

Pemanfaatan Instrumen Tes *Essay* Untuk *Mapping* Kemampuan *Critical Thinking* Siswa SMA

Elang Mulia Lesmana¹, Wawan Bunawan²

Universitas Negeri Medan

elanglesmana8@gmail.com , wawanbunawan@unimed.ac.id

ABSTRAK

Penelitian ini bertujuan untuk untuk mengembangkan instrumen tes berbasis berpikir kritis sebagai instrumen dalam pembelajaran di SMA, untuk mengetahui tingkat keefektifan yang digunakan untuk instrumen berbasis berpikir kritis, dan mengetahui karakteristik instrumen asesmen soal berbasis berpikir kritis yang dikembangkan. Penelitian ini merupakan penelitian Research & Development (R & D) dengan menggunakan model 4-D. Hasil validitas butir memiliki nilai r_{hitung} tertinggi yaitu 0,719 (valid) dan r_{hitung} terendah yaitu 0,345 (cukup valid), hasil uji reliabilitas memiliki nilai 0,696 (reliabilitas sedang), Hasil uji tingkat kesukaran memiliki nilai tertinggi yaitu 0,817 (Mudah) dan nilai terendah yaitu 0,242 (Sukar). Hasil uji daya beda memiliki nilai tertinggi yaitu 0,561 (Cukup) dan nilai terendah yaitu 0,199 (Rendah). Hasil uji regresi memiliki nilai r_{hitung} tertinggi yaitu 0,804 (Sangat Kuat) dan r_{hitung} terendah yaitu 0 (Lemah). Dan mengetahui respon siswa terhadap penggunaan Instrumen tes berpikir kritis. Berdasarkan keterangan tersebut, maka kesimpulan dari instrumen tes mendapatkan hasil yang layak digunakan.

Kata kunci : Pemanfaatan instrumen tes, berpikir kritis

ABSTRACT

This study aims to develop a critical thinking-based test instrument as an instrument in high school learning, to determine the level of effectiveness used for critical thinking-based instruments, and to determine the characteristics of the critical thinking-based question assessment instrument developed. This research is a Research & Development (R & D) research using a 4-D model. The results of item validity have the highest r_{count} value, namely 0.719 (valid) and the lowest r_{count} is 0.345 (quite valid), the reliability test results have a value of 0.696 (medium reliability), the difficulty level test results have the highest value, namely 0.817 (Easy) and the lowest value is 0.242 (Hard). The results of the different power test have the highest value of 0.561 (Enough) and the lowest value of 0.199 (Low). The results of the regression test have the highest r_{count} value, namely 0.804 (Very Strong) and the lowest r_{count} , 0 (Weak). And knowing student responses to the use of critical thinking test instruments. Based on this information, the conclusion of the test instrument is that the results are suitable for use.

Keywords: Use of test instruments, critical thinking

PENDAHULUAN

Pembelajaran fisika sebagai halnya kerangka dari pendidikan nasional memegang peranan signifikan dalam perkembangan ilmu dan teknologi, karena fisika menggambarkan ilmu yang melandasi ilmu-ilmu lainnya. Dengan demikian pengkajian fisika diperlukan peserta didik sekolah dasar sampai sekolah menengah untuk menyuplai peserta didik dengan keterampilan berpikir logis, analitis, sistematis, kritis, dan kreatif, serta keahlian bekerja sama. Keterampilan ini perlu bagi peserta didik untuk dapat mengumpulkan, memahami, mengolah dan menggunakan informasi, serta mengikuti perkembangan zaman teknologi pada saat ini.

Realisasi aktivitas pembelajaran di dalam kelas dilaksanakan untuk mendapatkan suatu tujuan pembelajaran. Salah satu tujuan pembelajaran adalah mengembangkan harapan yang berubah dari tidak diketahui menjadi diketahui (area of knowledge) yaitu kemampuan penalaran yang tinggi dalam berpikir analitis induktif dan deduktif dengan menggunakan

konsep fisika oleh siswa untuk menjelaskan dan menjelaskan berbagai peristiwa alam yang mungkin menjadi prinsip pemecahan masalah secara kualitatif dan kuantitatif (Anderson & Krathwohl, 2001) dari tidak memenuhi syarat untuk memenuhi syarat (area of competence), memiliki sikap yang buruk memiliki sikap yang baik (area of attitude). Selain itu, yang dilakukan guru dalam proses pembelajaran adalah untuk mengatasi tiga hal pada diri siswa, yaitu (aspek pengetahuan atau kognitif), (aspek sikap atau afektif) dan (aspek keterampilan atau psikomotor).

Keberhasilan suatu proses pembelajaran dipengaruhi oleh sistem evaluasi pada program studi 2013. Sistem penilaian tidak diarahkan pada penilaian akhir belaka, tetapi penilaian proses pembelajaran harus dipertimbangkan dengan cermat, sebagai halnya dicantumkan dalam Peraturan Menteri Pendidikan dan Kebudayaan No. 20 Tahun 2016 mengenai Standar Kompetensi Lulusan Pendidikan Dasar dan Menengah. Kemampuan berpikir lebih jauh

didefinisikan oleh Kementerian Pendidikan dan Kebudayaan untuk menimbang aspek berpikir peserta didik, sebagai halnya dicantumkan pada Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia No. 22 Tahun 2016 mengenai Standar Proses Pendidikan Dasar dan Menengah dimana pengetahuan dicapai lewat kegiatan “mengingat, memahami, menerapkan, menganalisis, mengevaluasi dan mencipta”.

Berbagai dimensi perilaku yang mencerminkan sikap dalam Permendikbud Nomor 20 Tahun 2016: (1) memiliki iman dan beribadat kepada Tuhan YME, (2) berakhlak mulia, jujur, dan memeka, (3) bertanggung jawab, (4) pelajar sejati selamanya, dan (5) sehat jasmani dan rohani sejalan dengan kemajuan anak di sekitarnya. Aspek ilmu, mempunyai ilmu yang praktis, konseptual, prosedural, dan metakognitif akan fase dasar: (1) sains, (2) teknologi, (3) seni, dan (4) budaya. Aspek kompetensi, Mampu berefleksi dan bertindak: (1) kreativitas, (2) produktivitas, (3) berpikir kritis, (4) kemandirian, (5) kerjasama dan (6) komunikasi berkelanjutan berkat pendekatan saintifik yang sinkron dengan tahapan perkembangan anak sesuai pekerjaan yang dibutuhkan, sehingga diperlukan penilaian semester dalam proses pembelajaran.

Penilaian yang dipergunakan pada program 2013 yaitu penilaian autentik alias authentic assessment yang targetnya memberikan data autentik/valid dan sah tentang hal-hal yang sebenarnya terlihat dan bisa diaplikasikan peserta didik. Dalam menilai kemajuan dan hasil belajar peserta didik di sekolah, dimensi yang berkaitan pada pemilihan alat penilaian yaitu; Menyiapkan kuesioner, menganalisis item supaya mendapatkan pertanyaan berkualitas yang sesuai, dan memproses serta menginterpretasikan statistik penilaian. Menurut (Arikunto, 2012), sebuah tes yang baik wajib memiliki persyaratan sebagai berikut: 1) Validitas, 2) Reliabilitas, 3) Objektivitas, 4) Kepraktisan, 5) Ekonomis. Oleh karena itu, demi mendapatkan tes yang baik, tes itu semestinya terlebih dahulu diperiksa dan dianalisis agar memenuhi persyaratan tes yang baik dan benar.

Kemampuan berpikir kritis (critical thinking) tetap menjadi prioritas utama untuk pendidikan. Kebiasaan utama yang membentuk pemikiran kritis secara konsisten dan tegas telah diidentifikasi oleh sekolah menengah. Peserta didik diharapkan mampu memecahkan semua persoalan secara kritis dan kreatif. Mereka harus mampu memecahkan masalah secara kritis dengan kemampuan analitis dengan mempertimbangkan berbagai faktor yang berkaitan dengan alasan rasional dan kreatif (Bialik & Fadel, 2015). Pembelajaran hanya

menghafal teori kurang tepat dengan kebutuhan dan kondisi terkini. Setiap dimensi kehidupan berkembang, akhirnya setiap manusia perlu kritis dan kreatif saat menanggapi keadaan dinamis seperti ini. Berdasarkan hasil penelitian oleh (Jiwandono, 2019) menyiratkan bahwa keterampilan berpikir kritis masih minim karena konten pada instrument soal masih tidak memiliki indikator berpikir kritis sehingga tidak sesuai instrumen asesmen yang digunakan khusus untuk melatih berpikir kritis. Dalam hasil riset juga didapatkan masih banyak responden memberi jawaban kurang tepat karena jawaban yang peserta didik berikan terlalu singkat dan tidak memiliki indikator berpikir kritis. Oleh karena itu menjadi alasan bahwa instrumen tes keterampilan berpikir kritis perlu dikembangkan. Penyusunan perangkat asesmen membahas keterampilan berpikir kritis yang sangat diperlukan untuk dikembangkan dan diujicoba pada pembelajaran, khususnya di tingkat Sekolah Menengah Atas (SMA). Sampai saat ini, jarang ditemukan alat penilaian yang membahas bidang kemampuan berpikir tingkat tinggi, sehingga tidak jelas bagaimana kemampuan berpikir tingkat tinggi peserta didik pada tingkat Sekolah Menengah Atas (SMA).

Untuk kemampuan berpikir tingkat tinggi, data memperlihatkan bahwa hasil fisika yang diperoleh dalam dimensi penalaran Indonesia berada pada ranking 40 dari 42 negara (Istiyono et al., 2014). Selanjutnya sesuai dengan hasil TIMSS maka disimpulkan sebagai berikut: (1) prestasi belajar IPA peserta didik Indonesia secara normal dari segi intelektualitas (mengetahui, menerapkan, berpikir) masih rendah; (2) kecenderungan prestasi belajar fisika peserta didik Indonesia untuk terus mengalami penurunan di setiap sudut pandang intelektual sehingga kemampuan peserta didik Indonesia dalam sains fisika harus digarap di semua sudut, terutama perspektif berpikir dengan membekali siswa dengan pemikiran permintaan yang lebih tinggi. Sejalan dengan itu, tingginya permintaan tersebut mengingat kapasitas siswa sekolah menengah (SMA) di Indonesia masih rendah di kancah global. Rendahnya prestasi belajar IPA materi dapat disebabkan oleh sistem Hasil wawancara dengan pengajar bidang studi fisika menunjukkan bahwa permasalahan di sekolah adalah instrument evaluasi kognitif yang dipergunakan berupa soal pilihan ganda yang umumnya akan menguji lebih banyak bagian memori, pemahaman, dan aplikasinya, sehingga soal yang dimuat pada pekerjaan peserta didik masih termasuk ke dalam soal LOTS yang hanya sekedar untuk dihafal, dipahami dan diterapkan. Bahkan industri era keempat yang sering dikenal industri 4.0 mengharapkan peserta didik untuk berfikir lebih kreatif dan

kritis. Karena soal-soal yang melatih siswa untuk berpikir kritis masih minim dan belum adanya alat tes yang dirancang khusus untuk melatih berpikir kritis maka penting untuk mengembangkan instrumen berbasis berpikir kritis untuk mengukur kemampuan berpikir kritis peserta didik. Selanjutnya, instrumen berbasis berpikir kritis yang dikembangkan akan membantu siswa melatih kemampuan berpikir kritis.

METODE PENELITIAN

Penelitian ini adalah penelitian pengembangan atau Research and Development (R&D). Model pada penelitian adalah model penelitian pengembangan yang diterapkan mengukur pada model 4-D yang sudah dikualifikasi dan penyesuaian. Model ini dikemukakan oleh Thiagarajan, dkk (1974) Model pengembangan ini terbagi atas 4 tahapan, yaitu: defines, designs, development, dan Dessiminated. Penelitian ini dilakukan di SMA di Kota Medan, pada bulan November 2021. Subjek dalam penelitian ini adalah peserta didik di kelas XI MIPA yang terdiri atas 30 orang siswa.

Teknik analisis data yang digunakan dalam penelitian ini adalah analisis kuantitatif dengan menghitung validitas, reliabilitas soal, daya pembeda, tingkat kesukaran, respon siswa, serta uji keefektifan regresi.

HASIL DAN PEMBAHASAN

Hasil Penelitian

Hasil penelitian ini mendeskripsikan tentang proses Pengembangan Instrumen Berbasis berpikir kritis Pada Materi Gerak Melingkar Beraturan. Penelitian ini dikembangkan dengan menggunakan model pengembangan 4-D yaitu tahap Define, Design, Develop hingga Disseminate, oleh karena itu hasil dari penelitian kelayakan instrumen Berdasarkan hasil analisis dari 3 validator ahli untuk kelayakan instrumen soal berbasis berpikir kritis di ujikan kepada peserta didik adapun hasilnya sebagai berikut ini :

Tabel 1 Validasi Ahli

No.	Nama Validator Ahli	Nilai	Hasil Analisis
1	Validator I	98 %	Valid
2.	Validator 2	93 %	Valid
3.	Validator 3	83 %	Valid

1. Analisis Data

a. Validitas butir soal

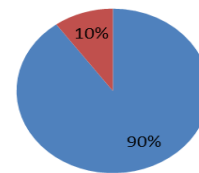
Berdasarkan korelasi Skor dari 10 butir soal, diperoleh hasil bahwa sebanyak 9 soal dinyatakan valid dan 1 soal dinyatakan cukup valid. Rekapitulasi hasil perhitungan validitas

instrumen tes menggunakan aplikasi SPSS

Gambar 1 Analisis Uji Validitas

Uji Validitas

■ valid ■ cukup valid



validitas berkenaan dengan ketepatan alat penilaian terhadap konsep yang dinilai sehingga betul-betul menilai apa yang seharusnya dinilai. Untuk mengetahui apakah tes itu valid atau tidak, harus dilakukan melalui penelaahan kisi-kisi tes untuk memastikan bahwa soal-soal tes itu sudah mewakili atau mencerminkan keseluruhan konten atau materi yang seharusnya dikuasai secara proporsional. Berdasarkan korelasi Skor, diperoleh hasil bahwa sebanyak 90 % dinyatakan valid dan 10 % dinyatakan cukup valid. Instrument yang dikembangkan disimpulkan bahwa instrument dapat dan layak digunakan serta sudah mewakili atau mencerminkan keseluruhan konten atau materi yang seharusnya dikuasai secara proporsional (Azwar, 2015; Mardapi, 2018; Matondang, 2009).

b. Reliabilitas

suatu tes dikatakan reliabel jika selalu memberikan hasil yang sama bila diteskan pada kelompok yang sama. Dengan menggunakan rumus Reliability Statistic cronbach alpha menggunakan aplikasi SPSS, mendapatkan hasil uji reliabilitas sebesar 0,696. berarti berada pada kategori Sedang. Dari hasil reliabilitas dapat disimpulkan bahwa tanggapan terhadap tes dari peserta didik memiliki konsistensi sehingga memberikan hasil yang sama (Mardapi, 2018; Mas'ula & Rokhis, 2020; Matondang, 2009).

Tabel 2 Reliabilitas Instrumen Tes

Reliability Statistics	
Cronbach's Alpha	N of Items
.696	10

c. Uji Tingkat Kesukaran

Analisis tingkat kesukaran dimaksudkan untuk mengetahui apakah soal tersebut tergolong mudah atau sukar. Dengan menggunakan rumus uji tingkat kesukaran pada aplikasi MS. Excel, mendapatkan hasil uji tingkat kesukaran dengan 20 % tingkat kesukaran dengan kategori mudah, 40 % tingkat kesukaran dengan kategori sedang dan 40 % tingkat kesukaran dengan kategori sukar. Dari

hasil analisis tingkat kesukaran dapat disimpulkan bahwa instrumen tes layak digunakan karena memiliki tingkat kesukaran yang tergolong mudah dan sukar (Arikunto, 2012; Bagiyono, 2017; Salmina & Adyansyah, 2017).

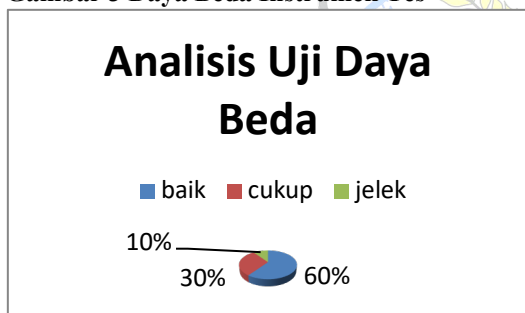
Gambar 2 Tingkat Kesukaran Instrumen



d. Uji Daya Beda

Uji daya beda instrument tes bertujuan untuk melihat sejauh mana item mampu membedakan antara individu atau kelompok individu yang memiliki atribut dengan yang tidak memiliki atribut yang akan diukur (Azwar, 2015). Dengan menggunakan rumus uji daya beda pada aplikasi SPSS, mendapatkan hasil uji tingkat kesukaran dengan 6 butir soal dengan kategori Baik, 3 butir soal dengan kategori cukup dan 1 butir soal dengan kategori rendah. Dari hasil analisis uji daya beda instrument tes dapat disimpulkan bahwa instrumen tes dapat digunakan karena dikatakan baik dan mampu membedakan antara individu (Azwar, 2015; Bagiyono, 2017; Sani et al., 2019)

Gambar 3 Daya Beda Instrumen Tes



e. Respon siswa terhadap soal

Respon peserta didik yang dihasilkan dari angket respon siswa yang dibagikan kepada siswa diukur dengan menghitung persentase yang di dapatkan. Berdasarkan hasil penelitian disimpulkan bahwa respon peserta didik terhadap soal yaitu sebesar 92,7% yaitu dikatakan pada kategori baik.

Pembahasan

Persyaratan Instrumen tes yang baik meliputi validitas, realibilitas, dan regresi linear. Analisis butir soal juga diperlukan untuk mengkaji pertanyaan – pertanyaan tes agar di peroleh perangkat pertanyaan yang memiliki

kualitas yang memadai dan handal. Tujuan khusus dari analisis butir soal yaitu mencari soal tes yang baik maupun memadai. Sehingga, kelayakan instrumen tes ini ditinjau berdasarkan validitas dan analisis butir soal yang meliputi realibilitas, dan regresi tiap butir soal.

Uji efektivitas regresi

Tujuan penerapan uji regresi ini adalah untuk meramalkan atau memprediksi nilai variabel tak bebas (dependen) yang dipengaruhi oleh variabel bebas (independen) yaitu memprediksi hubungan dan pengaruh antara hasil skor dari siswa terhadap indikator instrumen berbasis berpikir kritis.

Pada aspek regresi diperoleh dapat dideskripsikan bahwa ada hubungan searah dari variabel X dan Y bahwa apabila X meningkat maka Y juga meningkat. Dari lima indikator keterampilan berpikir kritis yang memiliki hubungan dengan hasil belajar siswa adalah pada indikator 1, 2, 3, dan 5 yaitu memberikan penjelasan sederhana, membangun keterampilan dasar, menarik kesimpulan, dan membangun strategi dan taktik. Sedangkan indikator yang tidak memiliki hubungan dengan hasil belajar siswa adalah pada indikator ke 4 yaitu memberikan penjelasan lebih lanjut.

Tabel 3 Uji Keefektifan Regresi

Model Summary					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Kategori
Indikator 1	.679 ^a	0.461	0.442	4.227	Kuat
Indikator 2	.748 ^a	0.559	0.543	3.824	Kuat
Indikator 3	.677 ^a	0.459	0.44	4.236	Kuat
Indikator 5	.804 ^a	0.646	0.634	3.424	Kuat

Uji regresi bertujuan untuk mengukur tingkat keefektifan suatu instrument tes. Instrumen tes yang diukur tingkat keefektifannya adalah instrument tes berbasis berpikir kritis. Indikator berpikir kritis yang diukur adalah indikator berpikir kritis FRISCO yang dikemukakan oleh Ennis. Untuk melihat suatu instrument tes dikatakan efektif apabila indikator berpikir kritis sebagai X memenuhi atau memiliki hubungan searah dengan hasil belajar sebagai Y. Berdasarkan hasil dari uji regresi dihasilkan bahwa indikator berpikir kritis FRISCO signifikan dengan instrument tes berpikir kritis yang diawali dari indikator Focus (X1) sebagai tergolong kuat dengan instrument tes berpikir kritis, indikator Reason (X2) tergolong kuat dengan instrument tes berpikir kritis, Inference (X3) tergolong kuat dengan

instrument tes berpikir kritis. Clarity (X1), tergolong kuat dengan instrument tes berpikir kritis. Namun pada indikator situasion (X4) dan Overview (X4), tergolong lemah karena tidak efektif digunakan pada instrument tes berbasis berpikir kritis. Dari hasil tersebut dapat disimpulkan bahwa instrument tes berpikir kritis sudah dikatakan efektif Karena signifikan dengan indikator berpikir kritis(L, 2016) (Affandy et al., 2019; Ennis, 1995; Jiwandono, 2019; L, 2016; Zubaidah et al., 2018).

Pengembangan alat penilaian pertanyaan pilihan ganda dua tingkat yang mampu mengukur keterampilan berpikir tingkat tinggi dan menentukan kelayakan " alat penilaian pertanyaan pilihan ganda dua tingkat sebagai formatif penilaian. Hasil penelitian menunjukkan bahwa fitur-fitur alat penilaian soal pilihan ganda dua tingkat dikembangkan berdasarkan indikator kemampuan berpikir tingkat tinggi menurut Anderson dan Krathwohl (2001), yaitu analisis, evaluasi dan kreatif; valid dengan interpretasi minimal "cukup"; dan keandalan "tinggi". Kelangsungan produk alat evaluasi dipastikan dengan kandungan nilai "baik"; efek konstitutif "baik"; validitas butir soal dengan pemahaman minimal "cukup", tingkat kesukaran soal dengan rasio 15% mudah : 80% sedang : 5% sulit; membedakan masalah daya dengan interpretasi minimal "cukup" dan kepraktisan penggunaan "baik". Alat penilaian pilihan ganda dua tingkat dapat digunakan sebagai alternatif alat penilaian formatif di sekolah dengan aplikasi yang lebih dalam (Affandy et al., 2019; Amalia & Susilaningsih, 2014; Istiyono et al., 2014; Nofiana et al., 2014; Zubaidah et al., 2018).

Dalam melakukan penelitian ini, tidak dapat dipungkiri bahwa terjadi banyak kendala yang dialami. Beberapa kendala yang dihadapi oleh saya pada saat melakukan penelitian ini seperti siswa kesulitan dalam mengerjakan soal dikarenakan mereka telah lupa materi tersebut. Sehingga harus melakukan review materi sebelum siswa mengerjakan soal tersebut. Selain itu kendala lainnya yang dihadapi oleh saya yaitu kesulitan mengkondusifkan kelas agar siswa mengerjakan soal dalam keadaan tenang dan mampu mengerjakan soal dengan tepat. Hal ini disebabkan oleh ketidakbiasaan siswa dalam menghadapi dan mengerjakan soal – soal Berpikir Kritis (Critical Thinking). Sehingga, siswa mengalami kesulitan dalam memahami kalimat soal dan siswa menjadi tidak kondusif karena selalu bertanya mengenai pertanyaan di dalam soal. Oleh karena itu, saya harus membimbing dan menggiring siswa dalam memahami kalimat soal dan memaknai setiap butir pertanyaan di dalam soal agar dapat mengerjakannya dengan benar dan tepat.

KESIMPULAN DAN SARAN

Untuk Mengukur tingkat Keefektifan Instrument Tes Berbasis Berpikir Kritis (Critical Thinking) telah dikembangkan pada materi Gerak Melingkar Beraturan di Kelas XI MIA 3 SMAN 5 Medan yang terdiri dari 10 (Sepuluh) butir soal dalam bentuk essay sudah efektif dari setiap butir soal dikatakan kuat karena sesuai dengan indikator berpikir kritis FRISCO. Untuk mengukur karakteristik instrumen tes yang dikembangkan, dilakukan uji Validitas, uji Reliabilitas, uji Tingkat Kesukaran, dan uji Daya Beda. Berdasarkan dari ujicoba butir soal, instrument tes berpikir kritis dapat digunakan. Hasil dari angket respon siswa disimpulkan bahwa instrumen tes yang telah dikembangkan sudah sangat layak dilihat dari hasil angket dari segi materi, kesesuaian instrumen tes dengan instrumen tes berbasis Berpikir Kritis (Critical Thinking), kejelasan perintah pada soal, kejelasan pertanyaan pada soal, berada di atas rata-rata. Hal itu memperlihatkan respon siswa terhadap instrument tes yang dikembangkan sudah sangat baik.

Diharapkan pada saat mengembangkan alat tes berbasis berpikir kritis jika dalam penelitiannya menggunakan sistem online (online) memang harus menyesuaikan waktu untuk waktu d Evaluasi dan waktu pengumpulan dan pengolahan dapat seefisien dan seakurat mungkin. Pada saat melakukan penelitian alat tes berbasis berpikir kritis diharapkan melakukan penelitian di sekolah yang telah menerapkan pembelajaran berbasis berpikir kritis.

DAFTAR PUSTAKA

- Affandy, H., Aminah, N. S., & Supriyanto, A. (2019). Analisis Keterampilan Berpikir Kritis Siswa Pada Materi Fluida Dinamis Di SMA Batik 2 Surakarta. *Jurnal Materi Dan Pembelajaran Fisika (JMPF)*, 9(1), 25–33.
- Amalia, N. F., & Susilaningsih, E. (2014). Pengembangan Instrumen Penilaian Keterampilan Berpikir Kritis Siswa SMA Pada Materi Asam Basa. *Jurnal Inovasi Pendidikan Kimia*, 8(2), 1380–1389.
- Anderson, L. W., & Krathwohl, D. R. (2001). *A Taxonomi For Learning , Teaching, aand Assessing A Revision Of Bloom’s Taxonomi Of Educational Objectives*. Longman.
- Arikunto, S. (2012). *Dasar-dasar Evaluasi Pendidikan Edisi ke-2*. Bumi Aksara.

- Azwar, S. (2015). *Validitas dan Reliabilitas*. Pustaka Belajar.
- Bagiyono. (2017). Analisis Tingkat Kesukaran dan Daya Pembeda Sial Ujian Pelatihan Radiografi Tingkat 1. *Widyanuklida*, 16(No. 1), 1–12.
- Basuki, Ismet, & Hariyanto. (2014). *Asesmen Pembelajaran*.
- Bialik, M., & Fadel, C. (2015). Skills for the 21st Century: What should students learn? *Center for Curriculum Redesign*, 3(4), 29.
- Brookhart, S. M. (2010). *How To Assess Higher-Order Thinking Skills In Your Classroom*. ASCD.
- Budiman, A., & Jailani, J. (2014). Pengembangan Instrumen Asesmen Higher Order Thinking Skill (Hots) Pada Mata Pelajaran Matematika Smp Kelas Viii Semester I. *Jurnal Riset Pendidikan Matematika*, 1(2), 139.
- Conklin, W. (2012). Higher order thinking skills to develop 21st century learners. Shell Educational Publishing, Inc.
- Ennis, R. H. (1995). *Critical Thinking*. University of Illinois.
- Giancoli. (2005). *Physics*. Upper Saddle River.
- Istiyono, E., Mardapi, D., & Suparno, S. (2014). Pengembangan Tes Kemampuan Berpikir Tingkat Tinggi Fisika (PysTHOTS) Peserta Didik SMA. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 18(1), 1–12.
- Jiwandono, N. R. (2019). Kemampuan Berpikir Kritis (Critical Thinking) Mahasiswa Semester 4 (Empat) Pada Mata Kuliah Psikolinguistik. *Ed-Humanistics : Jurnal Ilmu Pendidikan*, 4(1).
- L, F. A. (2016). Efektifitas Penggunaan Media Pembelajaran dan Model Pembelajaran Jigsaw Dalam Meningkatkan Pretasi Belajar PPKn Pada SMPN 3 Kalidawir Tulungagung. 1–14.
- Mardapi, D. (2018). Teknik Penyusunan Instrumen Tes dan Non Tes. *Parama Publisihing*.
- Mas'ula, N., & Rokhis, T. A. (2020). Pengembangan Instrumen Tes Kemampuan Berpikir. *Sap*, 4(3), 177–185.
- Matondang, Z. (2009). Validitas dan reliabilitas suatu instrumen penelitian. 6(1), 87–97.
- Nofiana, M., Sajidan, S., & Puguh, P. (2014). Pengembangan Instrumen Evaluasi Two-Tier Multiple Choice Question untuk Mengukur Keterampilan Berpikir Tingkat Tinggi pada Materi Kingdom Plantae. *Jurnal Inkuiri*, 3(2), 60–74.
- Raudhah, S., Hartoyo, A., & Nursangaji, A. (2019). Analisis berpikir kritis siswa dalam menyelesaikan soal spltv di sma negeri 3 pontianak. *Jurnal Pendidikan Dan Pembelajaran Khatulistiwa*, 4, 1–8.
- Reta, I. (2012). Pengaruh Model Pembelajaran Berbasis Masalah Terhadap Keterampilan Berpikir Kritis Ditinjau Dari Gaya Kognitif Siswa. *Jurnal Pendidikan Dan Pembelajaran IPA Indonesia*, 2(1), 1–17.
- Riyani, R., Maizora, S., & Hanifah. (2017). *Jurnal Penelitian Pembelajaran Matematika Sekolah (JP2MS)*, Vol. 1, No. 1, Agustus 2017 eISSN 2581-253X. 1(1), 60–65.
- Salmina, M., & Adyansyah, F. (2017). Analisis Kualitas Soal Ujian Metematika Semester Genap Kelas XI SMA Inshafuddin Kota Banda Aceh. *Journal of Chemical Information and Modeling*, 4(1), 37–47.
- Sani, R. A., Rahmatsyah, & Bunawan, W. (2019). *Soal Fisika HOTS*. Bumi Aksara.
- Sudirman, S., Kistiono, K., Akhsan, H., & Ariska, M. (2020). Pengembangan Instrumen Penilaian Pengetahuan, Sikap Dan Keterampilan Ipa Berbasis Berpikir Kritis Pada Konsep Listrik Siswa Smp. *Jurnal Inovasi Dan Pembelajaran Fisika*, 7(1), 28–40. <https://doi.org/10.36706/jipf.v7i1.10903>
- Thiagarajan. (2015). *Model Pengembangan 4-D*.
- Widana, I. W. (2016). *Penulisan Soal HOTS untuk Ujian Sekolah*. Direktorat Pembinaan SMA.
- Widoyoko, E. P. (2014). *Penilaian Hasil Pembelajaran di Sekolah*. Pustaka Pelajar.
- Zubaidah, S., Malang, U. N., & Aloysius, D. C. (2018). *Asesmen Berpikir Kritis Terintegrasi Tes Essay Asesmen Berpikir Kritis Terintegrasi Tes Essay*. UM.