

PENGEMBANGAN INSTRUMEN TES FISIKA SEMESTER GANJIL KELAS X SMA RSBI

Latifah N. Qomariyatuzzamzami¹⁾, Sarwanto²⁾, Sri Budiawanti²⁾
Program Studi Pendidikan Fisika UNS Surakarta
latifah_zamzami@yahoo.co.id

Abstrak

Telah dilakukan penelitian untuk mengembangkan kisi-kisi dan instrumen tes Fisika Semester Ganjil Kelas X SMA RSBI berdasarkan KTSP dan *A Level* yang memiliki karakteristik internal serta tingkat validitas dan reliabilitas yang baik.

Penelitian *Research and Development (R&D)* ini dilakukan dengan melibatkan beberapa komponen meliputi pakar/ *expert* bidang studi Fisika, mahasiswa, guru Fisika RSBI dan siswa Kelas X di SMA Negeri 1 Surakarta menggunakan metode 4-D yang dikembangkan oleh S. Thiagarajan, Dorothy S. Semmel, dan Melvyn I. Semmel. Data dikumpulkan dengan teknik dokumentasi dan tes. Analisis data menggunakan analisis butir meliputi uji validitas, daya beda, tingkat kesukaran, efektivitas distraktor, dan reliabilitas.

Hasil penelitian berupa deskripsi prosedural proses pengembangan dan produk yang dikembangkan melalui tahap *Define* dengan analisis kurikulum dan analisis kebutuhan, *Design* meliputi penyusunan silabus dan kisi-kisi dengan metode adopsi berdasarkan kurikulum KTSP dan silabus *A Level*, serta penyusunan instrumen tes yang terdiri dari instrumen A dan instrumen B yang teruji validitasnya melalui *expert judgement*, *Develop* dengan uji coba, dan pelaksanaan *Disseminate*. Uji coba dilakukan melalui uji coba *peer* untuk memperoleh validasi *peer reviewer* dan uji coba terbatas dengan mengujikan masing-masing instrumen pada 15 siswa, sedangkan diseminasi dilakukan pada 61 siswa. Hasil diseminasi menunjukkan bahwa tes yang dikembangkan telah memenuhi karakteristik internal tes yang baik berdasarkan analisis butir soal berupa daya beda, tingkat kesukaran, dan efektivitas distraktor, tes yang dikembangkan memiliki validitas isi baik berdasarkan hasil telaah pakar dan reliabilitas yang tinggi ($r_{11}=0,781$), produk akhir tes hasil penelitian pengembangan berupa 26 soal tes yang dapat dijadikan bank soal yang siap diimplementasikan, dan meskipun tes telah dikembangkan sesuai dengan standar, tes masih memiliki kelemahan, yaitu Bahasa Inggris yang dianggap sulit oleh siswa dan beberapa siswa menjawab soal dengan menebak dan mengandalkan keberuntungan.

Kata Kunci: RSBI, pengembangan tes, 4-D, valid dan reliabel

- 1) Mahasiswa Peneliti
- 2) Dosen Pembimbing

PENDAHULUAN

Rintisan Sekolah Bertaraf Internasional atau RSBI merupakan realisasi Undang-Undang Nomor 20 tahun 2003 tentang Sistem Pendidikan Nasional Pasal 50 ayat 3 dalam rangka memenuhi kebutuhan sumber daya manusia yang unggul. Pembelajaran di sekolah RSBI juga tidak lepas dari komponen pembelajaran yang meliputi tujuan, proses, dan penilaian pembelajaran. Tercapai tidaknya tujuan pembelajaran dapat diketahui melalui penilaian. Salah satu cara untuk melakukan penilaian adalah melalui tes. Tes digunakan untuk menilai dan mengukur beberapa aspek baik kognitif, afektif, dan psikomotor, di mana aspek kognitif lebih dominan. Meskipun demikian, hasil belajar psikomotor dan afektif juga harus menjadi bagian dari hasil penilaian dalam proses pembelajaran. Pencapaian tingkatan pada aspek kognitif harus disesuaikan dengan jenjang pendidikan serta distribusi butir soal dalam tes untuk penilaian aspek tersebut harus merata dan berimbang. Akan tetapi, menurut Widiyanti (2008), soal-soal tes formatif dan sumatif yang dilakukan untuk melakukan penilaian di SMA berwawasan internasional (RSBI) mengandung jenjang kognitif yang tidak merata dan belum berimbang. Sehingga hasil penilaian belum dapat mengungkapkan ketercapaian kemampuan kognitif yang seharusnya dikuasai oleh siswa.

Di samping itu, sesuai dengan perkembangan pendidikan, instrumen penilaian yang digunakan harus mencerminkan kurikulum yang berlaku saat ini. Untuk RSBI, tes juga harus diperkaya dengan sistem ujian internasional yang mengacu pada negara maju yang memiliki keunggulan tertentu atau sering disebut dengan rumusan $SNP + X$ (OECD). Dengan adanya sistem ini, guru dituntut membuat instrumen tes berstandar internasional. Selain itu, harapan untuk dapat bersaing dengan komunitas internasional juga ditunjukkan dengan penggunaan bahasa internasional dalam instrumen tes untuk RSBI. Hasil *Test of English for International Communication* (TOEIC) yang diselenggarakan oleh Departemen Pendidikan Nasional tahun 2010 untuk mengetahui kemampuan Bahasa Inggris guru-guru RSBI menunjukkan hasil yang kurang memuaskan karena masih banyak guru RSBI yang mendapatkan nilai kurang dari nilai minimum TOEIC yaitu 245 (Anonim, 2010). Selain itu, menurut Danny Whitehead, *Head of English Development British Council* berdasarkan hasil penelitian Stephen Bax University of Bedfordshire Inggris menyatakan bahwa guru sekolah RSBI yang menguasai bahasa Inggris dengan baik, dalam arti mampu berbahasa Inggris

dengan baik dan menyampaikan materi pelajaran dalam bahasa Inggris tidak mencapai 25% (Ardiansyah: 2011). Dengan demikian, instrumen tes untuk sekolah RSBI penting dikembangkan untuk memperkaya perbendaharaan soal guru-guru RSBI dan dapat membantu sekolah RSBI menuju sekolah SBI.

Terkait permasalahan di atas, penelitian ini bertujuan untuk mengembangkan kisi-kisi dan instrumen tes Fisika Semester Ganjil Kelas X SMA RSBI berdasarkan KTSP dan *A Level* yang memiliki karakteristik internal (daya beda, tingkat kesulitan, efektivitas distraktor) serta tingkat validitas dan reliabilitas yang baik.

TINJAUAN PUSTAKA

Penelitian pengembangan merupakan penelitian yang bertujuan untuk mengembangkan dan memvalidasi produk pendidikan dan pembelajaran untuk meningkatkan dan mengembangkan mutu pendidikan dan pembelajaran secara efektif dan adaptable, hal ini dikemukakan oleh Anik Ghufro, dkk (2007). Penelitian pengembangan ini merupakan penelitian pengembangan tes yang bertujuan untuk mengembangkan dan memvalidasi produk tes untuk Kelas X SMA RSBI.

Penelitian sebelumnya dilakukan oleh Deni Widyanoro dkk (2009) mengembangkan 40 soal pilihan ganda, soal yang dapat dikategorikan baik hanya 45%, sedangkan 55 % soal dikategorikan jelek. Namun, pengembangan instrumen tes ini belum memenuhi prosedur pengembangan tes yang baik karena setelah dilakukan *expert judgement*, instrumen tersebut langsung didiseminasikan tanpa melalui uji coba terbatas. Sehingga hasil pengembangan tes kurang sesuai dengan yang diharapkan, karena lebih dari 50 % butir tes masih dikategorikan jelek. Penelitian yang dilakukan oleh Widowati Pusporini (2009), pengembangan instrumen tes dilakukan dengan prosedur yang hanya cukup sampai merakit tes dan tidak melakukan diseminasi atau uji coba tes dalam lingkup yang lebih luas.

Berdasarkan hal tersebut, maka pengembangan instrumen tes Ulangan Akhir Semester Ganjil Kelas X SMA RSBI dilaksanakan melalui tahap 4-D yang dikembangkan oleh S. Thagarajan, Dorothy S. Semmel, dan Melvyn I. Semmel (Thiagarajan, 1974) sebagai berikut: (1) Pendefinisian (*Define*), meliputi studi literatur untuk menganalisis kesesuaian indikator dengan tujuan dan proses pembelajaran serta kurikulum yang digunakan, sehingga dapat tersusun silabus yang baru sebagai pemetaan indikator-indikator yang dapat diukur dengan tes maupun non-tes, (2) Perencanaan (*Design*) meliputi penyusunan spesifikasi tes berupa pembuatan kisi-kisi dan penyusunan instrumen tes, (3) Pengembangan (*Develop*) meliputi penelaahan butir soal oleh pakar, analisis validitas, uji coba terbatas, analisis butir soal uji

coba, dan revisi butir soal (4) Penyebaran (*Disseminate*) yang meliputi implementasi instrumen tes dalam lingkup yang lebih luas.

METODE PENELITIAN

Penelitian ini menggunakan model pengembangan prosedural yaitu bersifat deskriptif dan menggariskan langkah-langkah sistematis yang harus diikuti untuk menghasilkan produk. Instrumen ini disusun dalam bentuk tes pilihan ganda (*multiple choice*) dengan melibatkan beberapa komponen meliputi *expert*, mahasiswa, guru Fisika RSBI dan siswa Kelas X SMA Negeri 1 Surakarta karena merupakan salah satu SMA RSBI yang memenuhi standar mutu sebagai penyelenggara RSBI (milist RSBI, 2011). Pengembangannya dilakukan melalui R & D (*Research and Development*) yang dikembangkan oleh S. Thiagarajan, Dorothy S. Semmel, dan Melvyn I. Semmel. Model pengembangan 4-D terdiri atas 4 tahap utama yaitu: (1) *Define* (Pendefinisian), (2) *Design* (Perancangan), (3) *Develop* (Pengembangan) dan (4) *Disseminate* (Penyebaran) (Thiagarajan, 1974).

Analisis soal dilakukan melalui dua cara, yaitu analisis kualitatif dan analisis kuantitatif. Analisis kualitatif dengan validitas isi/ validitas logis (*logical validity*), yaitu menganalisis soal ditinjau dari segi materi, konstruk, dan bahasa. Analisis kuantitatif dilakukan dengan teori Tes Klasik. Adapun uji coba yang dilakukan terhadap instrumen meliputi daya pembeda, tingkat kesukaran, efektivitas distraktor, dan reliabilitas item tes.

HASIL PENELITIAN DAN PEMBAHASAN

Penyiapan Data

Hasil penelitian berupa deskripsi prosedural proses pengembangan dan produk yang dikembangkan melalui tahap-tahap sesuai dengan metode 4-D (Thiagarajan, 1974). Produk dari tahap pengembangan meliputi analisis kesesuaian kurikulum, indikator hasil belajar, silabus pengembangan, kisi-kisi instrumen tes. Produk akhir berupa instrumen tes Fisika untuk Semester Ganjil Kelas X SMA RSBI.

Tahap *Define* meliputi analisis kurikulum dan analisis kebutuhan. Analisis kurikulum dilakukan dengan mempelajari kurikulum RSBI yaitu kurikulum SNP+X. Sistem atau model penilaian yang digunakan untuk RSBI seharusnya telah memenuhi Standar Penilaian sesuai dengan SNP dan telah diperkaya dengan sistem penilaian dari sekolah unggul dari salah satu negara OECD (Haryana, 2007). Hasil analisis menunjukkan bahwa penilaian yang dilakukan di SMA terkait, ditinjau dari segi penilaian berdasarkan SNP+X, sudah sesuai dengan Kurikulum KTSP yaitu memenuhi Standar Isi dan Standar Kompetensi Lulusan, tetapi belum memenuhi standar plus X. SMA ini menggunakan sistem

adaptasi sistem adopsi. Sistem adopsi yang dilakukan dengan penambahan jam tambahan untuk pengayaan *A Level* dan tidak diwajibkan kepada siswa tetapi pada kenyataannya siswa banyak yang tidak berminat. Sedangkan sistem adaptasi yang dilakukan belum jelas mengacu pada standar salah satu negara OECD, karena tingkat kemampuan yang diukur masih setara dengan SMA non-RSBI. Hal yang membedakan hanyalah pada penggunaan dwi bahasa (*bilingual*) pada instrumen penilaian. Jika dilakukan penyetaraan tingkat kemampuan yang sesuai dengan *A Level*, terkendala pada kemampuan siswa yang rata-rata berada pada tingkat mengingat (*remember/ C1*) sampai (*analyze/ C4*). Sehingga, cara yang efektif untuk diimplementasikan adalah cara adopsi yaitu penambahan atau pengayaan. Cara ini dilakukan dengan penambahan indikator dari *A Level* yang belum ada pada indikator yang dikembangkan pada silabus, tetapi dengan tidak meninggalkan komponen belajar lain sebagai acuan. Adopsi indikator hasil belajar dalam Silabus *A Level Physics* dapat dilakukan dengan mempertimbangkan komponen pembelajaran lain yaitu tujuan dan proses pembelajaran. Hasil analisis ini digunakan sebagai dasar untuk merumuskan silabus yang memuat indikator-indikator ketercapaian pembelajaran. Analisis kebutuhan dilakukan berdasarkan hasil Ulangan Harian dan Ulangan Tengah Semester dapat diketahui bahwa kemampuan kognitif siswa rata-rata mencakup kemampuan mengingat (*remember/ C1*) hingga menganalisis (*analyze/ C4*). Hasil ini menjadi dasar perumusan kisi-kisi instrumen tes yang disusun.

Berdasarkan hasil analisis di atas, dilakukan penyusunan indikator-indikator pembelajaran dan *expert judgement* terhadap indikator tersusun dengan *interview* satu per satu oleh dua pakar yaitu dosen Pendidikan Fisika dan seorang praktisi/ guru Fisika. Indikator tersusun menjadi dasar terbentuknya silabus penelitian pengembangan sebagai acuan penyusunan kisi-kisi instrumen tes yang dikembangkan. Penyusunan kisi-kisi mempertimbangkan kemampuan siswa yang akan diukur dan proporsi tiap kemampuan yaitu kemampuan mengingat (*remember/ C1*) 10%, memahami (*understand/ C2*) 16,67%, menerapkan (*apply/ C3*) 43,33%, dan analisis (*analyze/ C4*) 30%.

Penyusunan Instrumen Tes

Setelah kisi-kisi instrumen dirumuskan, mulai dilakukan tahap perancangan (*Design*) instrumen tes yang terdiri dari dua paket instrumen masing-masing berjumlah 30 soal dengan alternatif lima jawaban, yaitu paket A dan paket B. Paket B disusun berdasarkan paket A, artinya paket B identik dengan paket A yang telah diuji validitasnya oleh pakar dan praktisi.

Expert judgement dilakukan untuk memperoleh masukan untuk improvisasi instrumen dengan melakukan telaah butir soal oleh pakar dan praktisi baik dari segi materi/ isi, konstruksi, dan bahasa. Berdasarkan masukan dari para ahli, instrumen direvisi agar lebih tepat, efektif, dan memiliki kualitas yang baik (Thiagarajan, 1974). *Expert judgement* dilakukan oleh tiga pakar dan satu praktisi. Salah satu dari ketiga pakar Fisika merupakan pakar Bahasa Inggris Fisika. Hasil *expert judgement* menunjukkan bahwa terdapat 8 soal pada instrumen A yang perlu dilakukan revisi. Revisi meliputi konstruk soal yang kurang sesuai dengan indikator, kesalahan penulisan kunci jawaban dan penggantian alternatif jawaban, kekurangsesuaian ilustrasi, serta kesalahan dalam pengetikan.

Beberapa tes yang salah kunci dan penulisan dibetulkan, sedangkan tes yang perlu diganti, disusun ulang. Instrumen yang baru tersusun dan hasil revisi dilakukan *expert judgement* ulang sehingga dapat dijadikan tes dengan validitas yang baik. Hasil *expert judgement* pada revisi ini berupa teknis penulisan di mana sebaiknya kata negatif diperjelas. Selanjutnya dari segi Bahasa Inggris, terdapat beberapa koreksi yaitu penggunaan Bahasa Inggris yang kurang komunikatif sehingga dimungkinkan siswa tidak mudah mengerti.

Instrumen A direvisi dijadikan acuan untuk penyusunan instrumen B. Hasil *expert judgement* menyatakan bahwa instrumen B tidak perlu dilakukan revisi sehingga kedua instrumen A dan B siap diujicobakan.

Pelaksanaan Uji Coba

Sebelum diujicobakan ke siswa, kedua paket diujicobakan secara *peer* untuk memperoleh validasi *peer reviewer*. Langkah ini merupakan tahap pengembangan (*Develop*). Uji coba lapangan dilakukan untuk memperoleh masukan langsung berupa respon, reaksi, komentar para pengamat terhadap perangkat pembelajaran yang telah disusun (Thiagarajan, 1974). Masing-masing instrumen tes diujicobakan secara *peer* dengan cara mengerjakan instrumen secara essay sesuai dengan alokasi waktu yang ditetapkan, yaitu 60 menit.

Berdasarkan hasil uji coba, diketahui bahwa dengan alokasi waktu 60 menit, rata-rata responden menjawab 24 sampai 28 butir dari 30 butir soal yang diteskan. Hasil ini mengindikasikan bahwa apabila instrumen tersebut diaplikasikan ke siswa SMA, alokasi waktu 60 menit belum cukup untuk mengerjakan semua soal yang diujikan. Dengan hasil ini, alokasi waktu ditambah menjadi 90 menit.

Sedangkan ditinjau dari kesesuaian dengan kunci jawaban, diketahui bahwa untuk masing-masing instrumen A dan B terdapat 11 dan 8 butir soal dari 30 soal yang diujikan berhasil dijawab dengan benar sesuai dengan kunci jawaban oleh responden. *Interview*

dilakukan dengan responden yang hasilnya menunjukkan bahwa hal tersebut terjadi karena kekurangtelitian responden, lupa dan salah mengenali konsep materi yang diteskan. Berdasarkan hasil tersebut, dapat diketahui bahwa terdapat beberapa kesalahan penulisan pada instrumen sehingga dilakukan pengecekan ulang tes sebelum diujicobakan. Hasil ini menunjukkan bahwa semua instrumen A dan B telah valid dan dapat digunakan karena jawaban peserta tes sesuai dengan yang diharapkan pada penelitian pengembangan ini, artinya instrumen tersusun sudah sesuai untuk mengukur kemampuan siswa berdasarkan indikator-indikator pembelajaran yang akan dicapai dan faktor penyebab terjadinya kesalahan bukan pada instrumen, tetapi pada individu responden.

Uji coba terbatas dilakukan untuk mengetahui karakteristik internal dan reliabilitas soal. Uji coba terbatas dilakukan dengan mengujikan instrumen A dan B pada masing-masing 15 siswa RSBI. Hasil tes dilakukan analisis daya beda, tingkat kesukaran, dan efektivitas distraktor instrumen A dan B yang menjadi karakteristik internal instrumen yang dikembangkan

Hasil analisis daya beda menunjukkan bahwa untuk daya beda instrumen A dan B, masing-masing terdapat 20% dan 23,33% soal memiliki daya beda lebih dari 0,30, sehingga dapat dikategorikan baik.. Selanjutnya, terdapat 23,33% pada instrumen A dan 26,67% pada instrumen B yang mempunyai daya beda $0,2 \leq D < 0,3$ artinya cukup. Sedangkan 33,33% dari instrumen A dan 30% dari instrumen B yang memiliki daya beda yang jelek yaitu $< 0,2$. Soal yang memiliki daya beda jelek tidak langsung ditolak tetapi masih dipertimbangkan dengan melihat karakteristik internal lainnya. Untuk soal daya beda negatif, masing-masing instrumen terdapat 20% soal. Menurut Arikunto (2005), soal yang memiliki daya beda negatif sebaiknya dibuang, sehingga soal-soal tersebut tidak digunakan.

Hasil analisis kuantitatif untuk tingkat kesukaran menunjukkan bahwa 13,33% instrumen A dan 23,33% instrumen B soal tergolong mudah. Soal yang tergolong sedang sebanyak 66,67% dan 56,67% untuk masing-masing instrumen A dan B. Sedangkan soal yang dapat dikategorikan sulit pada instrumen A dan B adalah 20,00%. Soal-soal yang memiliki tingkat kesukaran tinggi menunjukkan bahwa untuk menjawab soal tersebut diperlukan tingkat kemampuan yang tinggi. Soal memiliki variasi tingkat kesukaran karena kemampuan yang diukur juga berbeda.

Efektivitas distraktor dianalisis untuk mengetahui berfungsi tidaknya distraktor/ pengecoh pada tiap butir item. Sebanyak 70% dari instrumen A dan 56,67% dari instrumen B soal memiliki distraktor yang berfungsi baik. Jika di antara soal-soal tersebut terpilih dalam

pemilihan soal dari instrumen A dan B, maka perlu dilakukan revisi. Namun, apabila soal tersebut tidak terpilih atau dibuang, maka tidak dilakukan revisi.

Berdasarkan hasil ini, reliabilitas soal dicari dan hasilnya menunjukkan tingkat reliabilitas yang rendah karena $0,20 \leq r_{11} < 0,40$ baik untuk instrumen A maupun B yaitu masing-masing $r_{11} = 0,294$ dan $r_{11} = 0,102$ (Arikunto, 2005). Rendahnya reliabilitas instrumen mengakibatkan perlunya dilakukan pembuangan soal yang menyebabkan reliabilitas instrumen menjadi rendah. Instrumen yang memiliki daya beda negatif (Arikunto, 2005) dan tidak banyak menyumbangkan skor terhadap skor keseluruhan (Sudijono, 2005) tidak dipakai. Setelah dilakukan eliminasi terhadap butir-butir soal tersebut, reliabilitas soal menjadi $r_{11} = 0,660$ dan $r_{11} = 0,672$ untuk masing-masing instrumen A dan B sehingga dapat dikategorikan instrumen tersebut telah memiliki reliabilitas yang tinggi.

Pemilihan soal dilakukan berdasarkan analisis di atas. Berdasarkan analisis, terdapat beberapa soal yang tidak tersedia baik dari instrumen A maupun instrumen B, yaitu sejumlah 4 soal. Dengan demikian perlu dilakukan penyusunan instrumen baru untuk nomor-nomor soal tersebut. Sedangkan butir soal yang digunakan adalah masing-masing 13 soal dari paket A dan B.

Soal dengan indikator merumuskan suatu besaran dengan analisis dimensi dari instrumen A memiliki daya beda negatif dan tingkat kesukaran sedang sedangkan pada instrumen B daya bedanya jelek dan sukar. Sehingga instrumen baru disusun berdasarkan instrumen B tetapi dengan sedikit menurunkan tingkat kesukaran soal. Hal ini dilakukan dengan menggunakan besaran-besaran yang sering dijumpai oleh siswa. Soal dengan indikator menerapkan penulisan notasi ilmiah dalam operasi menggunakan aturan angka penting dari instrumen A dan B memiliki daya beda yang jelek dan sukar sehingga instrumen baru dibuat lebih mudah, yaitu dengan menyajikan soal dalam bentuk π supaya memudahkan siswa dalam menghitung. Soal berindikator menyelesaikan permasalahan menggunakan persamaan gerak lurus berubah beraturan termasuk gerak benda jatuh bebas tanpa hambatan udara dari kedua instrumen A dan B menghasilkan daya beda yang negatif tetapi soal dapat dikategorikan sedang, sehingga instrumen baru yang disusun identik dengan instrumen sebelumnya dan lebih memperhatikan kategori kemampuan kognitif yang akan diukur. Sedangkan soal dengan indikator menjelaskan arti fisis rumus hukum 2 Newton tentang gerak menghasilkan daya beda yang negatif dan dikategorikan mudah, sehingga instrumen baru yang disusun sedikit ditingkatkan kesukarannya.

Instrumen baru juga diujicobakan melalui *peer* yaitu enam mahasiswa dengan alokasi waktu yang ditetapkan, yaitu 12 menit. Hasil uji coba menunjukkan bahwa ditinjau dari kesesuaian dengan kunci jawaban, diketahui soal berindikator menerapkan penulisan notasi ilmiah dalam operasi menggunakan aturan angka penting dan menyelesaikan permasalahan menggunakan persamaan gerak lurus berubah beraturan termasuk gerak benda jatuh bebas tanpa hambatan udara berhasil dijawab dengan benar oleh sebagian responden dan dijawab salah oleh responden yang lain. Dengan hasil ini, dilakukan *interview* dengan responden yang hasilnya menunjukkan bahwa responden kurang teliti dan lupa tentang konsep yang ditekankan.

Pelaksanaan Diseminasi

Kegiatan diseminasi dilaksanakan dengan mengimplementasikan hasil uji coba dengan peserta berskala lebih besar yaitu 61 siswa. Instrumen yang digunakan dalam tahap diseminasi ini merupakan instrumen hasil pemilihan soal dari uji coba terbatas. Hasil analisis pelaksanaan diseminasi menunjukkan karakteristik internal berupa daya beda, tingkat kesukaran, dan efektivitas distraktor.

Analisis daya beda menunjukkan bahwa 23,33% memiliki daya beda lebih dari 0,30, sehingga dapat dikategorikan baik, artinya soal tersebut dapat dengan baik membedakan antara kelompok atas dan kelompok bawah. Selanjutnya, sebanyak 63,33% soal mempunyai daya beda $0,2 \leq D < 0,3$ yang artinya cukup. Hal ini menunjukkan bahwa soal tersebut sudah cukup membedakan antara kelompok atas dan kelompok bawah. Persentase soal yang memiliki daya beda negatif adalah 13,33%. Soal yang memiliki daya beda negatif perlu dilakukan kajian ulang untuk mengetahui penyebab daya beda soal menjadi negatif. Hal ini dilakukan dengan melakukan *interview* terhadap peserta tes yang tergolong dalam kelompok atas dan kelompok bawah.

Hasil *interview* disimpulkan bahwa soal dengan indikator merumuskan suatu besaran dengan analisis dimensi dirasa cukup sulit karena siswa berpikir bahwa kecepatan gelombang air laut belum diajarkan, padahal soal tersebut dapat dikerjakan hanya mengandalkan analisis dimensi. Hal ini menyebabkan siswa menjawab secara acak dan disimpulkan bahwa inilah yang menjadi penyebab kenegatifan daya beda. Sedangkan soal dengan indikator menyelesaikan permasalahan menggunakan persamaan gerak lurus berubah beraturan termasuk gerak benda jatuh bebas tanpa hambatan udara, sebagian besar siswa sudah lupa dengan materinya sehingga jawaban siswa menjadi kurang objektif. Hal serupa terjadi pada indikator menjelaskan karakteristik Gerak Melingkar Beraturan dan

mengidentifikasi penerapan hukum 1 Newton dalam kehidupan sehari-hari. Kemungkinan inilah yang menyebabkan daya beda dari masing-masing soal menjadi negatif. Selain itu, secara umum Bahasa Inggris yang digunakan dalam soal dianggap sulit oleh siswa.

Berdasarkan analisis tingkat kesukaran terdapat soal yang dapat dikategorikan mudah sebanyak 16,67%. Hasil ini menunjukkan bahwa terdapat banyak siswa yang dapat menjawab soal tersebut dan bukan hanya siswa pada kelompok atas saja. Soal yang tergolong sedang sebanyak 70%. Hal itu menunjukkan bahwa soal tersebut dapat dijawab oleh kelompok atas dan kelompok bawah dengan proporsi yang tepat artinya kelompok atas lebih banyak yang dapat menjawab jika dibandingkan dengan kelompok bawah. Sedangkan soal yang dapat dikategorikan sulit 13,33%. Soal-soal yang memiliki tingkat kesukaran tinggi menunjukkan bahwa untuk menjawab soal tersebut diperlukan tingkat kemampuan yang tinggi. Distraktor berfungsi dengan baik apabila dipilih oleh paling sedikit 2% dari peserta tes. Hasil analisis menunjukkan bahwa hanya distraktor pada 2 soal yang tidak berfungsi dengan baik.

Reliabilitas soal yang diperoleh dengan menggunakan metode Kuder-Richardson (KR-20) menunjukkan tingkat reliabilitas yang tinggi ($r_{11} = 0,781$) karena $0,70 \leq r_{11} < 0,80$ (Arikunto, 2005: 100). Hal ini menunjukkan bahwa instrumen yang digunakan dalam diseminasi ini reliabel untuk diteskan.

Hasil pengujian pada tahap diseminasi menunjukkan bahwa skor penguasaan materi Fisika Semester Ganjil yang diperoleh dari 61 peserta tes cukup memuaskan. Berdasarkan jumlah jawaban benar menunjukkan bahwa 54% testee mendapat skor minimal 15 dari total jawaban benar 30. Ukuran sampel yang digunakan pada tahap diseminasi menjadi 61 peserta yang ditingkatkan dari 15 responden pada tahap uji coba terbatas.

Hasil pada tahap ini menunjukkan bahwa pada indikator merumuskan suatu besaran dengan analisis dimensi dan menjelaskan karakteristik gerak melingkar beraturan yang terdapat pada masing-masing soal nomor 4 dan 18 memiliki nilai tingkat kesukaran paling rendah yang berarti bahwa soal tergolong sulit. Soal nomor 4 merupakan soal analisis (*analyze/ C4*) sehingga diperlukan tingkat kemampuan yang tinggi untuk merespon jawaban. Namun pada kenyataannya, justru kelompok bawah yang lebih banyak menjawab benar, begitu pula untuk soal nomor 18. Soal nomor 18 merupakan soal untuk mengukur kemampuan memahami (*understand/ C2*). Dengan hasil ini, menunjukkan bahwa soal kurang dapat mengukur kemampuan siswa yang sebenarnya karena siswa yang menjawab ternyata hanya menjawab secara acak. Terjadinya hal seperti ini kemungkinan disebabkan dalam

proses pembelajaran kemampuan menentukan suatu rumus atau besaran berdasarkan analisis dimensi kurang ditekankan sehingga ketika materi tersebut diteskan siswa merasa belum diajarkan. Sedangkan untuk pemahaman karakteristik gerak melingkar beraturan kemungkinan disebabkan karena siswa kurang paham dengan besaran-besaran pada gerak melingkar tersebut dan kurang paham tentang besaran vektor dan skalar. Karena pada dasarnya karakteristik yang diteskan dalam soal ini membahas tentang konstan atau tidaknya besaran dalam gerak melingkar beraturan. Jadi, besaran tersebut selain ditentukan oleh besar (*magnitude*), perlu juga diperhatikan arah (*direction*). Selanjutnya soal yang memiliki tingkat kesukaran rendah adalah soal nomor 6 dengan indikator menerapkan penulisan notasi ilmiah dalam operasi menggunakan aturan angka penting. Soal ini merupakan soal C3. Dalam hal ini, hanya sedikit responden yang menjawab dengan benar yang kemungkinan disebabkan oleh kurangnya pemahaman tentang operasi bilangan khususnya perkalian dengan menggunakan aturan angka penting. Sehingga ketika diberikan soal penerapan, siswa hanya melakukan perhitungan tanpa memperhatikan aturan angka pentingnya.

Berdasarkan hasil tersebut dapat diketahui bahwa meskipun soal merupakan soal-soal tersebut untuk mengukur kemampuan memahami, menerapkan, dan menganalisis, tetapi diperlukan kemampuan yang tinggi. Sehingga agar indikator-indikator tersebut dapat tercapai, perlu dilakukan penekanan pada pokok bahasan yang memuat indikator tersebut. Sedangkan pada indikator menentukan percepatan benda berdasarkan grafik hubungan kecepatan dengan waktu yang terdapat pada nomor 15 memiliki nilai tingkat kesukaran paling besar yang artinya soal mudah. Soal nomor 15 merupakan soal penerapan. Dengan hasil ini menunjukkan bahwa konsep siswa dalam memahami grafik untuk menentukan percepatan sudah baik dan tidak diperlukan tingkat kemampuan tinggi untuk dapat merespon soal tersebut. Untuk soal nomor 1, 3, 5, dan 9 juga dikategorikan mudah. Indikator-indikator pada nomor-nomor tersebut masing-masing adalah menyatakan satuan SI besaran turunan dalam satuan baku, menentukan besaran-besaran yang memiliki dimensi yang sama, membaca skala yang ditunjukkan suatu alat ukur besaran panjang, waktu dan massa, menerapkan metode komponen vektor untuk memperoleh resultan vektor.

Sedangkan soal-soal selain yang disebutkan di atas mempunyai tingkat kesukaran yang sedang. Hasil diseminasi ini menunjukkan terjadinya perbedaan dalam sebaran tingkat kesukaran pada tahap diseminasi dibandingkan dengan tahap uji coba terbatas. Salah satu penyebab terjadinya hal tersebut adalah perbedaan jumlah responden.

Tingkat kesukaran dan daya beda soal akan selalu berbeda dengan kelompok siswa yang berbeda. Sehingga tidak bijaksana apabila menentukan secara mutlak daya beda minimum suatu item. Hal yang perlu diperhatikan adalah daya beda soal positif atau tidak dan setiap soal mampu mengukur hasil belajar atau tidak. Apabila semua hal tersebut terpenuhi maka soal hendaknya dipertahankan (Wayan Nurkancana, 1983). Dengan demikian, terdapat 26 soal yang sebaiknya dipertahankan untuk dapat diimplementasikan, sedangkan 4 soal yang lain perlu dipertimbangkan.

SIMPULAN DAN SARAN

Kesimpulan penelitian adalah kisi-kisi instrumen tersusun berdasarkan silabus penelitian pengembangan dengan menggunakan metode adopsi sesuai dengan kurikulum KTSP dan silabus *A Level*, tes yang dikembangkan telah memenuhi karakteristik internal tes yang baik berdasarkan analisis butir soal berupa daya beda dan tingkat kesukaran soal, tes yang dikembangkan memiliki validitas isi yang baik berdasarkan hasil telaah pakar dan memiliki reliabilitas tes yang tinggi ($r_{11}=0,781$), produk akhir tes hasil penelitian pengembangan berupa 26 soal tes yang dapat dijadikan bank soal yang siap diimplementasikan, dan meskipun tes telah dikembangkan sesuai dengan standar, tes masih memiliki kelemahan, yaitu Bahasa Inggris yang dianggap sulit oleh siswa dan beberapa siswa menjawab soal dengan mengandalkan keberuntungan.

DAFTAR PUSTAKA

- Anonim. 2010. "Peringkat Sekolah RSBI se-Indonesia", (on line), <http://jedi-indonesia.blogspot.com/2011/03/peringkat-sekolah-rsbi-se-indonesia.html>, diakses Agustus 2011.
- Ardiansyah, M. Asrori. 2010. "Efektifitas Pengajaran Bahasa Inggris di RSBI", (on line), <http://history55education.wordpress.com/2011/07/23/efektifitas-pengajaran-bahasa-inggris-di-rsbi/>, diakses Agustus 2011.
- Arikunto, Suharsimi. 2005. *Dasar-Dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.
- Haryana, Kir. 2007. *Konsep Sekolah Bertaraf Internasional*. Jakarta: Direktorat Pembinaan Sekolah Menengah.
- Milist RSBI. 2011. Diakses tanggal 06 Agustus 2011, (on line), www.sman1-solo.sch.id.
- Nurkancana, Wayan, PPN. Sumartana. 1983. *Evaluasi Pendidikan*. Surabaya: Usaha Nasional.
- Pusporini, Widowati. 2009. Tesis. Pengembangan Tes Kompetensi Calistung Siswa Sekolah Dasar Provinsi DIY. Yogyakarta: UNY.
- Sudijono, Anas. 2005. *Pengantar Evaluasi Pendidikan*. Jakarta: PT. Raja Grafindo Persada.
- Thiagarajan, S., Doroty S. Semmel, dan Melvyn I. Semmel. 1974. *Instructional Development for Training Teachers of Exceptional Children*. Source Book. Bloomington: Center for Innovation on Teaching the Handicapped.

- Widianti, Evi. 2008. *Analisis Domain Kognitif dan Tipe Permasalahan dalam Soal Tes Formatif dan Sumatif pada Materi Kimia di Kelas XI Semester II SMA Berwawasan Internasional*. Skripsi pada Jurusan Pendidikan Kimia FPMIA UPI:tidak diterbitkan.
- Widyantoro,Deni dkk. 2009. “Pengembangan Soal Tes Pilihan Ganda Kompetensi Sistem Starter dan Pengisian Program Keahlian Teknik Mekanik Otomotif Kelas XII”. *Jurnal PTM*, 9(1): 14-21.