

PENGEMBANGAN INSTRUMEN TES HOTS FISIKA PADA POKOK BAHASAN FLUIDA UNTUK SMA/MA

Triwahyu Ramadhani¹, Sahyar²

(¹Mahasiswa jurusan Fisika UNIMED, ²Guru besar UNIMED)

Jurusan Fisika FMIPA Universitas Negeri Medan

triwahyuramadhani@mhs.unimed.ac.id

ABSTRAK

Penelitian ini bertujuan untuk mengembangkan soal-soal tes objektif *Higher Order Thinking Skill* (HOTS) pada materi fluida di SMA/MA yang memenuhi kualifikasi baik, meliputi aspek validitas, reliabilitas, daya pembeda, tingkat kesukaran, dan efektivitas pengecoh. Jenis penelitian ini adalah penelitian dan pengembangan (R&D), dengan model Borg dan Gall. Hasil penelitian menunjukkan bahwa validitas isi dari tes objektif HOTS Fisika materi Suhu dan Kalor di SMA/MA adalah Sangat Valid dengan persentase 98,7%. Ditinjau dari validitas empiris, diperoleh 29 butir soal (58%) valid dan 21 butir soal (42%) tidak valid. Reliabilitas tes tinggi, yaitu 0,89. Dari penelitian ini diperoleh tes objektif HOTS materi Fluida sebanyak 29 soal, meliputi 10 soal (34,6%) mengenai prinsip Archimedes, 6 soal (20,7%) mengenai Hidrostatik, 6 soal (20,7%) mengenai persamaan Bernoulli, 5 soal (17,2%) mengenai persamaan Kontinuitas, 1 soal (3,4%) mengenai Hukum Pascal, dan 1 soal (3,4%) mengenai prinsip Tegangan Permukaan. Tes HOTS ini dapat digunakan sebagai tes standar untuk mengukur kemampuan berpikir tingkat tinggi siswa pada materi Fluida di SMA/MA.

Kata Kunci : Tes Objektif HOTS, Fluida

ABSTRACT

This research aims to develop objective physics test questions on Higher Order Thinking Skill (HOTS) in fluid in SMA/MA that meet good qualifications, including aspects of validity, reliability, differentiation, level of difficulty, and effectiveness of deception. This type of research is research and development (R&D), with a model of Borg and Gall. The results of the research showed that the content validity of the objective test HOTS Physics of Temperature and Heat in SMA/MA was Very Valid with a percentage of 98.7%. Judging from empirical validity, 29 items (58%) were valid and 21 items (42%) were invalid. Test reliability is high, which is 0.89. From this research, objective test HOTS was obtained for the material Fluid as many as 29 questions, including 10 questions (34,6%) about Archimedes principle, 6 questions (20,7%) about hydrostatic, 6 questions (20,7%) about Bernoulli equation, 5 questions (17,2%) Continuity equation, 1 question (3,4%) about Pascal's Law, and 1 question (3,4%) about equation Surface Tension. This HOTS test can be used as a standard test to measure the students' higher order thinking skill in the material Fluid in SMA/MA.

Keywords: Objective Test HOTS, Fluid

PENDAHULUAN

Kemampuan berpikir tingkat tinggi atau *Higher Order Thinking Skill* (HOTS) merupakan kemampuan menghubungkan, memanipulasi, dan mentransformasi pengetahuan serta pengalaman yang sudah dimiliki untuk berpikir secara kritis dan kreatif dalam upaya menentukan keputusan dan memecahkan masalah pada situasi baru. Brookhart (2010) dalam bukunya mengategorikan HOTS menjadi tiga, yaitu HOTS sebagai transfer (*HOTS as transfer*), HOTS sebagai berpikir kritis (*HOTS as critical thinking*), dan HOTS sebagai pemecahan masalah (*HOTS as problem solving*). HOTS sebagai transfer didefinisikan sebagai keterampilan untuk mengaplikasikan pengetahuan dan keterampilan yang sudah dikembangkan dalam pembelajaran pada konteks yang baru.

Merujuk pada Anderson dan Krathwohl (2010), HOTS sebagai transfer ini meliputi tiga proses kognitif, yaitu kemampuan menganalisis (C₄), mengevaluasi (C₅), dan mencipta (C₆). HOTS sebagai berpikir kritis didefinisikan sebagai keterampilan memberikan penilaian yang bijak dan mengkritisi sesuatu menggunakan alasan logis dan ilmiah. HOTS sebagai pemecahan masalah didefinisikan sebagai keterampilan mengidentifikasi masalah dan menyelesaikan masalah menggunakan strategi yang non-otomatis.

Pada tingkat internasional, ada beberapa tes yang digunakan untuk mengukur kemampuan peserta didik, misalnya yang diselenggarakan oleh PISA, TIMSS, dan PIRLS. Tes ini mencakup kemampuan dalam membaca, matematika, dan IPA yang ditujukan untuk peserta didik di sekolah dasar dan sekolah

menengah pertama. Capaian peserta didik Indonesia pada umumnya tidak memuaskan untuk semua tes tersebut. Secara rata-rata, peserta didik di negara kita hanya mencapai level dua dari enam level yang ada pada tes PISA dan TIMSS. Berdasarkan PISA yang dilaporkan oleh *Organization Economic Co-Operation Development* (OECD), Indonesia berada pada peringkat 64 dari 65 negara pada PISA 2012. Dan mengalami kenaikan enam peringkat yaitu peringkat 62 dari 70 negara pada PISA 2015. Meski peningkatan capaian Indonesia tahun 2015 cukup memberikan optimisme, skor ini masih tergolong sangat rendah dibanding rerata negara OECD lainnya. Rendahnya capaian HOTS ini disebabkan peserta didik Indonesia belum terbiasa dengan soal-soal HOTS karena kurang tersedianya soal-soal HOTS di lapangan. Seminar hasil TIMSS dan PISA yang digelar Puspendik 2016 lalu menyatakan bahwa peningkatan pada PISA 2015 harus tetap dipertahankan dengan membiasakan siswa dengan soal-soal HOTS agar pada tahun 2030 nanti Indonesia mampu menyamai OECD. Oleh karenanya, perlu dikembangkan instrumen tes HOTS yang memenuhi aspek validitas, reliabilitas, daya pembeda, tingkat kesukaran, dan efektivitas pengecoh (distraktor).

Sesuai dengan masalah dan pembatasan yang telah dikemukakan, maka rumusan masalah dari penelitian ini ialah bagaimana validitas, reliabilitas, daya pembeda, tingkat kesukaran, dan efektivitas pengecoh (distraktor) dari tes objektif HOTS yang telah dikembangkan?

Penelitian ini bertujuan untuk mengetahui validitas, reliabilitas, daya pembeda, tingkat kesukaran, dan efektivitas pengecoh (distraktor) dari instrumen tes objektif HOTS pada materi Fluida di SMA/MA yang telah dikembangkan.

Hasil penelitian ini nantinya diharapkan dapat memberikan manfaat sebagai berikut: (1) Soal-soal tes objektif yang telah dikembangkan diharapkan dapat digunakan oleh siswa sebagai latihan guna mengembangkan serta mengoptimalkan kemampuan berpikir tingkat tinggi mereka; (2) Soal-soal tes objektif yang telah dikembangkan dapat menjadi salah satu referensi guru dalam menilai kemampuan berpikir tingkat tinggi siswa; (3) Soal-soal tes objektif yang telah dikembangkan dapat menjadi sarana dalam meningkatkan kualitas tes kemampuan berpikir tingkat tinggi siswa sehingga dapat membantu peningkatan mutu sekolah; (4) Penelitian ini menjadi bahan referensi yang dapat digunakan apabila hendak melakukan penelitian serupa.

METODE PENELITIAN

Penelitian ini dilaksanakan di MAN 02 Model Medan pada Maret-April 2019. Populasi dalam penelitian ini adalah seluruh siswa MAN 02 Model Medan yang telah mengikuti pembelajaran materi Fluida dengan jumlah populasi 756 siswa. Sampel penelitian ini berjumlah 100 siswa yang diambil dari teknik *random class*.

Jenis penelitian yang digunakan dalam penelitian ini adalah penelitian pengembangan atau *research and development (R&D)*, sedangkan model pengembangan yang digunakan menggunakan langkah riset Borg & Gall (dalam Sugiyono, 2018) yang terdiri dari sepuluh langkah, namun dalam penelitian ini dibatasi hanya sampai pada tahap ketujuh, yaitu; (1) tahap studi pendahuluan (*research and information collecting*), (2) merencanakan penelitian (*planning*), (3) pengembangan desain (*develop preliminary of product*), (4) uji lapangan terbatas (*preliminary field testing*), (5) revisi hasil uji lapangan terbatas (*main products revision*), (6) uji lapangan lebih luas (*main field test*), dan (7) evaluasi hasil uji lapangan lebih luas (*operational product revision*).

Adapun teknik analisis data yang digunakan dalam penelitian ini adalah analisis secara kualitatif dan secara kuantitatif. Analisis kualitatif dilakukan dengan meminta pertimbangan para ahli, meliputi telaah aspek materi, konstruksi, dan bahasa. Analisis kualitatif dilakukan berdasarkan validitas isi. Validitas isi sendiri dibagi menjadi dua, yakni validitas tampang (*face validity*) dan validitas logis (*logical validity*). Validitas tampang bersifat kualitatif dan judgemental karena berasal dari expert judgment. Sedangkan validitas logis bersifat kuantitatif, yang dilakukan dengan menghitung seberapa tinggi kesepakatan para expert. Hal ini dapat dilakukan dengan mencari rasio validitas isi-Lawshe's CVR. Validator diminta menilai apakah suatu item esensial (yang diperlukan dan sangat penting bagi tujuan pengukuran yang bersangkutan) dalam tiga tingkatan esensialitas yaitu 'Esensial', 'Berguna tapi tidak esensial', dan 'Tidak diperlukan.' Sedangkan analisis secara kuantitatif meliputi telaah validitas, reabilitas, taraf kesukaran, daya pembeda, dan efektivitas distraktor (pengecoh).

HASIL DAN PEMBAHASAN

Hasil Penelitian

(1) Studi Pendahuluan (*Research and Information Collecting*)

Dari tahap ini diketahui bahwa kemampuan berpikir tingkat tinggi siswa Indonesia masih tergolong rendah. Siswa di Indonesia kurang terlatih dalam menyelesaikan soal-soal

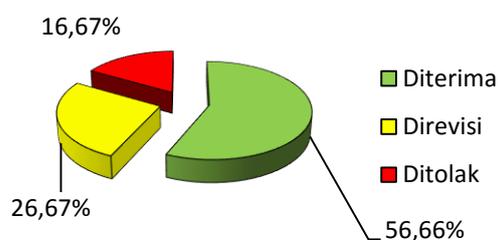
kontekstual yang mengukur kemampuan berpikir tingkat tinggi mereka, hal ini dikarenakan instrumen penilaian yang didesain khusus untuk melatih HOTS juga masih kurang tersedia di sekolah-sekolah, sehingga perlu dibuat pengembangan instrumen penilaian HOTS.

(2) Membuat Perencanaan (*Planning*)

Hasil dari tahap perencanaan ini berupa kisi-kisi soal dan lembar validasi instrumen tes. Kisi-kisi soal memuat soal-soal objektif pilihan ganda pada materi fluida yang mencakup pilihan ganda biasa, pilihan ganda asosiasi, dan hubungan antarhal (sebab-akibat) dengan sebaran kompetensi 42 butir soal C₄, 5 butir soal C₅, dan 3 butir soal C₆. Adapun lembar validasi instrumen tes berupa angket yang digunakan validator untuk menilai validitas instrumen tes berdasarkan aspek materi, konstruksi, dan bahasa.

(3) Pengembangan Desain (*Develop Preliminary of Product*)

Hasil validasi oleh ahli terhadap butir soal tes kemampuan berpikir tingkat tinggi (HOTS) pada materi Fluida di SMA/MA ditinjau dari aspek materi, konstruksi dan bahasa diketahui ada 16 item yang tidak memenuhi aspek-aspek tersebut. Ada 6 item tidak memenuhi aspek materi dan konstruksi, dan 10 item tidak memenuhi aspek materi. Dan telah dilakukan revisi terhadap ke-16 item tersebut sesuai dengan saran dari ahli.



Berdasarkan telaah butir soal tes kemampuan berpikir tingkat tinggi (HOTS) pada materi Fluida melalui proses validasi pakar menggunakan indeks CVR menunjukkan bahwa hasil analisis memiliki 34 butir soal dapat diterima, 16 butir soal direvisi dan 10 butir soal ditolak dengan persentase keidealan 98,7%. Ini berarti bahwa para ahli sepakat instrumen tes objektif kemampuan berpikir tingkat tinggi pada materi Fluida telah relevan dengan indikator dan kompetensi dasar.

(4) Uji Lapangan Terbatas (*Preliminary Field Testing*)

Uji coba ini dilaksanakan di kelas XI IPA 4 MAN 02 Model Medan dengan jumlah sampel

12 orang. Dari hasil uji coba lapangan terbatas ini dilakukan analisis data secara kuantitatif terhadap validitas, reliabilitas, daya pembeda, tingkat kesukaran, dan efektifitas pengecoh (distraktor).

(a) Validitas

Berdasarkan hasil analisis ke-50 butir soal, dapat diketahui bahwa soal yang masuk kategori valid berjumlah 20 butir soal (40%). Sedangkan soal yang termasuk kategori tidak valid berjumlah 30 butir soal (60%).

(b) Reliabilitas

Teknik analisis reliabilitas yang digunakan dalam penelitian ini adalah teknik formula *Kuder-Richardson-20* atau KR-20 diketahui bahwa reliabilitas soal sebesar 0,60 berarti tes yang digunakan sudah memiliki reliabilitas yang cukup baik.

(c) Tingkat Kesukaran

Untuk butir soal yang tergolong baik adalah pada taraf sedang. Hasil analisis taraf kesukaran instrumen tes dari hasil uji lapangan terbatas diperoleh 9 butir soal (18%) pada kategori mudah ($p > 0,7$), 15 butir soal (30%) pada kategori sedang ($0,3 \leq p \leq 0,7$), dan 26 butir soal (52%) pada kategori sulit ($p < 0,3$).

(d) Daya Pembeda

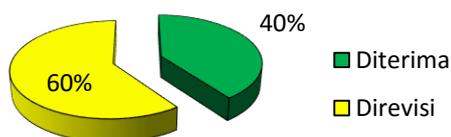
Hasil analisis daya beda instrumen tes pada lapangan terbatas didapatkan bahwa 7 butir soal (14%) memiliki daya pembeda baik sekali (0,50 – 1,00), 18 butir soal (36%) baik (0,30 – 0,49), dan 25 butir soal (50%) jelek (0,00 – 0,19).

(e) Efektivitas Pengecoh (Distraktor)

Pada uji lapangan terbatas jumlah peserta tes terdiri dari 12 siswa, sehingga pengecoh dikatakan efektif apabila dipilih minimal 1 siswa dan jawaban omit minimal 2 siswa. Dari analisis efektivitas pengecoh yang dilakukan, diketahui bahwa 18 butir soal (36%) dalam kategori baik sekali, 20 butir soal (40%) dalam kategori baik, 8 butir soal (16%) dalam kategori cukup, 2 butir soal (4%) dalam kategori kurang baik, dan 2 butir soal (4%) dalam kategori tidak baik.

(f) Interpretasi Uji Lapangan Terbatas

Setelah didapatkan hasil analisis validitas, reliabilitas, tingkat kesukaran, daya beda, dan efektivitas pengecoh, kemudian dari masing-masing soal diputuskanlah 20 butir soal sudah baik dan bisa diterima untuk kemudian disimpan di bank soal, sementara 30 butir soal harus direvisi sebelum dibawa ke uji lapangan lebih luas.



(5) Revisi Hasil Uji Lapangan Terbatas (*Main Products Revision*)

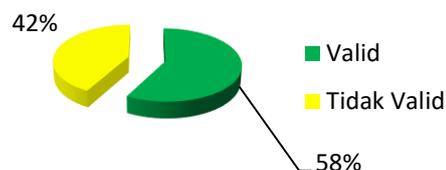
Dari hasil analisis yang didapat tampak bahwa belum semua instrumen memenuhi kriteria tes yang baik, sehingga setelah berkonsultasi peneliti memutuskan perlu dilakukan revisi untuk soal yang belum baik. Secara umum, ketiga puluh item soal hasil uji lapangan terbatas tersebut direvisi dikarenakan: (1) Tingkat kesukaran soal yang terlalu tinggi atau terlalu rendah, (2) Daya pembeda yang jelek, dan (3) Efektivitas pengecoh yang kurang/tidak baik. Dari ketiga puluh soal yang telah direvisi tersebut, 18 soal di antaranya memiliki tingkat kesukaran yang tinggi, 5 soal memiliki tingkat kesukaran yang rendah (mudah), dan 7 soal sisanya memiliki tingkat kesukaran sedang. Juga diketahui bahwa 20 soal memiliki daya pembeda yang jelek, serta pengecoh yang belum berfungsi maksimal. Oleh karenanya, dilakukan perbaikan/revisi terhadap soal-soal tersebut yakni dengan cara meng-hilangkan kalimat yang dirasa ambiguitas, meringkas kalimat yang bertele-tele, membenarkan indikator, menggunakan angka-angka yang lebih sederhana, dan memberikan informasi tambahan untuk mempermudah siswa dalam pengerjaan soal-soal tersebut.

(6) Uji Lapangan Lebih Luas (*Main Field Test*)

Pada uji coba ini menggunakan 50 butir soal yang telah direvisi berdasarkan hasil uji lapangan terbatas. Uji lapangan lebih luas dilaksanakan pada empat kelas XI IPA MAN 02 Model Medan dengan subyek sebanyak 100 siswa.

(a) Validitas

Berdasarkan hasil analisis ke-50 butir soal, dapat diketahui bahwa soal yang masuk kategori valid berjumlah 29 butir soal (58%). Sedangkan soal yang termasuk kategori tidak valid berjumlah 21 butir soal (42%). Adapun distribusi ke-50 butir soal tersebut berdasarkan validitas sebagai berikut:

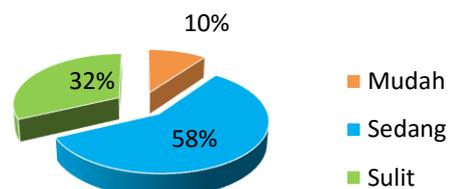


(b) Reliabilitas

Teknik analisis reliabilitas yang digunakan dalam penelitian ini adalah teknik formula *Kuder-Richardson-20* atau KR-20 diketahui bahwa reliabilitas soal sebesar 0,89 berarti tes yang digunakan sudah memiliki reliabilitas yang baik.

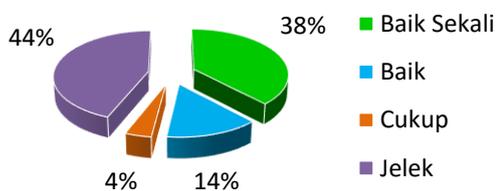
(c) Tingkat Kesukaran

Hasil analisis taraf kesukaran instrumen tes dari hasil uji lapangan lebih luas diperoleh 5 butir soal (10%) pada kategori mudah ($p > 0,7$), 29 butir soal (56%) pada kategori sedang ($0,3 \leq p \leq 0,7$), dan 16 butir soal (34%) pada kategori sulit ($p < 0,3$). Hasil analisis taraf kesukaran instrumen tes dari hasil uji lapangan lebih luas dapat dilihat pada tabel berikut:



(d) Daya Pembeda

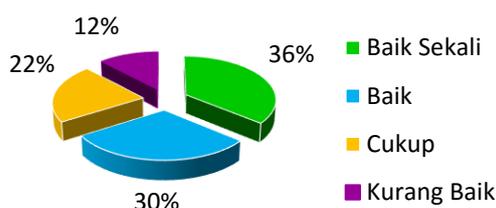
Hasil analisis daya beda instrumen tes pada lapangan luas didapatkan bahwa 19 soal (38%) memiliki daya pembeda baik sekali ($0,50 - 1,00$), 7 soal (14%) baik ($0,30 - 0,49$), 2 soal (4%) cukup ($0,20 - 0,29$), dan 22 soal (44%) jelek ($0,00 - 0,19$).



(e) Efektivitas Pengecoh (Distraktor)

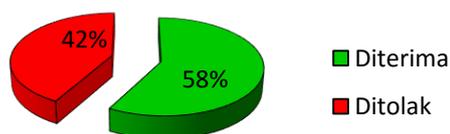
Peserta tes pada uji lapangan lebih luas ini sebanyak 100 siswa. Jadi jika dihitung 5% dari 100 siswa adalah 5 yang artinya dalam tes ini suatu pengecoh dikatakan efektif apabila dipilih minimal 5 orang siswa. Pengecoh yang baik adalah jika dipilih banyak oleh kelompok

bawah. Pengecoh dikatakan berfungsi apabila semua pengecoh pada tiap soal berfungsi baik, jika pengecoh ternyata belum berfungsi atau menyesatkan maka pengecoh tersebut perlu direvisi. Hasil analisis rekapitulasi efektivitas pengecoh memperlihatkan data jumlah pengecoh yang berfungsi maupun yang belum berfungsi. 18 butir soal (36%) termasuk dalam kategori sangat baik, 15 butir soal (30%) kategori baik, 11 butir soal (22%) kategori cukup, dan 6 butir soal (12%) kategori kurang baik.



(f) Interpretasi Hasil Uji Lapangan Lebih Luas

Setelah didapatkan hasil validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh, kemudian dari masing-masing soal diputuskanlah soal yang diterima dan soal yang ditolak. Hasil analisis empiris uji lapangan lebih luas dapat dilihat pada tabel berikut:



(7) Evaluasi Hasil Uji Lapangan Lebih Luas (*Operational Product Revision*)

Hasil analisis yang didapat belum semua instrumen tes yang valid memenuhi kriteria tes yang baik sehingga ada soal yang diterima dan ada soal yang ditolak. Selanjutnya soal yang telah dievaluasi disimpan pada bank soal Tes Objektif *Higher Order Thinking Skills* (HOTS) pada Materi Fluida di SMA/MA dengan jumlah 29 butir soal.

Pembahasan

Hasil validasi ahli terhadap butir soal HOTS pada Materi Fluida di SMA/MA memiliki keidealannya rata-rata sebesar 98,7% yang artinya hampir semua item memenuhi kriteria yang sesuai. Ditinjau dari aspek materi,

konstruksi, dan bahasa diketahui ada 16 butir soal yang tidak memenuhi aspek materi dan konstruksi. Butir soal yang tidak memenuhi kriteria aspek materi dan konstruksi berjumlah 6 butir soal. Butir soal tersebut tidak memenuhi aspek soal atau isi materi yang sesuai dengan indikator dan tujuan pengukuran, juga belum memenuhi pokok soal yang dirumuskan dengan singkat, jelas, dan tegas. Sedangkan butir soal yang hanya tidak memenuhi aspek materi berjumlah 10 butir soal. Setelah dilakukan perbaikan diperoleh sebanyak 50 butir soal sudah memenuhi ketiga aspek yaitu aspek materi, konstruksi, dan bahasa.

Hasil penelitian menyatakan bahwa tes objektif *HOTS* pada Materi Fluida di SMA/MA sudah baik. Reliabilitas tes dihitung menggunakan KR-20, pada uji coba lapangan terbatas reliabilitas yang diperoleh sebesar 0,60, sementara pada uji lapangan lebih luas reliabilitas mencapai 0,89, berarti tes yang digunakan sudah memiliki reliabilitas yang baik.

Untuk butir soal yang taraf kesukarannya tergolong baik adalah pada taraf sedang. Dari analisis data, pada uji lapangan terbatas diperoleh 9 butir soal (18%) pada kategori mudah ($p > 0,7$), 15 butir soal (30%) pada kategori sedang ($0,3 \leq p \leq 0,7$) dan 26 butir soal (52%) pada kategori sulit ($p < 0,3$). Sementara pada uji lapangan lebih luas diperoleh 5 butir soal (10%) pada kategori mudah, 28 butir soal (56%) pada kategori sedang, dan 17 butir soal (34%) pada kategori sulit. Dengan demikian, taraf kesukaran tes objektif *HOTS* pada materi Fluida mendeskripsikan taraf kesukaran dari rentang yang tinggi, sedang, dan rendah.

Dari uji lapangan terbatas diperoleh 7 butir soal (14%) memiliki daya pembeda sangat baik, 18 butir soal (36%) baik, dan 25 butir soal (50%) jelek. Sedangkan pada uji lapangan lebih luas diperoleh 19 butir soal (38%) memiliki daya pembeda baik sekali, 7 butir soal (14%) baik, 2 butir soal (4%) cukup dan 22 butir soal (44%) jelek.

Peserta tes pada uji lapangan terbatas sebanyak 12 siswa, sehingga batas minimal pengecoh dikatakan efektif apabila dipilih 1 siswa dengan omit maksimal 2 siswa. Hasil rekapitulasi efektivitas pengecoh menunjukkan bahwa 18 butir soal (36%) termasuk dalam kategori sangat baik, 20 butir soal (40%) kategori baik, 8 butir soal (16%) kategori cukup, 2 butir soal (4%) kategori kurang baik, dan 2 butir soal (4%) kategori tidak baik. Sedangkan peserta tes pada uji lapangan lebih luas adalah sebanyak 100 siswa, sehingga batas minimal pengecoh dikatakan efektif apabila dipilih minimal 5 siswa dengan omit maksimal 10

siswa. Hasil rekapitulasi efektivitas pengecoh menunjukkan bahwa 18 butir soal (36%) dalam kategori sangat baik, 15 butir soal (30%) baik, 11 butir soal (22%) cukup, dan 6 butir soal (12%) kurang baik.

Berdasarkan data di atas diketahui bahwa 29 butir soal dapat diterima dan disimpan dalam bank soal tes objektif *HOTS* pada materi Fluida di SMA/MA karena telah memenuhi validitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh yang baik. Terdapat 21 butir soal yang ditolak dan tidak dapat digunakan karena sama sekali tidak memenuhi satu pun kriteria.

Lewy, Zukardi, dan Nyimas Aisyah (2009) melakukan penelitian guna mengembangkan soal untuk mengukur kemampuan berpikir tingkat tinggi pada pokok bahasan barisan dan deret bilangan tingkat SMP di Palembang, kemudian memperoleh hasil bahwa *prototype* perangkat soal yang dikembangkan telah memiliki *potensial efek* dengan nilai 35,59 yang termasuk kategori baik, dimana penelitian tersebut dilakukan dengan tiga kali uji coba, yaitu uji coba pertama terdiri dari seorang siswa (*one-to-one*), uji coba kedua terdiri dari lima orang siswa, dan uji coba ketiga terdiri dari satu kelas, sedangkan penelitian ini melakukan uji coba dua kali sesuai langkah penelitian Borg & Gall, yaitu uji lapangan terbatas dengan jumlah maksimal 12 sampel dan uji lapangan lebih luas dengan jumlah maksimal 100 sampel. Selanjutnya, penelitian yang dilakukan oleh Lindawaty, Antomi Siregar, dan Yuberti (2016) bertujuan untuk mengembangkan instrumen *authentic assesment* untuk mengukur *HOTS* peserta didik memperoleh hasil validasi oleh ahli asesmen dengan nilai “layak”, oleh ahli materi dengan nilai “sangat layak”, oleh ahli media dengan nilai “sangat layak”, dan oleh ahli bahasa dengan nilai “sangat layak”, dimana penilaian kualitatif pada penelitian tersebut menggunakan skala Likert, sedangkan dalam penelitian ini item soal dianalisis dengan menggunakan CVR oleh ahli dan memperoleh hasil sangat baik dengan persentase keidealan 98,7%.

Penelitian ini menggunakan teori tes klasik dimana kondisi dan karakteristik peserta tes akan mempengaruhi hasil penelitian. Kondisi yang dimaksud yaitu faktor internal dan eksternal peserta tes. Faktor internal meliputi tingkat kecerdasan, motivasi, kesehatan dan sebagainya. Sedangkan konduktivitas ruang ujian merupakan salah satu faktor eksternal yang juga mempengaruhi kondisi peserta tes. Semakin konduktivitas ruang ujian maka hasilnya akan semakin baik pula. Semakin rendah kemampuan kelompok peserta tes maka semakin banyak juga golongan item tes yang sukar dan begitu

pula sebaliknya. Adapun daya pembeda tergantung pada homogenitas peserta tes. Keterbatasan inilah yang menyebabkan adanya perbedaan hasil jika penelitian ini digunakan pada kelompok peserta tes dengan karakteristik yang berbeda subjek coba.

KESIMPULAN

Kesimpulan yang dapat diambil dari penelitian pengembangan ini adalah sebagai berikut: (1) Telah dikembangkan instrumen tes *HOTS* pada Materi Fluida untuk SMA/MA. Pengembangan ini menggunakan model penelitian dan pengembangan Borg & Gall, yang direvisi berdasarkan masukan para ahli yaitu 2 orang dosen serta dilakukan uji lapangan terbatas untuk 12 orang siswa dan uji lapangan lebih luas untuk 100 orang siswa. Ditinjau dari analisis kualitatif, kualitas tes objektif adalah sangat baik (SB) dengan persentase keidealan 98,7%. Ditinjau dari analisis kuantitatif diperoleh 29 butir soal (58%) valid, sedangkan 21 butir soal (42%) ditolak karena tidak memenuhi kriteria yang ada, yakni validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh, (2) Diperoleh instrumen tes *HOTS* Fisika pada materi Fluida sebanyak 29 soal, meliputi 10 soal (34,6%) mengenai prinsip Archimedes, 6 soal (20,7%) mengenai Hidrostatika, 6 soal (20,7%) mengenai persamaan Bernoulli, 5 soal (17,2%) mengenai persamaan Kontinuitas, 1 soal (3,4%) mengenai Hukum Pascal, dan 1 soal (3,4%) mengenai prinsip Tegangan Permukaan, (3) Instrumen tes yang telah dikembangkan memiliki reliabilitas baik yaitu 0,89, (4) Berdasarkan daya pembeda, diperoleh 19 soal dengan kategori baik sekali (38%), 7 soal dengan kategori baik (14%), 2 soal dengan kategori cukup (4%), dan 22 soal dengan kategori jelek (44%), (5) Berdasarkan taraf kesukaran, sebanyak soal mudah 5 (10%), 28 soal sedang (56%), dan 17 soal sulit (34%), dan (6) Berdasarkan keefektifan pengecoh, terdapat 18 soal (36%) dengan kategori sangat baik, 15 soal (30%) baik, 11 soal (22%) cukup, dan 6 soal (12%) kurang baik.

SARAN

Berdasarkan hasil penelitian pengembangan instrumen tes *HOTS* Fisika pada materi Fluida untuk SMA/MA maka saran yang dapat diajukan sebagai berikut: (1) Instrumen tes yang sudah diujicobakan dan dianalisis ini dapat dimanfaatkan sebagai bank soal, (2) Penelitian ini, menunjukkan bahwa untuk membuat suatu instrumen tes yang baik perlu dilakukan analisis butir soal baik secara kualitatif maupun kuantitatif sehingga memenuhi kriteria instrumen tes yang baik, agar menumbuhkan

motivasi para pelaku dalam bidang pendidikan untuk membuat instrumen tes Fisika yang baik pada materi lainnya, (3) Hendaknya peneliti lain yang ingin melakukan penelitian sejenis sedapat mungkin memilih materi fisika yang berbeda, sehingga soal-soal HOTS untuk semua materi fisika tersedia di lapangan, (4) Pengolahan data menggunakan Ms. Excel dalam menganalisis butir soal secara kuantitatif memerlukan waktu yang cukup lama sehingga disarankan menggunakan aplikasi lain agar waktu yang digunakan lebih efektif.

DAFTAR PUSTAKA

- Anderson, Lorin W. & David R. Krathwohl, (2010), *Kerangka Landasan Untuk Pembelajaran, Pengajaran, dan Asesmen: Revisi Taksonomi Pendidikan Bloom*, Pustaka Pelajar, Yogyakarta.
- Brookhart, (2010), *How to Assess Higher-Order Thinking Skills*, ASCD, Alexandria.
- Direktorat Pembinaan SMA, (2010), *Juknis Analisis Butir Soal di SMA*, Kemendikbud, Jakarta.
- Kusuma Wardany, Sajidan, Murni Ramli, (2017), *Pengembangan Penilaian Untuk Mengukur High Order Thinking Skills Siswa*, UNS Surakarta, **Vol. 6, No.2**.
- Lewy, Zulkardi, Nyimas Aisyah, (2009), *Pengembangan Soal Untuk Mengukur Kemampuan Berpikir Tingkat Tinggi Pokok Bahasan Barisan dan Deret Bilangan di Kelas IX Akselerasi SMP Xaverius Maria Palembang*, UNSRI Palembang, **Vol. 3, No. 2**.
- Lindawati, Antomi Siregar, Yuberti, (2016), *Pengembangan Instrumen Authentic Assesment Untuk Mengukur Higher Order Thinking Skills Peserta Didik*, Seminar Nasional Pendidikan, **140-149**, IAIN Raden Intan Bandar Lampung.
- Sugiyono, (2018), *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*, Alfabeta, Bandung.
- Wulaningtyas, Ratna Candra & Sukanti, (2016), *Analisis Butir Soal Ujian Akhir Semester Gasal Mata Pelajaran Pengantar Akuntansi dan Keuangan*, UNY Yogyakarta, **Edisi 7**.