

PENGEMBANGAN TES BENTUK SOAL PILIHAN GANDA DAN URAIAN BERBASIS *HIGHER ORDER THINKING SKILLS* (HOTS) PADA RANAH KOGNITIF MATA PELAJARAN PEMROGRAMAN DASAR (STUDI PADA SMK NEGERI 12 MALANG)

Meilita Dwi Aryati¹, Satrio Hadi Wijoyo^{*2}, Admaja Dwi Herlambang³

^{1,2,3}Universitas Brawijaya, Malang

Email: ¹meilitaaryati@student.ub.ac.id, ²satriohadi@ub.ac.id, ³herlambang@ub.ac.id

^{*}Penulis Korespondensi

(Naskah masuk: 23 April 2021, diterima untuk diterbitkan: 27 Februari 2023)

Abstrak

Hasil belajar siswa dapat diketahui dengan tes. Dilihat dari survei internasional pada tahun 2018 (PISA), negara Indonesia menempati nilai terendah di Dunia karena soal yang digunakan masih belum mengukur *Higher Order Thinking Skills* (HOTS). Pada penelitian ini dilakukan pengembangan tes dengan bentuk soal pilihan ganda dan uraian yang menggunakan indikator HOTS pada ranah kognitif seperti menganalisis (C4), mengevaluasi (C5), dan mencipta (C6). Metode yang digunakan pada penelitian ini yaitu metode *research and development* (R&D) dengan model *formative research* dari Tessmer, yang terdiri dari 4 tahap pengembangan yaitu *preliminary*, *self evaluation*, *prototyping*, dan *field test*. Hasil tahap *preliminary* yaitu telah menentukan subjek dan tempat penelitian yang berada di SMK Negeri 12 Malang, namun pelaksanaan pengujian dilakukan secara daring karena adanya Pandemi COVID-19. Tahap *self evaluation* menghasilkan desain soal yang disebut dengan *prototype* I. Tahap *Prototyping* menghasilkan soal yang telah diperbaiki pada uji ahli dan uji *one-to-one* yang disebut dengan *prototype* II, pada Uji *small group* tidak terdapat perbaikan karena mendapatkan respon positif dari peserta didik. Hasil pengujian *field test* yaitu terkumpul data sebanyak 29 orang, lalu dilakukan penskoran dan analisis kualitas soal. Hasil analisis kognitif peserta didik memiliki rata-rata cukup baik, sedangkan hasil analisis HOTS pada indikator menganalisis, mengevaluasi, dan mencipta adalah sangat kurang. Hal tersebut disebabkan karena peserta didik tidak mampu menjawab soal dengan baik dan banyak menjawab secara asal-asalan. Oleh karena itu perlu adanya pembiasaan menjawab soal dengan kriteria HOTS agar dapat melatih kognitif dan kemampuan berpikir tingkat tinggi.

Kata kunci: pengembangan tes, kemampuan berpikir tingkat tinggi (HOTS), *formative research*

DEVELOPMENT OF TEST TYPE MULTIPLE CHOICE QUESTION AND ESSAY QUESTION BASED ON HIGHER ORDER THINKING SKILLS (HOTS) IN THE COGNITIVE DOMAIN OF BASIC PROGRAMMING SUBJECTS (STUDY AT SMK NEGERI 12 MALANG)

Abstract

The result of student learning can be known by the test. From the international survey in 2018 (PISA), Indonesia has lowest score in the world, because the questions used still not measured *Higher Order Thinking Skills* (HOTS). In this study, the development of test was carried out with the type of multiple choice questions and descriptions, using HOTS indicators in the cognitive domain such as analyzing (C4), evaluating (C5), and creating (C6). The method used in this study is *research and development* (R&D) method with the *formative research* model from Tessmer, which consist of 4 development stages namely *preliminary*, *self evaluation*, *prototyping*, and *field test*. The result of the *preliminary* stage has been determined the subject and place of research at SMK Negeri 12 Malang, but the testing was carried out online because of the COVID-19 pandemic. The *self-evaluation* stage produces questions design called *prototype* I. The *prototyping* stage produces questions that have been corrected in the expert test process and *one-to-one* test was called *prototype* II, in the *small group* test there's no improvement because it gets a positive response from students. The results of the *field test* were 29 people data was collected then scored and analyzed the quality of the questions. The results of the cognitive analysis of students have a fairly good average, while the results of HOTS analysis on indicators of analyzing, evaluating, and creating are very lacking. This is because students are not able to answer questions well and many answer carelessly. Therefore it is necessary to get used to answering questions with HOTS criteria in order to train cognitive and higher-order thinking skills.

Keywords: *test development, higher order thinking skills (HOTS), formative research*

1. PENDAHULUAN

Proses penilaian atau evaluasi merupakan proses yang terletak di akhir pembelajaran di mana pada proses ini digunakan untuk mengetahui rancangan kegiatan selanjutnya melalui hasil yang diterima setelah kegiatan pembelajaran selesai. Pada proses evaluasi guru atau dosen atau pendidik biasanya menggunakan alat ukur berupa instrumen tes. Instrumen tes tersebut dijadikan sebagai alat ukur untuk mengetahui dan mengumpulkan informasi mengenai pengetahuan, minat, motivasi, dan sikap dari peserta didik.

Negara Indonesia memiliki nilai terendah dan menempati urutan dari bawah di dunia berdasarkan pada hasil survei pendidikan dunia pada tahun 2018 yang diselenggarakan oleh PISA (*Programme for International Student Assessment*). PISA sendiri merupakan suatu program yang diselenggarakan setiap 3 (tiga) tahun sekali dengan diikuti oleh peserta didik usia 15 tahun atau usia pendidikan kelas menengah. Program tersebut bertujuan untuk melakukan evaluasi terhadap hasil belajar peserta didik. Indonesia melakukan perbaikan kurikulum dan evaluasi belajar berdasarkan pada laporan penilaian PISA (Hewi dan Shaleh, 2020). Hasil PISA tahun 2018 menunjukkan bahwa Indonesia mendapatkan peringkat bawah karena instrumen tes yang digunakan di Indonesia masih berada dibawah PISA, sedangkan dalam PISA tes yang digunakan sudah memiliki kriteria mengukur *higher order thinking skills* (HOTS) peserta didik (Angriani, dkk., 2018), sehingga kegiatan pembelajaran dan proses penilaian diharuskan memuat aspek HOTS seperti yang tertera pada kurikulum yang telah direvisi.

Dari hasil wawancara kepada salah satu guru mata pelajaran Pemrograman Dasar yang dilakukan oleh peneliti, menyatakan bahwa soal yang biasa digunakan di SMK Negeri 12 Malang yaitu soal bentuk pilihan ganda biasa dan bentuk uraian. Dalam pembuatan soal bentuk HOTS, guru mengaku tidak mengalami kesulitan, tetapi mengalami kesulitan ketika peserta didik tidak memahami soal dengan tipe HOTS yang bersifat analisis. Namun setelah dilakukan analisis soal HOTS yang digunakan di SMK masih banyak menggunakan tipe uraian dibandingkan dengan pilihan ganda dalam pembuatan soal HOTS, lalu soal uraian yang dibuat masih belum memenuhi kriteria HOTS seperti pemberian alasan, kesimpulan, penjelasan atas pernyataan, kasus yang diberikan, dan penggunaan kata atau kalimat masih belum menyinggung kriteria HOTS (Kependidikan, 2018). Pada penelitian ini akan dikembangkan tes pilihan ganda dan uraian yang memuat kriteria soal HOTS atau kemampuan berpikir tingkat tinggi.

Kemampuan berpikir yang berada di atas taksonomi kognitif Bloom mampu mengembangkan proses belajar yang telah didapat kedalam konteks

baru dan nyata merupakan ciri dari Kemampuan berpikir tingkat tinggi (Brookhart, 1918). Kemampuan berpikir tersebut melakukan proses berpikir seperti menganalisis, mengevaluasi, dan menciptakan. Kemampuan berpikir tingkat rendah merupakan proses kemampuan berpikir mengingat, memahami, dan menerapkan. Kedua kemampuan berpikir tersebut berdasarkan pada taksonomi kognitif Bloom yang telah direvisi (Anderson, Krathwohl, & Bloom, 2001).

Berdasarkan penelitian sebelumnya mengenai pengembangan instrumen HOTS dengan bantuan aplikasi Kahoot (Jannah dan Pahlevi, 2020), bahwa tes yang dikembangkan memuat indikator C4, C5, dan C6. Bentuk soal yang digunakan adalah pilihan ganda. Instrumen tes yang dikembangkan akan diujikan dengan menggunakan bantuan aplikasi Kahoot. Berdasarkan penelitian tersebut diketahui bahwa aplikasi Kahoot memiliki kekurangan seperti keterbatasan pengisian kolom jawaban dan nilai yang didapatkan berdasarkan jawaban dan kecepatan peserta didik dalam menjawab, sehingga kurang adil jika nilai yang didapatkan sebagai hasil akhir untuk mengetahui kemampuan berpikir peserta didik.

Peneliti menggunakan media berupa Moodle dalam mengujikan soal yang dikembangkan kepada peserta didik, karena pada penelitian sebelumnya tentang pengembangan instrumen tes dengan menggunakan Moodle (Apganendaria and Sujatmiko, 2019), bahwa Moodle dapat mengatur penilaian secara otomatis saat siswa menjawab salah dan benar, serta tidak adanya batasan siswa dalam menjawab soal bentuk uraian.

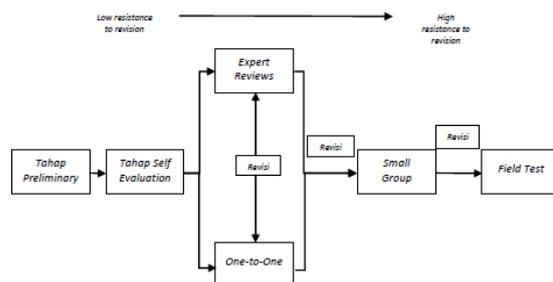
Moodle juga merupakan *e-learning* yang digunakan oleh peserta didik SMK Negeri 12 Malang ketika pelaksanaan pembelajaran dilaksanakan secara daring akibat pandemi, sehingga akan lebih efektif jika menggunakan Moodle sebagai media dalam penyampaian soal. Soal yang akan dikembangkan adalah soal bentuk pilihan ganda dan uraian yang memuat indikator kognitif C4 (menganalisis), C5 (mengevaluasi), dan C6 (mencipta). Aspek kognitif C4, C5, dan C6 merupakan aspek yang berada di atas taksonomi Bloom yaitu kemampuan berpikir tingkat tinggi. Aspek C4 (analisis) merupakan aspek yang melakukan penguraian dari suatu materi menjadi komponen yang lebih jelas. Aspek C5 (evaluasi) merupakan aspek yang melakukan pembuatan kesimpulan terhadap nilai, kasus, dan situasi. Aspek C6 (mencipta) merupakan aspek tertinggi dari ranah kognitif taksonomi Bloom dengan melibatkan keterampilan yang kompleks seperti berfikir kritis dalam memecahkan permasalahan (Ruwaida dan Selatan, 2019).

Metode yang digunakan pada penelitian ini yaitu metode pengembangan dan penelitian (R&D) dengan menggunakan model *formative research* dari Tessmer. Penggunaan metode ini berdasarkan pada

penelitian sebelumnya tentang pengembangan instrumen tes untuk mengukur pemecahan masalah siswa pada matematika (Angriani dkk., 2018), bahwa metode yang digunakan sudah sesuai dengan situasi pada penelitian ini yaitu tidak ada uji kepraktisan yang dilakukan oleh guru lain pada mata pelajaran yang sama, seperti pada penggunaan model ADDIE pada penelitian sebelumnya dari Eka Rachma dan Ayen Arsisari (Kurniasari dan Arsisari, 2020). Uji kepraktisan tersebut prosesnya sama dengan uji ahli namun dilakukan oleh guru lain pada mata pelajaran yang sama, sedangkan pada penelitian ini guru mata pelajaran Pemrograman Dasar di SMK Negeri 12 Malang hanya terdiri dari 2 orang yang juga merupakan ahli pada penelitian ini, sehingga peneliti tidak menggunakan model ADDIE karena adanya keterbatasan guru mata pelajaran Pemrograman Dasar di SMK Negeri 12 Malang untuk melakukan uji kepraktisan yang sifat dan tahapannya sama dengan validasi para ahli.

Soal yang dibuat diujikan oleh para ahli (*expert review*) yang terdiri dari 4 orang, 4 orang tersebut merupakan guru mata pelajaran pemrograman dasar SMK Negeri 12 Malang sebanyak 2 orang, dan dosen Pendidikan Teknologi Informasi Universitas Brawijaya yang terdiri dari 2 orang. Subjek pada penelitian adalah peserta didik dari SMK Negeri 12 Malang yang hanya menggunakan 2 kelas yaitu kelas X MM 1 dan X MM 3 dari SMK Negeri 12 Malang. Materi yang akan dijadikan sebagai bahan dalam pembuatan soal terdapat pada Kompetensi Dasar 3.6 dan Kompetensi Dasar 3.7 mata pelajaran Pemrograman Dasar. Tujuan penelitian ini adalah mengetahui hasil pengembangan tes kognitif HOTS soal pilihan ganda dan uraian serta mengetahui hasil pengukuran kualitas tes yang dikembangkan.

2. METODE PENELITIAN



Gambar 1 Model *Formative Research* Dari Tessmer.

Gambar 1 merupakan metode yang digunakan pada penelitian yaitu dari Tessmer dengan model *formative research* yang merupakan metode pengembangan (R&D) (Angriani dkk., 2018). Model tersebut terdiri dari 4 tahap pengembangan yaitu *preliminary*, *self evaluation*, *prototyping*, dan *field test*. Berikut merupakan penjelasan dari keempat tahap tersebut:

2.1 Tahap *Preliminary*

Melakukan kajian literatur pada referensi sebelumnya yang telah dicari untuk dijadikan sebagai pedoman dalam menyelesaikan penelitian ini merupakan kegiatan yang dilakukan pada tahap *preliminary*, seperti referensi dalam membuat soal yang baik dan benar, lalu pada tahap ini juga terdapat proses penentuan subjek penelitian dan lokasi penelitian.

2.2 Tahap *Self Evaluation*

Pada tahap *self evaluation* terdapat 4 proses kegiatan seperti analisis kurikulum, analisis peserta didik, analisis materi, dan desain tes. Pada analisis kurikulum dilakukan proses pengkajian kurikulum yang digunakan di Sekolah Menengah Kejuruan (SMK) terutama pada mata pelajaran Pemrograman Dasar. Kurikulum yang akan dianalisis yaitu kurikulum 2013.

Analisis peserta didik berkaitan dengan jumlah dan karakteristik dari peserta didik untuk mengetahui perkembangan peserta didik yang diuji coba dilihat dari kognitif peserta didik. Perkembangan kognitif dapat dilihat dari informasi hasil belajar pemrograman dasar berupa kemampuan berpikir rendah, sedang, dan tinggi.

Analisis materi dilakukan berdasar pada kurikulum 2013 dengan melakukan identifikasi pada materi yang digunakan untuk dijadikan sebagai dasar dalam pembuatan soal. Materi pada KD 3.6 tentang penerapan struktur kontrol percabangan dan materi pada KD 3.7 tentang penerapan struktur kontrol perulangan dalam bahasa pemrograman merupakan materi yang akan dianalisis pada penelitian ini.

Pada proses desain dilakukan dengan membuat desain instrumen tes yang dikembangkan. Seperti proses pembuatan kisi-kisi dan soal beserta jawaban. Lalu dilakukan pembuatan lembar validitas untuk mengetahui kevalidan instrumen tes dan pembuatan lembar respon peserta didik untuk mengetahui tingkat keterbacaan instrumen tes. Hasil dari instrumen tes yang telah didesain disebut dengan *prototype I*.

2.3 Tahap *Prototyping*

Pada tahap ini terdiri dari 3 tahapan yaitu *expert review* atau uji ahli, uji *one-to-one*, dan uji *small group*. Uji ahli atau uji validitas dilakukan dengan membagikan lembar validitas berisi pernyataan yang disesuaikan dengan instrumen tes yang dikembangkan dan dinilai oleh para ahli. Pada penelitian ini terdiri dari 4 ahli atau validator, 2 Dosen Pendidikan Teknologi Informasi yang merupakan ahli instrumen tes, dan 2 Guru mata pelajaran Pemrograman Dasar yang merupakan ahli materi. Jika terdapat saran perbaikan yang diminta oleh para ahli maka dilakukan perbaikan (*revisi*).

Pada tahap ini uji *One-to-One* dilakukan oleh 3 peserta didik yang bukan merupakan subjek penelitian yaitu dari kelas X TKJ 1 yang memiliki

tingkat kemampuan berpikir rendah, sedang, dan tinggi. Peserta didik mengisi lembar respon berupa *google form* berdasarkan pada hasil pengerjaan instrumen tes (*prototype I*). Hasil pengisian respon peserta didik bertujuan untuk mengetahui respon pada tes yang dibuat, apakah sudah sesuai dengan bahasa yang dipahami peserta didik atau belum. Jika terdapat soal yang belum difahami maka perlu dilakukan perbaikan. Proses perbaikan dengan melihat hasil respon peserta didik, jika terdapat respon yang kurang positif atau tidak positif maka perlu dilakukan perbaikan, dan jika terdapat soal yang tidak dapat dipahami seperti penggunaan kata atau kalimat yang ambigu atau berbelit-belit juga dilakukan perbaikan (*revisi*). Hasil revisi dari uji coba ahli dan uji *one-to-one* disebut dengan *Prototype II*.

Tahap terakhir yaitu uji *small group* yang dikerjakan oleh 6 orang peserta didik yang bukan merupakan subjek penelitian. Peserta didik tersebut dipilih berdasarkan kognitif yang ditentukan yaitu kemampuan berpikir rendah, kemampuan berpikir sedang, dan kemampuan berpikir tinggi yang masing-masing tingkat kognitifnya terdiri dari 2 orang dan dilihat berdasarkan hasil belajar pemrograman dasar. Proses uji coba yang dilakukan sama dengan uji coba *one-to-one* yaitu melakukan proses pengisian lembar respon terhadap instrumen tes yang telah direvisi (*Prototype II*). Setelah melakukan pengujian *small group*, jika terdapat respon yang kurang baik maka dilakukan perbaikan sehingga dihasilkan soal yang telah diperbaiki disebut dengan *prototype III*.

2.4 Tahap Field Test

Proses uji coba hasil perbaikan dari tahap *Prototyping* yang telah menghasilkan *prototype III* merupakan kegiatan yang dilakukan pada tahap ini. Hasil perbaikan diujikan kepada subjek pada penelitian ini yaitu peserta didik SMK Negeri 12 Malang dari kelas X MM 1 dan X MM 3 dengan menggunakan Moodle sebagai media dalam penyampaian soal.

3. HASIL

3.1 Hasil Tahap Preliminary

Hasil dari tahap ini telah melakukan kajian literatur dalam pembuatan soal dengan berdasar pada kata kerja operasional (KKO) aspek kognitif C4, C5, dan C6. Tahap ini juga telah menentukan subjek penelitian dan tempat penelitian. Subjek penelitian ini adalah peserta didik dari SMK Negeri 12 Malang pada kelas X MM 1 dan X MM 3. Tempat penelitian dilakukan di SMK Negeri 12 Malang, namun untuk pelaksanaan pengerjaan instrumen tes dilakukan secara daring dengan menggunakan Moodle, karena proses pembelajaran di Sekolah dilaksanakan secara daring akibat pandemi COVID-19.

3.2 Hasil Tahap Self Evaluation

Hasil analisis kurikulum yaitu telah melakukan analisis pada kurikulum 2013 dengan mengarah pada mata pelajaran Pemrograman Dasar jurusan Multimedia di SMK. Analisis tersebut dilihat dari segi aspek Kompetensi Dasar (KD), Kompetensi Inti (KI), dan Standar Kompetensi Lulusan (SKL). Pada aspek Standar Kompetensi Lulusan (SKL), diketahui bahwa lulusan SMK harus memiliki pengetahuan, sikap, dan keterampilan sesuai dengan standar kompetensi lulusan pada Permendikbud No. 54 tahun 2013. Aspek Kompetensi Inti meliputi 4 kompetensi yang dirancang sesuai dengan jenjang peserta didik di kelas tertentu (Kompetensi Keahlian Multimedia (C2)). Keempat aspek kompetensi tersebut terdiri dari Aspek sikap spiritual, sikap sosial, pengetahuan, dan keterampilan. Selain itu hasil dari analisis kurikulum pada aspek Kompetensi Dasar (KD) topik Pemrograman Dasar terdiri dari 14 KD. Pada penelitian ini menggunakan 2 Kompetensi Dasar yaitu KD 3.6 tentang penerapan struktur kontrol percabangan dan KD 3.7 tentang penerapan struktur kontrol perulangan.

Pada hasil analisis peserta didik menggunakan subjek penelitian dari kelas X MM 1 yang terdiri dari 35 orang dan X MM 3 yang terdiri dari 38 orang. Dilihat dari hasil pengerjaan tugas mata pelajaran Pemrograman Dasar peserta didik pada kelas tersebut rata-rata memiliki tingkat kemampuan berpikir rendah, sedang, dan tinggi.

Hasil analisis materi yaitu materi yang digunakan pada penelitian ini yaitu materi struktur kontrol percabangan (KD 3.6) dan struktur kontrol perulangan. Pada materi struktur kontrol percabangan peserta didik mempelajari struktur percabangan 1 kondisi, 2 kondisi, dan percabangan bersarang. Pada materi struktur kontrol perulangan peserta didik mempelajari materi tentang struktur perulangan, perulangan *for*, *while*, *do-while*, dan perulangan bersarang.

Hasil desain pada tahap ini yaitu rancangan instrumen tes baik soal bentuk pilihan ganda dan bentuk uraian yang berjumlah 20 soal. Soal pilihan ganda terdiri dari 15 soal, sedangkan soal bentuk uraian terdiri dari 5 soal. Setiap soal dirancang sesuai dengan indikator keterampilan berpikir tingkat tinggi (HOTS) yaitu C4 (analisis), C5 (penilaian), dan C6 (cipta). Pada tahap ini menghasilkan soal dengan sebutan *prototype I*. Pada tahap desain juga telah dihasilkan kisi-kisi instrumen tes, soal beserta jawaban, dan rancangan tampilan soal pada Moodle. Pada pengujian validitas peneliti juga merancang lembar validitas yang terdiri dari 5 aspek yaitu aspek isi, kognitif, tata bahasa, aspek waktu, tampilan, serta saran dan komentar. Untuk mengetahui respon peserta didik dibuat lembar respon untuk mengetahui tingkat keterbacaan instrumen tes yang dibuat apakah sudah mudah dipahami oleh peserta didik kelas X SMK atau belum. Peneliti merancang lembar respon baik dalam bentuk *file* dan *google form*.

3.3 Hasil Tahap *Prototyping*

Tahap *prototyping* dilakukan setelah menghasilkan *prototype* I pada tahap *self evaluation*, yang terdiri dari uji ahli, uji *one-to-one*, dan uji *small group*. Pada hasil uji ahli, telah dilakukan pengujian validitas oleh para ahli yang terdiri dari 4 ahli atau validator, yaitu guru mata pelajaran pemrograman dasar SMK Negeri 12 Malang dan 2 dosen dari Pendidikan Teknologi Informasi Universitas Brawijaya. Para ahli atau validator mengisi lembar validitas sesuai dengan instrumen tes pada *prototype* I. Setelah dilakukan pengisian lembar validitas, maka dilakukan perhitungan validitas untuk mengetahui kriteria yang diperoleh dari tiap-tiap butir soal dengan menggunakan rumus Aiken's V. Diketahui dari hasil perhitungan validitas 14 soal merupakan soal yang sangat valid dan 6 soal merupakan soal yang memiliki kriteria sedang. Soal dengan kriteria sedang masih dapat diterima dan dapat dilanjutkan pada proses pengujian.

Pada hasil uji *one-to-one* telah dilakukan proses pengisian lembar respon oleh peserta didik yang bukan merupakan subjek pada penelitian. Peserta didik yang dijadikan responden terdiri dari 3 orang dan memiliki kriteria kemampuan berpikir rendah, sedang, dan tinggi sesuai dengan capaian belajar pemrograman dasar. Peserta didik mengisi respon pada *goggle form* karena pembelajaran masih dilaksanakan secara daring, lalu dari hasil pengisian respon pada *google form* direkap pada lembar respon peserta didik oleh peneliti. Setelah direkap maka dilakukan perhitungan respon. Hasil perhitungan respon dapat dilihat pada Tabel 1.

Tabel 1 Hasil Respon Peserta Didik Uji *One-to-One*

No.	Responden		
	R1	R2	R3
1	3	3	3
2	2	2	3
3	4	2	3
4	4	3	3
5	3	2	3
6	3	3	3
7	4	2	3
Jumlah	23	17	21
Persentase (P)	82%	60%	75%
Ket.	Positif	Kurang positif	Positif

Dari tabel 1 hasil respon peserta didik uji *one-to-one*, responden 1 memiliki persentase sebesar 82%, yang artinya memiliki tingkat keterbacaan baik dan respon yang positif, pada responden 2 memiliki persentase 60%, yang artinya memiliki tingkat keterbacaan belum cukup atau respon kurang positif, sedangkan pada responden 3 memiliki persentase sebesar 75% yang artinya tingkat keterbacaan pada soal sudah baik dan sudah sesuai dengan bahasa yang dipahami oleh peserta didik. Instrumen tes (*prototype* I) yang telah diujikan pada uji *one-to-one* dan uji *expert* dilakukan perbaikan, sehingga pada tahap ini menghasilkan *prototype* II.

Setelah dilakukan perbaikan maka dilakukan uji *small group* yang dilaksanakan dengan peserta didik yang bukan merupakan subjek penelitian berjumlah 6 orang yang tiap 2 orang memiliki kemampuan berpikir rendah, sedang, dan tinggi. Teknik pengujian sama dengan uji sebelumnya yaitu dengan mengisikan lembar respon. Dari respon tersebut direkap dan dihitung lalu dianalisis. Diketahui bahwa hasil uji *small group* terdapat 4 responden yang masing-masing memiliki respon sangat positif, jadi dapat disimpulkan bahwa instrumen tes pada uji *small group* memiliki tingkat keterbacaan yang sangat baik dan positif, sehingga tidak dilakukan perbaikan dan *prototype* II. *Prototype* II diujikan ke subjek penelitian pada tahap selanjutnya yaitu *field test*.

3.4 Hasil Tahap *Field Test*

Pada tahap *field test* dilakukan pengujian kepada subjek penelitian dengan total 73 peserta didik dengan waktu pengerjaan 90 menit. Proses pengujian dilaksanakan dengan menggunakan media Moodle (e-learning SMK Negeri 12 Malang). Proses pengerjaan dilaksanakan dengan rentang 5 hari untuk mengantisipasi adanya peserta didik yang tidak memiliki paket data atau sinyal dirumahnya.

Setelah dilaksanakan pengujian, diketahui terkumpul 43 peserta didik yang mengerjakan dari total 73 orang subjek penelitian, lalu hasil jawaban tersebut dianalisis oleh peneliti. Hasil jawaban yang dianalisis dilihat dari jawaban peserta didik pada soal bentuk uraian, diketahui bahwa terdapat soal yang dijawab secara asal-asalan oleh peserta didik sehingga peneliti membuang data peserta didik yang menjawab secara asal-asalan, sehingga terkumpul data sebanyak 29 orang, 14 orang dari kelas X Multimedia 1 dan 15 orang dari kelas X Multimedia 3. Data yang terkumpul dilakukan pemberian nilai atau skor. Pada soal pilihan ganda, skor 3 didapatkan peserta didik jika jawaban yang diperoleh adalah benar dan skor 1 jika jawaban yang diperoleh adalah salah. Selain itu soal uraian disesuaikan dengan pedoman penskoran yang telah disusun sebelumnya. Setelah dilakukan penskoran pada hasil jawaban peserta didik, maka dilakukan analisis kualitas dari tes yang dikembangkan seperti Reliabilitas, Daya Pembeda, Indeks Pengecoh, Tingkat Kesukaran, dan Analisis Kemampuan Berpikir Tingkat Tinggi.

3.5 Hasil Analisis Kualitas Instrumen Tes

Setelah dilakukan pengujian, maka dilakukan analisis kualitas tes, yang terdiri dari analisis reliabilitas, daya pembeda, indeks pengecoh, tingkat kesukaran, dan analisis kemampuan berpikir tingkat tinggi. Analisis reliabilitas bertujuan untuk mengetahui soal mana yang reliabel atau dapat dipercaya meskipun dilakukan pada kondisi dan waktu yang berbeda namun dengan kelompok yang sama. Analisis reliabilitas perlu melakukan perhitungan dengan menggunakan KR-21 untuk soal

pilihan ganda sedangkan untuk soal uraian dengan menggunakan perhitungan Alfa Croanbach. Hasil dari pengukuran reliabilitas soal pilihan ganda dan soal uraian dapat dilihat pada Tabel 2.

Tabel 2 Hasil Perhitungan Reliabilitas

Reliabilitas	Nilai	Kriteria
Soal Pilihan Ganda (KR-21)	5,1	Sangat Tinggi
Soal Uraian (Alfa Croanbach)	0,54	Sedang

Pada Tabel 2 diketahui bahwa soal pilihan ganda memiliki reliabilitas sangat tinggi dan soal uraian memiliki reliabilitas sedang. Reliabilitas sangat tinggi berarti soal tersebut memiliki kriteria soal yang dapat dipercaya karena memberikan hasil yang sama jika dilakukan dengan waktu dan tempat yang beda namun dikerjakan oleh kelompok yang sama. Reliabilitas sedang berarti cukup *reliable* atau cukup dipercaya dalam memberikan hasil yang sama meskipun dalam kondisi yang berbeda.

Selain itu untuk mengetahui soal yang dapat membedakan kemampuan berpikir peserta didik dengan melakukan analisis daya pembeda agar diketahui peserta didik mana saja yang memiliki kemampuan berpikir tinggi, sedang, dan rendah. Dalam melakukan analisis daya pembeda perlu melakukan perhitungan DP pada masing-masing bentuk soal yaitu soal uraian dan soal pilihan ganda. Perhitungan DP pada masing-masing bentuk soal memiliki rumus yang berbeda. Setelah dilakukan perhitungan maka dilakukan perbandingan pada kriteria yang telah ditentukan. Rekapitulasi soal dengan kriteria daya pembeda sangat baik, baik, cukup, dan tidak baik dapat dilihat pada Tabel 3.

Tabel 3 Rekapitulasi Hasil Daya Pembeda

Kriteria DP	Nomor Soal
Tidak Baik	11 dan 17
Cukup Baik	14 dan 19
Baik	2, 7, 12, 13, 16 dan 18
Sangat Baik	1, 3, 4, 5, 6, 8, 9, 10, 15, dan 20

Pada Tabel 3 Rekapitulasi diketahui bahwa soal dengan kriteria tidak baik atau tidak mampu membedakan kemampuan berpikir rendah dan tinggi adalah pada nomor 11 dan 17, oleh karena itu harus dibuang. Pada kriteria cukup baik membedakan kemampuan berpikir rendah dan tinggi adalah pada nomor 14 dan 19 sehingga perlu diperbaiki. Lalu kriteria daya pembeda baik dan sangat baik berarti mampu menentukan peserta didik yang memiliki tingkat berpikir rendah dan tinggi sehingga dapat terus digunakan.

Selanjutnya dilakukan analisis tingkat kesukaran untuk mengetahui soal dengan kriteria sukar dan mudah. Dalam melakukan analisis tingkat kesukaran dilakukan perhitungan Tingkat Kesukaran

(TK) pada masing-masing bentuk soal yaitu soal bentuk pilihan ganda dan uraian. Hasil perhitungan TK dibandingkan dengan kriteria tingkat kesukaran seperti sukar, sedang, dan mudah, hasilnya dapat dilihat pada Tabel 4.

Tabel 4 Rekapitulasi Hasil Tingkat Kesukaran

Kriteria TK	Nomor Soal
Mudah	5, 6, dan 15
Sedang	1, 3, 7, 8, 11, 13, dan 20
Sukar	2, 4, 9, 10, 12, 14, 16, 17, 18, dan 19

Pada Tabel 4 diketahui hasil tingkat kesukaran pada tiap soal yang memiliki kriteria sukar, sedang, dan mudah. Soal yang mempunyai ciri kesukaran tidak terlalu sukar dan tidak terlalu mudah (sedang) merupakan ciri soal yang baik (Fatimah dan Alfath, 2019). Oleh karena itu, soal yang mempunyai kriteria sukar dan mudah perlu dicek daya pembeda dan indeks pengecoh (Arifin, 2009). Jika daya pembeda dan indeks pengecoh baik maka soal tersebut masih ditema, jika tidak memenuhi maka perlu diperbaiki. Dari hasil pengecekan soal dengan kriteria mudah dan sukar jika dilihat dari perhitungan daya pembeda dan indeks pengecoh, soal yang diterima adalah soal dengan nomor 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, 18, dan 20. Soal yang perlu diperbaiki adalah soal nomor 14 dan 19. Untuk soal yang perlu dibuang adalah soal nomor 17.

Berikutnya analisis indeks pengecoh dilakukan untuk mengetahui pengecoh mana yang dapat mengetahui peserta didik mampu dan kurang mampu dalam menjawab soal. Analisis indeks pengecoh dilakukan pada soal dengan kriteria soal pilihan ganda karena terdapat beberapa pengecoh, soal pilihan ganda terdiri dari 15 soal. Dalam melakukan analisis indeks pengecoh perlu dilakukan perhitungan IP untuk mengetahui pengecoh dengan kriteria sangat baik, baik, jelek, dan sangat jelek. Perhitungan indeks pengecoh dilakukan untuk soal pilihan ganda, jadi dihitung IP yang diperoleh oleh masing-masing pengecoh pada masing-masing soal. rekapitulasi hasil indeks pengecoh dapat dilihat pada Tabel 5.

Pada Tabel 5 diketahui bahwa keterangan yang ada pada tabel menunjukkan pengecoh yang harus diganti atau diperbaiki (revisi) atau baik (diterima). Jika keterangan menyatakan diganti berarti pengecoh memiliki kriteria IP jelek atau sangat jelek, sedangkan jika keterangan menyatakan revisi berarti pengecoh memiliki kriteria IP kurang baik. Jika keterangan menyatakan baik maka seluruh pengecoh pada soal tersebut memiliki IP yang baik dan sangat baik. Soal yang baik hanya terdiri dari 3 soal yakni nomor 4, 10, dan 11.

Tabel 5 Rekapitulasi Hasil Indeks Pengecoh

Analisis kemampuan berpikir tingkat tinggi (HOTS) dilakukan untuk mengetahui kemampuan

Nomor Soal	Keterangan
1	2 Diganti
2	2 Direvisi
3	1 Revisi, 1 Diganti
4	Baik
5	2 Diganti
6	1 Direvisi
7	2 Direvisi
8	2 Direvisi
9	1 Direvisi, 1 Diganti
10	Baik
11	Baik
12	2 Diganti
13	1 Diganti, 1 Direvisi
14	2 Direvisi
15	1 Direvisi

berpikir dari peserta didik setelah mengerjakan instrumen tes yang dikembangkan. Dalam melakukan analisis, perlu melakukan perhitungan rata-rata nilai yang didapatkan dari hasil jawaban seluruh peserta didik yang berjumlah 29 orang pada tiap indikator masing-masing soal. Rekapitulasi hasil kemampuan berpikir tingkat tinggi (HOTS) dapat dilihat pada Tabel 6.

Tabel 6 Rekapitulasi Hasil Kemampuan Berpikir Tingkat Tinggi

Indikator	No. Soal	Rata-Rata	Kriteria
C4	1, 2, 5, 6, 7, 8, 9, 10, 13, 14, 15, dan 20	25,79	Sangat Kurang
C5	3, 4, 11, 12, 17, 18, dan 19	14,79	Sangat Kurang
C6	16	4,62	Sangat Kurang
Rata-Rata		15,07	Sangat Kurang

Pada Tabel 6 diketahui bahwa hasil rata-rata pada tiap indikator seperti indikator C4 (analisis), C5 (evaluasi), dan C6 (mengkreas) adalah sangat kurang karena peserta didik banyak yang tidak mampu menjawab soal dengan baik, tidak sesuai dengan perintah pengerjaan soal yang diberikan, dan banyak menjawab secara asal-asalan terutama pada soal uraian. Jadi dapat disimpulkan bahwa seluruh peserta didik yang melakukan pengujian dengan total 29 orang pada tiap-tiap indikator HOTS memiliki kriteria sangat kurang. Berikut merupakan contoh jawaban yang diperoleh dari peserta didik yang menjawab benar pada indikator C6 (mencipta), dapat dilihat pada Gambar 2.

Buatlah algoritma pemrograman pseudocode sesuai dengan alur flowchart diatas dan tuliskan maksud dari algoritma pemrograman tersebut ! (skor 15)

```

start
int nilai
input nilai
if (nilai >= 80) then
printf ("sangat baik")
else
if (nilai <= 80) then
printf ("cukup baik")
endif
end

```

maksud dari algoritma diatas adalah jika nilai lebih besar atau sama dengan 80 maka akan tampil tulisan "sangat baik", dan jika nilai kurang dari atau sama dengan 80 maka akan tampil tulisan "cukup baik".

Gambar 2 Contoh Jawaban Peserta Didik Menjawab Benar Pada Indikator C6

Dari Gambar 2 diketahui bahwa jawaban benar dan lengkap yang dihasilkan oleh peserta didik lengkap sesuai dengan kriteria pedoman penskoran, yaitu telah mampu membuat pemrograman *pseudocode* berdasarkan *flowchart* dan menjelaskan alur dari algoritma pemrograman tersebut sehingga skor yang didapatkan adalah 15. Selanjutnya merupakan contoh jawaban kurang tepat pada soal yang sama (C6) sehingga hanya mendapatkan skor 1. Berikut merupakan hasil jawaban salah pada indikator C6 yang dapat dilihat pada Gambar 3.

Buatlah algoritma pemrograman pseudocode sesuai dengan alur flowchart diatas dan tuliskan maksud dari algoritma pemrograman tersebut ! (skor 15)

Jawaban.

Algoritma merupakan sebuah alur atau urutan secara sistematis tujuannya untuk menyelesaikan suatu masalah dan salah satu metode penyampaianya melalui flowchart. Dalam hal ini mungkin cukuplah jelas dalam menjadi perbedaan yang paling mencolok terhadap algoritma dan flowchart.

Sedangkan flowchart merupakan sebuah penggambaran atau diagram yang mempunyai satu aliran atau lebih, namun dalam hal itu hanya berlaku secara sekuensial atau saling bisa berkesinambungan.

Selain itu juga umumnya terdapat simbol-simbol tertentu yang menjadi sebuah perwujudan dari setiap masing – masing alur, yang saling berhubungan dengan anak panah.

Gambar 3 Contoh Jawaban Peserta Didik Menjawab Salah Pada Indikator C6

Dari Gambar 3 diketahui bahwa peserta didik menjawab salah pada indikator C6. Pada soal dengan indikator C6 diminta untuk membuat pemrograman *pseudocode* sesuai dengan *flowchart* yang diberikan, namun pada gambar tersebut peserta didik menjawab pengertian algoritma dan *flowchart* yang tidak ada hubungannya dengan soal sehingga hanya mendapatkan skor 1 sesuai dengan pedoman penskoran. Disimpulkan bahwasanya peserta didik kurang mampu memahami soal dengan tipe HOTS, sehingga perlu dilakukan pembiasaan peserta didik menjawab soal kriteria HOTS pada materi lain agar peserta didik mampu melatih kemampuan berpikir mereka.

Analisis hasil kognitif peserta didik dilihat dari seluruh peserta didik yang mengerjakan soal dan jumlah soal tanpa dibedakan indikatornya. Jumlah peserta didik yang melakukan pengujian adalah 29 orang, sedangkan total soal yang dibuat adalah 25 soal. Hasil analisis kognitif peserta didik berdasarkan pada hasil jawaban seluruh soal yang dibuat dapat dilihat pada Tabel 7.

Tabel 7 Rekapitulasi Hasil Kognitif Peserta Didik

Kriteria	Total Peserta Didik	Total Skor Peserta Didik
Sangat Baik	6 orang	501
Baik	9 orang	648
Cukup Baik	6 orang	303
Kurang	8 orang	266
Sangat Kurang	0 orang	0
Jumlah Skor Total		1718
Rata-Rata		52,06
Kriteria		Cukup Baik

Pada Tabel 7 diketahui bahwa hasil kognitif peserta didik dengan kriteria sangat baik terdiri dari 6 orang, kriteria baik 9 orang, kriteria cukup baik 6 orang, kriteria kurang baik 8 orang, dan kriteria sangat kurang tidak ada atau nol. Sedangkan jika diambil rata-rata skor total, peserta didik memiliki tingkat kognitif yang cukup baik dan terbukti juga tidak ada yang memiliki kriteria sangat kurang dari total keseluruhan soal yang dikerjakan.

4. PENUTUP

4.1 Kesimpulan

Kesimpulan yang didapat adalah dalam mengembangkan instrumen tes menggunakan metode dari Tessmer dengan model *formative research* yang terdiri dari 4 tahapan. Hasil pengembangan instrumen tes pada tahap ini yaitu instrumen tes *prototype* II, karena pada tahap uji *small group* tidak ada respon yang kurang baik sehingga tidak dilakukan perbaikan, dan dilanjutkan pada tahap uji coba subjek penelitian. Subjek penelitian yaitu peserta didik dari SMK Negeri 12 Malang dengan menggunakan kelas X MM 1 dan X MM 3. Dari hasil pengerjaan instrumen tes yang dilakukan oleh peserta didik terkumpul 29 data final yang digunakan untuk dianalisis agar mengetahui kualitas dari instrumen tes. Analisis tersebut berupa analisis reliabilitas, daya pembeda, tingkat kesukaran, indeks pengecoh, dan analisis kemampuan berpikir tingkat tinggi.

Hasil analisis kualitas soal yang dikembangkan dapat disimpulkan bahwa reliabilitas soal pilihan ganda adalah sangat reliabel sedangkan reliabilitas soal uraian adalah cukup reliabel. Hasil analisis daya pembeda menyatakan bahwa diketahui soal yang baik membedakan peserta didik kelompok pandai dan tidak pandai berjumlah 16 soal, sedangkan 2 soal perlu diperbaiki, dan 2 soal lagi perlu dibuang. Hasil

analisis indeks pengecoh diketahui soal yang baik hanya terdiri dari 3 soal. Selanjutnya hasil analisis tingkat kesukaran, diketahui ciri soal yang baik memiliki kriteria kesukaran sedang dengan jumlah 7 soal, namun tingkat kesukaran pada kriteria mudah dan sedang dianalisis lagi dari hasil daya pembeda dan indeks pengecoh, jika hasil dari keduanya memenuhi syarat maka soal tersebut diterima. Jadi total soal yang diterima berjumlah 17 soal.

Hasil analisis kemampuan berpikir tingkat tinggi pada indikator menganalisis, mengevaluasi, dan mencipta memiliki kriteria sangat kurang. Hal itu disebabkan adanya peserta didik dalam menjawab soal secara singkat, tidak melakukan kegiatan analisis seperti yang diminta dalam soal, serta adanya peserta didik yang tidak faham dengan instruksi yang diberikan. Oleh karena itu perlu adanya pembiasaan untuk menggunakan soal berbasis HOTS dengan membuat soal-soal bentuk HOTS yang bervariasi dan menerapkan kriteria kemampuan berpikir tingkat tinggi.

4.2 Saran

Untuk penelitian selanjutnya diharapkan dalam membuat soal berbasis HOTS perlu menggunakan soal dengan jumlah yang sama pada tiap-tiap indikator yaitu indikator C4, C5, dan C6, atau dengan menambahkan lebih banyak soal C6, agar dalam melakukan perhitungan kualitas instrumen tes dapat seimbang dan sesuai. Serta diharapkan pada penelitian selanjutnya terdapat pengembangan tes pada mata pelajaran lain seperti mata pelajaran produktif di SMK dengan kompetensi keahlian Multimedia, Rekayasa Perangkat Lunak, dan Teknologi Komputer dan Jaringan.

DAFTAR PUSTAKA

- ANDERSON, L.W., KRATHWOHL, D.R. AND BLOOM, B.S., 2001. A taxonomy for learning, teaching, and assessing: a revision of Bloom's taxonomy of educational objectives. [online] 41(4), p.352. Tersedia Di: <<http://books.google.com/books?id=JPKX AQAAMAAJ&pgis=1>> [Diakses 18 Agustus 2020].
- ANGRIANI, A.D., NURSALAM, N., FUADAH, N. AND BAHARUDDIN, B., 2018. Pengembangan Instrumen Tes Untuk Mengukur Kemampuan Pemecahan Masalah Matematika Siswa. AULADUNA: Jurnal Pendidikan Dasar Islam, 5(2), p.211.
- APGANENDARIA, H. AND SUJATMIKO, B., 2019. Pengembangan Instrumen Tes Online Menggunakan High Order Thinking Skill (HOTS) Berbasis Moodle Pada Mata Pelajaran Rancang Bangun Jaringan di SMK Negeri 1 Surabaya. 04,

- Pp.150–155.
- ARIFIN, Z., 2009. *Evaluasi Pembelajaran*. 1st ed. Bandung: PT. REMAJA ROSDAKARYA.
- BROOKHART, S.M., 1918. *Assess Higher-Order Thingking Skills In Your Classroom*. Journal of Education, .
- FATIMAH, L.U. AND ALFATH, K., 2019. Analisis Kesukaran Soal, Daya Pembeda, Dan Fungsi Distraktor. *Jurnal Komunikasi dan Pendidikan Islam*, 8, pp.37–64.
- HEWI, L. AND SHALEH, M., 2020. Refleksi Hasil PISA (The Programme For International Student Assesment): Upaya Perbaikan Bertumpu Pada Pendidikan Anak Usia Dini). *Jurnal Golden Age*, 4(01), pp.30–41.
- JANNAH, K. AND PAHLEVI, T., 2020. Pengembangan Instrumen Penilaian Berbasis Higher Order Thinking Skills Berbantuan Aplikasi “ Kahoot !” Pada Kompetensi Dasar Menerapkan Penanganan Surat Masuk dan Surat Keluar Jurusan OTKP di SMK Negeri 2 Buduran. 8(1), pp.108–121.
- KEPENDIDIKAN, D. JENDERAL GURU DAN TENAGA, 2018. *Buku Pegangan Keterampilan Berpikir Tingkat Tinggi Berbasi Zonasi*. Direktorat Jendral Guru dan Tenaga Kependidikan.
- KURNIASARI, E.R. AND ARSISARI, A., 2020. Pengembangan Instrumen Pengukur Higher Order Thinking Skills (Hots) Matematika Pada Siswa Sekolah Menengah Pertama. 9(4), pp.1213–1222.
- RUWAIDA, H. AND SELATAN, K., 2019. Proses Kognitif Dalam Taksonomi Bloom Revisi : Analisis Kemampuan Mencipta (C6) Pada Pembelajaran Fikih Di Mi Miftahul Anwar Oleh : 4(1), pp.51–76.

Halaman ini sengaja dikosongkan