



## PENGEMBANGAN INSTRUMEN TES BERBASIS CRITICAL THINKING SKILL PADA MATERI POKOK FLUIDA SMA

Tri Aprilia dan Jonny Haratua Panggabean

Jurusan Fisika FMIPA Universitas Negeri Medan

[triaprilial2.ta@gmail.com](mailto:triaprilial2.ta@gmail.com)

Diterima: Maret 2021. Disetujui: April 2021. Dipublikasikan: Mei 2021

### ABSTRAK

Penelitian ini bertujuan untuk mengetahui kelayakan instrumen tes fisika berbasis *critical thinking skill* pada materi pokok fluida yang memenuhi kualifikasi baik, meliputi aspek validitas, reliabilitas, daya beda, tingkat kesukaran dan uji respon pada siswa. Penelitian ini merupakan penelitian pengembangan dengan pendekatan *ADDIE* yang merupakan perpanjangan dari *analysis, design, development, implementation and evaluation*. Hasil yang diperoleh berdasarkan validasi oleh validator adalah 86.50% dengan kategori sangat valid. Hasil implementasi menghasilkan nilai validitas pengguna 0,4-0,9. Reliabilitas instrumen 0,8, tingkat kesukaran 0,18-0,7, daya beda 0,2-0,83 dan hasil uji respon menunjukkan 77,58% responden yang menyatakan sangat setuju dengan instrumen tes jenis ini. Berdasarkan kriteria tersebut, maka 15 soal telah dinyatakan baik dan layak digunakan sebagai instrumen tes fisika untuk melatih keterampilan berpikir kritis bagi siswa.

**Kata Kunci:** pengembangan ADDIE, tes critical thinking, fluida

### ABSTRACT

*This study was designed to study the feasibility of a Critical Thinking-based physics test instrument on fluid base material that fulfills good qualifications, containing aspects of validity, reliability, different power, difficulty levels and response tests on students. This research is a research development by adding ADDIE which is an extension of analysis, design, development, implementation and evaluation. The results obtained based on validation by the validator are 86.50% with a very valid category. The results of the implementation, the value of user validity is 0.4-0.9. The reliability of the instrument is 0.8, the level of difficulty is 0.18-0.7, the difference is 0.2-0.83 and the results of the response test show 77.58% of respondents who agree strongly agree with this type of test instrument. Based on these criteria, then 15 questions were approved as good and suitable to be used as physics test instruments to practice critical thinking skills for students.*

**Keywords:** ADDIE development, critical thinking test, fluid

### PENDAHULUAN

Perkembangan ilmu pengetahuan dan teknologi pada abad 21 telah mengubah pola pikir masyarakat secara luas. Gaya hidup dan cara berinteraksi masyarakat, hingga persaingan kompetensi dalam dunia kerja menuntut adanya sumber daya manusia (SDM) yang memiliki cara berpikir dan cara bertindak

sesuai dengan tuntutan kebutuhan yang ada pada saat ini. Hal ini sesuai dengan pendapat Sani (2019) yang menyebutkan bahwa masyarakat pada abad 21 menyadari pentingnya mempersiapkan generasi muda yang kreatif, luwes, mampu berpikir kritis, dapat mengambil keputusan yang tepat, serta terampil dalam memecahkan masalah.

Kompetensi-kompetensi tersebut merupakan cakupan dari keterampilan berpikir tingkat tinggi (*higher order thinking skills (HOTS)*). Berdasarkan perkembangan zaman saat ini, maka kurikulum 2013 juga mengembangkan proses pembelajaran berbasis *HOTS*. Memfasilitasi proses pembelajaran *HOTS* adalah salah satu tuntutan yang harus dijalankan dalam proses pembelajaran di sekolah.

Permendikbud No.64 Tahun 2013 tentang standar isi menyatakan bahwa salah satu kompetensi yang harus dikembangkan dalam pelaksanaan kurikulum 2013 adalah kompetensi berpikir kritis (*critical thinking*). Keterampilan berpikir kritis merupakan salah satu jenis *HOTS*. Pencapaian tujuan pembelajaran berbasis *critical thinking* hanya dapat di ukur dengan instrumen penilaian yang juga memuat *critical thinking* di dalamnya. Salah satu instrumen penilaian yang dapat mengukur keterampilan berpikir kritis siswa adalah instrumen tes yang mencakup soal-soal yang sesuai dengan kriteria keterampilan berpikir tersebut. Penggunaan instrumen tes berbasis *HOTS* tersebut tentunya dapat melatih dan meningkatkan keterampilan berpikir siswa. Hal ini juga telah dibuktikan oleh peneliti sebelumnya yaitu Kusuma, dkk., (2017), yang menyatakan bahwa penilaian instrumen *HOTS* sebagai penilaian untuk belajar, efektif untuk melatih *HOTS* siswa serta efektif dalam mengukur kemampuan berpikir siswa berdasarkan pada tingkat *HOTS* setiap siswa.

Proses pembelajaran yang sedang berjalan di sekolah saat ini ternyata masih belum sesuai dengan tujuan kurikulum tersebut. Menurut hasil studi terakhir TIMSS pada tahun 2015, hasil skor IPA siswa Indonesia berada pada ranking 45 dari 48 negara. Khususnya pada ilmu fisika, rerata persentase jawaban benar siswa Indonesia adalah 32%, sedangkan persentase secara internasional adalah 48%. Berdasarkan hasil tersebut dapat dilihat bahwa daya berpikir siswa khususnya dalam ilmu fisika masih dalam kategori rendah.

Pembelajaran fisika merupakan ilmu dasar yang mempelajari tentang gejala-gejala alam serta seluruh interaksi yang ada di dalamnya. Ilmu dasar tersebut sangat berpengaruh dalam perkembangan teknologi yang menuntut sumber daya manusia yang mampu berpikir kritis, oleh sebab itu pembelajaran yang memfasilitasi pelatihan *HOTS* khususnya keterampilan berpikir kritis sangat berkaitan erat dalam mengembangkan ilmu fisika pada peserta didik. Hal ini juga diperkuat oleh pendapat Brookhart (2010), yaitu siswa yang sudah mampu berpikir tingkat tinggi berarti siswa tersebut mampu menerapkan pengetahuan dan keterampilan yang mereka peroleh selama belajar dan mengembangkannya dalam konteks baru. Selain itu, berdasarkan penelitian yang telah dilakukan oleh Irmawati dkk. (2018), diketahui bahwa siswa yang diajarkan dengan strategi belajar yang berbasis *HOTS* memiliki motivasi lebih tinggi daripada kelompok siswa dengan strategi konvensional.

Salah satu sekolah yang telah menerapkan kurikulum 2013 adalah SMA Negeri 3 Medan. Berdasarkan hasil observasi di sekolah tersebut diketahui bahwa instrumen tes yang diterapkan masih instrumen tes dengan mengacu pada level kognitif C1-C4 taksonomi bloom revisi, sedangkan soal dengan indikator berpikir kritis masih jarang diterapkan karena soal dengan indikator tersebut masih minim dilapangan.

Brookhart (2010), memaparkan definisi *HOTS* dalam hal pemikiran kritis adalah masuk akal, pemikiran reflektif yang difokuskan pada memutuskan apa yang harus dipercaya atau dilakukan. Selain itu Johnson (2002) dalam (Putri, dkk., 2016), juga menyatakan bahwa berpikir kritis merupakan sebuah proses berpikir untuk membangun perkembangan kepercayaan dan mengambil tindakan dengan cara terorganisasi dalam kegiatan mental seperti memecahkan masalah, mengambil keputusan, menyisati, menganalisis asumsi, dan melakukan penelitian. Seseorang dikatakan berpikir kritis dapat tercermin dari cara berpikir yang terarah

dan terorganisir dengan baik dalam menyelesaikan suatu permasalahan.

Keterampilan berpikir kritis bukanlah keterampilan yang dibawa oleh seseorang sejak ia lahir, sehingga keterampilan ini dapat diterapkan, dilatih, dan dikembangkan melalui proses pembelajaran dan evaluasi pembelajaran. dalam evaluasi pembelajaran, guru dapat memberikan latihan kepada siswa melalui pemberian masalah dalam bentuk soal yang bervariasi dan juga guru dapat mengembangkan instrumen yang mampu mengukur keterampilan berpikir kritis siswa (Ningsih, dkk., 2012).

Menurut Arikunto (2013), sebuah tes dapat dikatakan baik sebagai alat pengukur harus memenuhi persyaratan tes, yaitu memiliki validitas dan reliabilitas. Tes yang valid artinya benar-benar mengukur apa yang harus diukur. Tes tersebut benar-benar dapat memberikan gambaran tentang apa yang diinginkan untuk diukur. Sedangkan reliabilitas alat penilaian adalah ketetapan atau keteraturan soal tersebut dalam menilai apa yang dinilainya. Artinya, kapanpun alat penilaian tersebut digunakan akan memberikan hasil yang relatif sama.

Berdasarkan uraian tersebut, perlu dilakukan pengembangan instrumen tes fisika yang dapat melatih keterampilan berpikir kritis yang memenuhi kelayakan instrumen tes yang baik. Penelitian ini bertujuan untuk mengetahui validitas, reliabilitas, daya pembeda, tingkat kesukaran, dan respon siswa dari instrumen tes berbasis *critical thinking* pada materi fluida SMA yang dikembangkan.

#### METODE PENELITIAN

Penelitian dan pengembangan ini menggunakan model yang dikembangkan oleh Robert Maribe Branch dengan pendekatan *ADDIE* yang merupakan perpanjangan dari *analysis, design, development, implementation, and evaluation*. Penelitian ini dilaksanakan di SMAN 3 Medan yang beralamat di Jl. Budi Kemasyarakatan No. 3, Pulo Brayan Kota Medan, pada bulan Juli 2019. Subjek dalam penelitian ini adalah siswa di

kelas XII MIA 1 SMAN 3 Medan yang terdiri dari 29 orang siswa.

Adapun teknik analisis data yang digunakan dalam penelitian ini adalah analisis kuantitatif dengan menghitung validitas butir soal, reliabilitas soal, daya pembeda, tingkat kesukaran, dan respon siswa. Berikut penjelasan langkah-langkah yang dilakukan pada masing-masing tahapan:

Tahap *analysis* dilakukan untuk mengetahui informasi mengenai potensi dan masalah yang ada. Pada tahap analisis dilakukan tiga macam analisis, yaitu analisis kebutuhan, analisis kurikulum, dan analisis karakteristik siswa.

Tahap *design* dilakukan untuk mendesain instrumen dengan menentukan kompetensi, bentuk instrumen dan materi bahan ajar.

Tahap *development* dilakukan pengembangan pendahuluan (draft awal) instrumen yang selanjutnya dilakukan validasi oleh ahli.

Tahap *implementation* dilakukan dalam upaya mengujicobakan instrumen tes yang telah dikembangkan terhadap subjek penelitian guna mengetahui kelayakan instrumen tes tersebut. Adapun langkah-langkah yang dilakukan dalam tahapan ini yaitu mengujicobakan kelayakan instrumen dan melihat respon dari responden terhadap instrumen tes yang diujikan, kemudian menganalisis data yang diperoleh.

Tahap *evaluation* dilakukan dalam upaya meminimalisir terjadinya kesalahan dan kekeliruan dalam proses penelitian. Evaluasi dapat berupa Formatif atau Summatif. Evaluasi Formatif telah berlangsung selama dan di antara fase. Tujuan dari jenis evaluasi ini adalah untuk meningkatkan instruksi sebelum versi final diimplementasikan. Evaluasi Sumatif biasanya terjadi setelah versi akhir instruksi diimplementasikan (McGriff, 2000).

#### HASIL DAN PEMBAHASAN

##### a. Hasil Penelitian

Penelitian ini menghasilkan produk berupa tes uraian berbasis *critical thinking skill* pada materi Fluida. Berdasarkan penelitian yang telah dilakukan, diperoleh analisis data untuk menentukan kelayakan instrumen berdasarkan aspek validitas, reliabilitas, tingkat kesukaran, daya beda dan respon siswa.

Validasi soal pada penelitian ini dilakukan dengan 2 tahapan, yaitu validasi oleh ahli dan validitas butir soal hasil implementasi. Berdasarkan hasil validasi ahli didapatkan bahwa dari 20 butir instrumen tes *HOTS* yang telah dikembangkan semua soal dinyatakan valid dengan rerata presentasi hasil validitas instrumen adalah 86.50%, dengan kategori sangat valid. Hal ini menunjukkan bahwa ketiga validator setuju instrumen yang telah dikembangkan dilanjutkan ke tahap implementasi.

Sedangkan berdasarkan analisis validasi pengguna diketahui bahwa soal yang valid berjumlah 18, sedangkan soal yang tidak valid berjumlah 2 soal seperti pada Tabel 1.

**Tabel 1.** Hasil analisis validitas butir

Nomor Soal	Rentang Nilai	Kategori
2	0,0-0,19	sangat rendah
4	0,20-0,39	rendah
1, 3, 6, 8, 9, 10, 11, 12, dan 20	0,40-0,59	cukup
4, 7, 13, 14, 16, 17 dan 18	0,60-0,79	tinggi
15 dan 19	0,80-1,00	sangat tinggi

Berdasarkan data di atas diketahui bahwa soal yang tidak valid sebanyak 2 soal, hal tersebut kemungkinan terjadi karena siswa sulit memahami kalimat soal.

Analisis data selanjutnya adalah menghitung reliabilitas tes menggunakan formula *Cronbach's Alpha*. Diketahui reliabilitas instrumen tes dalam penelitian ini adalah 0,897 dan berada pada kategori reliabilitas sangat tinggi.

Aspek lain yang menentukan kelayakan soal lainnya adalah tingkat kesukaran masing-masing soal. Tingkat kesukaran dari 20 soal tersebut berada pada

kategori mudah, sedang hingga mudah, yang artinya semua soal telah memenuhi kriteria tingkat kesukaran yang baik, seperti pada Tabel 2.

**Tabel 2.** Tingkat kesukaran instrumen tes

Nomor Soal	Rentang Nilai	Kategori
16 dan 20	0,01- 0,30	sukar
1, 2, 3, 4, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 17, 18 dan 19	0,31-0,70	sedang
15	0,71-0,99	mudah

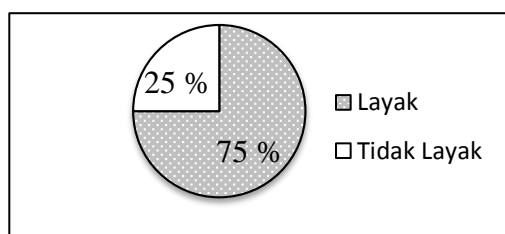
Kemudian untuk daya pembeda pada instrumen tes ini diketahui bahwa 5 soal terkategori daya bedanya jelek, 5 soal cukup, 8 soal baik dan 2 soal pada kategori sangat baik. Hasil analisis daya beda untuk 20 butir soal tersebut dapat dilihat pada Tabel 3.

**Tabel 3.** Daya beda instrumen tes

Nomor Soal	Rentang Nilai	Kategori
1, 2, 5, 6 dan 8	0,0-0,20	jelek
3, 7, 12, 16 dan 20	0,21-0,40	cukup
4, 9, 10, 11, 13, 14, 15 dan 18	0,41-0,70	baik
17 dan 19	0,71-1,00	sangat baik

Berdasarkan Tabel 3, soal yang tidak layak cukup banyak, yaitu 5 soal. Banyaknya soal yang berdaya beda jelek ini kemungkinan terjadi karena banyaknya siswa yang kurang peduli dengan peraturan ujian saat tahap implementasi. Siswa-siswa tersebut banyak yang saling menukar jawaban sehingga daya bedanya menjadi rendah.

Berdasarkan analisis butir soal tersebut diketahui dari validitas pengguna terdapat 18 soal valid, reliabilitas 0,89 pada kategori sangat tinggi, dan tingkat kesukaran 0,18 sampai 0,7 pada kategori sukar hingga mudah, namun pada daya pembeda terdapat 5 soal yang daya bedanya jelek ( $DP < 0,20$ ), sehingga dari 20 soal yang telah dikembangkan, 15 soal telah dapat dinyatakan layak berdasarkan uji implementasi dilapangan. Persentase dari hasil uji tersebut dapat dilihat pada diagram *pie* di Gambar 1.



**Gambar 1.** Diagram pie hasil analisis uji implementasi

Respon siswa yang diperoleh berdasarkan angket respon yang telah diberikan kepada siswa dihitung dengan mencari persentase rata-rata nilai yang didapatkan. Berdasarkan hasil penelitian didapatkan respon siswa terhadap instrumen tes yaitu sebesar 77.58%, dengan kategori sangat baik atau setuju dengan pernyataan-pernyataan yang ditanyakan dalam angket respon.

#### b. Pembahasan

Pengembangan instrumen tes ini diperlukan suatu kriteria untuk menentukan kelayakan instrumen tes yang telah dikembangkan kayak atau tidak. Kriteria tersebut diperlukan sebagai patokan untuk menentukan sejauh mana proses pengembangan dilakukan. Pada penelitian ini, kelayakan instrumen diukur berdasarkan aspek validitas, reliabilitas, tingkat kesukaran, daya beda dan respon siswa.

Penelitian ini menggunakan dua teknik dalam mengukur validitas instrumen tes. Teknik pertama adalah validasi oleh ahli yang menilai kevaliditasan soal berdasarkan aspek isi, konstruk dan bahasa soal. Berdasarkan hasil analisis data diketahui bahwa rerata analisis validasi oleh ahli adalah 86.50% dengan kategori sangat valid, namun beberapa soal masih harus disempurnakan untuk memperbaiki penulisan, kalimat soal bermakna ganda dan penyesuaian dengan karakteristik soal berpikir kritis. Hasil tersebut hampir sama dengan penelitian Najihah, dkk., (2018), nilai presentase rata-rata validasi ahli adalah 81,5%, dengan kategori valid, dengan demikian instrumen tes berbasis *critical thinking skill* pada materi pokok fluida ini telah valid dan dapat digunakan berdasarkan validasi isi dan konstruksinya.

Teknik selanjutnya adalah validasi pengguna. Berdasarkan hasil analisis ke-20 butir soal, diketahui bahwa soal yang valid ( $r > 0,40$ ) berjumlah 18, sedangkan soal yang tidak valid berjumlah 2 soal. Masing-masing pada kategori validitas sangat rendah 5%, validitas rendah 5%, validitas cukup 45%, validitas tinggi 35% dan validitas sangat tinggi 10%. Dengan demikian 90% soal pada instrumen tes ini telah dapat dikatakan bisa mengukur keterampilan berpikir kritis dan layak dari aspek validitas.

Selanjutnya untuk menentukan koefisien reliabilitas dalam penelitian ini, digunakan rumus Cronbach's Alpha. Reliabilitas soal dalam penelitian ini adalah 0,897 dan berada pada kategori sangat tinggi. Sementara pada penelitian Malik, dkk., (2018) reliabilitas tes lebih dari 0,8 dan Yanto (2019) telah mengembangkan tes dengan reliabilitas 0,912 dengan interpretasi sangat tinggi. Soal yang memiliki koefisien reliabilitas sama dengan atau lebih dari 0,70 dikategorikan sebagai soal yang reliabel. Artinya, instrumen tes yang dikembangkan dapat dipercaya dan dan memberikan hasil yang sama apabila instrumen tes ini dilakukan pada subjek, tempat maupun kondisi yang berbeda.

Selanjutnya berdasarkan hasil analisis ke-20 butir soal, diketahui bahwa keseluruhan soal berkisar pada tingkat kesukaran mudah sampai dengan sukar. Masing-masing pada kategori mudah dengan  $0,70 < IK \leq 1,00$  adalah 5%, pada kategori sedang dengan  $0,30 < IK \leq 0,70$  adalah 85% dan pada kategori sukar dengan  $0,00 < IK \leq 0,30$  adalah 10%. Menurut Siswanto, 2017 dalam (Yanto, 2019), parameter yang digunakan untuk memperoleh kualitas soal yang baik, disamping memenuhi validitas dan reliabilitas, adalah adanya keseimbangan dari tingkat kesukaran soal tersebut. Soal yang baik adalah soal yang tidak terlalu mudah atau tidak terlalu sukar. Soal yang terlalu mudah tidak merangsang siswa untuk mempertinggi usaha untuk menjawabnya. Sebaliknya soal yang terlalu sukar akan menyebabkan siswa menjadi putus asa dan tidak mempunyai semangat untuk mencoba lagi, karena diluar jangkauannya.

Yanto (2019) menghubungkan antara tingkat kesukaran soal dengan *HOTS*. Keterampilan berpikir dan level kesulitan soal merupakan hal yang berbeda. *HOTS* mengacu pada tingkat penalaran yang digunakan, sehingga kita mengenal istilah *HOTS (high order thinking skills)* dan *LOTS (low order thinking skills)*, sedangkan level kesulitan atau tingkat kesukaran soal adalah bilangan yang menunjukkan sukar atau mudahnya soal ditunjukkan dengan berapa persen dari seluruh peserta tes yang menjawab soal tersebut dengan benar. Tingkat kesulitan soal dikaitkan dengan kerumitan pengerjaan soal, walaupun sebenarnya tidak memerlukan penalaran tingkat tinggi (*HOTS*). Berdasarkan penjelasan di tersebut, maka dapat disimpulkan bahwa instrumen tes berbasis *critical thinking skill* yang dikembangkan telah berada pada kriteria tingkat kesukaran yang sesuai dan layak untuk digunakan.

Aspek yang mendukung penentuan kelayakan instrumen tes selanjutnya adalah daya beda soal. Daya pembeda soal (indeks diskriminasi) adalah kemampuan suatu soal untuk membedakan antara siswa yang pandai (berkemampuan tinggi) dengan siswa yang bodoh (berkemampuan rendah). Soal memiliki daya pembeda bila tes diberikan kepada siswa yang berkemampuan tinggi menunjukkan hasil tinggi, sedangkan bila diberikan kepada siswa yang berkemampuan rendah, hasilnya akan rendah. Tes dikatakan tidak memiliki daya pembeda bila hasilnya menunjukkan kesamaan antara siswa berkemampuan tinggi dan berkemampuan rendah (Arikunto, 2013).

Berdasarkan analisis data, diketahui bahwa 25% soal terkategori daya bedanya jelek, 25% soal cukup, 40% soal baik dan 10% soal pada kategori sangat baik. 15 soal yang diujikan berada antara 0,2 sampai 0,83 pada kategori cukup hingga sangat baik karena berada pada interval  $0,20 < DP < 1,00$ , sedangkan 5 soal berdaya beda jelek atau  $DP < 0,20$ . Instrumen tes yang telah dikembangkan 75% soal telah dapat membedakan siswa yang berkemampuan tinggi dan berkemampuan rendah, sehingga dapat dikatakan layak dari aspek daya beda. 25% soal dengan daya pembeda jelek kemungkinan

dikarenakan para peserta tes melakukan kerja sama (saling menukar jawaban) ketika ujian berlangsung. Mereka tidak memerdulikan peraturan yang disampaikan oleh pengawas tes. Hal tersebut juga merupakan kendala bagi peneliti karena sulit untuk membuat suasana kondusif saat tes berlangsung.

Uji respon merupakan suatu proses lanjutan pada tahap implementasi yang bertujuan untuk mengetahui respon siswa terhadap instrumen tes yang telah diujikan. Hasil analisis uji respon berdasarkan angket tersebut, tingkat kelayakan berdasarkan uji respon adalah 77.58%, dengan kategori sangat baik. Berdasarkan pengamatan peneliti saat tahap implementasi siswa merasa lebih terbuka pemikirannya terhadap aplikasi konsep dasar fluida dalam kehidupan sehari-hari melalui soal-soal yang diberikan. Namun mereka masih merasa sulit untuk menganalisis jawabannya karena belum pernah diberikan soal dengan karakteristik *critical thinking skill*. Selain itu kemampuan dasar siswa terhadap keterkaitan konsep dasar fisika dengan kehidupan sehari-hari masih kurang, sehingga respon tersebut tidak sejalan dengan nilai-nilai siswa. Maka, perlu adanya latihan soal-soal berpikir kritis yang telah dikembangkan.

Selain kemampuan dasar tersebut, kurangnya pengalaman siswa dalam proses praktikum juga menjadi kendala, beberapa soal yang berkaitan dengan prosedur maupun analisis hasil praktikum juga sulit terjawab. Kendala-kendala tersebut mengakibatkan perencanaan alokasi waktu tidak sesuai dan harus ditambah. Walaupun ada kendala-kendala yang dialami peneliti, namun diharapkan instrumen tes ini telah dapat memenuhi keterbatasan instrumen tes berbasis *critical thinking skill* di lapangan agar siswa mulai terlatih dalam menyelesaikan soal-soal keterampilan berpikir kritis ini.

Menurut Arikunto (2013), Sebuah tes dapat dikatakan baik sebagai alat pengukur harus memenuhi persyaratan tes, yaitu memiliki validitas dan reliabilitas. Instrumen tes yang dikembangkan telah memenuhi aspek-aspek yang disyaratkan tersebut. Beberapa

aspek pendukung seperti tingkat kesukaran dan daya pembeda tes telah memenuhi kriteria yang layak atau baik. Respon dari siswa terkait dengan instrumen yang diujikan kepada mereka juga sangat baik, sehingga dapat dikatakan bahwa instrumen tes berbasis *critical thinking skill* pada materi pokok fluida telah dapat digunakan sebagai suatu alat ukur untuk keterampilan berpikir kritis siswa, melatih keterampilan berpikir kritis siswa dan membantu memenuhi kebutuhan instrumen tes yang masih minim di lapangan.

## KESIMPULAN DAN SARAN

### a. Kesimpulan

Intrumen tes yang dikembangkan telah masuk dalam kategori layak atau baik digunakan sebagai instrumen tes yang dapat melatih keterampilan berpikir kritis. Intrumen tes ini telah memenuhi kriteria validitas ahli dengan presentase 86,50% pada kategori sangat valid, 90% valid berdasarkan validitas pengguna pada kategori cukup valid hingga sangat valid ( $r_{hitung} \geq 0,40$ ), reliabilitas 0,89 pada kategori sangat tinggi karena berada pada interval  $0,80 < r_{11} < 1,00$ , daya beda 0,2 sampai 0,83 pada kategori cukup hingga sangat baik karena berada pada interval  $0,20 < DP < 1,00$  dan tingkat kesukaran 0,18 sampai 0,7 pada kategori sukar hingga mudah karena berada pada interval  $0,00 < IK < 1,00$ . Intrumen tes yang dikembangkan juga telah mendapat respon dengan presentase 77,58% pada kategori sangat baik dari responden.

### b. Saran

1. Respon siswa terhadap instrumen tes yang dikembangkan telah baik, namun nilai rata-rata keterampilan berpikir kritis siswa berdasarkan uji implementasi instrumen tes masih dibawah KKM, agar berhasil meningkatkan keterampilan berpikir kritis maka perlu banyak dilakukan latihan dan ulangan. Karena untuk mengubah kemampuan berpikir seseorang dibutuhkan sebuah proses dan latihan yang tidak singkat. Instrumen ini dapat dijadikan salah satu alternatif untuk itu.

2. Perlu dipertimbangkan rencana yang matang dalam pengalokasian waktu dan sebaiknya uji implementasi dilakukan tidak jauh setelah siswa menerima materi pelajaran, agar kemampuan dasar siswa memadai saat tes berlangsung.
3. Peneliti lain yang hendak melakukan penelitian serupa disarankan memilih materi lain agar instrumen tes berbasis critical thinking untuk semua materi fisika tersedia di lapangan.
4. Menganalisis butir soal secara kuantitatif, pengolahan data menggunakan Microsoft Excel memerlukan waktu yang relatif lama, sehingga disarankan kepada peneliti lain untuk menggunakan aplikasi lain agar waktu yang diperlukan menjadi lebih efektif.

## DAFTAR PUSTAKA

- Arikunto, (2013), *Dasar-dasar Evaluasi Pendidikan Edisi 2*, Bumi Aksara Jakarta.
- Brookhart, S. M., (2010), *How to Assess Higher Order Thinking Skills in Your Classroom*, ASCD, Alexandria.
- Irmawati, R.D., Yetti S., dan Muchlas Suseno., (2018), Pengaruh Strategi Pembelajaran dan Motivasi Belajar Terhadap Higher Order Thinking Skills (HOTS) Dalam Pembelajaran IPA Siswa Kelas IV Sekolah Dasar, *Jurnal Tunas Bangsa*, 5 (2), 143-156
- Kemendikbud, (2018), *Dinamika Perkembangan Kurikulum 2013*, Kemendikbud, Jakarta.
- Kusuma, M. D., Undang R., Abdurrahman dan Agus S. (2017), The Development of Higher Order Thinking Skill (HOTS) Instrument Assessment in Physics Study, *IOSR Journal of Research & Method in Education (IOSR-JRME)*, 7 (1), 1-7
- Malik, Abdul, Undang R, dan Chandra E., (2018), Pengembangan Instrumen Asesmen HOTS Fisika SMA

Menggunakan Model Inkuiri Terbimbing, *Jurnal Lentera Pendidikan Pusat Penelitian LPPM UM METRO*, 3 (1), 11-25

- McGriff, S. J., (2000), *Instructional System Design (ISD): Using the ADDIE Model*, Penn State University.
- Najihah, A. R., Vina S., Mutia S., (2018), The Development of High Order Thinking Skills (*HOTS*) Assessment Instrument for Temperature and Heat Learning, *JPPPF (Jurnal Penelitian dan Pengembangan Pendidikan Fisika)*, 4 (1), 19-26.
- Ningsih, S. M., Bambang S., A. Sopyan, (2012), Implementasi Model Pembelajaran Process Oriented Guided Inquiry Learning (POGIL) untuk Meningkatkan Kemampuan Berpikir Kritis Siswa, *Unnes Physics Education Journal*, 1 (2), 44-52
- Putri, F.S., E. Istiyono, dan E. Nurcahyanto, (2016), Pengembangan Instrumen Tes Keterampilan Berpikir Kritis dalam Bentuk Pilihan Ganda Beralasan (Politomus) di DIY, *Unnes Physics Education Journal*, Vol. 5 (2), 76-84
- Sani, R.A., (2019), *Pembelajaran Berbasis HOTS (Higher Order Thinking Skills)*, Tira Smart, Tanggerang.
- Yanto, S. H., (2019), *Pengembangan Two Tiers Multiple Choice Test untuk Mengidentifikasi Kemampuan Berpikir Tingkat Tinggi pada Materi Sistem Peredaran Darah pada Manusia*, Unimed, Medan.