

## ***Development Of A Test Instrument Based On Higher-Order Thinking Skills Using Wayground In The Element Of Archival Management***

### **Pengembangan Instrumen Tes Berbasis *Higher Order Thinking Skills* Berbantuan Wayground Pada Elemen Pengelolaan Kearsipan**

**Sovia Nurul Azizah<sup>1</sup>, Fitriana Rahmawati<sup>2</sup>**

Program Studi Pendidikan Administrasi Perkantoran, Fakultas Ekonomika dan Bisnis,  
Universitas Negeri Surabaya, Jawa Timur<sup>1,2</sup>

Email: [sovia.22053@mhs.unesa.ac.id](mailto:sovia.22053@mhs.unesa.ac.id), [fitriarahmawati@unesa.ac.id](mailto:fitriarahmawati@unesa.ac.id)

\*Corresponding Author

---

Received : 22 Maret 2026, Revised : 15 April 2026, Accepted : 2 Mei 2026

---

#### **ABSTRACT**

*The primary objective of this study encompasses four key aspects: the procedure for developing the evaluation instrument, the validity or feasibility of the instrument, students' perceptions, and the analysis of students' critical thinking skills. This research employs a Research and Development (R&D) method by adopting the 4D development model, which consists of the Define, Design, Develop, and Disseminate stages. The study involved 36 eleventh-grade students from the MPLB program at SMK Ketintang Surabaya. The instrument was deemed highly feasible by subject-matter experts (93.3%), evaluation experts (88%), and language experts (93.3%). Furthermore, the instrument underwent validity, reliability, discrimination index, difficulty level, and distractor analysis. The validity test results indicated that four items were invalid, leaving 26 items categorized as valid. The instrument was found to be reliable, with a Cronbach's Alpha coefficient of 0.879. The discrimination index analysis showed that 19 items were classified as very good, 3 as good, 5 as adequate, and 3 as poor. The difficulty level was considered appropriate, and all distractors functioned effectively. Students' responses to the instrument were positive. In terms of higher-order thinking skills, most students were classified in the "moderate" category.*

**Keywords:** *Development, HOTS, Wayground, Test Instruments, 4D, Archival Management.*

#### **ABSTRAK**

Tujuan utama dari penelitian ini mencakup empat aspek: prosedur pengembangan instrumen evaluasi, validitas atau kelayakan instrumen, persepsi peserta didik, serta analisis terhadap keterampilan berpikir kritis siswa. Penelitian ini menerapkan metode *Research and Development* (R&D) dengan mengadopsi prosedur pengembangan model 4D, yang meliputi tahap *Define, Design, Develop, dan Disseminate*. Adapun sasaran penelitian melibatkan 36 siswa kelas XI program keahlian MPLB di SMK Ketintang Surabaya. Instrumen ini dinyatakan sangat layak oleh ahli materi dengan presentase 93,3%, ahli evaluasi dengan presentase 88%, dan ahli bahasa dengan presentase 93,3%. Selanjutnya, instrumen akan dilakukan proses uji validitas, reliabilitas, daya beda, tingkat kesukaran, dan analisis pengecoh. Berdasarkan hasil uji validitas, 4 soal dinyatakan tidak valid sehingga tersisa 26 soal yang dinyatakan valid. Instrumen dinyatakan reliabel dengan nilai Cornbach's Alpha sebesar 0,879. Daya pembeda soal menunjukkan 19 sangat baik, 3 soal baik, 5 soal cukup, dan 3 soal kurang. Tingkat kesukaran instrumen dinyatakan ideal dan seluruh pengecoh berfungsi dengan baik. Respon peserta didik terhadap instrumen positif. Terkait keterampilan berpikir tingkat tinggi peserta didik mayoritas pada kategori "cukup".

**Kata Kunci:** *Pengembangan, HOTS, Wayground, Instrumen Tes, 4D, Pengelolaan Kearsipan.*

## **1. Pendahuluan**

Abad ke-21 yang menuju era 5.0 telah membawa transformasi yang cukup berpengaruh dalam dunia pendidikan. Dalam upaya menghadapi kompetensi global di masa depan, institusi pendidikan saat ini didorong untuk membekali siswa dengan keahlian yang relevan dengan

perkembangan zaman dan teknologi digital. Oleh karena itu, hal ini menuntut adanya individu berkualitas yang mampu menerapkan berpikir secara kritis (Suratman *et al.*, 2019). Kondisi tersebut dapat ditanamkan dengan adanya 4C (*Critical Thinking and Problem Solving, Creativity, Communication, dan Collaboration*) (Zamkabay & Wulandari, 2022). urgensi kemampuan berpikir kritis di masa kini terletak pada fungsinya sebagai alat penyelesaian masalah yang mendukung adaptabilitas peserta didik. Salah satu manifestasi nyata dari pengembangan kompetensi tersebut adalah melalui integrasi model pembelajaran berorientasi HOTS.

*HOTS* merupakan perangkat penilaian yang dirancang untuk mengevaluasi kecakapan kognitif kompleks siswa. Instrumen ini melampaui sekadar pengujian daya ingat, melainkan lebih menekankan pada kemampuan subjek dalam membedah informasi serta melakukan analisis mendalam terhadap data yang diterima. Saat ini kurikulum di Indonesia yakni kurikulum merdeka menerapkan pengembangan kemampuan berpikir logika, mengkomunikasikan, dan mencipta (Hamidah & Wulandari, 2021). Menurut Nurhayati (2020) melalui penerapan kurikulum merdeka, diharapkan potensi setiap peserta didik dapat terasah melalui pembelajaran yang kritis, progresif, variatif, dan aplikatif.

Salah satu permasalahan pada dunia pendidikan yakni rendahnya kualitas instrumen penilaian dalam mengukur kemampuan *HOTS* peserta didik. Oleh karena itu, instrumen evaluasi menjadi komponen esensial untuk mengalibrasi pemahaman peserta didik (Puspitasari & Wulandari, 2021). Pengembangan alat ukur yang tepat juga berfungsi sebagai sarana untuk mengasah nalar kritis siswa, terutama dalam memproses informasi dan mencari solusi atas permasalahan melalui kerangka berpikir yang logis. Melihat urgensi yang telah dipaparkan, instrumen penilaian yang andal sangat diperlukan untuk meninjau keberhasilan pembelajaran. Salah satu bentuk yang dapat diterapkan adalah melalui perangkat tes guna memfasilitasi kebutuhan evaluasi dalam kegiatan belajar mengajar (Desilva *et al.*, 2020).

Penggunaan instrumen tes dapat memudahkan guru dalam mengetahui sejauh mana kompetensi siswa sehingga guru dapat memilih strategi terbaik dalam pembelajaran selanjutnya. Penilaian berbasis *HOTS* yang mencakup dimensi berpikir dalam Taksonomi Bloom khususnya pada kemampuan menganalisis (C4), mengevaluasi (C5), dan menciptakan (C6) dapat membantu penalaran kritis, pemecahan masalah dan pemikiran kreatif peserta didik. Oleh sebab itu, guru diharapkan dapat menyusun instrumen penilaian berbasis *HOTS* untuk membantu proses evaluasi peserta didik (Latifah *et al.*, 2023). Menurut Widana (2017) penyusunan instrumen berbasis *HOTS* menuntut guru untuk merancang soal yang kreatif dan inovatif dalam penyusunan stimulus yang dikaitkan dengan masalah sehari – hari. Kemampuan berpikir kompleks siswa akan berkembang saat mereka dihadapkan pada tantangan nyata di sekitarnya. Maka dari itu, implementasi pembelajaran *HOTS* oleh guru harus dilakukan secara kontinu agar peserta didik terlatih untuk berpikir kritis dan solutif (Kurniawan, 2020).

Era serba digital saat ini menuntut guru untuk lebih kreatif dalam melaksanakan asesmen yang tidak hanya terpacu pada metode konvensional saja. Guru harus merancang lingkungan pembelajaran yang inovatif sehingga mampu memacu motivasi peserta didik dalam pembelajaran (Handayani & Wulandari, 2021). Salah satu inovasi dalam pembelajaran dapat memanfaatkan *Information and Communication Technology (ICT)* yang dapat meminimalkan pembelajaran konvensional. Hal tersebut sejalan dengan Purba (2021) yang menyatakan bahwa ICT memiliki fitur yang dapat memudahkan guru dalam penilaian dan tidak memerlukan kertas dalam pengerjaannya. Pemanfaatan *Wayground* merepresentasikan implementasi teknologi informasi dan komunikasi (ICT) dalam konteks instruksional, yang berfungsi sebagai media pendukung kegiatan belajar mengajar. *Wayground* adalah platform berbasis web dan gamifikasi yang dapat meningkatkan aktivitas belajar (Uktolseja *et al.*, 2025). Menurut Mufida *et al* (2025) platform seperti *Wayground* dapat menghadirkan suasana belajar yang menyenangkan sehingga mengurangi rasa bosan peserta didik di dalam kelas. Fitur evaluasi otomatis pada *Wayground* dapat mempermudah guru untuk memetakan capaian peserta didik secara *real-time* (Magfirah *et al.*, 2025). Adanya *Wayground* mendukung pemanfaatan ICT ditengah arus digitalisasi saat ini.

Studi terdahulu yang dilakukan oleh (Hamidah & Wulandari, 2021) menunjukkan bahwa kecakapan berpikir tingkat tinggi siswa kelas X OTKP di SMK Negeri 1 Lamongan berada pada level "cukup". Hal ini diperkuat oleh temuan (Puspitasari & Wulandari, 2021) yang mengategorikan keterampilan serupa dalam kriteria "kurang", sehingga urgensi pembelajaran berbasis HOTS sangat diperlukan. Merujuk pada penelitian sebelumnya, dapat dikatakan bahwa kemampuan berpikir tingkat tinggi peserta didik masih dalam kategori "cukup" dan "kurang". Kementerian Pendidikan dan Kebudayaan melalui Ditjen GTK menginisiasi program Peningkatan Kompetensi Pembelajaran (PKP) untuk meningkatkan mutu pendidikan melalui pengembangan instrumen HOTS untuk mengasah kemampuan HOTS peserta didik (Maulidia & Pahlevi, 2020).

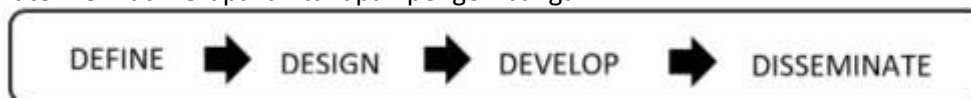
Merujuk pada data yang dihimpun melalui sesi tanya jawab dengan guru pengajar di SMK Ketintang Surabaya, diperoleh temuan bahwa kecakapan berpikir tingkat tinggi peserta didik sebagian besar masih berada pada level yang belum memadai. Asessmen yang diberikan guru tidak sepenuhnya berbasis HOTS karena masih bergabung dengan LOTS. Fenomena ini menuntut kreativitas guru dalam mengembangkan perangkat penilaian berpikir tingkat tinggi secara berkelanjutan, sebagai upaya sistematis untuk menstimulasi nalar kritis peserta didik. Selain itu, asessmen HOTS yang diberikan guru masih terbatas pada soal uraian dan terpacu pada LKS peserta didik.

Penelitian ini akan mengujikan instrumen pilihan ganda berbasis HOTS pada elemen pengelolaan kearsipan. Pemilihan elemen ini dikarenakan peserta didik belum sepenuhnya dapat mengimplementasikan teori mengenai pengelolaan kearsipan. Pada kenyataannya memang pengelolaan kearsipan cenderung mengandalkan praktik, namun jika peserta didik tidak memahami secara teori maka praktik juga tidak akan maksimal. Beberapa penelitian sebelumnya seperti Suhardini *et al* (2025) yang mengembangkan e-modul pada elemen pengelolaan kearsipan. Selain itu penelitian Yanti & Hasyim (2025) melakukan pengembangan LKPD pada mata pelajaran pengelolaan kearsipan. *Research Gap* ini ditemukan pada terbatasnya pengembangan perangkat evaluasi berupa tes pilihan ganda yang mengintegrasikan aspek HOTS. Atas dasar itulah, penelitian ini hadir sebagai upaya untuk menjawab kekosongan literatur dan praktik melalui pengembangan instrumen penilaian yang aplikatif.

Merujuk pada uraian sebelumnya, fokus utama studi ini adalah pengembangan instrumen tes pilihan ganda berorientasi HOTS dengan memanfaatkan Wayground pada elemen pengelolaan kearsipan. Integrasi teknologi dalam instrumen ini diharapkan dapat meningkatkan antusiasme belajar peserta didik dan memberikan kemudahan bagi guru dalam melakukan penelaahan hasil belajar secara sistematis.

## 2. Metodologi

Penelitian ini menggunakan jenis penelitian *Research and development (R&D)* dengan model pengembangan 4D. Model pengembangan 4D terdiri dari *Define, Design, Develop, and Disseminate*. Berikut merupakan tahapan pengembangan 4D :



**Gambar 1. Alur pengembangan 4D**

Merujuk pada konsep yang dikemukakan oleh Thiagarajan, model ini berfungsi sebagai prosedur metodologis dalam menghasilkan produk baru atau menyempurnakan kualitas instrumen yang telah tersedia agar lebih efektif dan relevan (Sugiyono, 2023). Tahap pendefinisian (*define*) dalam penelitian ini terdiri dari analisis awal akhir, analisis peserta didik, analisis tugas, analisis konsep dan spesifikasi tujuan. Tahap perancangan (*design*) terdiri dari penyusunan kisi-kisi dan rancangan awal. Selanjutnya pada tahap pengembangan (*Develop*) terdiri dari validasi ahli, revisi produk, uji coba produk, analisis produk dan revisi produk akhir. Terakhir pada tahap penyebaran (*Disseminate*) dilakukan secara terbatas.

Desain uji coba dalam penelitian ini yakni pengembangan instrumen pilihan ganda berbasis *HOTS* dengan jumlah 30 soal berbantuan platform gamifikasi *Waygorund*. Produk yang telah dikembangkan kemudian akan diverifikasi melalui validasi ahli dan analisis butir soal guna memastikan standarnya. Butir soal yang telah dinyatakan layak selanjutnya akan diujicobakan ke peserta didik dan untuk selanjutnya akan di analisis setiap butir soalnya. Subjek uji coba dalam penelitian ini yakni peserta didik kelas XI MP 3 dengan jumlah 36 peserta didik. Jumlah tersebut sudah termasuk layak dalam penelitian pengembangan dengan minimal total subjek 30 (Sugiyono, 2023). Penentuan subjek uji coba didasarkan pada kebutuhan kelas dengan adanya instrumen berbasis *HOTS* ini.

Data yang dihimpun dalam studi ini mencakup data bersifat kualitatif serta data kuantitatif yang diperoleh dari hasil validasi dan uji coba instrumen. Data kualitatif diperoleh dari hasil validasi ahli materi, ahli bahasa, dan ahli evaluasi. Validator akan diberikan lembar validasi yang berisikan beberapa pernyataan mengenai instrumen yang dikembangkan. Setiap pernyataan akan diberikan nilai dari 1 – 5 dengan keterangan dari sangat kurang hingga sangat baik. Selanjutnya, hasil dari setiap validator akan dihitung dan akan diinterpretasikan sesuai dengan tabel berikut :

**Tabel 1. Kriteria Kelayakan Hasil Validasi**

Penilaian	Kriteria
$81 \leq p \leq 100\%$	Sangat Layak
$61 \leq p \leq 81\%$	Layak
$41 \leq p \leq 61\%$	Cukup Layak
$21 \leq p \leq 41\%$	Tidak Layak
$0 \leq p \leq 21\%$	Sangat Tidak Layak

Sumber : (Arikunto, 2018)

Data kuantitatif dihimpun melalui serangkaian analisis statistik terhadap butir soal. Analisis tersebut meliputi uji validitas dan reliabilitas instrumen, serta peninjauan terhadap tingkat kesukaran, daya beda soal, dan fungsi pengecoh pada setiap butir tes. Setiap uji memiliki kriteria yang menunjukkan kualitas dari setiap soalnya. Uji yang pertama yakni uji validitas, butir soal akan dinyatakan valid jika  $r_{hitung} > r_{tabel}$  dimana  $r_{tabel}$  yang digunakan dalam penelitian ini sebesar 0,270. Selanjutnya yakni uji reliabilitas, instrumen akan dinyatakan reliabel apabila nilai *Cronbach's Alpha* melebihi angka 0,70.

Selanjutnya, dilakukan uji daya pembeda soal untuk melihat sejauh mana sebuah butir tes mampu membedakan kapasitas akademik antarsiswa dalam kelompok uji coba (Lestari & Yudhanegara, 2024). Berikut merupakan kriteria dari daya pembeda soal :

**Tabel 2. Kriteria Daya Pembeda Soal**

Penilaian	Kriteria
$\geq 0,40$	Sangat Baik
0,30 – 0,39	Baik
0,20 – 0,29	Cukup
$\leq 0,19$	Kurang

Sumber : (Basuki & Hariyanto, 2016)

Selanjutnya dilakukan uji tingkat kesukaran untuk memastikan kualitas butir soal. Sebagaimana ditegaskan oleh Lestari dan Yudhanegara (2024), soal yang berkualitas adalah soal yang berada pada rentang kesulitan ideal. Adapun standar kriteria tingkat kesukarannya adalah sebagai berikut :

**Tabel 3. Kriteria Tingkat Kesukaran**

Mean	Kriteria
1,00 – 0,90	Sangat Mudah
0,71 – 0,89	Mudah
0,31 – 0,70	Sedang
0,21 – 0,30	Sukar
0 – 0,20	Sangat Sukar

Sumber : (Basuki & Hariyanto, 2016)

Uji yang terakhir yakni mengenai analisis pengecoh, dimana hal tersebut penting untuk mengetahui keoptimalan pilihan jawaban. Pengecoh akan dianggap berfungsi apabila menarik minat peserta didik untuk memilih setidaknya 5% peserta didik (Arikunto, 2018). Selanjutnya terdapat pula respon peserta didik bagaimana mengenai pengembangan soal yang telah diujicobakan. Terdapat 2 indikator pada respon peserta didik yakni respon dan reaksi yang mana setiap indikatornya terdapat butir-butir pernyataan yang harus dijawab oleh peserta didik (Aisyah *et al.*, 2016). Setelah data uji coba terhimpun, selanjutnya dilakukan pemetaan terhadap kategori kemampuan berpikir tingkat tinggi peserta didik yang disesuaikan dengan skor hasil pengerjaan mereka.

**Tabel 4. Kriteria Keterampilan Berpikir Tingkat Tinggi**

Nilai	Kriteria
81 – 100	Sangat Baik
61 – 80	Baik
41 – 60	Cukup
21 – 40	Kurang
< 20	Sangat Kurang

Sumber : (Purbaningrum, 2017)

### 3. Literature Review

#### Teori Kognitivisme

*Cognitive* atau *cognition* memiliki persamaan dengan kata “*knowing*” yang berarti mengetahui. Kognitiv mengarah pada aktivitas yang melibatkan proses kognisi, yang dilandaskan pada pengetahuan dan bersifat empiris (Wahab & Rosnawati, 2021). Teori kognitif yang dikembangkan oleh Jean Piaget memandang bahwa kecerdasan tercipta melalui skema bagaimana individu memberikan persepsi lingkungan dan kemampuan secara mental (Hascan & Suyadi, 2021). Teori kognitif menganggap belajar sebagai proses kompleks yang melibatkan stimulus dan respon. Hal tersebut sejalan dengan pendapat Nurhadi (2020) yang memandang bahwa belajar juga melibatkan proses berpikir yang lebih kompleks. Oleh karena itu, pandangan kognitivisme menegaskan bahwa proses pembelajaran menjadi prioritas daripada pencapaian akhir.

#### Instrumen Penilaian

Pengukuran merupakan proses sistematis dalam pemberian nilai berupa angka dalam bidang tertentu (Ismail, 2018). Pengukuran dapat dilaksanakan menggunakan instrumen baik itu tes maupun non tes (Kurniawan *et al.*, 2022). Instrumen memiliki peran penting dalam meninjau proses pembelajaran di kelas (Rohmad, 2017). Instrumen penilaian merupakan sarana yang digunakan untuk mengukur, mengamati, mengumpulkan dan mengevaluasi hasil belajar peserta didik. Instrumen penilaian juga berfungsi untuk mengetahui seberapa efektif pembelajaran yang selama ini berlangsung di sekolah. Instrumen penilaian pada umumnya terbagi menjadi dua klasifikasi utama, yaitu instrumen tes serta non-tes. Penyusunan sebuah tes yang berkualitas menuntut serangkaian langkah prosedural yang saksama. Adapun jenis instrumen yang dikembangkan dalam studi ini adalah tes objektif yang disajikan melalui model pilihan ganda.

### **Higher Order Thinking Skills (HOTS)**

*HOTS* dipahami sebagai kecakapan berpikir tingkat lanjut yang tidak hanya bersifat teknis, melainkan melibatkan unsur penalaran, pemikiran kritis, aspek reflektif, serta metakognisi dalam pengolahan informasi (Supriadi, 2020). *HOTS* mencakup proses mengelola informasi misalnya ketika peserta didik memeriksa, menggabungkan atau menyatukan fakta dan gagasan serta menafsirkannya. Menurut Widana (2017) instrumen tes berbasis *HOTS* perlu dirancang untuk mengukur kompetensi peserta didik. Hal tersebut selaras dengan Setiawati *et al* (2019) yang menyatakan instrumen *HOTS* mengevaluasi kemampuan metakognitif peserta didik yang tidak hanya terbatas pada penguasaan informasi saja. Dimensi berpikir *HOTS* yang diklasifikasikan berdasarkan kategori keterampilan oleh Anderson & Krathwol termasuk ke dalam menganalisis (C4), mengevaluasi (C5), dan mengkreasi (C6). Menurut Supriadi (2020) soal *HOTS* memiliki karakteristik utama berupa pengukuran kapasitas kognitif tingkat lanjut, orientasi pada masalah kontekstual, dan keberagaman bentuk item tes. Penggunaan instrumen ini ditargetkan dapat memacu perkembangan kemampuan berpikir kritis dan analitis siswa ke tingkat yang lebih optimal.

### **3. Hasil dan Pembahasan**

#### **Pengembangan Instrumen Tes Berbasis HOTS Berbantuan Wayground Pada Elemen Pengelolaan Kearsipan**

Penelitian ini menggunakan jenis penelitian *R & D* dengan model 4D. Tahap *define* dimulai dengan melakukan analisis terkait permasalahan yang ada melalui wawancara dengan guru mata pelajaran produktif. Berdasarkan informasi yang ada, asesmen yang diberikan oleh guru belum sepenuhnya berbasis *HOTS*. Soal yang diberikan ke peserta didik mayoritas diambil dari LKS. Persoalan tersebut sejalan dengan argumentasi Serevina *et al* (2019) yang menyatakan bahwa salah satu permasalahan di dunia pendidikan yakni kurangnya kualitas dari soal *HOTS*. Selanjutnya yakni menganalisis peserta didik dimana subjek yang digunakan adalah peserta didik kelas XI MP 3 yang masih membutuhkan adanya soal berbasis *HOTS*. Selanjutnya akan dilakukan uji coba terbatas 30 soal pilihan ganda dengan 4 pilihan jawaban. Materi yang akan digunakan yakni pada elemen pengelolaan kearsipan dan soal yang dikembangkan akan disesuaikan dengan taksonomi bloom menganalisis (C4), mengevaluasi (C5), dan mengkreasi (C6).

Selanjutnya mengenai analisis konsep yakni materi pokok yang digunakan terbagi menjadi 3 dan setiap materi pokok akan dikembangkan menjadi 10 soal *HOTS*. Pada tahap *design* akan dibuat kisi – kisi berdasarkan indikator setiap materi pokoknya. Kisi – kisi akan memudahkan penyusun instrumen untuk menghasilkan tes yang sesuai harapan (Susiatin, 2019). Sumber materi yang digunakan oleh peneliti yakni LKS siswa dan *e-book* melalui platform online. Kisi – kisi tersebut selanjutnya akan dikembangkan menjadi sebuah butir soal. Seluruh butir soal dikembangkan berdasarkan spesifikasi dalam kisi-kisi yang sudah dipersiapkan guna memastikan setiap indikator kompetensi terwakili dengan baik. Lalu dilanjutkan dengan tahap *develop* dimana peneliti melakukan tahap validasi instrumen. Validator terdiri dari dosen ahli materi, ahli bahasa dan ahli evaluasi. Guru mata pengampu elemen pengelolaan kearsipan juga menjadi pendukung untuk memvalidasi materi pada instrumen. Jika validator telah memberikan revisi, selanjutnya akan dilakukan revisi dan instrumen siap untuk diujicobakan secara terbatas. Setelah melakukan uji coba, langkah selanjutnya yakni melakukan analisis setiap butir soalnya. Tahap terakhir yakni *disseminate* terbatas yang akan dilakukan dengan mengirimkan instrumen melalui google drive ke SMK Ketintang Surabaya.

#### **Kelayakan Instrumen Tes Berbasis HOTS Pada Elemen Pengelolaan Kearsipan**

Kelayakan instrumen pada penelitian ini akan dinilai berdasarkan hasil validasi ahli dan juga analisis setiap butir soal. Selajan dengan Wardany *et al* (2017) yang juga melakukan hal serupa untuk menilai kelayakan instrumennya. Validator terdiri dari validator ahli materi, ahli bahasa dan juga ahli evaluasi. Berdasarkan hasil validasi ahli materi 1 yakni dosen memperoleh

kriteria “sangat layak” dengan nilai sebesar 93,3% serta terdapat sedikit revisi mengenai kelengkapan materi yang ada. Masukan dari validator ahli akan dianalisis kembali untuk merevisi instrumen sesuai dengan kebutuhan (Hermanto *et al.*, 2022). Selanjutnya dari ahli materi 2 yakni guru pengampu mata pelajaran juga memperoleh hasil yang “sangat layak” dengan nilai sebesar 93,3%. Berdasarkan hasil validasi ahli bahasa, instrumen tersebut dinyatakan memenuhi kriteria “sangat layak” dengan perolehan nilai akumulatif sebesar 93,3%. Terakhir dari ahli evaluasi menyatakan bahwa instrumen ini “layak” dengan nilai sebesar 88%. Berdasarkan hasil penilaian di atas, maka dapat dikatakan bahwasannya instrumen ini sangat layak untuk diujicobakan secara terbatas kepada peserta didik.

Selanjutnya yakni melakukan analisis butir soal setelah melakukan uji coba secara terbatas kepada peserta didik. Sebagaimana dikemukakan oleh Warju *et al* (2020) bahwa analisis butir soal sangat diperlukan untuk mengukur sejauh mana keandalan suatu instrumen evaluasi dalam memberikan hasil yang akurat. Analisis butir soal akan dilakukan menggunakan *IBM SPSS Statistics* 23. Uji yang pertama yakni uji validitas, butir soal akan dinyatakan valid apabila apabila nilai  $r_{hitung} > r_{tabel}$ , dimana  $r_{tabel}$  untuk 36 responden adalah 0,270. Merujuk pada data uji validitas, ditemukan empat item soal yang dikategorikan tidak valid. Butir-butir soal yang tidak memenuhi standar validitas tersebut mencakup nomor soal 3, 10, 18, serta 26. Menurut Arikunto (2018) butir soal yang tidak valid tidak memiliki kontribusi yang memadai sehingga butir soal yang tidak valid harus dibuang. Hasil validitas dari soal nomor 3 yakni 0,159 lalu hasil validitas nomor 10 yakni -0,425, selanjutnya di nomor 18 hasil uji validitasnya yakni 0,269 dan terakhir pada nomor 26 sebesar -0,071. Angka – angka tersebut berada di bawah 0,270 yang artinya tidak dapat dinyatakan valid. Sehingga berdasarkan olah data yang diperoleh menunjukkan bahwa terdapat 26 soal (86,67%) yang dinyatakan “valid” dan 4 (13,33%) soal yang dinyatakan “Tidak Valid”.

Setelah uji validitas, butir soal akan diuji reliabilitasnya Menurut Heale & Twycross, (2015) sebuah instrumen dikatakan reliabel apabila melebihi angka 0,70. Hasil analisis reliabilitas menunjukkan nilai *Cornbach's Alpha* sebesar 0,879 yang artinya instrumen tes dikatakan memiliki reliabilitas yang tinggi. Tahap berikutnya adalah pengujian daya pembeda, yang bertujuan untuk mengukur efektivitas setiap butir soal dalam mengklasifikasikan tingkat kompetensi peserta didik secara akurat (Lestari & Yudhanegara, 2024). Dari hasil analisis daya pembeda, dapat disimpulkan bahwa sebanyak 19 soal (63,33%) dinyatakan “sangat baik”, 3 soal (10%) dinyatakan “baik”, 5 soal (16,67%) dinyatakan “cukup”, dan 3 soal (10%) dinyatakan “kurang”. 3 soal yang dinyatakan kurang daya pembedanya yakni pada nomor 3, 10, dan 26. Analisis tingkat kesukaran menjadi tahapan selanjutnya, di mana pengukuran dilakukan dengan meninjau persentase perolehan jawaban benar dari total populasi subjek yang mengikuti uji coba instrumen. Berdasarkan hasil uji diperoleh hasil bahwa sebanyak 1 soal (3,33%) dinyatakan “mudah”, 27 soal (90%) dinyatakan “sedang” dan 2 soal (6,67%) dinyatakan “sukar”. Hasil tersebut menunjukkan bahwa tingkat kesukaran berada dalam kriteria yang memadai. Terakhir yakni uji pengecoh soal, menurut Arikunto (2018) pengecoh soal dapat dikatakan berfungsi apabila dipilih peserta didik paling sedikit 5%. Berdasarkan hasil analisis dapat disimpulkan bahwa seluruh pengecoh soal berfungsi dengan baik karena setiap pilihan jawaban dipilih lebih dari 5% oleh peserta didik.

### **Respon Peserta Didik Terhadap Pengembangan Instrumen Tes Berbasis HOTS Berbantuan Wayground Pada Elemen Pengelolaan Kearsipan**

Respon peserta didik terhadap pengembangan instrumen ini dikumpulkan melalui *google form*. Menurut Aisyah *et al* (2016) respon peserta didik penting karena untuk mengetahui umpan balik mereka setelah uji coba dilakukan. Respon peserta didik berisikan 10 pernyataan dengan 2 indikator yakni respon dan reaksi yang mana nantinya setiap indikator akan terdapat beberapa pernyataan. Indikator respon meliputi empat pernyataan terkait penerimaan peserta didik terhadap instrumen. Hasil menunjukkan bahwa kesesuaian soal dengan materi elemen 3 memperoleh kriteria positif, menandakan keselarasan dengan

pembelajaran di sekolah. Kemudahan memahami materi dan perluasan pengetahuan melalui soal pilihan ganda memperoleh kriteria sangat positif, yang mengindikasikan bahwa soal berbasis HOTS mampu memperluas wawasan dan mendorong pemikiran di luar materi yang dipelajari (Fanani, 2018). Selain itu, aspek keterbacaan bahasa juga mendapat kriteria positif, sehingga soal dinilai cukup mudah dipahami.

Sementara itu, indikator reaksi terdiri dari enam pernyataan yang menggambarkan sikap dan perilaku peserta didik selama mengerjakan instrumen. Ketertarikan terhadap soal berbasis teknologi memperoleh kriteria sangat positif, menunjukkan bahwa penggunaan teknologi dapat meningkatkan partisipasi dan minat belajar (Noor, 2020). Dorongan untuk berpikir lebih mendalam melalui soal HOTS juga mendapat kriteria sangat positif, yang menegaskan perannya dalam mengembangkan kemampuan berpikir kritis dan kreatif (Fitriyani & Ghani, 2024). Aspek motivasi berada pada kriteria positif, mengindikasikan bahwa fitur gamifikasi Wayground cukup efektif meningkatkan motivasi belajar. Selain itu, peningkatan semangat belajar, kejujuran dalam mengerjakan soal, serta kehati-hatian dalam memilih jawaban semuanya memperoleh kriteria sangat positif, yang menunjukkan bahwa integrasi teknologi mampu meningkatkan antusiasme, ketelitian, dan kemampuan berpikir kritis peserta didik.

#### **Keterampilan berpikir tingkat tinggi peserta didik melalui pengembangan instrumen tes berbasis HOTS berbantuan Wayground pada elemen pengelolaan kearsipan kompetensi keahlian MPLB**

Hasil pengerjaan tes menjadi indikator utama untuk mengidentifikasi sejauh mana tingkat penguasaan keterampilan berpikir kritis dan analitis peserta didik. Menurut Purbaningrum (2017), kemampuan tersebut tercermin dari nilai yang diperoleh masing-masing peserta didik. Berdasarkan analisis hasil tes, diketahui bahwa 5 peserta didik (13,89%) berada pada kriteria sangat baik, 7 peserta didik (19,44%) pada kriteria baik, 11 peserta didik (30,55%) pada kriteria cukup, 9 peserta didik (25%) pada kriteria kurang, dan 4 peserta didik (11,11%) pada kriteria sangat kurang. Data tersebut menunjukkan bahwa secara keseluruhan keterampilan berpikir tingkat tinggi peserta didik kelas XI MP 3 berada pada kategori cukup. Temuan ini selaras dengan penelitian Maulidia & Pahlevi (2020) yang menyebutkan bahwa kemampuan berpikir tingkat lanjut peserta didik SMK mayoritas masuk dalam kategori "cukup", sehingga data ini dapat dijadikan indikator umum mengenai kondisi kemampuan nalar siswa di lingkungan SMK.

Kondisi tersebut tidak terlepas dari pola asesmen yang selama ini digunakan. Berdasarkan analisis awal, pembelajaran masih didominasi oleh soal berbasis LOTS, sehingga peserta didik masih belum adaptasi dengan soal HOTS. Hal ini sejalan dengan Wardany *et al* (2017) yang mengemukakan bahwa meskipun guru dituntut menyusun asesmen berbasis HOTS, praktik di lapangan masih didominasi LOTS. Oleh karena itu, wajar jika saat pertama kali dihadapkan pada soal HOTS, hasil peserta didik masih didominasi kategori cukup dan sebagian masih sangat kurang. Temuan ini menegaskan pentingnya penerapan instrumen berbasis HOTS secara berkelanjutan, mengingat keterampilan berpikir tingkat tinggi merupakan kompetensi krusial bagi peserta didik SMK untuk menghadapi dunia kerja. Hal ini didukung oleh Setiawati *et al* (2019) yang menegaskan bahwa kecakapan tersebut meliputi dimensi berpikir kritis, logis, kreatif, dan kemampuan dalam menyelesaikan masalah, yang mana seluruhnya hanya dapat terstimulasi secara maksimal melalui penggunaan instrumen yang dikonstruksi secara spesifik.

#### **4. Kesimpulan**

Pengembangan instrumen tes pilihan ganda berbasis HOTS ini menggunakan jenis penelitian R&D dengan model 4D (*define, design, develop, disseminate*). Setiap tahapannya memiliki beberapa langkah yang harus dilakukan untuk mengetahui lebih lanjut mengenai kebutuhan mengapa instrumen ini harus dikembangkan. Hasil validasi menunjukkan bahwa

instrumen yang dikembangkan dinyatakan "sangat layak" oleh ahli materi serta ahli bahasa dengan persentase penilaian sebesar 93,3%. Di sisi lain, tinjauan dari ahli evaluasi menghasilkan skor 88% dengan kategori "layak". Temuan tersebut mengonfirmasi bahwa perangkat tes pilihan ganda berbasis HOTS yang telah dirancang memiliki tingkat kelayakan yang sangat tinggi, sehingga siap untuk diimplementasikan pada tahap uji coba kepada siswa. Selanjutnya dilakukan uji coba terbatas kepada peserta didik kelas XI MP 3 SMK Ketintang Surabaya yang berjumlah 36 peserta didik. Setelah dilakukan uji coba, akan dilakukan analisis butir soal untuk mengetahui kualitas dari butir soalnya. Berdasarkan hasil analisis butir soal sebanyak 26 butir soal dinyatakan valid dan dikatakan reliabel dengan kategori "sangat baik". Dilihat dari daya pembeda soal, disimpulkan bahwa sebanyak 19 soal (63,33%) dinyatakan "sangat baik", 3 soal (10%) dinyatakan "baik", 5 soal (16,67%) dinyatakan "cukup", dan 3 soal (10%) dinyatakan "kurang". Tingkat kesukaran pada pengembangan instrumen ini yakni sebanyak 1 soal (3,33%) dinyatakan "mudah", 27 soal (90%) dinyatakan "sedang" dan 2 soal (6,67%) dinyatakan "sukar". Berdasarkan analisis pengecoh, didapatkan hasil bahwa seluruh soal pengecohnya berfungsi dengan baik. Dari hasil respon peserta didik diketahui bahwasannya mayoritas memiliki hasil yang "sangat positif", artinya instrumen ini memiliki umpan balik yang bagus dari peserta didik. Terkait kapasitas berpikir tingkat tinggi, setiap siswa tentu menunjukkan tingkat pencapaian yang beragam. Berdasarkan data penelitian, teridentifikasi bahwa mayoritas kemampuan berpikir kompleks peserta didik berada pada kategori "cukup".

## References

- Aisyah, Panjaitan, R., & Marlina, R. (2016). Respon Siswa Terhadap Mediae-Comic Bilingual Sub Materi Bagian-Bagian Darah. *Jurnal Pendidikan Dan Pembelajaran Khatulistiwa (JPPK)*, 5(3), 1–10. [https://doi.org/Respon,E-comic bilingual, Bagian darah](https://doi.org/Respon,E-comic%20bilingual,%20Bagian%20darah).
- Arikunto, S. (2018). *Dasar-Dasar Evaluasi Pendidikan*. Bumi Aksara.
- Basuki, I., & Hariyanto. (2016). *Asemen Pembelajaran*.
- Desilva, D., Sakti, I., & Medriati, R. (2020). Pengembangan Instrumen Penilaian Hasil Belajar Fisika Berorientasi Hots (Higher Order Thinking Skills) Pada Materi Elastisitas Dan Hukum Hooke. *Jurnal Kumparan Fisika*, 3(1), 41–50. <https://doi.org/10.33369/jkf.3.1.41-50>
- Fanani, M. Z. (2018). Strategi Pengembangan Soal Higher Order Thinking Skill (HOTS) Dalam Kurikulum 2013. *Journal of Islamic Religious Education*, 11(1), 57–76.
- Fitriyani, D., & Ghani, M. Al. (2024). Penerapan Soal HOTS dalam Meningkatkan Motivasi dan Kemampuan Belajar Siswa Kelas 8 . 1 SMPN 3 TANGSEL pada Materi Segitiga dan Segiempat. *Seminar Nasional Dan Publikasi Ilmiah 2024 FIP UMJ*, 1707–1714.
- Hamidah, M., & Wulandari, S. (2021). Pengembangan Instrumen Penilaian Berbasis Hots Menggunakan Aplikasi "Quizizz." *Efisiensi : Kajian Ilmu Administrasi*, 18(1), 105–124. <https://doi.org/https://doi.org/10.21831/efisiensi.v17i1.34895>
- Handayani, R., & Wulandari, D. (2021). Modern Assessment dalam Menyongsong Pembelajaran Abad 21 dan Hambatan di Negara Berkembang. *Jurnal Pendidikan Edutama*, 8(1), 13. <https://doi.org/10.30734/jpe.v8i1.1363>
- Hascan, M. A., & Suyadi, S. (2021). Penerapan Teori Belajar Kognitif Pada Mata Pelajaran PAI Tingkat SMP di SIT Bina Insan Batang Kuis. *Edumaspul: Jurnal Pendidikan*, 5(2), 138–146. <https://doi.org/10.33487/edumaspul.v5i2.1548>
- Heale, R., & Twycross, A. (2015). Validity and reliability in quantitative studies. *Evidence-Based Nursing*, 18(3), 66–67. <https://doi.org/10.1136/eb-2015-102129>
- Hermanto, F. Y., Wulandari, S. S., Muyassaroh, J., & Hidayati, B. (2022). Validitas Modul Praktik Perkantoran Menggunakan Metode Pembelajaran Role- Playing Berbasis Kebutuhan Industri. *Kajian Pendidikan Ekonomi Dan Ilmu Ekonomi*, 9(2), 103–111. <https://doi.org/https://doi.org/10.36706/jp.v9i2.19014>
- Ismail, F. (2018). *Pengantar Evaluasi Pendidikan* (M. Astuti (ed.)). Karya Sukses Mandiri (KSM).
- Kurniawan, A., Febrianti, A. N., Hardianti, T., Ichsan, Desy, Risan, R., Sari, D. M. M., Sitopu, J. W., Dewi, R. S., Sianipar, D., Fitriyah, L. A., Zulkarnaini, Jalal, N. M., Hasriani, & Hasyim, F.

- (2022). *Evaluasi Pembelajaran*. PT. Global Eksekutif Teknologi.
- Kurniawan, D. (2020). *Assesment For Learning (AFL) dalam Pendidikan Matematika*. CV Budi Utama.
- Latifah, A., Husaini, F., Nisa, A., & Shaleh. (2023). Pengembangan Instrumen Penilaian Berbasis Hots. *Didaktik : Jurnal Ilmiah PGSD FKIP Universitas Mandiri*, 09(02), 4486–4496.
- Lestari, K. E., & Yudhanegara, M. R. (2024). *Penelitian Pendidikan Matematika*. PT. Refika Aditama.
- Magfirah, I., Afiyati, F., & Bashith, A. (2025). Transformasi Evaluasi Pembelajaran PAI berbasis Digital: Optimalisasi Media Quizizz sebagai Alat Ukur Adaptif. *AL-QALAM : Jurnal Kajian Islam Dan Pendidikan Volume*, 17(1), 435–444. <https://doi.org/10.47435/al-qalam.v17i1.3949>
- Maulidia, F., & Pahlevi, T. (2020). Pengembangan Instrumen Penilaian Tes Soal Pilihan Ganda Berbasis HOTS Pada Mata Pelajaran Administrasi Umum Jurusan OTKP SMK Negeri 1 Lamongan. *Jurnal Pendidikan Administrasi Perkantoran (JPAP)*, 8(1), 136–145.
- Mufida, R. E., Lestari, A. D., Hikmah, A., Rosita, M. A., & Muhtarom, T. (2025). Jenis , Pengaruh , dan Permasalahan dalam Modernisasi Pembelajaran Tematik Sekolah Dasar dengan Pemanfaatan Media Pembelajaran Digital. *Jurnal Pengabdian Masyarakat Dan Riset Pendidikan*, 4(1), 3712–3722. <https://doi.org/https://doi.org/10.31004/jerkin.v4i1.1851>
- PENDAHULUAN
- Noor, S. (2020). Penggunaan Quizizz dalam penilaian pembelajaran pada materi ruang lingkup Biologi untuk meningkatkan hasil belajar siswa kelas X.6 SMA 7 Banjarmasin. *Jurnal Pendidikan Hayati*, 6(1), 1–7.
- Nurhadi. (2020). Teori kognitivisme Serta Aplikasinya dalam pembelajaran. *Jurnal Edukasi Dan Sains*, 2(1), 77–95.
- Nurhayati, E. (2020). Meningkatkan Keaktifan Siswa Dalam Pembelajaran Daring Melalui Media Game Edukasi Quiziz Pada Masa Pencegahan Penyebaran Covid-19. *Jurnal Penelitian Dan Pengembangan Pendidikan*, 7(3), 145–150. <https://doi.org/10.34012/bip.v2i2.1729>
- Purba, P. A. (2021). Penerapan aplikasi Quizizz sebagai media penilaian online pada pembelajaran matematika selama masa pandemi covid-19. *Prosiding Seminar Nasional Matematika Dan Pendidikan Matematika*, 6, 62–68.
- Purbaningrum, K. A. (2017). Kemampuan Berpikir Tingkat Tinggi Siswa Smp Dalam Pemecahanmasalah Matematika Ditinjau Dari Gaya Belajar. *JPPM*, 10(2), 40–49.
- Puspitasari, R. O., & Wulandari, S. S. (2021). Development of Higher Order Thinking Skills Based Assessment Instruments Using Live Worksheet Application. *Dinamika Pendidikan*, 16(2), 194–203. <https://doi.org/10.15294/dp.v16i2.33433>
- Retnawati, H., Djidu, H., Kartianom, K., Apino, E., & Anazifa, R. (2018). Teachers' knowledge about higher-order thinking skills and its learning strategy. *Problems of Education in the 21st Century*, 76, 215–230. <https://doi.org/10.33225/pec/18.76.215>
- Rohmad. (2017). *Pengembangan Instrumen Evaluasi dan Penelitian*. Kalimedia.
- Serevina, V., Sari, Y. P., & Maynastiti, D. (2019). Developing high order thinking skills (HOTS) assessment instrument for fluid static at senior high school. *Journal of Physics: Conference Series*, 1185(1), 1–10. <https://doi.org/10.1088/1742-6596/1185/1/012034>
- Setiawati, W., Asmira, O., Ariyana, Y., Bestary, R., & Pusjiastuti, A. (2019). *Buku Penilaian Berorientasi Higher Order Thinking SKills*. Direktorat Jenderal Guru dan Tenaga Kependidikan Kementerian Pendidikan dan Kebudayaan.
- Sugiyono. (2023). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. CV Alfabeta.
- Suhardini, A., Mutmainnah, S., & Rozaini, N. (2025). Pengembangan E-Modul Pengelolaan Kearsipan Berbasis Moodle Untuk Meningkatkan Hasil Belajar Siswa Manajemen Perkantoran Dan Layanan Bisnis SMK Swasta BinaSatria Medan. *JPEK (Jurnal Pendidikan Ekonomi Dan Kewirausahaan)*, 9(1), 127–138. <https://doi.org/10.29408/jpek.v9i1.29484>
- Supriadi, G. (2020). *Penilaian Berbasis Higher Order Thinking Skill ( HOTS ) Oleh : Aswaja Pressindo*.

- Suratman, B., Wulandari, S. S., & Nugraha, J. (2019). Relevance of office administrative education and vocational high school curriculum to improve teacher learning: empirical study from Indonesia. *International Journal of Applied Business and Economic Research*, 15(April), 263–272. <http://www.serialsjournal.com>
- Susiatin. (2019). Meningkatkan Kemampuan Guru Dalam Menyusun Kisi-Kisi Soal Dengan Metode Pendampingan Pola “ OCF .” *Jurnal Dinamika Manajemen Pendidikan*, 4(1), 17–24.
- Uktolseja, L. J., Boekorsyom, C. D., Manuhutu, M. A., Manurung, T., Maer, S., Salay, V. S., & Manutmasa, M. F. (2025). Implementasi CALL melalui Penggunaan Quizizz sebagai Media Pembelajaran Telling the Time di Kelas VII-B SMP Emeyodere Berbasis Pesantren (SBP) Kota Sorong. *Jurnal Pengabdian Masyarakat Bhinneka*, 3(4), 870–876. <https://doi.org/10.58266/jpmb.v3i4.280>
- Wahab, G., & Rosnawati. (2021). *Teori-Teori Belajar dan Pembelajaran* (H. Zanki Azmi (ed.)). Penerbit Adab.
- Wardany, K., Sajidan, & Ramli, M. (2017). Pengembangan Penilaian Untuk Mengukur Higher Order Thinking Skills Siswa. *Jurnal Inkuiri*, 6(2), 1–6.
- Warju, Ariyanto, S. R., Soeryanto, & Trisna, R. A. (2020). Analisis kualitas butir soal tipe hots pada kompetensi sistem rem siswa di sekolah menengah kejuruan. *Jurnal Pendidikan Teknologi Dan Kejuruan*, 17(1), 95–104.
- Widana, I. W. (2017). Modul Penyusunan Soal Higher Order Thinking Skill (HOTS). Jakarta: Direktorat Jenderal Pendidikan Dasar dan Menengah Departemen Pendidikan Dan Kebudayaan. *Direktorat Jendral Pendidikan Dasar Dan Menengah*.
- Yanti, S. F., & Hasyim. (2025). Pengembangan Lembar Kerja Peserta Didik (LKPD) Berbasis Aplikasi Corel Draw Melalui Pendekatan Problem Based Learning Untuk Meningkatkan Pemahaman Dan Kreativitas Siswa Pada Elemen Pengelolaan Arsip Di SMK Negeri 7 Medan. *Jurnal Pemasaran Bisnis*, 7(2), 14–27.
- Zamkay, Y., & Wulandari, S. S. (2022). Pengembangan Instrumen Evaluasi Berbasis HOTS Mata Pelajaran OTK Humas Dan Keprotokolan di SMK IPIEMS Surabaya. *Jurnal Pendidikan Administrasi Perkantoran (JPAP)*, 10(1), 67–80.