



Jurnal Inovasi Pendidikan Matematika

Volume 2, Nomor 2, Desember 2021

Pengembangan Alat Evaluasi Pembelajaran Matematika Berbasis *Two Tier Multiple Choice* Berlevel HOTS Menggunakan *Ispring Suite 9*

Wardhatus Afifah¹, Sarwo Edy², Syaiful Huda³

Pendidikan Matematika, FKIP, Universitas Muhammadiyah Gresik, Jl. Sumatera No. 101 GKB Gresik, Jawa Timur Indonesia 61121^{1,2,3}; wardhatusafifah123@gmail.com¹; sarwoumg@gmail.com²; syaifulhuda@umg.ac.id³

Abstract

The purpose of this study was to develop an evaluation tool based on two tier multiple choice with HOTS level using ispring suite 9 which can be used online and can familiarize and hone students' higher order thinking skills. Because considering the current curriculum is the 2017 revised K13 which integrates with 4C and HOTS. The research and development model used is 4D. The research instruments used were interview sheets, validation sheets, test instruments, and questionnaires. Data analysis techniques used qualitative and quantitative. The results of the study can be seen from: Valid criteria based on the percentage of media experts by 97% and material experts by 89%, both categories are very feasible. The effective criteria based on the percentage of students' questionnaire responses were 76% in the interesting category. As well as the quality of the items seen from the validity there are 9 valid questions, the reliability is 8 reliable questions, the level of difficulty is obtained a percentage of 80% moderate, 10% difficult and 10% easy, the distinguishing power is obtained 6 good questions, 2 good enough, 1 very good question, and 1 question was bad, the effectiveness of the trick was 7 options selected by less than 2 students.

Keywords: *Research Development (R&D), evaluation tools, two tier multiple choice, HOTS, ispring suite 9*

Abstrak

Tujuan penelitian ini adalah mengembangkan alat evaluasi berbasis *two tier multiple choice* berlevel HOTS menggunakan *ispring suite 9* yang dapat digunakan secara *online* serta dapat membiasakan dan mengasah kemampuan berpikir tingkat tinggi peserta didik. Karena mengingat kurikulum yang berlaku saat ini adalah K13 revisi 2017 yang mengintegrasikan dengan 4C dan HOTS. Model penelitian pengembangan yang digunakan adalah 4D. Instrumen penelitian yang digunakan adalah lembar wawancara, lembar validasi, instrumen tes, dan angket. Teknik analisis data menggunakan kualitatif dan kuantitatif. Hasil penelitian dapat dilihat dari: Kriteria valid didasarkan pada presentase ahli media sebesar 97% dan ahli materi sebesar 89%, keduanya kategori sangat layak. Kriteria efektif berdasarkan presentase angket respon peserta didik sebesar 76% dalam kategori menarik. Serta kualitas butir soal dilihat dari validitasnya terdapat 9 soal valid, reliabilitasnya ada 8 soal reliabel, tingkat kesukarannya diperoleh presentase 80% sedang, 10% sukar dan

10% mudah, daya pembedanya diperoleh 6 soal baik, 2 soal cukup baik, 1 soal sangat baik, dan 1 soal jelek, efektivitas pengecohnya ada 7 opsi yang dipilih kurang dari 2 peserta didik.

Kata Kunci : Penelitian Pengembangan (R&D), alat evaluasi, *two tier multiple choice*, HOTS, *ispring suite 9*

INFO ARTIKEL

<p>ISSN : 2733-0597 e-ISSN : 2733-0600 DOI : http://dx.doi.org/10.30587/postulat.v2i2.3835</p>	<p style="text-align: center;">Jejak Artikel</p> <p>Submit Artikel: 2 Juni 2021 Submit Revisi: 25 Oktober 2021 Upload Artikel: 26 Desember 2021</p>
--	---

A. PENDAHULUAN

Akibat pandemi COVID-19 banyak dampak yang dirasakan oleh pendidik dan peserta didik. Pendidik harus bisa melakukan proses pembelajaran yang efektif meskipun dilakukan secara *online*. Maka dari itu dibutuhkan pula sebuah alat evaluasi untuk mengukur kemampuan matematika, untuk melihat sejauh mana peserta didik mengetahui pemahaman terhadap materi yang disampaikan. Namun dalam melakukan evaluasi secara *daring* kebanyakan pendidik hanya memberikan soal mengenai pemahaman konsep untuk mengetahui kemampuan pemahaman konsepnya saja tanpa memperhatikan kemampuan yang lainnya.

Mengingat kurikulum yang berlaku saat ini adalah kurikulum 2013 revisi 2017 yang mengintegrasikan keterampilan abad 21 atau diistilahkan dengan 4C (*creative, critical thinking, communicative, and collaborative*) dan *High Order Thinking Skills* (HOTS) (Kemendikbud, 2019). *High Order Thinking Skills* (HOTS) diterapkan mengingat rendahnya peringkat (PISA) dan (TIMSS). Berdasarkan studi TIMSS tahun 2015, Indonesia berada pada peringkat ke-46 dari 51 negara. Sedangkan hasil studi tahun 2018, Indonesia berada pada peringkat 72 dari 78 negara (Permana, 2019). Berdasarkan fakta tersebut, merupakan bukti bahwa pembelajaran matematika yang belum menekankan kemampuan berpikir tingkat tinggi. Maka dari itu, kemampuan berpikir tingkat tinggi perlu diketahui sejak dini. Hal itu dapat dilakukan dengan, memberikan evaluasi pembelajaran dengan soal yang berlevel HOTS. Mengingat HOTS juga sekarang diterapkan pada program pemerintah

yang baru, yaitu Asesmen Kompetensi yang akan diterapkan di sekolah menengah sebagai pengganti UN. Sehingga untuk saat ini tentunya peserta didik sudah harus dibiasakan untuk mengerjakan soal bermodel HOTS.

Dalam proses pembelajaran, terdapat berbagai macam tes yang digunakan. Bentuk tes yang sering digunakan adalah essay (uraian) dan *multiple choice* (pilihan ganda). Masing-masing dari bentuk tes yang sering digunakan ini memiliki kelebihan dan kekurangan.

Bentuk tes essay, memberikan kebebasan kepada setiap penempuh tes untuk mengekspresikan daya nalarnya, sehingga jawaban yang diberikan oleh setiap penempuh tes akan menunjukkan kemampuan berpikir secara kompleks (Smith, 2020). Selain itu bentuk soal essay juga meminimalisir peserta didik yang asal menjawab. Karena kebanyakan peserta didik yang diberikan soal essay apabila tidak bisa menjawab maka akan mengosongi jawabnya. Namun demikian ada kelemahan bentuk tes essay. Bentuk tes essay dalam memberikan skor membutuhkan waktu yang lama dan relatif lebih sulit. Di samping itu, penskoran bentuk tes essay bersifat subjektif dan harus dilakukan oleh ahli atau yang berwenang sehingga tidak dapat dilakukan komputersasi dalam penskorannya (Smith, 2020).

Berbeda dengan bentuk tes essay, bentuk tes *multiple choice* lebih praktis dalam penskorannya. Pada bentuk tes *multiple choice* siapa pun yang memeriksa akan memberikan skor yang sama, sehingga kesalahan dalam penskoran dapat menjadi kecil. Hal ini juga ditegaskan oleh (Sudijono, 2015) bahwa bentuk tes *multiple choice* memberikan hasil yang objektif yang tinggi dan dapat dipercaya, dalam pengoreksian jawabannya tidak sulit, dan lebih representatif dalam hal mencakup dan mewakili materi yang telah diajarkan. Menurut sebagian banyak orang, tes bentuk seperti ini lebih efisien penggunaannya dalam proses evaluasi pembelajaran karena bisa mencakup materi pembelajaran yang lebih luas. Hal ini sejalan dengan sebuah pendapat yang dikemukakan oleh (Arikunto S. , 2016) bahwa tes bentuk pilihan ganda merupakan bentuk tes objektif yang paling banyak digunakan karena banyak sekali materi yang dapat dicakup.

Pada kondisi saat ini, dimana semua pembelajaran dilakukan secara *online*, maka model soal berbasis *two tier multiple choice* (TTMC) dapat menjadi alternatif solusinya.

Two tier multiple choice diagnostic tes, merupakan tes diagnostik yang efektif, yang terdiri dari dua tingkatan, dan dikembangkan oleh David F. Treaguast pada tahun 1988. Pada tingkat pertama merupakan pertanyaan yang diajukan dari materi yang telah diajarkan, dan tingkat kedua merupakan penjelasan atas alasan yang diambil dari jawaban ditingkat pertama. Dengan mengungkapkan alasan atas jawaban yang dipilih, maka memudahkan pendidik untuk mendiagnosa peserta didik yang mengetahui kemampuan peserta didik pada kategori tinggi, rendah dan sedang. Selain itu memudahkan pendidik untuk memberikan penilaian. Dengan penggunaan *two tier multiple choice diagnostic tes*, dapat meminimalisir peserta didik asal menjawab soal evaluasi yang telah diberikan.

Sedangkan menurut (Rositasari, 2014) *two tier multiple choice* memiliki kelebihan: (a) Menurunkan kemampuan menebak. Pada tes pilihan ganda biasa, apabila terdapat lima pilihan jawaban maka setiap pilihan jawaban memiliki 20% kemungkinan jawaban benar. Sedangkan pada tes pilihan ganda *Two-tier*, apabila terdapat lima pilihan memiliki 4%kemungkinan untuk mendapatkan jawaban benar. (b) Memungkinkan menggabungkan beberapa aspek dalam suatu fenomena. Pada *tier* pertama, siswa diminta untuk menjawab pertanyaan. Sedangkan pada *tier* kedua, peserta didik diminta untuk memberikan alasan dari jawaban tersebut. Oleh karena itu, dapat dikatakan bahwa *tier* pertama merupakan *menological domain* sedangkan *tier* kedua merupakan *conceptual doamain*. (c) *Two tier multiple choice* memiliki kelebihan dari pada pilihan ganda biasa pada umumnya. Tes diagnosis *two tier* lebih mudah dikelola dan dihitung dibandingkan dengan metode lainnya, dengan demikian sangat berguna digunakan dalam kelas.

Selain kelebihan-kelebihan yang sudah dijelaskan di atas, soal berbentuk *two tier multiple choice* dapat dikatakan lebih baik dari soal bentuk essay jika memiliki pengocoh yang baik dan juga dengan daya pembeda yang baik. Hal ini juga sependapat dengan (Wulandari, 2015) yang menyatakan bahwa kualitas suatu soal berbentuk *two tier multiple choice* dapat dikatakan baik dengan memperhatikan beberapa indikator kualitas soal yang digunakan dalam tes, terutama pada daya pembeda dan pengecoh.

Ada penelitian sebelumnya yaitu penelitian yang dilakukan (Girsang, 2019) yang mengembangkan tes *two-tier multiple choice* untuk mengukur kemampuan berpikir tingkat tinggi. Menunjukkan bahwa tes diagnostik pilihan ganda dua tingkat dengan soal bermodel

HOTS tersebut dapat mengidentifikasi kemampuan berpikir tingkat tinggi dari peserta didik dengan rincian; Kualitas soal tes sangat baik dengan presentase validasi 91,33%, tes memiliki reliabilitas yang cukup adalah 0,587. Diskriminasi indeks pertanyaan diperoleh oleh 58% pertanyaan di kategori sangat baik, 21% pertanyaan dalam kategori baik, 21% pertanyaan dalam kategori buruk. Dan keefektifannya sebesar 79%.

Namun penelitian yang dilakukan oleh Multia instrumen yang digunakan masih dalam bentuk konvensional. Padahal, peserta didik seringkali merasa takut, cemas, bosan, dan bahkan kurang antusias ketika guru melakukan evaluasi pembelajaran dengan alat yang konvensional menggunakan kertas. Di zaman yang serba modern ini, peserta didik lebih tertarik kepada sesuatu yang berhubungan dengan media teknologi informasi dan komunikasi sebagai alat evaluasi dalam pembelajaran.

Dengan demikian, dari penelitian tersebut pendidik perlu melakukan pembaruan dalam mengadakan evaluasi pembelajaran. Misalnya menggunakan evaluasi dengan sistem tes *online* menggunakan aplikasi pembuatan soal yang dapat digunakan secara *online*. Untuk memberikan rasa menarik pada peserta didik, maka alat evaluasi harus bisa didesain semenarik mungkin untuk menciptakan alat evaluasi yang efektif, interaktif, dan efisien. Terdapat banyak *software* yang bisa digunakan untuk membuat alat evaluasi pembelajaran, seperti *ispring suite 9*. Menurut (Cahyati, 2018) *ispring suite 9* merupakan salah satu tool yang mengubah file presentasi menjadi bentuk *flash* dan bentuk SCORM / AICC, yaitu bentuk yang biasa digunakan dalam pembelajaran dengan *e-learning* LMS (*Learning Management System*). Karena tampilan dari aplikasi ini mirip dengan power point maka setiap pendidik bisa menggunakannya, selain itu *ispring suite 9* merupakan salah satu aplikasi yang bisa digunakan untuk alat evaluasi yang dapat diintegrasikan menjadi aplikasi yang digunakan di *handphone* sehingga mempermudah peserta didik dan juga pendidik.

Sehingga berdasarkan dari uraian tersebut, peneliti ingin mengembangkan sebuah alat evaluasi yang dapat mengatasi permasalahan tersebut dan ingin melakukan sebuah penelitian yang berjudul “Pengembangan Alat Evaluasi Pembelajaran Matematika Berbasis *Two Tier Multiple Choice* Berlevel HOTS Menggunakan *Ispring Suite 9*”

B. METODE PENELITIAN

Jenis penelitian yang digunakan dalam penelitian ini adalah penelitian pengembangan (R&D). Dalam penelitian pengembangan ini, peneliti menggunakan model pengembangan 4D yang dikembangkan oleh Sivasailam Thiagarajan, Dorothy S. Semmel, dan Melvyn I. Semmel (1974). Model pengembangan ini terdiri dari beberapa tahapan yaitu *Define, Design, Development, dan Disseminate*. Namun, penelitian ini terbatas 3D atau pada tahap *development*, tidak melakukan tahap *disseminate*, hal ini seperti yang dilakukan oleh (Noto, 2014).

Subjek dari penelitian ini adalah kelas VIII-A UPT SMP Negeri 31 Gresik dengan jumlah peserta didik sebanyak 32 dengan rincian perempuan dan laki-laki masing-masing sebanyak 16 anak. Rancangan penelitian ini adalah mengembangkan sebuah alat evaluasi berupa alat evaluasi pembelajaran matematika berbasis *two tier multiple choice* berlevel HOTS menggunakan *ispring suite 9* dengan bentuk aplikasi online yang bisa di *install* di *handphone*. Dalam pengumpulan data menggunakan beberapa metode diantaranya adalah 1). Wawancara terstruktur dilakukan dengan secara langsung tatap muka kepada informan yaitu guru mata pelajaran matematika dan beberapa peserta didik, 2). Validitas instrumen dengan menunjukkan alat evaluasi yang dikembangkan serta memberikan lembar validasi, 3). Metode tes evaluasi *two tier multiple choice* berlevel HOTS dengan alat evaluasi yang dikembangkan untuk mendapatkan nilai dari peserta didik yang kemudian akan dianalisis datanya untuk melihat kualitas dari butir soal yang telah dibuat. Dilakukan dengan membagikan aplikasi alat evaluasi yang dikembangkan di grup *WhatsApp* dan meminta untuk menginstall dan mengerjakan soal tes evaluasi tersebut 4). Metode angket dengan menyebarkan *link bit.ly google form* yang sudah termuat di alat evaluasi yang dikembangkan yang berisi angket respon peserta didik terhadap alat evaluasi yang telah digunakan.

Instrumen yang digunakan dalam penelitian ini adalah lembar wawancara, lembar validasi, instrument tes, dan angket respon peserta didik. Untuk teknik analisis data dengan menggunakan kualitatif dan kuantitatif. Analisis data kualitatif diperoleh dari hasil wawancara secara langsung kepada guru mata pelajaran matematika dan beberapa peserta didik. Selain dari wawancara, data kualitatif juga diperoleh dari saran atau masukan para validator saat tahap validasi ahli. Sedangkan analisis data kuantitatif memaparkan hasil dari

pengembangan alat evaluasi pembelajaran matematika berbasis *two tier multiple choice* berlevel HOTS menggunakan *ispring suite 9* dan analisis butir soal.

Keberhasilan pengembangan alat evaluasi pembelajaran matematika berbasis *two tier multiple choice* menggunakan *ispring suite 9* jika memenuhi kriteria valid dan efektif.

C. HASIL PENELITIAN

Penelitian pengembangan ini dilakukan dengan mengacu pada tahap-tahap model 4D, yaitu model yang dikembangkan oleh Sivasailam Thiagarajan, Dorothy S. Semmel, dan Melvyn I. Semmel. Dimana tahapan-tahapannya yaitu *Define*, *Design*, *Development*, dan *Disseminate*. Penelitian ini dibatasi tidak melaksanakan tahap keempat, karena produk pengembangan tidak disebarluaskan dikarenakan keterbatasan waktu dan biaya. Tahapan penelitian ini sebagai berikut:

1. *Define* (Pendefinisian)

Pada tahap *define* ini ada beberapa langkah analisis yang dilakukan, yaitu analisis awal akhir, analisis peserta didik, analisis tugas, analisis konsep, dan perumusan tujuan pembelajaran sebagai berikut :

a. Analisis Awal Akhir

Di langkah ini dilakukan wawancara secara langsung dengan tatap muka kepada guru mata pelajaran matematika UPT SMP Negeri 31 Gresik. Dari hasil wawancara didapatkan masalah mendasar yaitu pendidik ingin menerapkan soal HOTS untuk membiasakan peserta didik dan melatih peserta didik mengerjakan soal HOTS namun di masa pandemi saat ini sulit untuk diterapkan. Maka dari itu peneliti mendapatkan alternatif penyelesaian dengan mengembangkan alat evaluasi *online* dengan soal-soal berbasis *two tier multiple choice* berlevel HOTS menggunakan *ispring suite 9* untuk mempermudah pendidik dalam memberikan soal berlevel HOTS dan membiasakan peserta didik dalam mengerjakan soal HOTS.

b. Analisis Peserta Didik

Pada langkah ini yang dilakukan adalah wawancara secara langsung kepada beberapa peserta didik yang telah dipilihkan oleh guru matematika. Wawancara tersebut dilakukan kepada 6 peserta didik. Dari keenam peserta didik yang telah diwawancarai,

Semuanya belum pernah mengerjakan soal berlevel HOTS dan belum pernah ditanya alasan memilih jawaban tersebut serta belum pernah mengetahui *two tier multiple choice*. Peserta didik menginginkan evaluasi pembelajaran matematika yang menarik, interaktif, dengan dilengkapi multimedia seperti visual, ataupun video. Selain itu peserta didik saat melakukan pembelajaran secara *online* ini biasanya menggunakan *handphone*. Sehingga dari hasil wawancara tersebut, dapat disimpulkan bahwa peserta didik menginginkan tes evaluasi yang interaktif dan mudah diakses. Maka alat evaluasi yang dikembangkan harus dirancang sesuai dengan kebutuhan dan keinginan peserta didik.

c. Analisis Tugas

Pada langkah ini yang dilakukan adalah mengidentifikasi keterampilan-keterampilan utama yang dilakukan peserta didik selama pembelajaran, dan mengetahui kemampuan peserta didik selama mengikuti proses tes evaluasi. Berdasarkan analisis didapatkan hasil bahwa peserta didik terkadang dalam menyelesaikan tes pilihan ganda tidak ditanyakan proses atau alasan memilih jawaban, serta lebih senang proses tes evaluasinya dengan hal baru, seperti dilengkapi animasi atau multimedia lainnya. Serta peserta didik lebih suka melakukan tes evaluasi secara online. Dalam pengembangan alat evaluasi pembelajaran matematika berbasis *two tier multiple choice* berlevel HOTS menggunakan *ispring suite 9* ini menggunakan materi yang ditetapkan yaitu materi teorema pythagoras.

d. Analisis Konsep

Pada pengembangan alat evaluasi pembelajaran matematika berbasis *two tier multiple choice* berlevel HOTS dengan aplikasi *ispring suite 9* menggunakan konsep-konsep pada materi teorema pythagoras yang berkaitan dengan panjang sisi pada segitiga dan sudut dalam segitiga. Selain itu, dalam alat evaluasi ini didukung fitur-fitur aplikasi untuk memudahkan dalam penggunaan alat evaluasi yang dikembangkan.

e. Perumusan Tujuan Pembelajaran

Berdasarkan hasil analisis KI dan KD yang telah didiskusikan bersama guru mata pelajaran matematika. Kemudian ditetapkan tujuan dilakukan evaluasi berbasis *two tier multiple choice* berlevel HOTS menggunakan *ispring suite 9*.

2. Design (Perancangan)

Pada tahap perancangan terdapat beberapa langkah, diantaranya yaitu penyusunan tes acuan, pemilihan media, pemilihan format dan rancangan awal media sebagai berikut:

a. Penyusunan Tes Acuan

Penyusunan tes acuan ini menghasilkan kisi-kisi soal, yang kemudian akan dijadikan patokan pembuatan soal. Soal yang akan dipakai dalam evaluasi ini adalah 20 soal pilihan ganda dengan rincian 10 soal di tingkat pertama, dan 10 alasan di tingkat kedua.

b. Pemilihan Media

Pada langkah pemilihan media, dilakukan untuk menetapkan media yang akan digunakan dalam pengembangan. Pada penelitian ini digunakan beberapa media pada proses penelitian. Seperti yang terlihat pada tabel berikut

Tabel 1 Media dan Kegunaanya

Media	Kegunaan
Microsoft Word 2013	Digunakan untuk penyusunan soal
FilmoraGo	Untuk pembuatan soal dalam bentuk video animasi
Corel Draw X7	Mendesain background, ikon aplikasi, dan gambar soal
Ispring Suite 9	Digunakan untuk pembuatan alat evaluasi
Web2Apk	Untuk mengubah file html5 dalam bentuk aplikasi android

c. Pemilihan Format

Pada langkah ini dilakukan pemilihan format yang akan digunakan dalam rancangan awal. Penjabaran pemilihan format bisa dilihat pada tabel berikut:

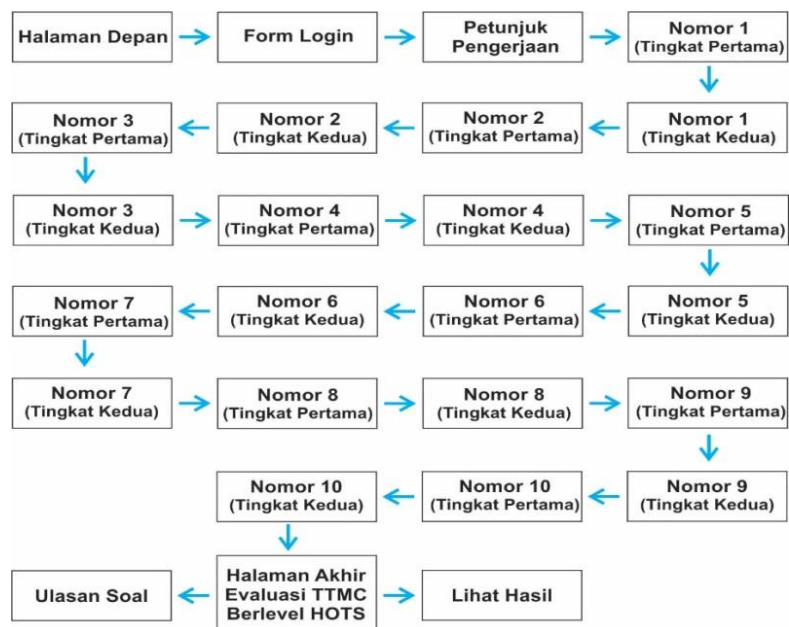
Tabel 2 Pemilihan Format Pada Aplikasi Yang Digunakan

Aplikasi	Format Yang Digunakan
Microsoft Word 2013	Dalam pembuatan soal ini digunakan format size font 12, jenis font Time new Roman dengan spasi 1,5 dan memberi warna kuning untuk menandai jawaban.

FilmoraGo	Memberikan animasi kartun agar menarik dan menambahkannya dengan audio, untuk durasi videonya adalah 30 detik, dan format penyimpanannya MP4.
Corel Draw X7	Untuk background alat evaluasi didesain dengan posisi portrait. Desain icon aplikasi berisi nama dari alat evaluasi tersebut, dengan ukuran 5 x 5 cm. Desain gambar penunjang soal dibuat sesuai kebutan dari soal tersebut
Ispring Suite 9	File ispring ini disimpan dalam bentuk html5 agar bisa diubah ke aplikasi android.
Web2Apk	Dipilih format fleksibel disemua ukuran handphone dan full screen agar tampilan dari aplikasi alat evaluasi ini bisa menyesuaikan ukuran handphone dari peserta didik

d. Rancangan Awal



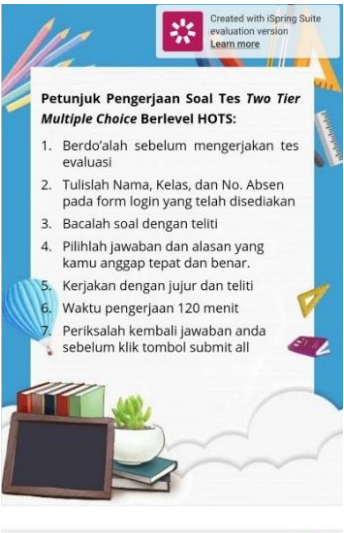
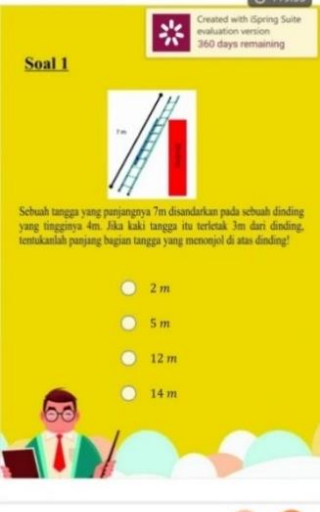
Dalam tahap ini akan menghasilkan desain awal terkait alat evaluasi yang dikembangkan. Dari rancangan awal juga menghasilkan diagram alur mengenai alat evaluasi yang dikembangkan, diagram alur ini menjelaskan proses jalannya alat evaluasi ketika di install di handphone. Berikut adalah diagram alurnya.

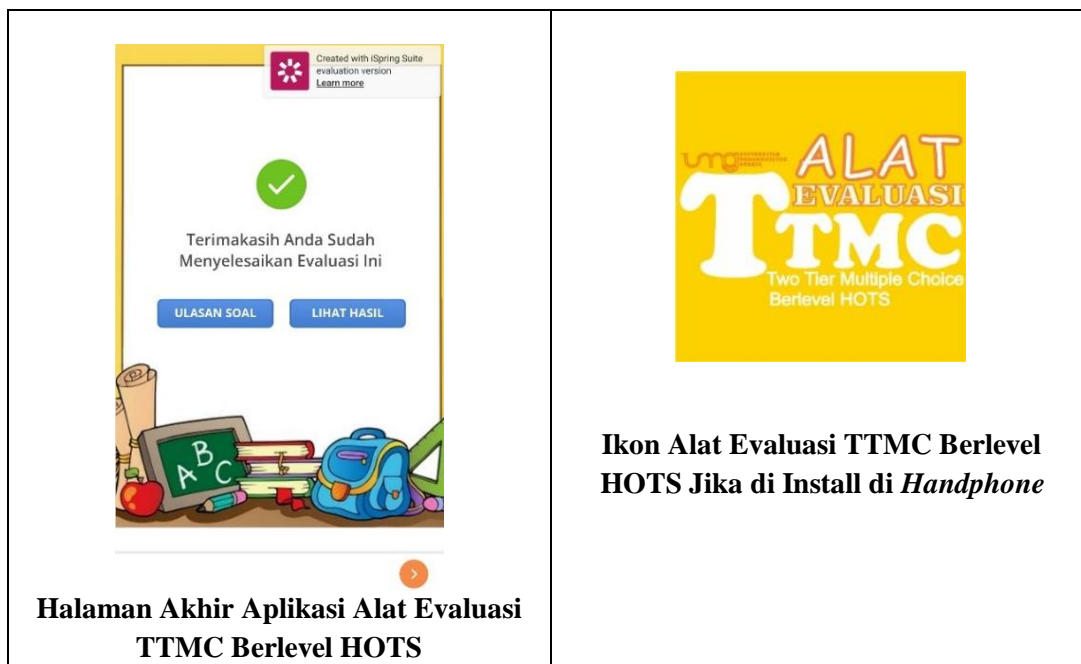


Gambar 1 Diagram Alur Alat Evaluasi TTMC Berlevel HOTS

Adapun desain awal dari alat evaluasi *two tier multiple choice* berlevel HOTS dapat dilihat pada tabel dibawah ini :

Tabel 3 Desain Awal Alat Evaluasi TTMC Berlevel HOTS

 <p>Halaman Depan Pada Aplikasi Alat Evaluasi TTMC Berlevel HOTS</p>	 <p>Form Login Pada Aplikasi Alat Evaluasi TTMC Berlevel HOTS</p>
 <p>Petunjuk Pengerjaan Evaluasi TTMC Berlevel HOTS</p>	 <p>Tampilan Soal Alat Evaluasi TTMC Berlevel HOTS</p>



3. *Development* (Pengembangan)

Pada tahap pengembangan ada 2 langkah, diantaranya adalah uji kelayakan atau validasi, dan uji coba sebagai berikut :

a. Uji Kelayakan/ Validasi Ahli

Validasi ahli terdiri ahli media, dan ahli materi. Untuk nama validator disajikan pada tabel 4 berikut :

Tabel 4 Nama Validator Ahli

Validator	Ahli Media	1. Umi Chotijah, S. Kom, M. Kom (Dosen Teknik Informatika) 2. Heri Ardiansyah, S.T, M.T. (Dosen D1 Komputer)
	Ahli Materi	1. Fatimatul Khikmiyah, M.Sc. (Dosen Pendidikan Matematika) 2. Anam Widodo, S.Pd., M.Pd. (Guru Matematika UPT SMPN 31 Gresik)

Adapun hasil dari validasi atau penilaian oleh para ahli dapat disimpulkan pada tabel 5 berikut :

Tabel 5
Hasil Validasi

	Ahli Media	Ahli Materi
Skor total	175	177
Persentase	97%	89%
Kualifikasi	Sangat layak	Sangat layak

Kesimpulan

Dari hasil validasi para ahli yang disajikan pada tabel 5 diatas dapat diketahui bahwa alat evaluasi pembelajaran matematika berbasis *two tier multiple choice* berlevel HOTS menggunakan *ispring suite 9* memenuhi kriteria valid. Sebab hal ini hasil validasi ahli media dan ahli materi mendapatkan persentase $> 61\%$ yakni sebesar 97% dan 89%.

b. Uji Coba Pengembangan

Setelah alat evaluasi yang dikembangkan dikatakan layak oleh para ahli. Maka alat evaluasi tersebut diuji cobakan. Pada tahap uji coba ini, peneliti melakukan beberapa langkah diantaranya :

1. Langkah Uji Coba Produk

Peneliti membagikan aplikasi alat evaluasi TTMC berlevel HOTS ke pendidik melalui *WhatsApp*, dan pendidik menginstruksikan kepada peserta didik untuk mendownload aplikasi tersebut dan menginstall di *handphone* masing-masing.

Untuk mengetahui hasil tes evaluasi berbasis *two tier multiple choice* berlevel HOTS menggunakan *ispring suite 9* pertama kali harus menentukan jumlah jawaban benar dari peserta didik. Untuk menentukan jumlah jawaban benar dari setiap soal jika pada tingkat pertama dan tingkat kedua benar, maka dianggap benar dengan mendapat skor 1. Tapi kalau pada tingkat pertama benar dan tingkat kedua salah ataupun sebaliknya, maka dianggap salah dan mendapat skor 0. Begitujuga pada tingkat pertama dan tingkat kedua salah, maka dianggap salah dan diberi skor 0.

Selain, skor yang diperoleh dari peserta didik, peneliti juga harus mengetahui jawaban yang dipilih oleh peserta pada setiap tingkat dalam masing-masing soal, yang mana hasilnya akan digunakan untuk analisis butir soal. Hal ini dapat dilihat pada tabel berikut :

Tabel 6 Jawaban Peserta Didik Pada Setiap Butir Soal Tes Evaluasi

	Nomor 1								Nomor 2							
	Tingkat Pertama				Tingkat Kedua				Tingkat Pertama				Tingkat Kedua			
Alternatif Jawaban	A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D
Frekuensi	30	1	0	1	3	2	26	1	5	19	2	6	8	5	7	12

	Nomor 3								Nomor 4							
	Tingkat Pertama				Tingkat Kedua				Tingkat Pertama				Tingkat Kedua			
Alternatif Jawaban	A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D
Frekuensi	3	5	21	3	4	7	14	7	24	2	4	2	2	26	3	1

	Nomor 5								Nomor 6							
	Tingkat Pertama				Tingkat Kedua				Tingkat Pertama				Tingkat Kedua			
Alternatif Jawaban	A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D
Frekuensi	20	3	1	8	15	7	4	6	14	5	6	7	5	7	6	14

	Nomor 7								Nomor 8							
	Tingkat Pertama				Tingkat Kedua				Tingkat Pertama				Tingkat Kedua			
Alternatif Jawaban	A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D
Frekuensi	18	5	6	3	6	6	14	6	2	4	25	1	24	4	2	2

	Nomor 9								Nomor 10							
	Tingkat Pertama				Tingkat Kedua				Tingkat Pertama				Tingkat Kedua			
Alternatif Jawaban	A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D
Frekuensi	3	23	3	3	3	3	23	3	12	6	8	6	16	5	8	3

Keterangan : = Kunci Jawaban

2. Langkah Meminta Respon Peserta Didik

Membagikan angket yang sudah dimuat di alat evaluasi dan sudah disambungkan ke *google form*. Dari hasil angket penilaian peserta didik terhadap alat evaluasi tersebut didapatkan hasil sebagai berikut.

Tabel 7 Hasil Angket Respon Peserta Didik

No	Indikator Penilaian	Persentase Skor	Kriteria
1	Kemenarikan	81 %	Sangat Menarik
2	Kualitas Isi	73 %	Menarik
3	Kebahasaan	76 %	Menarik
4	Kemudahan	74 %	Menarik
Rata – Rata Persentase		76 %	Menarik

3. Langkah Analisis Butir Soal

Dari hasil tes *two tier multiple choice* berlevel HOTS kemudian dianalisis untuk melihat kualitas butir soal yang ada pada tes evaluasi jika dilihat dari validitas, reliabilitas,

tingkat kesukaran, daya pembeda dan efektivitas pengecohnya. Dimana hasil kualitas butir soal tersebut adalah :

a. Validitas

Pada uji validitas, butir soal dapat dikatakan valid jika memiliki nilai validitas r_{pbis} lebih dari r_{α} . Berdasarkan hasil uji validitas ada 1 soal dikatakan tidak valid

b. Reliabilitas

Kualitas butir soal dapat dikatakan baik jika memiliki derajat reliabilitas tinggi (lebih dari r_{α}). Berdasarkan uji reliabilitas diperoleh $r_{hitung}(KR - 20) = 0,817 > r_{\alpha} = 0,70$ Maka soal yang digunakan reliabel. Berdasarkan hasil uji reliabilitas ada 2 soal dikatakan tidak reliabel

c. Tingkat Kesukaran

Pada analisis tingkat kesukaran, kualitas butir soal lebih baik menghasilkan tingkat kesukaran 0,31 – 0,70 atau dalam kategori sedang. Berdasarkan analisis tingkat kesukaran diperoleh bahwa persentase tingkat kesukaran 80% dikatakan sedang, 10% sukar, dan 10% mudah.

d. Daya Pembeda

Kualitas butir soal harus memiliki daya pembeda minimal cukup atau $DP > 0,2$. Berdasarkan hasil uji daya pembeda yang dilakukan diperoleh bahwa persentase butir soal yang memiliki $DP > 0,2$ sebesar 90%. Sedangkan butir soal yang memiliki $DP < 0,2$ adalah 10%.

e. Efektivitas Pengecoh

Kualitas butir soal yang digunakan dalam tes juga dianalisis efektivitas pengecohnya, pengecoh (*distractor*) dapat dikatakan baik jika dipilih oleh minimal 5% dari seluruh peserta. Berdasarkan hasil analisis efektivitas pengecoh didapatkan dari 40 pengecoh jawaban di tingkat pertama ada 35 pengecoh yang berfungsi dengan baik, sedangkan pada tingkat kedua yakni pengecoh alasan menunjukkan bahwa dari 40 pengecoh alasan terdapat 38 pengecoh yang berfungsi dengan baik.

D. KESIMPULAN, DISKUSI DAN REKOMENDASI

Dari penelitian pengembangan alat evaluasi pembelajaran matematika berbasis *two tier multiple choice* berlevel HOTS dapat disimpulkan sebagai berikut :

1. Dengan kriteria yang diberikan, pengembangan alat evaluasi pembelajaran matematika berbasis *two tier multiple choice* berlevel HOTS menggunakan *ispring suite 9* telah berhasil dilakukan dengan kriteria valid dan efektif yang mencapai lebih dari 60%.
2. Dengan uji statistik yang dilakukan, pada instrumen telah memenuhi kriteria yang ditetapkan, sehingga alat evaluasi pembelajaran matematika berbasis *two tier multiple choice* berlevel HOTS menggunakan *ispring suite 9* dapat digunakan.

E. DAFTAR PUSTAKA

- Arikunto, S. (2016). *Dasar-Dasar Evaluasi Pendidikan*. Jakarta: PT Bumi Aksara.
- Cahyati, A. D. (2018). Pengembangan Alat Evaluasi Pembelajaran Matematika dengan Ispring Suite 8. *Jurnal UIN Raden Intan Lampung*, 24.
- Girsang, R. A. (2019). Development Two-Tier Multiple Choices Instrument to Measure Higher-Order Thinking Skills. *ATLANTIS PRESS*, 1.
- Kemendikbud. (2019). *Buku Penilaian Berorientasi Higher Order Thinking Skills*. Jakarta: Direktorat Jendral Guru Dan Tenaga Kependidikan.
- Noto, M. S. (2014). Perangkat Pembelajaran Matematika Berbasis SMART. *Jurnal Ilmiah Program Studi Pendidikan Matematika STKIP Siliwangi Bandung*, Vol 3, No 1.
- Permana, R. H. (2019). *PISA dan TIMSS, 2 Acuan Mendikbud Untuk Hapus Ujian Nasional*. Jakarta: DetikNews.
- Rositasari, D. (2014). Pengembangan Tes Diagnostik Two Tier Untuk Mendeteksi Miskonsepsi Siswa SMA Pada Topik Asam Basa. *Jurnal UIN Jakarta*, 36-37.
- Smith, W. (2020). Keunggulan dan Kelemahan Tes Hasil Belajar. *Jurnal BDK Palembang*, 1.
- Sudijono, A. (2015). *Pengantar Evaluasi Pendidikan*. Jakarta: PT Raja Grafindo Persada.
- Wulandari, R. R. (2015). Instrumen Two Tier Tes Aspek Pengetahuan Untuk Mengukur Keterampilan Proses Sains (KPS) Pada Pembelajaran kimia Untuk Siswa SMA/MA Kelas XI. *Jurnal Pendidikan Kimia(JPK)*, 151.