa yang hendak dinilai dengan tepat.
- b. Reliabilitas, berasal dari bahasa Inggris yaitu *reliable* yang artinya dapat dipercaya. Tes dikatakan reliabel jika diteskan secara berulang akan memperoleh hasil yang sama atau tetap. Dengan kata lain hasil yang diperoleh dari tes tersebut memperlihatkan suatu ketetapan. Kriteria interpretasi koefisien reliabilitas dipaparkan pada table 1.

c.

Tabel 1. Kriteria Pengujian Reliabilitas Instrumen

Reliabilitas Soal	Keterangan
$0,80 < r_{11} < 1,00$	Sangat tinggi
$0,60 < r_{11} \leq 0,80$	Tinggi
$0,40 < r_{11} \leq 0,60$	Sedang
$0,20 < r_{11} \leq 0,40$	Rendah
$r_{11} \leq 0,20$	Sangat Rendah

- d. Tingkat Kesukaran. Alat penilaian yang baik memuat item soal yang tidak terlalu gampang dan tidak terlalu sulit. Item soal yang terlalu gampang tidak dapat mendorong siswa untuk meningkatkan upayanya dalam memecahkan masalah. Item soal yang terlalu sulit dapat menjadikan siswa berputus asa dan menyerah sehingga semangat untuk meneruskan mengerjakan soal tidak ada. Klasifikasi indeks kesukaran dapat dilihat pada tabel 2 berikut.

Tabel 2. Klasifikasi Indeks Kesukaran

Indeks Kesukaran (P)	Klasifikasi	Keterangan
0,00 – 0,30	Soal Sukar	Ditolak/direvisi
0,31 – 0,70	Soal Sedang	Diterima
0,71 – 1,00	Soal Mudah	Ditolak/direvisi

- e. Daya Pembeda merupakan kesanggupan instrumen tersebut membedakan/ mengelompokkan siswa yang memiliki kemampuan tinggi, sedang, dan rendah. Kriteria daya pembeda soal dapat dilihat pada tabel 3 berikut.

Tabel 3. Kriteria Daya Pembeda

Range Daya Pembeda	Klasifikasi	Keterangan
0,00 – 0,19	Sangat tidak memuaskan	Direvisi total
0,20 – 0,29	Tidak memuaskan	Ditolak/direvisi
0,30 – 0,39	Memuaskan	Diterima
0,40 – 1,00	Sangat Memuaskan	Diterima

Teknik analisis data pada proses pengembangan yaitu: (1) mengubah hasil lembar catatan lapangan kedalam bentuk deskriptif kualitatif. (2) Analisis nilai akhir peserta didik diperoleh dengan rumus jumlah nilai yang didapatkan dibagi nilai maksimum, selanjutnya dikalikan 100. (3) analisis data validasi instrumen penilaian dilakukan setelah mencari rata-rata validasi masing-masing soal (V_a) berdasarkan penilaian validator; (4) Analisis hasil uji coba instrumen penilaian yang meliputi: (a) uji reliabilitas untuk tiap item soal konsisten atau tidak. Kategori instrumen penilaian reliabel adalah reliabilitasnya minimal 0,60; (b) tingkat kesukaran, diperoleh dengan cara mencari rerata dari setiap butir soal (jumlah nilai peserta tes pada item soal tertentu dibagi banyaknya siswa yang melaksanakan tes) kemudian mencari tingkat kesukaran soal dengan membagi nilai rerata setiap butir soal dengan skor maksimum. Soal tes dapat diterima jika indeks kesukarannya antara 0,31 – 0,70; (c) daya pembeda yang diperoleh dengan cara mengurutkan nilai siswa dari tinggi sampai rendah, membagi tiga kelompok yaitu kelompok atas untuk nilai tinggi, kelompok bawah untuk yang rendah, dan kelompok sedang, kemudian mencari rerata dari kelompok atas dan kelompok bawah, selanjutnya mencari daya pembeda setiap soal dengan mengurangkan rerata kelompok atas dan bawah lalu membaginya dengan skor maksimum soal. Soal tes dapat diterima jika memiliki daya pembeda minimal 0,30.

HASIL DAN PEMBAHASAN

Peneliti mengembangkan instrumen penilaian dengan indikator menurut Facione. Indikator inferensi tercantum pada butir pertama soal unjuk kerja serta indikator interpretasi tercantum pada nomor 2. Soal uraian nomor 1 dikembangkan berdasarkan indikator evaluasi dan nomor 2 mewakili indikator analisis.

Pengembangan instrumen penilaian *critical thinking skill* dalam hal ini menggunakan model *formative research* Tessmer yang terdiri dari (1) *preliminary*, yaitu melakukan pengkajian terhadap referensi pendukung untuk mengembangkan instrumen penilaian; (2) *self evaluation*, yaitu melakukan analisis terhadap kurikulum, materi, siswa serta merancang *prototype* I; (3) *prototyping*, dilakukan dengan memvalidasi instrumen dengan pakar yang ahli dalam pengembangan instrumen penilaian (*expert review*), mengujicobakan instrumen penilaian pada kelompok kecil non subjek penelitian (*one - to - one* dan *small group*); (4) *field test*, mengujicobakan instrumen penilaian kepada subjek penelitian.

Tahap *preliminary*, dilakukan pengkajian berbagai referensi yang terkait pengembangan instrumen penilaian untuk mengukur kecakapan siswa berpikir kritis khususnya untuk pokok bahasan relasi dan fungsi.

Tahap *self evaluation*, didesain prototype I dengan melakukan analisis kurikulum untuk pemilihan materi, analisis materi untuk menentukan kompetensi dasar, serta analisis siswa untuk mengetahui karakteristik siswa terlebih dahulu. Pemilihan materi relasi dan fungsi berdasarkan hasil analisis kurikulum dan dijadikan pedoman dalam pembuatan instrumen penilaian. Berdasarkan analisis materi, peneliti menetapkan kompetensi dasar yang akan diukur adalah kompetensi dasar 3.3 dan 4.3. Tingkat kemampuan siswa beragam, beberapa siswa tingkat kemampuannya tinggi, beberapa yang tingkat kemampuannya sedang, dan sebagian tingkat kemampuannya rendah. Selama pandemi, proses pembelajaran dilaksanakan secara daring. Tetapi penelitian dilakukan secara langsung dengan mematuhi protokol kesehatan dengan subjek penelitian sebanyak 30 siswa. Tahap awal perancangan desain prototype I adalah penyusunan kisi-kisi (peta distribusi soal) sesuai dengan kompetensi dasar yang telah ditetapkan pada analisis materi. Rancangan instrumen berupa tes dalam bentuk uraian dan dalam bentuk unjuk kerja. Masing-masing bentuk soal tes memuat 2 item soal yang masing-masing mewakili satu aspek *critical thinking skill* menurut Facione sebagai karakteristik dari instrumen penilaian tersebut.

Tahap *prototyping* diawali dengan tahap *expert review*, yaitu melakukan validasi terhadap instrumen yang telah didesain. Tiga orang validator yang diantaranya dua orang dosen dan seorang guru mata pelajaran matematika pada sekolah tersebut melakukan validasi. Hasil validasi menunjukkan ada beberapa hal yang perlu direvisi. Uji coba *prototype* I dilakukan pada tahap *One - to - one* terhadap tiga siswa yang bukan merupakan subjek penelitian dan merupakan wakil dari masing-masing tingkat kemampuan. *Prototype* II diperoleh melalui pertimbangan yang didasarkan pada pendapat ketiga siswa tersebut. Dari beberapa komentar siswa yang diperoleh peneliti selama ujicoba dan saran dari validator dilakukan beberapa revisi misalnya pada kisi-kisi tes perlu dipertimbangkan alokasi waktu yang digunakan, dan soal yang termasuk kategori mudah diubah menjadi kategori sedang. Hal tersebut sejalan dengan pendapat Arikunto (2012) yang menjelaskan tes yang memiliki tingkat kesulitan tidak terlalu gampang dan tidak terlalu sukar merupakan salah satu syarat tes yang baik.

Tahap *small group*, melakukan uji coba *prototype* II pada kelompok kecil yaitu terhadap 6 orang yang bukan merupakan subjek. Keenam siswa ini dipilih berdasarkan pertimbangan dari guru dan tingkat kemampuannya dimana 2 orang siswa mewakili tiap tingkatan. Hasil tahapan ini menunjukkan bahwa reliabilitas soal unjuk kerja sangat tinggi (0,83), demikian pula reliabilitas soal tes uraian termasuk kategori tinggi (0,94).

Tingkat kesukaraan dan daya pembeda (indeks diskriminasi) dapat dilihat pada tabel 4 berikut.

Tabel 4. Analisis Tingkat Kesukaran dan Daya Pembeda

Instrumen Tes	Unjuk Kerja		Uraian	
	1 (Infrensi)	2 (Interpretasi)	1 (Evaluasi)	2 (Analisis)
Tingkat Kesukaran	0,58	0,5	0,54	0,41
Kriteria	Sedang	Sedang	Sedang	Sedang
Daya Pembeda	0,5	0,5	0,63	0,5
Kriteria	Sangat Memuaskan	Sangat Memuaskan	Sangat Memuaskan	Sangat Memuaskan

Berdasarkan hasil uji coba di atas maka dapat dinyatakan bahwa ujicoba prototype pada kelompok yang lebih besar dapat dilakukan tanpa ada revisi. Berdasarkan hal tersebut maka dibuatlah prototype III yang selanjutnya akan dipakai pada tahap berikutnya.

Tahap *field test*, dilaksanakan dengan melakukan ujicoba pada subjek penelitian dengan menggunakan prototype yang sebelumnya telah divalidasi dan diperbaiki. Hasil ujicoba selanjutnya dianalisis untuk memperoleh informasi mengenai tingkat kecakapan siswa dalam berpikir kritis dan juga kualitas tes (validitas, reliabilitas, tingkat kesukaran dan daya pembeda). Kecakapan tingkat berpikir kritis siswa dapat dilihat pada tabel 5 berikut.

Tabel 5. Dstribusi Frekuensi dan Presentase Skor Hasil Tes Kemampuan Berpikir Kritis Siswa

Nilai Siswa	Unjuk Kerja		Uraian		Kategori
	Frekuensi	Presentase	Frekuensi	Presentase	
$80 < x \leq 100$	16	53,33	9	30	Sangat Baik
$60 < x \leq 80$	9	30	16	53,33	Baik
$40 < x \leq 60$	3	10	2	6,67	Cukup
$20 < x \leq 40$	2	6,67	2	6,67	Kurang
$0 < x \leq 20$	0	0	1	3,33	Sangat Kurang

Hasil analisis data pada tahap *field test* menunjukkan bahwa pada umumnya kecakapan berpikir kritis siswa adalah baik, tetapi beberapa siswa terbilang masih kategori kurang kemampuan berpikir kritisnya. Sehingga proses belajar mengajar yang dilaksanakan oleh guru harus merangsang dan mendorong serta membiasakan siswa berpikir kritis. Hal tersebut selaras dengan pandangan Susanti (2019) yang mengemukakan bahwa ada baiknya proses belajar mengajar di sekolah dapat membiasakan siswa untuk menggali kecakapan dan keterampilan dalam menyelidiki, mengolah, dan mengevaluasi beragam informasi secara analitis. Ketika peserta didik diperhadapkan pada suatu masalah, terkadang kemampuan berpikir kritisnya akan muncul. Sehingga instrumen yang dikembangkan ini diharapkan dapat mengasah kecakapan berpikir kritis yang dimiliki siswa dalam mencari solusi dari suatu masalah.

Berdasarkan data dari tahap *field test* menunjukkan bahwa nilai validitas masing-masing 4,12 untuk soal unjuk kerja dan 4,18 untuk soal tes uraian. Ini berarti instrumen valid untuk diterapkan. Nilai reliabilitas soal unjuk kerja termasuk kategori tinggi dengan nilai reliabilitas sebesar 0,71. Demikian halnya dengan soal tes uraian yang juga menunjukkan nilai reliabilitas 0,68 dan termasuk kategori tinggi. Berdasarkan hal tersebut maka instrumen penilaian tidak ada revisi. Hasilnya dapat dilihat pada table 6,

Tabel 6. Analisis Tingkat Kesukaran dan Daya Pembeda Tes *Critical Thinking Skill*

Instrumen Tes	Unjuk Kerja		Uraian	
Nomor Soal	1	2	1	2
Tingkat Kesukaran	0,84	0,68	0,70	0,65
Kriteria	Mudah	Sedang	Sedang	Sedang
Daya Pembeda	0,37	0,4	0,37	0,4
Kriteria	Memuaskan	Sangat Memuaskan	Memuaskan	Sangat Memuaskan

Kualitas instrumen yang telah dikembangkan dilihat berdasarkan nilai validitas, reliabilitas, tingkat kesulitan soal, serta daya pembeda. Hasil analisis menunjukkan bahwa indikator validitas telah terpenuhi. Instrumen penilaian yang reliabel yaitu yang mampu menilai apa yang akan dinilai relative tetap secara konsisten. Purwanto (2011) mengemukakan bahwa instrumen penilaian dikatakan reliabel apabila koefisien reliabilitas yang didapatkan melebihi batas minimal dari nilai reliabilitas. Nilai koefisien batas reliabel yang menjadi dasar pengambilan keputusan adalah sebesar 0,60. Hal ini berarti instrumen yang dikembangkan memenuhi indikator reliabel dengan nilai reliabilitas untuk soal unjuk kerja yaitu 0,71 dan tes uraian 0,68.

Kriteria tingkat kesukaran tes yang bagus adalah antara 0,31 – 0,70 atau tes dengan kategori “sedang”. Item dengan kategori mudah membuat peserta didik tidak berusaha meningkatkan upayanya dalam memecahkan masalah tersebut, dan item yang sangat sulit akan membuat siswa pesimis serta tidak bersemangat mengerjakan soal sebab di luar kemampuannya (Arikunto, 2012). Analisis kesukaran soal menunjukkan terdapat satu soal pada tes unjuk kerja yang termasuk soal yang mudah. Walaupun demikian, bukan berarti soal tersebut tidak bisa digunakan sebab soal tersebut dapat digunakan untuk membangkitkan semangat belajar kepada siswa yang termasuk kategori lemah kemampuan berpikir kritisnya. Apabila siswa dengan kemampuan yang lemah mampu menjawab soal tersebut maka akan menjadi motivasi tersendiri bahwa mereka mampu apabila berusaha.

Item soal yang baik adalah yang memiliki daya pembeda minimal 0,30 atau kategori cukup. Analisis daya pembeda menunjukkan bahwa instrumen yang dikembangkan mempunyai daya pembeda “memuaskan” yang berarti perbedaan kecakapan siswa dalam berpikir kritis dapat dinilai melalui

instrumen tersebut. Item soal dengan daya pembeda “sangat memuaskan” berarti dapat membedakan kemampuan berpikir kritis dengan sangat baik.

KESIMPULAN DAN SARAN

Instrumen penilaian *critical thinking skill* yang dikembangkan memiliki karakteristik dapat mendorong dan menilai keterampilan berpikir kritis siswa pada segi inferensi, interpretasi, evaluasi, serta analisis (Teori Facione). Indikator inferensi dan interpretasi diukur melalui soal unjuk kerja. Indikator evaluasi dan analisis diukur melalui soal tes uraian.

Kualitas instrumen yang dikembangkan layak untuk digunakan karena dapat mengukur dengan tepat dan dapat dipercaya (reliabel), memiliki tingkat kesulitan serta daya pembeda yang baik. Item soal tes pada prototype final yang layak dan berkualitas terdiri dari 4 item soal (2 item soal tes unjuk kerja dan 2 item soal tes uraian).

DAFTAR PUSTAKA

- Arikunto, S. 2012. *Dasar-Dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.
- Ashim, M., Asikin, M., & Kharisudin, I. 2019. *Perlunya Komunikasi Matematika dan Mobile Learning Setting Problem Based Learning untuk Meningkatkan Kemampuan 4C di Era Disrupsi*. Prosiding Seminar Nasional Matematika, Semarang, (Online), (<https://journal.unnes.ac.id/sju/index.php/prisma/>, diakses 2020).
- Langrehr, Jhon. 2006. *Thinking Skills*. Edisi Pertama. Jakarta: PT Gramedia.
- Muthmainnah, M., Hapizah, H., & Somakim, S. (2019). *Penerapan Strategi Probing Prompting Dalam Pembelajaran Matematika Materi Relasi Dan Fungsi Di Smp*. *Lentera Sriwijaya: Jurnal Ilmiah Pendidikan Matematika*, 1(1), 27–37. <https://doi.org/10.36706/jls.v1i1.9567>.
- Nur Kholidah. 2019. *Pengembangan Instrumen Penilaian Untuk Mengukur Critical Thinking Skill Siswa Madrasah Tsanawiyah Pada Mata Pelajaran Matematika*. Surabaya: UIN Sunan Ampel Surabaya.
- Permendikbud104-2014 *Tentang Pedoman Penilaian Hasil Belajar Oleh Pendidik Pada Pendidikan Dasar Dan Pendidikan Menengah*. 2014. Jakarta: Permendikbud.
- Purwanto. 2011. *Evaluasi Hasil Belajar*. Yogyakarta: Pustaka Pelajar
- Susanti, E., Sutisnawati, A., & Nurasiah, I. (n.d.). *Penerapan Model Group Investigation (Gi) Untuk Meningkatkan Kemampuan Berpikir Kritis Siswa di Kelas Tinggi. 11*. Sukabumi: Jurnal Kependidikan Utile.
- Siswono, T. Y. E. 2018. *Pembelajaran Matematika Berbasis Pengajaran dan Pemecahan Masalah pada Berpikir Kritis dan Berpikir Kreatif*. Bandung: PT Remaja Rosdakarya