

## Penilaian *Higher Order Thinking Skills* (HOTS) Menggunakan Quizizz Pada Materi Usaha Dan Energi

<sup>1</sup>Nabila Ayunda Putri dan <sup>2</sup>Mukhayyarotin Niswati Rodliyatul Jauhariyah

Prodi S1 Pendidikan Fisika, Jurusan Fisika, FMIPA, Universitas Negeri Surabaya Kampus Ketintang, Jalan Ketintang Gedung C3 Lt.1 Surabaya 60231

<sup>1</sup>[Nabila.17030184040@mhs.unesa.ac.id](mailto:Nabila.17030184040@mhs.unesa.ac.id) <sup>2</sup>[mukhayyarotinjauhariyah@unesa.ac.id](mailto:mukhayyarotinjauhariyah@unesa.ac.id)

Article Info	Abstract (10pt italic)
<p><b>Article History</b>            Received: 14 June 2021            Revised: 30 June 2021            Published: 30 June 2021</p> <p><b>Keywords</b>            Assessment;            HOTS;            Quizizz;            Work and Energy</p>	<p><i>The aim of this research was to produce a web based instrument test to assess higher order thinking skills that are feasible and to describe the profile of students higher order thinking skills. This research using Research and Development (R&amp;D) model from Borg and Gall with ADDIE (Analysis, Design, Development, Implementation, and Evaluation ). Data collection techniques in this study are the validation method that done by three lecturers from the physic department of State University of Surabaya and test method that done by 29 students of class X MIPA-2 SMA At-Ta'miriyyah using Quizizz. The data that has been collected were analyzed using quantitative descriptive analysis method. The percentage of average theoretical validity in the realms of construction, language, and media is is 97.03% or in the very feasible category. The result of product trials on 29 students analyzed the level of distinguishing power, reliability, and item validity. From the analysis obtained empirical validity as many as 10 questions declared empirically feasible. The appropriate test instrument is then used to describe the profile of students higher order thinking skills abilities. A viable test instrument is then used to describe the student's high-level thinking ability profile. 44.83% of students in very high categories, 13.79% with good categories, 0% with sufficient categories, and 41.38% with less categories. The ability to think high level in the competency of analyzing (C4) is lower at 67.49% compared to the evaluating competency (C5) of 80.46%.</i></p>
Informasi Artikel	Abstrak
<p><b>Sejarah Artikel</b>            Diterima: 16 Juni 2021            Direvisi: 30 Juni 2021            Dipublikasi: 30 Juni 2021</p> <p><b>Kata kunci</b>            Penilaian;            HOTS;            Quizizz;            Usaha dan Energi.</p>	<p>Penelitian ini bertujuan menghasilkan instrumen untuk menilai kemampuan berpikir tingkat tinggi yang layak, serta mendeskripsikan profil kemampuan berpikir tingkat tinggi siswa. Penelitian ini menggunakan model Research &amp; Development (R&amp;D) dari Borg dan Gall dengan desain pengembangan ADDIE (<i>Analysis, Design, Development, Implementation, and Evaluation</i>). Instrumen tes yang dikembangkan divalidasi oleh dosen ahli instrumen, Teknik pengumpulan data pada penelitian ini yaitu dengan metode validasi yang dilakukan oleh tiga dosen dari jurusan Fisika Universitas Negeri Surabaya dan metode tes menggunakan Quizizz yang dilakukan terhadap 29 siswa kelas X MIPA-2 SMA At-Ta'miriyyah Surabaya. Data yang telah terkumpul dianalisis dengan metode analisis deskriptif kuantitatif. Persentase validitas teoritis rata-rata pada ranah konstruksi, bahasa, dan media sebesar 97,03% atau berada pada kategori sangat layak. Hasil ujicoba produk pada 29 siswa dianalisis daya pembeda, reliabilitas, dan validitas itemnya. Dari analisis tersebut diperoleh validitas empiris sebanyak 10 soal dinyatakan layak secara empiris. Instrumen tes yang telah layak kemudian digunakan untuk mendeskripsikan profil kemampuan berpikir tingkat tinggi siswa. 44.83% siswa dengan kategori sangat tinggi, 13.79% dengan kategori baik, 0% dengan kategori cukup, dan 41.38% dengan kategori kurang. Kemampuan berpikir tingkat tinggi pada kompetensi menganalisis (C4) lebih rendah yaitu 67.49% dibandingkan pada kompetensi mengevaluasi (C5) sebesar 80.46%.</p>

**Sitasi:** Putri, N. A. dan Jauhariyah, M. N. R. (2021), Penilaian *Higher Order Thinking Skills* (HOTS) Menggunakan *Quizizz* Pada Materi Usaha Dan Energi. *Kappa Journal*. 5(1), 88-101.

## PENDAHULUAN

Proses penilaian dalam suatu pembelajaran sangat penting dilakukan untuk mengukur kemampuan berpikir siswa dan untuk mengetahui ketercapaian dari tujuan suatu pembelajaran. Penilaian dapat dilaksanakan selama kegiatan belajar atau setelah pembelajaran. Arifin (2013) memaparkan penilaian adalah suatu proses atau kegiatan yang sistematis dan berkesinambungan untuk mengumpulkan data tentang proses dan hasil belajar siswa dalam rangka membuat keputusan-keputusan berdasarkan kriteria dan pertimbangan tertentu. Penilaian kompetensi siswa yang dituntut dalam kurikulum 2013 salah satunya adalah kompetensi pengetahuan. Kesuksesan pendidikan tercapai dengan baik apabila didukung penilaian yang baik pula dan memberikan dampak dalam proses pembelajaran selanjutnya (Desiriah & Setyarsih, 2021). Setiadi (2016) menyatakan ketepatan dalam menentukan metode penilaian akan berdampak pada objektivitas dan validitas hasil suatu penilaian sedangkan kesalahan dalam memilih metode penilaian dapat memberikan informasi yang tidak tepat terhadap kualitas pembelajaran.

Penilaian yang dilakukan oleh guru dalam kelas juga berfungsi untuk mengembangkan kemampuan berpikir tingkat tinggi siswa. Penilaian pembelajaran dapat dilaksanakan melalui berbagai macam instrumen tes berupa soal-soal maupun instrumen non tes berupa angket dan lembar observasi. Penyusunan instrumen penilaian yang mengacu pada kemampuan berpikir tingkat tinggi sangat perlu dikembangkan dan diujikan pada siswa terutama dijenjang Sekolah Menengah Atas karena kebanyakan sekolah masih belum memperkenalkan soal HOTS kepada siswa, sehingga profil HOTS yang dimiliki siswa masih lemah. Hal tersebut terlihat dari hasil Ujian Nasional (UN) fisika tahun 2019 mengalami penurunan rerata nilai dibandingkan dengan tahun 2018. Dilansir dari *website* Pusat Penilaian Pendidikan bahwa rerata nilai UN pada tahun 2018 sebesar 51,00 sedangkan pada tahun 2019 sebesar 45,80 (Puspendik, 2019). Hal ini menjadikan Kementerian Pendidikan dan Kebudayaan melalui Direktorat Jendral Guru dan Tenaga Kependidikan berupaya untuk meningkatkan kualitas pembelajaran yang berujung pada peningkatan kualitas siswa dengan menyelenggarakan Program Peningkatan Kompetensi Pembelajaran (PKP), salah satu materi yang dikembangkan yaitu Penilaian Berbasis HOTS (Nugroho, 2018).

Dalam kaitannya dengan pengembangan kemampuan berpikir tingkat tinggi, pembelajaran kontekstual menekankan pembelajaran di kelas dengan kehidupan nyata, yang ada di sekitar siswa. Berpikir tingkat tinggi menuntut proses berpikir non rutin, yang akan mudah ditumbuhkembangkan jika pembelajarannya berkaitan dengan kasus nyata yang ada dalam kehidupan sehari-hari. Sehingga Dari pembelajaran tersebut diharapkan siswa dapat terlibat dalam permasalahan yang menuntut untuk berfikir kritis, kreatif, dan komunikatif (Kivunja, 2015). Sehingga salah satu tuntutan guru saat ini adalah berupaya membantu siswa memiliki kemampuan berpikir tingkat tinggi atau *Higher Order Thinking Skill* (HOTS) (Nisa, 2018). Berpikir tingkat tinggi merupakan berpikir yang lebih dari sekedar mengingat fakta dan menekankan pada aplikasi sehingga siswa mengkonstruksi pengetahuannya (Angraini & Sriyati, 2019). Wijaya dan Suyono (2019) menyebutkan berpikir tingkat tinggi adalah pola berpikir yang menekankan pada makna yang terdapat di dalamnya tidak hanya hafalan saja. Menurut Witri et al., (2019) berpikir tingkat tinggi/HOTS didefinisikan sebagai berpikir yang lebih menyeluruh dan kompleks dengan tujuan untuk memperoleh solusi dari permasalahan. Berpikir tingkat tinggi meliputi kemampuan menganalisis, mengevaluasi, dan menciptakan (Nurhayati & Angraeni, 2017). Dengan demikian dapat disimpulkan bahwa kemampuan berpikir tingkat tinggi (HOTS) adalah kemampuan berpikir yang tidak sekedar mengingat fakta tetapi juga menekankan pada makna untuk memperoleh solusi dari permasalahan

dengan menganalisis (C4), mengevaluasi(C5), dan atau mencipta (C6) suatu ide atau konsep secara logis, kritis, evaluatif, kreatif, dan solutif.

*Problem solving* atau kemampuan pemecahan masalah adalah salah satu bagian dari kemampuan berpikir tingkat tinggi (HOTS) (Alfika & Mayasari, 2018). Selain itu ada pendapat seorang ahli mengenai hal serupa, yaitu *characteristics of higher order thinking skills (HOTS): higher order thinking skills encompass both critical thinking and creative thinking* (Conklin, 2011). Kemampuan berpikir kritis dan kreatif merupakan HOTS. Oleh karena itu, pembelajaran abad 21 dengan kondisi ideal pada pelajaran fisika akan terwujud jika HOTS diterapkan (Rohmah & Sunarti, 2020). Dari data data hasil evaluasi *Programme for International Students Assessment (PISA) 2012* kemampuan sains siswa Indonesia berada pada posisi 64 dari 65 negara yang dievaluasi dengan skor 382 (OECD, 2019). Selanjutnya pada tahun 2015 kemampuan Indonesia dalam sains pada peringkat 64 dari 72 negara peserta dengan skor 403. Pada tahun 2018 kemampuan sains siswa Indonesia berada pada posisi 70 dari 78 negara peserta dengan skor 396 (OECD, 2019). Dari hasil tiga survei terakhir, nilai siswa Indonesia masih jauh dari skor rata-rata internasional yang ditetapkan oleh PISA yaitu 500. Gardner (2013) memiliki pandangan bahwa pembelajaran fisika seharusnya mendorong dan memotivasi siswa agar belajar aktif dan kreatif.

Berdasarkan hasil penelitian yang dilakukan oleh Hidayah et al., (2018) tentang “*Pengembangan Tes Higher Order Thinking Skill (HOTS) Pada Materi Getaran Harmonis*” menyatakan bahwa penilaian HOTS merupakan bentuk penilaian yang menyajikan soal tes pada level kognitif tingkat tinggi kepada siswa sehingga dapat meningkatkan kemampuan berpikir kritis dan berpikir kreatif siswa. Desiriah & Setyarsih (2021) telah menjabarkan penilaian HOTS memiliki tiga prinsip utama yaitu (1) diberikan stimulus baik dalam bentuk narasi maupun bentuk lainnya, (2) diberikan permasalahan baru yang belum diberikan di kelas, (3) diberikan soal dengan tipe kesukaran yang berbeda dan level kognitif yang berbeda-beda. Dalam merumuskan indikator soal HOTS biasanya menggunakan kata kerja operasional (KKO) berdasarkan taksonomi Bloom untuk mengetahui dimensi pengetahuan yang diukur oleh soal tersebut (Permendikbud No 23, 2016).

Seiring berkembangnya teknologi, pelaksanaan tes dapat dilakukan dengan memanfaatkan internet salah satunya berbantuan media *website* atau *web*. Berbagai jenis penilaian juga dapat dikembangkan untuk memudahkan guru dan lebih menarik kemauan siswa untuk mengikuti proses penilaian pembelajaran. Adanya pembelajaran daring selama pandemi juga menunjukkan bahwa guru harus dapat memanfaatkan teknologi agar proses pembelajaran tetap berlangsung dengan baik pada masa pandemi saat ini, penilaian yang dilakukan secara daring dengan memanfaatkan *web* lebih memungkinkan dan lebih praktis untuk dilakukan dibandingkan dengan berbasis kertas.

Terdapat banyak *web* maupun aplikasi yang dapat digunakan untuk tes, salah satunya adalah menggunakan *Quizizz*. *Quizizz* adalah sebuah *web tool* untuk membuat game kuis interaktif. *Quizizz* menyediakan beberapa fasilitas yaitu membuat kuis dimana materi soal dan tingkat kesukarannya dapat disesuaikan dengan kemampuan siswa. Setelah guru membuat kuis, siswa diminta untuk memasukkan kode yang telah diberikan oleh guru dimana kode tersebut digunakan untuk masuk ke kelas yang sudah ditentukan. Menggunakan media *Quizizz*, guru dapat mengetahui sejauh mana kemampuan dan pengetahuan para siswa karena guru dapat mengontrol dan mengetahui tingkat pengetahuan dan kemampuan siswa pada setiap soal kuis, setelah itu guru juga dapat mengunduhnya sebagai lembar kerja Excel.

*Quizizz* dapat digunakan sebagai strategi pembelajaran yang baik dan menyenangkan tanpa menghilangkan esensi belajar. Pemanfaatan *Quizizz* membantu guru dalam melakukan penilaian tanpa dibatasi tempat, tampilan menarik, dan pengaturan waktu yang diatur menuntut siswa untuk konsentrasi. Berdasarkan penelitian yang dilakukan oleh Ju dan Adam (2018) terhadap keterampilan bahasa arab, siswa sangat aktif menjawab pertanyaan yang diberikan dan lebih berkonsentrasi pada topik tersebut. Hal ini menunjukkan bahwa

penggunaan *Quizizz* dapat menjadi penilaian yang bersifat menyenangkan tapi tetap belajar yang dapat menyegarkan ingatan, menarik, dan memberikan kesan positif dalam memori otak siswa (Noor, 2020)

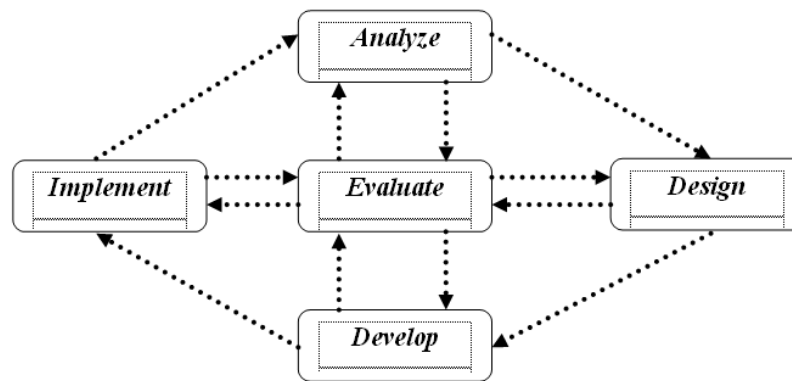
Hasil wawancara kepada siswa kelas X MIPA-2 SMA At-Ta'miriyah Surabaya menunjukkan bahwa sebagian besar siswa telah mengetahui bentuk soal sesuai indikator berpikir tingkat tinggi, namun bentuk soal-soal tersebut jarang mereka temui di kelas. Hal ini sesuai dengan informasi yang diberikan oleh guru bahwa belum ada penilaian secara khusus terhadap kemampuan berpikir tingkat tinggi siswa. Adapun tes yang dilakukan biasanya berbentuk uraian dan berbasis kertas, sehingga membutuhkan waktu cukup lama dalam pemberian skor. Sebelumnya telah banyak dilakukan penelitian terkait Pengembangan Instrumen Penilaian HOTS yang diantaranya dilakukan oleh Rohmah & Sunarti (2020), Ayumniyya & Setyarsih (2021), dan Yuliantaningrum & Sunarti (2020). Namun penelitian yang akan di lakukan ini memiliki kelebihan yakni instrumen penilaian yang dikembangkan menggunakan *Quizizz*.

Berdasarkan pemaparan di atas, penulis melakukan penelitian dengan tujuan menghasilkan sebuah instrumen tes menggunakan *Quizizz* yang layak untuk menilai kemampuan berpikir tingkat tinggi siswa SMA serta mendeskripsikan profil kemampuan berpikir tingkat tinggi siswa tersebut.

## METODE

Jenis penelitian ini merupakan penelitian pengembangan menggunakan model Research & Development (R&D) dari Borg dan Gall dengan desain pengembangan ADDIE (*Analysis, Design, Development, Implementation, Evaluation*). Tujuan utama model pengembangan ini digunakan untuk mendesain dan mengembangkan sebuah produk yang efektif dan efisien (Priyadi, 2016).

Gambar 1 menunjukkan ilustrasi desain pengembangan instrumen penilaian berbasis HOTS menggunakan ADDIE. Pada tahap *analysis* dilakukan identifikasi masalah dari laporan Puspendik mengenai hasil UN 2019 (Puspendik, 2019) bahwa siswa mampu menjawab dengan benar soal level penalaran sebesar 28%. Hal ini menunjukkan bahwa siswa harus dilatih soal berpikir tingkat tinggi (HOTS). Analisis KI dan KD untuk indikator soal juga dilakukan, yaitu KD 3.9 Menganalisis konsep energi, usaha (kerja), hubungan usaha (kerja) dan perubahan energi, hukum kekekalan energi, dan penerapannya dalam peristiwa sehari-hari. Melakukan wawancara salah seorang pengajar fisika pada semester genap tahun pelajaran 2020/2021 di SMA At-Ta'miriyah Surabaya. Selanjutnya, metode *design* adalah perencanaan instrumen penilaian HOTS meliputi perumusan kompetensi HOTS fisika yang akan dicapai dengan menganalisis soal tes asal sekolah, perumusan indikator materi usaha dan energi, dan perumusan kisi-kisi serta penyusunan soal dan kunci jawaban soal HOTS materi usaha dan energi. Rancangan produk berupa instrumen tes berpikir tingkat tinggi berbentuk 10 soal pilihan ganda. Kemudian dilanjutkan metode *development*, soal yang telah disusun selanjutnya dilakukan uji validitas kepada validator, yaitu tiga dosen ahli fisika Universitas Negeri Surabaya. Hasil validasi dosen ahli selanjutnya dianalisis dengan teknik analisis deskriptif kuantitatif berupa validitas teoritik. Validitas teoritik untuk sebuah instrumen yaitu apabila instrumen tersebut sudah dirancang dengan baik sesuai teori maupun ketentuan yang sudah berlaku dan dinyatakan valid apabila persentase validitas mencapai  $\geq 61\%$  (Riduwan, 2012). Untuk mendeskripsikan kelayakan secara teoritis terhadap perangkat tes yang dikembangkan, maka dilakukan validasi pada ranah materi, konstruksi, bahasa, dan media oleh tiga dosen dari jurusan fisika Universitas Negeri Surabaya. Pengumpulan data pada tahap validasi ini menggunakan lembar validasi.



Gambar 1. Desain Pengembangan ADDIE

Setelah dinyatakan valid secara teoritis, kemudian dilakukan tahap *implementation* yaitu hasil pengembangan diujikan terhadap 29 siswa kelas X MIPA-2 SMA At-Ta'miriyah Surabaya yang sebelumnya telah menerima materi usaha dan energi untuk mendeskripsikan kelayakan empiris yang meliputi daya pembeda, reliabilitas, dan validitas item dari instrumen tes yang dikembangkan. Tabel 1 menunjukkan kriteria validitas teoritis instrumen tes. Tabel 2 menunjukkan kriteria kategori daya pembeda. Sedangkan Tabel 3 menunjukkan kriteria kategori validitas item soal.

Tabel 1. Kriteria Validitas Teoritis Instrumen Tes

Persentase (%)	Kriteria
0-20	Sangat Lemah
21-40	Lemah
41-60	Cukup
61-80	Layak
81-100	Sangat Layak

(Riduwan, 2012)

Tabel 2. Kriteria Kategori Daya Pembeda

Nilai Daya Pembeda	Kriteria
$\geq 0,40$	Sangat Baik
0,30 - 0,39	Baik
0,20 - 0,29	Cukup
$\leq 0,19$	Rendah

(Arikunto, 2012)

Tabel 3. Kriteria Kategori Validitas Soal

Nilai $r_{xy}$	Kriteria
$0,81 < r \leq 1,00$	Sangat Tinggi
$0,61 < r \leq 0,80$	Tinggi
$0,41 < r \leq 0,60$	Cukup
$0,21 < r \leq 0,40$	Rendah
$0,00 < r \leq 0,20$	Sangat Rendah

(Arifin, 2013)



Butir soal yang sudah valid secara empiris kemudian digunakan untuk mendeskripsikan profil kemampuan berpikir tingkat tinggi siswa. Analisis profil kemampuan berpikir tingkat tinggi siswa untuk masing-masing kompetensi HOTS menggunakan rumus sebagai berikut:

$$\text{nilai siswa} = \frac{\text{Total Skor Kompetensi}}{\text{Skor maksimum item kmpetensi}} \times 100$$

Selanjutnya dipetakan ke dalam kriteria pencapaian dalam Kurikulum 2013 untuk menentukan kriteria profil HOTS siswa sesuai Tabel 5.

Tabel 5. Kriteria Profil HOTS siswa

Nilai	Kriteria
86 – 100	Sangat Baik
71 – 85	Baik
56 – 70	Cukup
0 – 55	Kurang

(Permendikbud No 53, 2015)

Tahap terakhir adalah *evaluation* dilakukan di setiap tahap model pengembangan ADDIE.

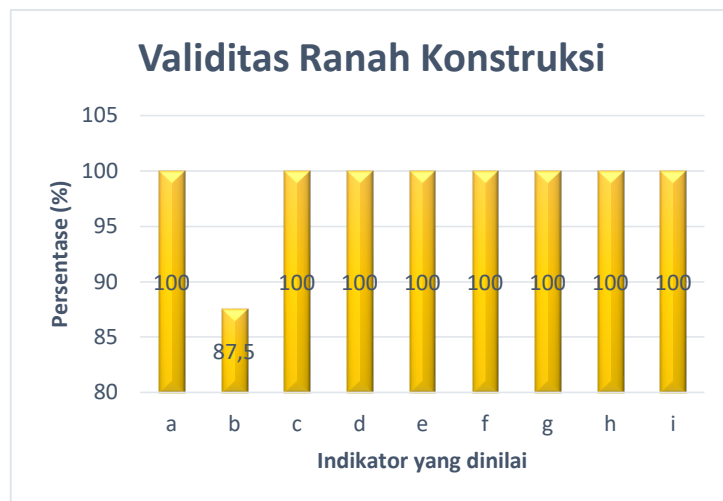
## HASIL DAN PEMBAHASAN

Latar belakang pengembangan instrumen penilaian HOTS fisika menggunakan *Quizizz* disebabkan oleh tuntutan kompetensi Kurikulum 2013 yang selaras dengan kompetensi HOTS di samping itu menggunakan *Quizizz* merujuk pada pentingnya penilaian yang efisien terutama dimasa pandemi. Berbagai tes seperti tes masuk perguruan tinggi sekarang mulai tidak menggunakan kertas (*PBT/Paper Based Test*) tetapi menggunakan computer atau *website* (*CBT/Computer Based Test*). Instrumen ini mengembangkan 10 soal pilihan ganda yang disusun sesuai kisi-kisi setiap butir soal. Diantara kelebihan tes pilihan ganda yaitu materi yang diujikan dapat mencakup lebih banyak materi pembelajaran, jawaban siswa dapat dikoreksi guru dengan cepat dan mudah, jawaban setiap pertanyaan sudah pasti benar atau salah, sehingga penilaian objektif (Sudjana, 2009). Selain itu, tes pilihan ganda sudah banyak digunakan untuk mengukur kemampuan berfikir tingkat tinggi siswa yang lebih praktis dan objektif (Istiyono et al., 2014). Butir soal yang dikembangkan berdasarkan ranah kognitif HOTS yang meliputi tujuh soal menganalisis (C4) pada nomor (1,2,4,5,6,8,10) dan tiga soal mengevaluasi (C5) pada nomor (3,7,9). Pada pengembangan instrumen penilaian ini tidak mengembangkan soal mencipta (C6) dikarenakan untuk mengembangkan soal C6 menggunakan tes uraian. Soal HOTS yang telah disusun kemudian divalidasi oleh tiga validator baik secara konstruksi, bahasa, dan media. Revisi desain dilakukan sesuai saran dan masukan validator hingga menghasilkan *draft* final yang dapat diterapkan pada 29 siswa yang kemudian diberikan penilaian untuk validitas teoritisnya.

Indikator penilaian pada ranah konstruksi yang dinilai oleh validator yaitu: (a) Butir soal yang dibuat sesuai dengan materi yang ada dan tidak mengandung materi lain yang tidak relevan; (b) Butir soal sesuai dengan tujuan High Order Thinking Skills (HOTS); (c) Pernyataan (pertanyaan dan jawaban) sesuai dengan indikator yang digunakan; (d) Pernyataan (butir soal) dirumuskan dengan singkat; (e) Kalimat yang digunakan bebas dari pernyataan yang tidak relevan; (f) Kalimat yang digunakan bebas dari pernyataan yang dapat

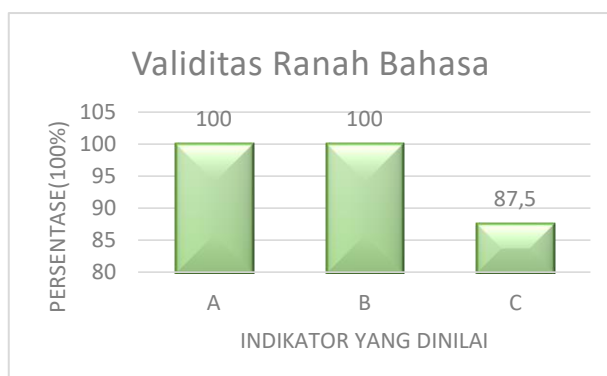
diinterpretasikan lebih dari satu makna; (g) Jumlah butir soal tidak menjemukan responden; (h) Sesuai dengan level siswa; dan (i) Kejelasan gambar atau tabel yang ditampilkan

Rata-rata hasil penilaian validator dalam persentase untuk masing-masing indikator pada ranah materi disajikan oleh diagram pada Gambar 2. Dari keseluruhan indikator yang dinilai pada ranah konstruksi diperoleh persentase rata-rata sebesar 98,61% dengan kriteria sangat layak.



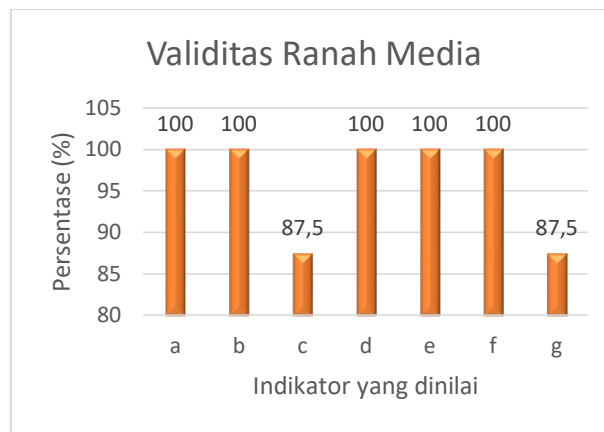
Gambar 2. Diagram Persentase Validitas pada Ranah Konstruksi

Indikator penilaian pada ranah bahasa yang dinilai oleh validator yaitu: (a) Ragam bahasa yang digunakan komunikatif dan sesuai jenjang pendidikan responden; (b) Bahasa yang digunakan adalah Bahasa Indonesia baku dan mudah dipahami; dan (c) Kata-kata yang digunakan singkat dan lugas. Rata-rata hasil penilaian validator dalam persentase untuk masing-masing indikator pada ranah bahasa disajikan oleh diagram pada Gambar 3. Dari keseluruhan indikator yang dinilai pada ranah konstruksi diperoleh persentase rata-rata sebesar 95,83% dengan kriteria sangat layak.



Gambar 3. Diagram Persentase pada Ranah Bahasa

Selanjutnya indikator penilaian pada ranah media yang dinilai oleh validator yaitu: (a) Kontribusi dalam memudahkan penilaian; (b) Kualitas dalam memotivasi siswa; (c) Kualitas tes dan penilaian yang digunakan; (d) Keterbacaan; (e) Mudah digunakan; (f) Kualitas tampilan dan (g) Kemenarikan. Rata-rata hasil penilaian validator dalam persentase untuk masing-masing indikator pada ranah media disajikan oleh diagram pada Gambar 4. Dari keseluruhan aspek yang dinilai pada ranah bahasa diperoleh persentase rata-rata sebesar 96,63 % dengan kriteria sangat layak.



Gambar 4. Diagram Persentase Validitas pada Ranah Media

Riduwan (2012) menyatakan bahwa instrumen tes dinyatakan valid secara teoritis apabila memiliki persentase kelayakan  $>61\%$ . Dengan demikian, instrumen tes yang telah dikembangkan dapat dinyatakan sangat layak, karena dari keseluruhan hasil penilaian validator pada ranah konstruksi, bahasa, dan media memiliki persentase kelayakan rata-rata sebesar  $97,02\%$  atau berada pada kategori sangat layak. *Draft* final yang telah layak secara teoritis tersebut selanjutnya diujicobakan kepada siswa.

Dari hasil ujicoba kepada siswa, dapat dianalisis validitas empiris dari instrumen tes yang dikembangkan yang meliputi daya pembeda, reliabilitas, dan validitas item. Tabel 6 adalah tabel hasil perhitungan daya pembeda soal. Berdasarkan Tabel 6, untuk mengetahui besar nilai dari daya pembeda soal dapat dilihat pada kolom *Corrected Item-Total Correlation*. Pada kolom tersebut nilai daya pembeda soal sebesar  $\geq 0,5$  yang berarti kriteria daya pembeda soal sangat baik untuk keseluruhan soal. Soal yang memiliki daya pembeda dengan kriteria baik memiliki kemampuan yang baik untuk membedakan antara siswa yang telah memahami kompetensi dengan siswa yang belum memahami (Adawiyah & Wisudawati, 2017). Sedangkan soal dengan daya pembeda yang kurang baik tidak mampu membedakan antara siswa yang telah memahami kompetensi dengan yang belum memahami. Indrawati (2018) menjelaskan bahwa soal yang memiliki daya pembeda tidak baik, maka hasil pengerjaan soal antara siswa kelompok atas dan siswa kelompok bawah akan menghasilkan skor yang sama-sama tinggi atau sama-sama rendah.

Tabel 6. Perhitungan Daya Pembeda Item Menggunakan SPSS

Item-Total Statistics				
	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
SOAL1	13.59	33.966	.837	.739
SOAL2	13.93	35.352	.555	.754
SOAL3	13.48	35.259	.684	.751
SOAL4	13.45	35.256	.739	.751
SOAL5	13.55	34.399	.782	.744
SOAL6	13.45	35.113	.772	.749
SOAL7	13.48	35.544	.623	.754
SOAL8	13.52	35.401	.615	.753
SOAL9	13.45	35.828	.609	.756
SOAL10	13.72	34.778	.627	.749
TOTALSKOR	7.14	9.695	1.000	.894

Menurut Arikunto (2012) untuk dinyatakan reliabel instrumen tes tidak harus sama persis pada setiap pengukuran, melainkan tes harus mampu menunjukkan persamaan



kedudukan siswa diantara anggota kelompok yang lain. Tabel 7 menunjukkan hasil dari perhitungan reliabilitas instrumen.

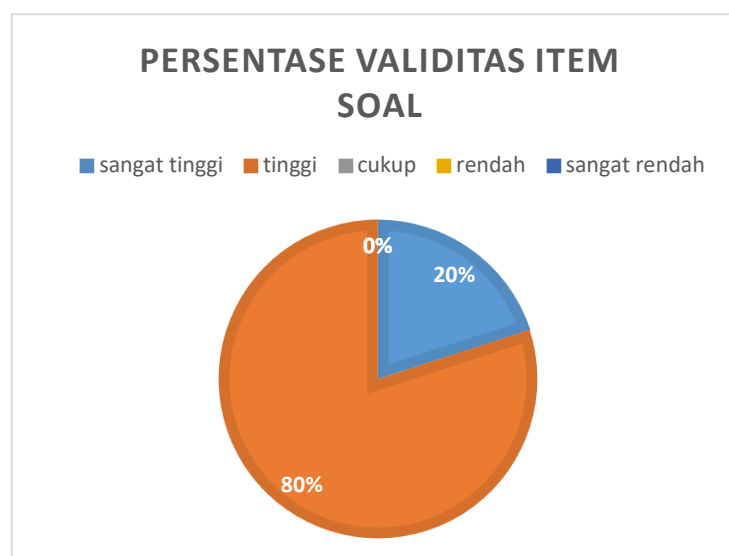
Tabel 7. Perhitungan Reliabilitas Item Menggunakan SPSS  
**Reliability Statistics**

Cronbach's Alpha	Part 1	Value	.836
		N of Items	5 <sup>a</sup>
	Part 2	Value	.762
		N of Items	5 <sup>b</sup>
	Total N of Items		10
Correlation Between Forms			.838
Spearman-Brown Coefficient	Equal Length		.912
	Unequal Length		.912
Guttman Split-Half Coefficient			.909

a. The items are: SOAL1, SOAL2, SOAL3, SOAL4, SOAL5.

b. The items are: SOAL6, SOAL7, SOAL8, SOAL9, SOAL10.

Reliabilitas dihitung menggunakan *split-half* karena soal berupa pilihan ganda dengan skor benar 1 dan salah 0 kemudian dikonsultasikan dengan tabel r produk moment. Berdasarkan Tabel 7 di atas nilai koefisien *Guttman Split-Half* dari instrumen yang sudah dikembangkan sebesar 0.909. Nilai tersebut kemudian dibandingkan dengan r tabel taraf signifikansi 5% untuk N=29, yang mana diperoleh nilai r tabel sebesar 0,367. Sugiyono (2013) menjelaskan bahwa instrumen tes dinyatakan reliabel jika nilai r hitung lebih besar dari r tabel. Instrumen yang telah dikembangkan memiliki nilai r hitung= 0.909 yang lebih besar dari nilai r tabel= 0,367, sehingga dapat disimpulkan bahwa instrumen yang telah dikembangkan dinyatakan reliabel. Melalui perhitungan menggunakan teknik *pearson correlation* diperoleh nilai validitas item setiap butir soal. Persentase validitas item butir soal berdasarkan kriterianya direpresentasikan oleh diagram pada Gambar 5.



Gambar 5. Diagram Persentase Validitas Item Soal

Dari Gambar 5 dapat dilihat bahwa sebanyak 20% dari 10 soal yang telah dikembangkan memiliki validitas item dengan kategori sangat tinggi, kemudian 80% berada pada kategori tinggi, dan tidak ada soal yang memiliki validitas item cukup atau rendah atau

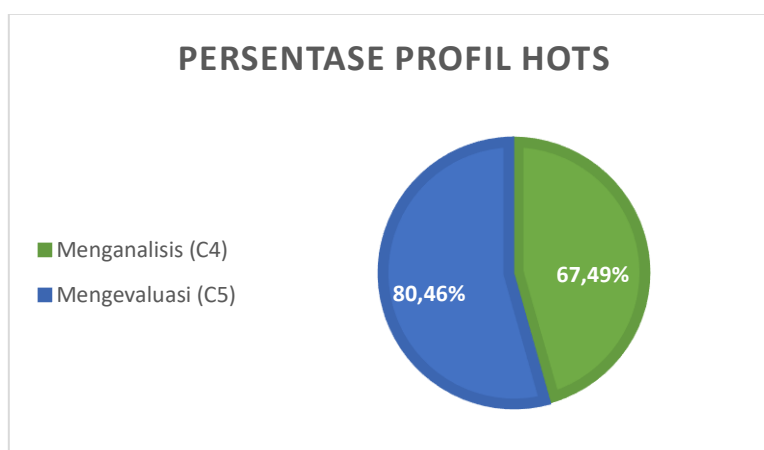
sangat rendah. Butir soal yang valid atau memiliki validitas tinggi menurut Rahmadani et al., (2018) adalah soal yang dapat mengukur kompetensi yang diharapkan. Sedangkan soal dikatakan tidak valid atau memiliki validitas rendah merupakan soal yang tidak bisa mengukur kompetensi yang diharapkan.

Berdasarkan analisis daya pembeda dan validitas item diperoleh 100% dari 10 soal yang dinyatakan valid secara empiris. Butir soal tersebut yaitu kemampuan menganalisis (C4) pada nomor 1, 2, 4, 5, 6, 8, dan 10 dan kemampuan mengevaluasi (C5) pada nomor 3, 7, dan 9. Butir soal yang telah valid secara empiris maupun teoritis tersebut kemudian digunakan untuk mengidentifikasi profil kemampuan berpikir tingkat tinggi siswa pada kelas ujicoba. Dimana hasilnya disajikan pada Tabel 8 berikut ini.

Tabel 8. Identifikasi Profil HOTS Secara Umum

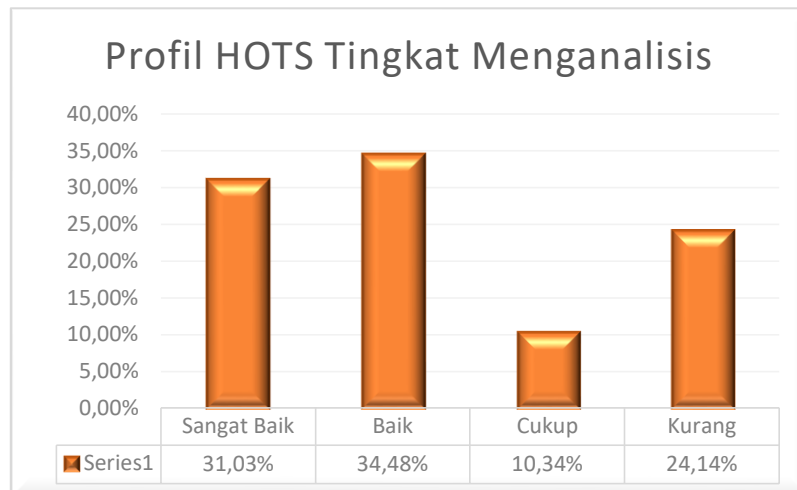
Kriteria	Jumlah Siswa	Persentase (%)
Sangat Baik	13	44.83%
Baik	4	13.79%
Cukup	0	0
Kurang	12	41.38%

Pada Tabel 8 dapat dilihat bahwa kemampuan berpikir tingkat tinggi siswa 44.8% siswa berada pada kategori sangat baik, 13.79% berada pada kategori baik, dan 41.38% siswa berada pada kategori kurang. Hal ini menunjukkan bahwa kemampuan berpikir tingkat tinggi siswa mayoritas baik, akan tetapi masih ada beberapa siswa yang masih kurang dalam menyelesaikan soal HOTS sehingga kemampuan berpikir tingkat tinggi siswa masih belum merata dalam kategori baik. Hal tersebut dikarenakan kelas yang digunakan dalam uji coba terdiri dari siswa yang memiliki kemampuan kognitif heterogen. Sehingga siswa yang memiliki kemampuan kognitif tinggi mampu untuk mengerjakan soal HOTS yang diberikan meskipun mereka belum terbiasa mengerjakan soal HOTS sebelumnya. Perbedaan ini juga dapat dikarenakan beberapa hal yaitu cara berpikir siswa berbeda dan strategi belajar siswa yang berbeda. Hal ini sesuai pernyataan dari Rosidah (2017) dimana dinyatakan bahwa siswa yang memiliki kemampuan tingkat kognitif tinggi akan mampu untuk menalar dan berfikir kritis yang lebih tinggi dalam menyelesaikan suatu permasalahan yang diberikan.



Gambar 6. Persentase Profil HOTS C4 dan C5

Berdasarkan Gambar 6, dapat diketahui nilai yang diperoleh siswa dalam mengerjakan soal HOTS didapatkan rata-rata persentase profil HOTS pada tingkat menganalisis (C4) sebesar 67,69% dengan kriteria cukup sedangkan untuk mengevaluasi (C5) persentase rata-ratanya adalah 80,46% dengan kriteria baik. Kenyataan ini menunjukkan bahwa profil HOTS siswa dalam menyelesaikan soal pada materi usaha dan energi tergolong cukup baik. Hasil perhitungan siswa didasarkan pada pedoman penilaian dalam Kurikulum 2013 Kemendikbud (2015) yakni digunakan penilaian dengan rentang 0-100. Interpretasi profil HOTS siswa pada tingkat menganalisis (C4) disajikan pada Gambar 7.

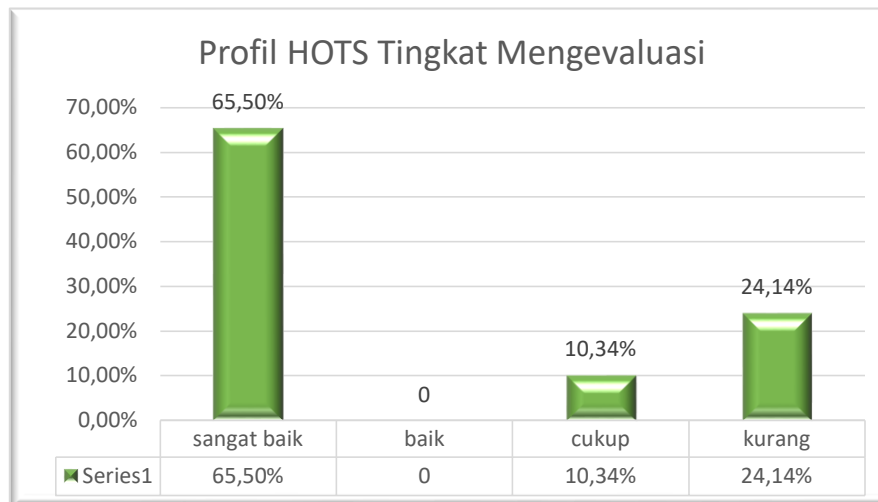


Gambar 7. Persentase Profil HOTS Tingkat C4

Gambar 7 menyajikan hasil analisis lembar jawaban dari tujuh soal ranah kognitif menganalisis (C4) yang diberikan kepada siswa, rata-rata kemampuan dalam menyelesaikan soal sebesar 67.49% dengan kriteria cukup berdasarkan kriteria profil HOTS oleh Kemendikbud 2015. Jumlah siswa yang mampu menjawab pada kriteria sangat baik yaitu 9 siswa (31,03%), 10 siswa pada kriteria baik (34.48%), 3 siswa pada kriteria cukup (10.34%), dan 7 siswa pada kriteria kurang (24.14%). Pada indikator menganalisis siswa diharapkan mampu menuliskan dan mengaitkan hubungan antar konsep. Dari hasil persentase tersebut artinya sebagian siswa telah mampu melihat hubungan antar konsep dan menggunakannya untuk menganalisis permasalahan dalam soal. Meskipun demikian, sebagian siswa lain masih kesulitan dan belum dapat menghubungkan konsep-konsep yang terkait untuk melakukan penyelesaian soal. Pernyataan ini sesuai dengan pendapat Suryani et al., (2016) bahwa siswa yang mengalami kesulitan dalam kemampuan menganalisis dikarenakan kurang adanya representasi dalam menyelesaikan permasalahan.

Butir soal dengan nilai terendah pada ranah kognitif menganalisis adalah soal nomor 2 dan 10. Adapun indikator soalnya untuk nomor 2 yaitu menganalisis konsep usaha terhadap perpindahan benda dengan benar. Pada indikator ini menganalisis dapat dilakukan apabila siswa mampu menganalisis pernyataan usaha yang dihasilkan saat benda (ayunan) sedang bergerak berdasarkan gambar dan narasi bacaan yang sudah disediakan. Akan tetapi, faktanya keterampilan siswa dalam menganalisis usaha dan perpindahan benda belum sepenuhnya sesuai dengan konsep yang ada. Siswa hanya mendefinisikan arti dari usaha dan perpindahan tanpa memahami konsep sebenarnya dari usaha dan perpindahan. Siswa juga belum berpikir logis terhadap gerakan ayunan antara gambar yang diberikan dengan kenyataan sebenarnya jika fenomena tersebut terjadi. Sedangkan indikator untuk nomor 10 yaitu menganalisis dan membandingkan besar energi kinetik dengan tepat. Pada indikator ini menganalisis dapat dilakukan apabila siswa mampu menganalisis dan membandingkan besar energi kinetik suatu benda dengan massa yang berbeda tetapi gaya dan lintasannya sama. Siswa masih

menganggap massa adalah satu-satunya faktor yang mempengaruhi energi yang dibutuhkan suatu benda untuk bergerak. Akan tetapi sesuai konsep yang ada jika gaya dan usaha yang diberikan pada benda sama, maka energi kinetik saat *finish* besarnya adalah sama.



Gambar 8. Persentase Profil HOTS Tingkat C5

Gambar 8 menyajikan profil HOTS pada ranah kognitif mengevaluasi (C5), rata-rata kemampuan menyelesaikan soal sebesar 80.46% dengan kriteria baik berdasarkan kriteria profil HOTS oleh Kemendikbud 2015. Ada 19 siswa (65.50%) mampu menjawab soal pada kriteria sangat baik, 3 siswa (10.34%) mampu menjawab soal pada kriteria cukup, 7 siswa (24.14%) dapat menjawab soal pada kriteria kurang, dan tidak ada siswa yang masuk pada kriteria baik.

Dari 3 soal ranah kognitif mengevaluasi yang diberikan kepada peserta didik, nomor soal dengan nilai terendah adalah nomor 9. Adapun indikator soalnya adalah siswa dapat memutuskan terkait benar atau salah dari pernyataan dan alasan konsep usaha dengan tepat. Pada indikator ini, banyak dari siswa kurang dapat mengevaluasi hubungan antara pernyataan dan alasan. Kebanyakan siswa mampu memutuskan mana pernyataan yang benar atau salah tetapi kurang tepat dalam memberikan alasan. Kemampuan siswa dalam menyelesaikan soal pada kriteria mengevaluasi lebih tinggi dibandingkan dengan kemampuan siswa dalam menyelesaikan soal pada kriteria menganalisis. Akan tetapi siswa masih mengalami kesulitan dalam mengidentifikasi hubungan dari beberapa pernyataan, konsep, dan berbagai model yang digunakan untuk membuat keputusan serta memberikan alasan terkait pernyataan yang ada. Kesulitan siswa kemungkinan dikarenakan kurangnya latihan mengevaluasi atau memberikan keputusan dan argumentasi atas suatu permasalahan karena terbiasa mendapatkan informasi secara langsung dari guru. Kesulitan tingkat mengevaluasi juga dipaparkan oleh Priyadi et al., (2018) bahwa siswa cenderung kesulitan dalam mengevaluasi dikarenakan masih belum mampu mengidentifikasi asumsi yang salah. Evaluasi merupakan kemampuan untuk menguji kebenaran dan membuat suatu keputusan.

Berdasarkan pembahasan di atas, instrumen penilaian HOTS menggunakan *Quizizz* yang dikembangkan pada materi usaha dan energi menggambarkan instrumen yang layak digunakan untuk mengidentifikasi konsep penerapan usaha dan energi pada siswa, sehingga instrumen ini dapat digunakan untuk memotret profil keterampilan berpikir tingkat tinggi siswa.

## KESIMPULAN

Instrumen penilaian HOTS yang telah dikembangkan dinyatakan layak digunakan. Berdasarkan hasil analisis, dapat disimpulkan bahwa profil HOTS peserta didik dalam menyelesaikan soal fisika materi usaha dan energi di SMA At-Ta'miriyah Surabaya

tergolong cukup baik. Hal ini sesuai dengan hasil penelitian berupa persentase peserta didik yang sudah memiliki HOTS cukup baik dalam menyelesaikan soal menganalisis (C4) sebesar 67.49% dan soal mengevaluasi (C5) sebesar 80.46%.

## SARAN

Guru disarankan untuk menggunakan instrumen HOTS yang dikembangkan untuk mengukur keterampilan berpikir tingkat tinggi siswa dan juga melatih siswa dalam meningkatkan kemampuan berpikir tingkat tinggi. Sebaiknya dilakukan pengembangan soal HOTS pada materi fisika lainnya juga sehingga dapat digunakan untuk mengukur HOTS siswa.

## UCAPAN TERIMA KASIH

Ucapan terimakasih disampaikan kepada seluruh komponen di SMA At-Ta'miriyyah Surabaya yang telah membantu proses pengambilan data penelitian pengembangan instrumen penilaian HOTS ini.

## DAFTAR PUSTAKA

- Adawiyah, R., & Wisudawati, A. (2017). Pengembangan instrumen tes berbasis literasi sains. *Indonesian Journal of Curriculum and Educational Technology Studies*, 5(2), 112–121.
- Alfika, Z. A., & Mayasari, T. (2018). Profil kemampuan memecahkan masalah pelajaran fisika siswa MTs. *Quantum: Seminar Nasional Fisika, Dan Pendidikan Fisika*, 583–589.
- Angraini, G., & Sriyati, S. (2019). Analisis Kemampuan Berpikir Tingkat Tinggi Siswa SMAN Kelas X di Kota Solok pada Konten Biologi. *Jurnal of Education Informatic Technology Aand Science*, 1(1), 114–124.
- Arifin, Z. (2013). *Evaluasi pembelajaran: Prinsip, teknik, prosedur [Learning evaluation: Principles, techniques, and procedure]*. Bandung, Indonesia: Remaja Rosdakarya.
- Arikunto, S. (2012). *Dasar-dasar evaluasi pendidikan edisi 2*. Jakarta: Bumi Aksara.
- Ayumniyya, L., & Setyarsih, W. (2021). Profil Kemampuan Berpikir Tingkat Tinggi Siswa SMA dalam Pemecahan Masalah pada Materi Hukum Newton. *IPF : Inovasi Pendidikan Fisika*, 10(1), 50–58.
- Conklin, W. (2011). *Higher-order thinking skills to develop 21st century learners*. Teacher Created Materials.
- Desiriah, E., & Setyarsih, W. (2021). Tinjauan Literatur Pengembangan Instrumen Penilaian Kemampuan Berpikir Tingkat Tinggi (HOTS) Fisika Di SMA. *ORBITA: Jurnal Kajian, Inovasi Dan Aplikasi Pendidikan Fisika*, 7(1), 79–89.
- Gardner, H. (2013). *Multiple Intelligences (terjemahan Yelvi Andri Zaimul)*. Jakarta: Daras Book.
- Hidayah, N., Silitonga, H. T. M., & Mahmuda, D. (2018). Pengembangan Tes Higher Order Thinking Skills (HOTS) Pada Materi Getaran Harmonis Untuk SMA. *Jurnal Pendidikan Dan Pembelajaran Khatulistiwa*, 7(7).
- Indrawati, M. D. (2018). Pengembangan Instrumen Penilaian Literasi Sains Fisika Peserta Didik Pada Bahasan Gelombang Bunyi di SMA Negeri 1 Gedangan Sidoarjo. *IPF : Inovasi Pendidikan Fisika*, 7(1).
- 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.
- Ju, S. Y., & Adam, Z. (2018). Implementing Quizizz as game based learning in the Arabic classroom. *European Journal of Social Science Education and Research*, 5(1), 194–198.
- Kivunja, C. (2015). Why Students Don't Like Assessment and How to Change Their Perceptions in 21 st Century Pedagogies. *Creative Education*, 6(20), 2117.
- Nisa, S. K. (2018). Analisis dan Pengembangan Soal High Order Thinking Skills (HOTS)



- Mata Pelajaran Fisika Tingkat Sekolah Menengah Atas (SMA). *IPF : Inovasi Pendidikan Fisika*, 7(2).
- Noor, S. (2020). Penggunaan quizz dalam penilaian pembelajaran pada materi ruang lingkup biologi untuk meningkatkan hasil belajar siswa kelas X. 6 SMAN 7 Banjarmasin. *Jurnal Pendidikan Hayati*, 6(1), 1–7.
- Nugroho, R. A. (2018). *HOTS (Kemampuan Berpikir Tingkat Tinggi: Konsep, Pembelajaran, Penilaian, dan Soal-soal)*. Jakarta: PT Gramedia Widiasarana Indonesia.
- Nurhayati, N., & Angraeni, L. (2017). Analisis kemampuan berpikir tingkat tinggi mahasiswa (higher order thinking) dalam menyelesaikan soal konsep optika melalui model problem based learning. *Jurnal Penelitian & Pengembangan Pendidikan Fisika*, 3(2), 119–126.
- OECD. (2019). *PISA 2018 Assessment and Analytical Framework*. Paris : OECD Publishing.
- Permendikbud No 23. (2016). *Peraturan Menteri Pendidikan dan Kebudayaan No. 23 Tahun 2016 tentang Standar Penilaian Pendidikan*. Jakarta: Badan Standar Nasional Pendidikan Kementerian Pendidikan Dan Kebudayaan RI, NO 23.
- Permendikbud No 53. (2015). *Peraturan Menteri Pendidikan dan Kebudayaan No. 53 Tahun 2015 tentang Penilaian Hasil Belajar Pendidik pada Pendidikan Dasar dan Pendidikan Menengah*. Jakarta: Kementerian Pendidikan Dan Kebudayaan, No. 53.
- Pribadi, B. A. (2016). *Desain dan pengembangan program pelatihan berbasis kompetensi implementasi model ADDIE*. Kencana.
- Priyadi, R., Mustajab, A., Tatsar, M. Z., & Kusairi, S. (2018). Analisis kemampuan berpikir kritis siswa SMA kelas X MIPA dalam pembelajaran fisika. *JPFT (Jurnal Pendidikan Fisika Tadulako Online)*, 6(1), 53–55.
- Puspendik. (2019). *Rekap Hasil Ujian Nasional Tahun Pelajaran 2018/2019 SMA/MA*. Hasil UN SMA. <https://hasilun.puspendik.kemdikbud.go.id/>
- Rahmadani, Y., Fitakurahmah, N., Funky, N., Prihatin, R., Majid, Q., & Prayitno, B. A. (2018). Profil Keterampilan Literasi Sains Siswa di Salah Satu Sekolah Swasta di Karanganyar. *Jurnal Pendidikan Biologi*, 7(3), 183–190.
- Riduwan, M. B. A. D. (2012). *Skala Pengukuran Variabel-variabel Penelitian*. Bandung: Alfabeta.
- Rohmah, U. A., & Sunarti, T. (2020). Profil Higher Order Thinking Skills (HOTS) Peserta Didik SMA Pada Materi Kalor. *IPF: Inovasi Pendidikan Fisika*, 09(03), 466–472.
- Rosidah, F. E. (2017). Pengembangan Tes Literasi Sains Pada Materi Kalor Di Sma Negeri 5 Surabaya. *IPF : Inovasi Pendidikan Fisika*, 6(3).
- Setiadi, H. (2016). Pelaksanaan penilaian pada Kurikulum 2013. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 20(2), 166–178.
- Sudjana, N. (2009). *Penilaian hasil belajar mengajar*. Bandung: Remaja Rosdakarya.
- Sugiyono, P. D. (2013). *Metode penelitian manajemen*. Bandung: Alfabeta, CV.
- Suryani, I., Yolanda, Y., & Ariani, T. (2016). Keterampilan Berpikir Kritis Siswa dalam Menyelesaikan Soal Fisika Tentang Impuls dan Momentum. *Jurnal Fisika*, 1(10), 21–27.
- Wijaya, P. A., & Suyono, A. (2019). Profil Kemampuan Mahasiswa dalam Mengembangkan Instrumen Tes Mengacu Standar Hots pada Mata Kuliah Evaluasi dan Teknik Pencapaian Hasil Belajar. *PEKA : Pendidikan Matematika*, 7(2).
- Witri, S., Febrian, F., & Tambunan, L. R. (2019). Deskripsi Kemampuan Berpikir Tingkat Tinggi Siswa dalam Menyelesaikan Soal Fungsi Kelas X SMAN 2 Tanjungpinang. *Jurnal Gantang*, 4(2), 155–160.
- Yuliantaningrum, L., & Sunarti, T. (2020). Pengembangan Instrumen Soal HOTS untuk Mengukur Keterampilan Berpikir Kritis, Berpikir Kreatif, dan Pemecahan Masalah Materi Gerak Lurus pada Peserta Didik SMA. *IPF : Inovasi Pendidikan Fisika*, 9(2).