

PENGEMBANGAN INSTRUMEN TES *MULTIPLE CHOICE HIGH ORDER THINKING* PADAPEMBELAJARAN FISIKA BERBASIS *E-LEARNING* DI SMA

¹⁾Dian Ratih Utama Sari, ¹⁾Sri Wahyuni, ¹⁾Rayendra Wahyu Bachtiar

¹⁾Program Studi Pendidikan Fisika FKIP Universitas Jember

Email: dianratih.dr@gmail.com

Abstract

The purpose of this research was to develop and described the validity, reliability, and analysis of items on each test package and know the ability to high order thinking of the students. The design of this research was 4D development. Based on the results of the analysis and discussion can be concluded that the validation results of experts meet the valid criteria. Empirical validity obtained meet the criteria of validity is very high. Reliability test results on the product indicate that the category was quite reliable. The item analysis was the difficulty level of packets A and B still require improvement. In the item discrimination analysis, there was 13 problem categories not feasible. In the evaluating item distractor, there was 16 questions of packet A and 10 matter of package B with distractor accepted because it is good enough and there are 14 problems with distractor revised. High order thinking skill of students after using multiple choice test instrument high order thinking on e-learning based physics learning in SMA N 1 Glenmore can be categorized quite good.

Keyword: *multiple choice, e-learning, high order thinking*

PENDAHULUAN

Kemajuan teknologi informasi banyak membawa dampak positif bagi kemajuan dunia pendidikan, khususnya teknologi komputer dan teknologi internet, baik dalam perangkat keras maupun perangkat lunak, memberikan banyak tawaran dan pilihan bagi dunia pendidikan dalam menunjang proses pembelajaran (Uji dan Nur, 2014). Menghadapi hal tersebut dunia pendidikan harus senantiasa siap untuk menyesuaikan perkembangan teknologi terhadap peningkatan mutu pendidikan, terutama penyesuaian dalam proses pembelajaran di sekolah. Dalam proses pembelajaran ada tiga komponen yaitu tujuan, kegiatan pembelajaran, dan evaluasi. Penilaian harus mampu memberikan informasi menyeluruh yang membantu guru meningkatkan kemampuan mengajarnya dan membantu siswa mencapai perkembangan pendidikannya secara optimal (Wahyuningsih *et al.*, 2016). Terbukti dengan adanya rintisan

pelaksanaan ujian nasional (UN) menggunakan komputer atau *Computer Based Test* (CBT) pada tahunpelajaran 2014/2015(Kompas, 2015).

Salah satu tujuan mata pelajaran fisika di SMA agar peserta didik memiliki kemampuan mengembangkan kemampuan bernalar dalam berpikir analisis induktif dan deduktif dengan menggunakan konsep dan prinsip fisika untuk menjelaskan berbagai peristiwa alam dan menyelesaikan masalah, baik secara kualitatif maupun kuantitatif (BSNP, 2006). Oleh karena itu, guru harus mampu menyiapkan generasi yang memiliki kemampuan berpikir tingkat tinggi (*higher order thinking*) sehingga mereka mampu berpikir secara kritis, meneliti, memecahkan sebuah masalah, membuat keputusan, dan memiliki karakter yang baik (*good character*) secara tepat dan arif (Widihastuti dan Suyata, 2014). Kemampuan berpikir

tingkat tinggi didefinisikan sebagai penggunaan pikiran secara lebih luas untuk menemukan tantangan baru. Kemampuan berpikir tingkat tinggi ini menghendaki seseorang untuk menerapkan informasi baru atau pengetahuan sebelumnya dan memanipulasi informasi untuk menjangkau kemungkinan jawaban dalam situasi baru (Heong *et al.*, 2011). Hal ini sejalan dengan keputusan menteri pendidikan mengenai pencapaian standart penilaian yang didukung dengan adanya perubahan soal-soal untuk ujian nasional menurut rencana akan berubah mulai 2016 guna menguji kemampuan berfikir tingkat tinggi murid. Hal ini dikemukakan oleh kementerian pendidikan dan kebudayaan, Anies Rasyid Baswedan (Kompas, 2015).

Berdasarkan observasi dan wawancara yang dilakukan di SMA Negeri 1 Glenmore dapat disimpulkan bahwa pembelajaran yang dilakukan disana sudah baik. Namun untuk evaluasi belum pernah menggunakan *computer-based test* namun guru masih menggunakan *pencil test*. Selain itu, guru membuat soal-soal evaluasi hanya sampai indikator soal analisis dan guru masih belum melakukan penilaian yang dapat menguji kemampuan berfikir tingkat tinggi siswa. Format tes yang masih bersifat konvensional memiliki beberapa masalah. Dalam satu analisis menyebutkan, format tes yang masih bersifat konvensional ini memiliki beberapa masalah, salah satunya yaitu guru terkadang kurang teliti dalam proses penilaian sehingga terjadi kesalahan penilaian. Instrumen tes yang masih bersifat konvensional tentu saja kurang efektif, efisien, tidak menarik, dan tidak *up to date* (Himah *et al.*, 2016).

Guru harus memanfaatkan perkembangan teknologi saat ini untuk melakukan tes berbasis komputer sehingga guru dapat memberikan *feed back* ke siswa dengan segera. Umpan balik yang diperoleh mahasiswa dari hasil tes akan menimbulkan motivasi untuk segera memperbaiki cara belajar, menggunakan waktu belajar secara efektif, dan

sebagainya (Wijayanti dan Mundilarto, 2015). Tes pilihan ganda lebih banyak digunakan dibandingkan dengan jenis tes lainnya. Hal ini karena tes pilihan ganda memiliki kelebihan antara lain, materi yang diujikan dapat mencakup sebagian besar bahan pembelajaran, jawaban siswa dapat dikoreksi dengan mudah dan cepat, jawaban setiap pertanyaan sudah pasti benar atau salah, sehingga penilaian objektif (Sudjana, 1990:49). Selain itu, tes pilihan ganda banyak digunakan untuk mengukur kemampuan berfikir tingkat tinggi siswa yang lebih praktis dan objektif (Istyono, 2014).

Menurut Wahyuni (2016), pengembangan materi pengajaran berbasis komputer yang memanfaatkan komputer lebih fleksibilitas untuk memecahkan masalah dan bisa digunakan analisis evaluasi pembelajaran. Selain itu dengan adanya bentuk tes yang dikemas dengan memanfaatkan teknologi komputer dan jaringan internet dapat digunakan untuk membiasakan siswa dalam penggunaan evaluasi yang berbasis *e-learning* sehingga siswa akan dengan mudah dalam menghadapi ujian nasional yang menggunakan *computer-based test*.

Beberapa penelitian yang relevan dengan penelitian ini yaitu penelitian yang dilakukan oleh Rofiah (2013) tentang Penyusunan Instrumen Tes Kemampuan Berfikir Tingkat Tinggi Fisika Pada Siswa SMP yang menyatakan bahwa telah disusun instrumen tes kemampuan berfikir tingkat tinggi Fisika pada siswa SMP dengan dua paket tes, yaitu paket A dan paket tes, Budi dan Ashari (2013) tentang pengembangan sistem evaluasi kegiatan belajar mengajar Berbasis Web diperoleh hasil yaitu dapat memecahkan masalah evaluasi, salah satunya dalam proses pengolahan nilai ulangan harian sehingga guru dapat melihat ketercapaian kompetensi siswa, dan Kawuwung (2011) tentang Profil Guru, Pemahaman Kooperatif NHT, dan Kemampuan Berpikir Tingkat Tinggi Di SMP Kabupaten Minahasa Utara.

Berdasarkan wawancara yang dilakukan di SMA N 1 Glenmore, maka penelitian ini mengembangkan “Instrumen Tes *Multiple Choice High Order Thinking* Pada Pembelajaran Fisika Berbasis *E-Learning* di SMA”. Instrumen tes ini dapat melatih siswa untuk banyak belajar melakukan tes berbasis komputer dan mampu membantu dan membangun pengetahuan siswa dalam menjawab soal-soal tes. Tujuan dari penelitian ini untuk mengembangkan dan mendeskripsikan validitas, reliabilitas, analisis butir soal serta mengetahui kemampuan berfikir tingkat tinggi siswa.

METODE

Penelitian ini adalah penelitian dan pengembangan (*R&D*) menggunakan desain pengembangan 4-D oleh Thiagarajan *et al.* (1974). Penelitian ini dilaksanakan pada bulan Oktober-November 2016 di SMA Negeri 1 Glenmore. Subjek penelitian ini adalah siswa kelas X-MIPA 4 berjumlah 39 orang.

Tahap Pendefinisian (*Define*) yaitu meliputi analisis ujung depan, analisis siswa, analisis tugas, analisis konsep, dan perumusan tujuan pembelajaran yang dapat dijadikan dasar untuk perumusan soal tes. Dalam penelitian pengembangan ini, batasan materi yang ditetapkan yaitu “gerak lurus beraturan dan gerak lurus berubah beraturan”.

Tahap Perancangan (*Design*) yaitu meliputi penyusunan tes hasil belajar yang beracuan pada kemampuan berfikir tingkat tinggi pokok bahasan gerak lurus, pemilihan media, pemilihan format, dan rancangan awal.

Tahap Pengembangan (*Develop*) yaitu meliputi validitas oleh ahli dan validitas empiris, uji reliabilitas, analisis butir soal dan mengetahui kemampuan berfikir tingkat siswa. Validasi ahli dilakukan oleh dua dosen Program Studi Pendidikan Fisika FKIP Universitas Jember dan 2 orang guru SMA N 1

Glenmore. Validitas empiris dilakukan dengan membandingkan nilai tes hasil penelitian dengan nilai tes dari guru. Uji reliabilitas tes yang digunakan yaitu analisis KR-20. Analisis butir soal meliputi tingkat kesukaran soal, daya pembeda soal dan analisis pengecoh soal. Analisis kemampuan berfikir tingkat tinggi siswa dianalisis melalui hasil tes yang telah dilakukan oleh siswa.

Teknik analisis data untuk validitas ahli dihitung menggunakan persamaan: Rata-rata validator semua aspek

$$V_a = \frac{\sum_{i=1}^n A_i}{n}$$

Kriteria validitas untuk menentukan tingkat validitas instrumen tes *high order thinking* pada pembelajaran fisika berbasis *e-learning* sesuai Tabel 1.

Tabel 1. Kriteria Validitas

Kategori Validitas	Interval
Tidak Valid	$1 \leq V_a < 2$
Kurang Valid	$2 \leq V_a < 3$
Cukup Valid	$3 \leq V_a < 4$
Valid	$4 \leq V_a < 5$
Sangat Valid	$= 5$

(Hobri, 2010:52)

Teknik analisis data untuk validitas empiris dengan metode *concurrent validity*:

$$r_{xy} = \frac{\Sigma_{xy}}{\sqrt{(\Sigma_x^2)(\Sigma_y^2)}}$$

Teknik analisis data validitas item menggunakan rumus korelasi:

$$\gamma_{pbi} = \frac{M_p - M_t}{S_t} \sqrt{\frac{p}{q}}$$

Teknik analisis data reliabilitas menggunakan rumus KR-20 sebagai berikut:

$$r_{11} = \left(\frac{n}{n-1} \right) \left(\frac{s_t^2 - \sum p_i q_i}{s_t^2} \right)$$

Teknik analisis data tingkat kesukaran soal sebagai berikut:

$$P = \frac{\Sigma B}{N}$$

Teknik analisis data daya pembeda soal sebagai berikut:

$$D = \frac{B_A}{J_A} - \frac{B_B}{J_B} = P_A - P_B$$

Teknik analisis data indeks pengecoh soal sebagai berikut:

$$IP = \frac{P}{(N-B)(n-1)} \times 100\%$$

Teknik analisis kemampuan berfikir tingkat tinggi yaitu nilai maksimum adalah skor tertinggi (nilai benar = 4) dikalikan dengan jumlah soal seluruhnya (20 soal), skor maksimumnya adalah $20 \times 4 = 80$. Nilai minimumnya adalah 0, sehingga peneliti membagi interval menjadi 4 selang dengan rentang 20. Kriteria kemampuan berfikir tingkat tinggi siswa sesuai dengan Tabel 2.

Tabel 2. Kriteria HOT

Nilai siswa	Kriteria
61-80	Sangat baik
41-60	Baik
21-40	Cukup
0-20	Kurang

(Lewy, 2009)

Tahap Penyebaran (Disseminate) yaitu oleh peneliti hanya dilakukan secara terbatas di sekolah SMA N 1 Glenmore di kelas X-MIPA 4 dan sosialisasi keberadaan program *e-learning* untuk melakukan tes fisika menggunakan komputer ke kelas XMIPA 1, X MIPA 2 dan X MIPA 4 (Isworini *et al.*, 2015).

HASIL DAN PEMBAHASAN

Produk dari pengembangan pengembangan instrumen tes *multiple choice high order thinking* pada pembelajaran fisika berbasis *e-learning* di SMA terdiri dari homepage awal website, halaman registrasi, form kerjakan soal, halaman kunci jawaban, halaman untuk melihat nilai yang diperoleh, halaman untuk melihat kriteria berfikir tingkat tinggi siswa.

Validitas ahli terhadap soal yang dikembangkan berupa data kuantitatif dan data kualitatif. Data hasil validitas ahli dapat dilihat pada Tabel 3.

Tabel 3. Hasil Validitas Ahli

Aspek Validitas	Tingkat Validitas
Kajian Instruksional	
Kesesuaian	
Keefektifan	3,9
Kelayakan	CukupValid
Kajian teknis	
Format	4,1
Bahasa	Valid
Rata-rata	4,0

Berdasarkan masukan yang diterima dari para validator data validitas diperoleh nilai sebesar 4,0 sehingga instrumen tes *multiple choice high order thinking* memenuhi kriteria valid. Soal dikatakan valid atau memiliki validitas tinggi merupakan soal yang dapat mengukur kompetensi yang diharapkan. Sedangkan soal yang tidak valid atau memiliki validitas rendah artinya soal tersebut tidak bisa mengukur kompetensi yang diharapkan (Rahmani *et al.*, 2015).

Hasil dari validasi empirik dalam penelitian ini dijadikan sebagai acuan untuk melihat keterlaksanaan ujian menggunakan komputer dengan instrumen tes *multiple choice high order thinking* yang telah dikembangkan. Hasil dari validitas empirik yang dihasilkan adalah $r_{xy} = 0,92$ memenuhi kriteria validitas tinggi. Uji validitas empiris yang digunakan adalah validitas ada sekarang (*concurrent validity*) yaitu membandingkan hasil evaluasi yang dilakukan oleh guru dengan hasil tes berfikir tingkat tinggi. Menurut Anas (2003:177), tinggi rendahnya koefisien korelasi yang diperoleh menunjukkan tinggi rendahnya validitas tes yang akan kita nilai kualitasnya.

Data hasil validitas item terdapat kriteria sangat valid, valid, cukup valid dan

kurang valid. Pada paket A terdapat kategori sangat valid dan sangat valid 25%, cukup valid 30% dan kurang valid 20%. Pada paket B terdapat 20% kategori sangat valid, 20% kategori valid, 30% kategori cukup valid dan 30% kategori kurang valid. Menurut Marthunis *et al.* (2015), butir soal yang tidak valid sebaiknya diperbaiki dan butir soal yang valid dapat digunakan kembali. Soal yang telah dinyatakan valid harus dipertahankan dengan cara mendokumentasikan soal tersebut kedalam bank soal. butir soal yang tidak valid sebaiknya diperbaiki dengan cara meningkatkan penguasaan teknik peneliti dalam penyusunan butir soal. Berdasarkan data hasil validitas item dapat disimpulkan bahwa soal pada penelitian ini termasuk soal yang cukup baik ditinjau dari tingkat validitasnya.

Uji reliabilitas soal dengan menggunakan uji reliabilitas KR-20. Data hasil reliabilitas ini memenuhi kriteria reliabilitas dengan kategori sedang atau cukup reliabel yaitu pada paket soal A dengan nilai 0,6 dan pada soal paket B 0,7. Nilai $0,6 < 0,70$ maka tes dikatakan mendekati reliabel (Suharti, 2017). Menurut Marthunis *et al.* (2015), instrumen yang reliabel mengandung makna bahwa instrumen yang digunakan cukup mantap untuk digunakan dalam mengambil data penelitian, sehingga mampu mengungkap data yang dapat dipercaya hasilnya.

Selanjutnya analisis butir soal meliputi analisis tingkat kesukaran soal, daya pembeda soal, dan analisis pengecoh. Data tingkat kesukaran soal di paket A memiliki 25% soal mudah, 75% soal kategori sedang. Data tingkat kesukaran soal di paket B memiliki 50% soal kategori mudah, 5% soal kategori sukar dan 45% soal kategori sukar. Dari hasil data tingkat kesukaran soal memiliki kualitas soal kurang baik karena terdapat proporsi soal mudah lebih banyak dibandingkan soal sedang dan sulit sehingga belum memiliki keseimbangan. Menurut Diana *et al.*

(2017), taraf kesukaran yang seimbang berarti 25% untuk soal dengan kriteria mudah, 50% untuk soal sedang, dan 25% untuk soal sukar. Skala penyusunan tersebut dijadikan salah satu instrument evaluasi. Butir soal yang baik merupakan butir soal yang tidak terlalu sulit dan juga tidak terlalu mudah atau soal yang baik merupakan soal yang mempunyai tingkat kesukaran sedang (Arikunto, 2011:207).

Data soal tes pada paket A dengan kategori daya pembeda sangat baik terdapat 10 %, Daya pembeda dengan kategori cukup terdapat 40%, dan daya pembeda soal dengan kategori jelek terdapat 45%. Soal tes pada paket B dengan kategori daya pembeda sangat baik terdapat 5%, daya pembeda dengan kategori baik terdapat 45%, daya pembeda dengan kategori baik terdapat 25% dan daya pembeda soal dengan kategori jelek terdapat 25%. Daya pembeda soal adalah kemampuan suatu soal untuk membedakan antara siswa yang pandai (berkemampuan tinggi) dengan siswa yang berkemampuan rendah, (Arikunto, 2011:211). Untuk menganalisis daya pembeda soal siswa dibedakan menjadi 2 kelompok, yaitu kelompok atas dan kelompok bawah. Berdasarkan analisis data daya pembeda soal yang telah dilakukan didapatkan kategori soal baik, cukup dan jelek. Daya diskriminasi yang baik memang pada umumnya terdapat pada item yang tidak terlalu mudah dan juga tidak terlalu sukar item yang memiliki indeks diskriminasi jelek dapat langsung dibuang, sedangkan item lainnya dapat ditelaah lebih lanjut untuk direvisi, (Alwi, 2012). Menurut Sarea *et al.* (2015), ada beberapa penyebab suatu butir memiliki daya pembeda yang rendah antara lain; soal yang mengandung bias, soal yang terlalu sulit, dan distraktor yang tidak masuk akal.

Data hasil analisis pengecoh soal menunjukkan kriteria baik, cukup, dan kurang baik. Kriteria baik jika 3 distraktor berfungsi dengan baik terdapat 25% pada paket A dan 25% pada paket B, kriteria

cukup jika 2 distraktor berfungsi yaitu 55% pada paket A dan 25% pada paket B. Jika distraktor yang berfungsi hanya 1 tergolong kriteria kurang baik, pada paket A terdapat 20% dan pada paket B 50%. Suatu distraktor dapat dikatakan berfungsi baik jika paling sedikit dipilih oleh 1 orang dari 19 orang atau 5% dari pengikut tes. Menurut Elvira dan Hadi (2016), pengecoh yang berfungsi dengan baik yaitu sekurangnya dipilih sebanyak 5% dari seluruh peserta tes. Dari hasil analisis data banyak terdapat banyak soal dengan pengecoh yang dipilih secara tidak merata. Efektivitas distraktor yang dapat berfungsi dengan baik akan menjadikan butir soal juga baik. Soal yang baik jika besarnya persentasi distraktor adalah 1 per jumlah pilihan alternatif yang disediakan (Sarea *et al.*, 2015). Butir soal yang masuk dalam kategori efektifitas pengecoh kurang baik sebaiknya harus memperbaiki dengan memperhatikan pengecoh agar dapat berfungsi dengan baik.

Tabel 4. Hasil Kemampuan Berfikir Tingkat Tinggi

Interval Skor	Frekuensi	Presentasi	Kategori
61-80	6	15,8%	Sangat Baik
41-60	16	42,1%	Baik
21-40	16	42,1%	Cukup

Tabel 4 merupakan hasil analisis data kemampuan berfikir tingkat tinggi siswa yang terbagi menjadi 3 kategori yaitu sangat baik, baik dan cukup. Menurut Lewy (2009:19), siswa dapat dikategorikan memiliki kemampuan berfikir tingkat tinggi yang sangat baik jika siswa dapat mengerjakan soal C4, C5 dan C6 maka siswa mendapat poin 4 untuk setiap soal yang dijawab dengan benar. Kategori kemampuan berfikir tingkat tinggi baik jika siswa dapat mengerjakan C4 dan C5 maka siswa mendapat poin 3 untuk setiap soal yang dijawab dengan benar. Sedangkan kategori kemampuan berfikir tingkat tinggi siswa adalah cukup jika siswa hanya dapat mengerjakan soal C4

saja dan memperoleh poin 2 untuk setiap soal yang dijawab benar. Jadi dapat disimpulkan siswa di SMA N 1 Glenmore memiliki kemampuan berfikir tingkat tinggi cukup baik.

SIMPULAN DAN SARAN

Berdasarkan hasil penelitian dan pembahasan pengembangan instrumen tes *multiple choice high order thinking* pada pembelajaran fisika berbasis *e-learning* di SMA dapat disimpulkan: (1) data hasil validasi ahli memenuhi kriteria valid. Validitas empirik yang didapat memenuhi kriteria validitas sangat tinggi. Hasil uji reliabilitas pada produk menunjukkan bahwa paket soal A dan paket soal B menunjukkan kategoricukup reliabel; (2) analisis butir soal yaitu tingkat kesukaran paket A dan B masih memerlukan perbaikan. Pada analisis daya pembeda terdapat 27 soal layak digunakan dan 13 soal kategori tidak layak karena memiliki daya pembeda jelek. Pada analisis pengecoh ada 16 soal paket A dan 10 soal paket B dengan distraktor diterima karena cukup baik dan ada 14 soal dengan distraktor direvisi karena kurang baik; dan (3) kemampuan berfikir tingkat tinggi siswa setelah menggunakan Instrumen tes *multiple choice high order thinking* pada pembelajaran fisika berbasis *e-learning* di SMA dapat dikategorikan cukup baik.

Berdasarkan penelitian dan pengembangan yang dilakukan, saran yang dapat diberikan antara lain: (1) sebelum pelaksanaan uji coba produk perlu dilakukan pengecekan terhadap setiap komputer; (2) instrumen yang akan digunakan, diucicobakan ke banyak sekolah dengan materi yang berbeda pula untuk mengetahui tingkat keefektifan produk; dan (3) bagi peneliti lain, sebaiknya penelitian pengembangan ini juga disertai pembahasan pada setiap butir soal agar produk instrumen tes *multiple choice high order thinking* pada pembelajaran fisika berbasis *e-learning*

di SMA dapat digunakan sebagai fasilitas untuk mengembangkan kemampuan berfikir tingkat tinggi siswa.

DAFTAR PUSTAKA

- Alwi, I. 2012. Kriteria Empirik Dalam Menentukan Ukuran Sampel Pada Pengujian Hipotesis Statistika Dan Analisis Butir. *Jurnal Formatif*. 2(2):140-148.
- Anas, S. 2003. *Pengantar Evaluasi Pendidikan*. Jakarta: PT Raja Grafindo Persada.
- Arikunto, S. 2011. *Dasar-dasar Evaluasi Pendidikan*. Jakarta : Bumi aksara.
- Budi, E. S. & Ashari, A. 2013. Pengembangan Sistem Evaluasi Kegiatan Belajar Mengajar Berbasis Web Studi Kasus:di SMA Negeri 1 Surakarta. *Indonesian Journal of Computing and Cybernetics Systems*. 7(2):199-208.
- Diana, G. U. S., Bagus,I. P. dan Nengah, I. M. 2017. Kualitas Butir Soal Ulangan Akhir Semester Ganjil Bahasa Indonesia Kelas Ix Smp Negeri 2 Singaraja Tahun Pelajaran 2016/2017 Ditinjau Dari Segi Taraf Kesukaran, Daya Beda, Dan Fungsi Pengecoh. *Jurnal Jurusan Pendidikan Bahasa dan Sastra Indonesia, Undiksha*. 7(2):1-12.
- Elvira, M. & Hadi, S. 2016. Karakteristik Butir Soal Ujian Semester Dan Kemampuan Siswa SMA Di Kabupaten Muaro Jambi. *Jurnal Evaluasi Pendidikan*. 4(1):58-68.
- Heong, Y.M., Othman, W.D., Md Yunos, J., Kiong, T.T., Hasan, R., & Mohamad, M.M. 2011. The Level of Marzano Higher Thinking Skills Among Technical Education Students. *International Journal Of Social and Humnity*. 1(2):121-125.
- Himah, F., Sudarti dan Subiki. Pengembangan Instrumen Tes *Computer Based Test Higher Order Thinking (Cbt-Hot)* Pada Mata Pelajaran Fisika Di SMA. *Jurnal Pembelajaran Fisika*. 5(1):89-95.
- Istyono, E. 2014 . Pengembangan Tes Kemampuan Berpikir Tingkat Tinggi Fisika (Pysthots) Peserta Didik SMA. *Jurnal penelitian dan evaluasi pendidikan*. (1):1-12.
- Isworini, W. Sunarmo, dan S. Saputro. 2015. Pengembangan Modul Pembelajaran Hidrolisis Garam Berbasis Model Inkuiri Terbimbing (*Guidedinquiry*) untuk Siswa Madrasah Aliyah Kelas XI. *Jurnal Inkuiri*. 4(3):9-20.
- Kawuwung, F. 2011. Profil Guru, Pemahaman Kooperatif NHT, dan Kemampuan Berpikir Tingkat Tinggi Di SMP Kabupaten Minahasa Utara. *Jurnal El-hayah*. 1(4):157-166.
- Kompas. 2015. *Soal UN Berubah Tahun 2016*.
- Lewy. 2009. Pengembangan Soal Untuk Mengukur Kemampuan Berpikir Tingkat Tinggi Pokok Bahasan Barisan Dan Deret Bilangan Di Kelas IX Akselerasi Smp Xaverius Maria Palembang. *Jurnal Pendidikan matematika*. 3(2):14-28.
- Marthunis, Khaldun, I & Zulfadli. 2015. Analisis Kualitas Butir Soal Ujian Semester Genap Mata Pelajaran Kimia Kelas X MAN Model Banda Aceh Tahun Pelajaran 2014/2015 Menggunakan Program Proanaltes. *Jurnal Ilmiah Mahasiswa Pendidikan Kimia (JIMPK)*. 1(4):70-78.

- Munir. 2012. *Kurikulum Berbasis Teknologi Informasi Dan Komunikasi*. Bandung: Alfabeta.
- Rofiah, E. 2013. Penyusunan Instrumen Tes Kemampuan Berpikir Tingkat Tinggi Fisika Pada Siswa Smp. *Jurnal pendidikan indonesia*. 1(2):17-22.
- Rahmani, M. N., Kurnia dan Nurdini A. 2015. Analisis Kualitas Butir Soal Buatan Guru Biologi Kelas X SMA Negeri 1 Tanah Pinoh. *Jurnal Pendidikan dan Pembelajaran*. 4(2):1-16.
- Sarea, M., Sayahrul & Hadi, S. 2015. Analisis Kualitas Soal Ujian Akhir Semester Mata Pelajaran Kimia SMA Di Kabupaten Gowa. *Jurnal Evaluasi Pendidikan*. 3(1):35-43.
- Sudjana, N. 1990. *Penilaian Hasil Belajar Mengajar*. Bandung: PT Remaja Rosdakarya.
- Suharti, S. 2017. Kualitas Tes Bahasa Arab dan Prestasi Peserta Didik Madrasah Tsanawiyah Kabupaten Bantul (Analisis Butir Soal UAMBN Tahun Ajaran 2013/2014). *Jurnal Pendidikan Madrasah*. 2(1):185-195.
- Thiagarajan, S., D. S. Semmel, dan M. I. Semmel. 1974. *Instructional Development for Training Teacher of Exceptional Children A Sourcebook*. Bloomington: Indiana University.
- Uji, R. C. & Nur B. 2014. Pengembangan Media Evaluasi Pembelajaran Dalam Bentuk Online Berbasis E-Learning Menggunakan Software Wondershare Quiz Creator Dalam Mata Pelajaran Akuntansi Sma Brawijaya Smart School (BSS). *Jurnal Pendidikan Akuntansi Indonesia*. 12(1):41-48.
- Wahyuni, S. 2016. Development of Computer Assisted Instruction (CAI) Based Teaching Materials in Junior High School. *International Journal of Learning and Teaching*. 1(2):117-120.
- Wahyuningsih, R., S. Wahyuni, dan A. D. Lesmono. 2016. Pengembangan Instrumen *Self Assessment* Berbasis Web untuk Menilai Sikap Ilmiah pada Pembelajaran Fisika di SMA. *Jurnal Pembelajaran Fisika*. Vol. 4(4):338-343.
- Widihastuti & Suyata. 2014. Model Afl Untuk Meningkatkan Pemahaman Dan Higher Order Thinking Skills Mahasiswa Vokasi Bidang Busana. *Jurnal penelitian dan evaluasi pendidikan*. (2):275-289.
- Wijayanti, E. & Mundilarto. 2015. Pengembangan Instrumen Asesmen Diri Dan Teman Sejawat Kompetensi Bidang Studi Pada Mahasiswa. *Jurnal Penelitian dan Evaluasi Pendidikan*. 19(2):129-144.