

PENGEMBANGAN INSTRUMEN TES GEOMETRI DAN PENGUKURAN PADA JENJANG SMP

Titin Mulyani, Dina Huriaty
Homeschooling Primagama Banjarmasin, STKIP PGRI Banjarmasin
dina_rty@yahoo.co.id

Abstrak: Penelitian ini bertujuan: (1) untuk mengembangkan instrumen tes pada aspek materi geometri dan pengukuran di jenjang SMP, serta (2) menetapkan karakteristik butir soal pada instrumen tes yang dikembangkan. Penelitian pengembangan ini menggunakan model pengembangan tes hasil belajar. Uji coba terhadap butir soal yang dikembangkan dilaksanakan di kelas IX SMP di Kota Banjarmasin pada 5 sekolah berbeda. Analisis butir dilakukan dengan program Iteman. Kesimpulan: (1) penelitian ini berhasil mengembangkan instrumen tes hasil belajar geometri dan pengukuran di jenjang SMP dengan memperhatikan prosedur-prosedur pengembangan yang berlaku yaitu, menyusun spesifikasi tes, menulis butir soal, menelaah butir soal, melakukan uji coba, dan menganalisis butir soal. Analisis butir dilakukan untuk mendapatkan informasi tentang tingkat kesukaran, daya beda, dan efektivitas pengecoh dari butir-butir soal yang telah dibuat, (2) dari 200 butir soal yang dikembangkan berdasarkan hasil uji coba dan analisis butir diperoleh 23 butir soal diterima, 42 butir soal yang harus direvisi, dan 135 butir soal ditolak.

Kata Kunci: pengembangan, tes, geometri, pengukuran

Evaluasi merupakan salah satu aspek penting dalam proses pembelajaran untuk mengetahui ketercapaian suatu tujuan pembelajaran dalam pendidikan. Evaluasi adalah dasar untuk menarik kesimpulan tentang apa yang harus dilakukan pada tahapan selanjutnya dalam menjalankan program pembelajaran. Mardapi (2008: 8) menyatakan, bahwa “Evaluasi merupakan salah satu rangkaian kegiatan dalam meningkatkan kualitas, kinerja, atau produktivitas suatu lembaga dalam melaksanakan programnya”. Cross dalam Sukardi (2012: 1) menyatakan,

“Evaluasi merupakan proses yang menentukan kondisi, dimana suatu tujuan telah dapat dicapai” sedangkan menurut Arifin (2012: 6), “Evaluasi merupakan salah satu komponen penting dan tahap yang harus ditempuh oleh guru untuk mengetahui keefektivan pembelajaran”. Berdasarkan definisi dari beberapa ahli tersebut dapat disimpulkan bahwa evaluasi harus dilakukan secara sistematis dan terus menerus agar dapat menggambarkan kemampuan peserta didik yang dievaluasi.

Evaluasi yang dilakukan menggunakan suatu instrumen atau alat

ukur yang berbentuk tes maupun nontes. Instrumen bentuk tes adalah pertanyaan secara tertulis yang disusun sesuai kaidah penyusunan tes yang benar, sedangkan instrumen bentuk nontes terdiri atas wawancara, pengamatan secara sistematis, atau penyebaran angket untuk mengukur suatu aspek tertentu pada peserta didik.

Mardapi (2008: 67) menyatakan, "Tes merupakan pertanyaan yang memiliki jawaban benar atau salah. Tes diartikan juga sebagai sejumlah pertanyaan yang membutuhkan jawaban, atau sejumlah pernyataan yang harus diberikan tanggapan dengan tujuan mengukur kemampuan seseorang atau mengungkap aspek tertentu dari orang yang dikenai tes". Hasan dalam Arifin (2012: 6) menyatakan, "Tes adalah alat pengumpulan data yang dirancang secara khusus yang kekhususan tersebut terlihat dari konstruksi butir soal yang digunakan". Dengan demikian, dapat diartikan bahwa tes pada hakikatnya adalah suatu alat ukur yang berisi serangkaian pertanyaan yang harus dijawab oleh peserta didik untuk mengetahui tingkat kemampuan peserta didik dalam menguasai materi pelajaran yang telah disampaikan, atau untuk mengukur perkembangan peserta didik dari aspek tingkah laku.

Menurut Mardapi (2012: 179), ada dua acuan yang digunakan dalam menyiapkan tes dan menafsirkan hasil tes, yaitu acuan norma dan acuan kriteria. Kedua acuan ini menggunakan asumsi yang berbeda tentang kemampuan seseorang. Teknik analisis butir pada kedua acuan ini ada yang sama dan ada yang berbeda. Pemilihan acuan yang tepat ditentukan oleh karakteristik bidang studi yang akan diukur dan tujuan yang akan dicapai.

Dilihat dari perencanaan tes dan penafsiran hasil tes, pengukuran dalam bidang pendidikan dapat berdasarkan acuan norma atau acuan kriteria. Acuan norma dan kriteria dalam memilih bahan tes pada prinsipnya tidak berbeda.

Perbedaan ini disebabkan oleh asumsi yang berbeda.

Bentuk tes yang digunakan dalam pendidikan terdiri atas tes objektif dan tes non objektif. Tes objektif memiliki bentuk penilaian yang sifatnya objektif dan hanya memiliki satu jawaban benar, sedangkan tes non objektif merupakan tes dengan jawaban yang dapat bervariasi sesuai dengan argumentasi yang diberikan oleh peserta didik dalam tes tersebut. Tes objektif terdiri dari beberapa bentuk, diantaranya bentuk benar-salah, pilihan ganda, menjodohkan, dan melengkapi isian, sedangkan tes non objektif berbentuk uraian.

Arifin (2012: 156) menyatakan, "Tes bentuk pilihan ganda dapat digunakan untuk mengukur hasil belajar yang lebih kompleks dan berkenaan dengan aspek ingatan, pengertian, aplikasi, analisis, sintesis, dan evaluasi". Menurut Mardapi (2008: 71), "Tes bentuk pilihan ganda adalah tes yang jawabannya dapat diperoleh dengan memilih alternatif jawaban yang telah disediakan". Sukardi (2012: 125) menyatakan, "Item tes pilihan ganda dapat digunakan untuk mengevaluasi aplikasi pengetahuan hasil belajar yang telah diberikan kepada siswa selama satu semester atau kuartal".

Tes bentuk pilihan ganda sering digunakan oleh para guru di sekolah untuk mengukur tingkah pemahaman peserta didik terhadap materi pelajaran. Tes bentuk pilihan ganda memiliki kelebihan berupa karakteristik yang baik untuk suatu alat pengukur hasil belajar peserta didik. Tes bentuk pilihan ganda lebih fleksibel dalam implementasi evaluasi dan efektif untuk mengukur tercapai tidaknya tujuan pembelajaran. Tes bentuk pilihan ganda juga mudah dalam pengoreksian, sehingga waktu yang digunakan lebih efektif dan efisien.

Selain beberapa kelebihan yang telah disebutkan, tes bentuk pilihan ganda juga memiliki kelemahan dalam hal mengembangkan kreativitas peserta didik. Tes bentuk pilihan ganda hanya berfokus

pada satu jawaban sehingga kemampuan berargumentasi peserta didik cenderung tidak berkembang. Kecurangan juga sering terjadi diantara peserta didik ketika menyelesaikan tes bentuk pilihan ganda.

Ketika akan melaksanakan ulangan harian atau bulanan, kecenderungan guru tidak membuat sendiri perangkat tes yang akan diadministrasikan kepada peserta didik. Kebanyakan guru hanya mengambil butir-butir soal dari buku atau LKS tanpa tahu apakah butir-butir soal tersebut baik atau tidak, serta mampu atau tidak mampu mengukur kompetensi peserta didik terhadap sejumlah materi pelajaran tertentu. Hal tersebut tentunya menjadi masalah dalam dunia pendidikan saat ini. Terlebih lagi jika hasil jawaban peserta didik tidak dianalisis dengan baik oleh guru.

Analisis butir soal yang dilakukan dapat menggunakan salah satu dari dua teori tes yang ada. Teori tersebut adalah teori tes klasik dan teori tes modern. Purwanto (2013: 98) menyatakan, "Teori tes klasik adalah teori mengenai analisis butir tes dimana analisis dilakukan dengan memperhitungkan kedudukan butir dalam suatu kelas atau kelompok". Karakteristik butir soal yang diuji dalam teori tes klasik meliputi tingkat kesukaran, daya beda, dan efektivitas pengecoh.

Crocker dan Algina dalam Purwanto (2013: 99) menyatakan, "Tingkat kesukaran atau (difficulty index) atau disingkat TK dapat didefinisikan sebagai proporsi siswa peserta tes yang menjawab benar". Aiken dalam Wahidmurni, dkk (2010: 131) mendefinisikan, "Tingkat kesukaran soal adalah peluang untuk menjawab benar suatu soal pada tingkat kemampuan tertentu yang biasanya dinyatakan dalam bentuk indeks". Indeks tingkat kesukaran tersebut pada umumnya dinyatakan dalam proporsi yang berkisar dari 0,00 sampai dengan 1,00. Semakin besar indeks tingkat kesukaran, semakin mudah soal yang dibuat.

Wahidmurni, dkk (2010: 134) menyatakan, "Daya pembeda soal adalah kemampuan suatu butir soal dapat membedakan antara warga belajar/peserta didik yang telah menguasai materi yang ditanyakan dan warga belajar/peserta didik yang tidak/kurang/belum menguasai materi yang ditanyakan". Indeks daya beda juga dinyatakan dalam bentuk proporsi yang berkisar antara -1,00 sampai dengan +1,00. Semakin tinggi daya beda, semakin baik soal yang dibuat.

Analisis butir juga dilakukan dengan memperhatikan efektivitas pengecoh. Pengecoh (distractor) adalah pilihan jawaban yang bukan merupakan kunci jawaban. Pengecoh bukan sekedar pelengkap pilihan. Pengecoh diadakan untuk menyesatkan peserta didik agar tidak memilih kunci jawaban. Pengecoh menggoda peserta didik yang kurang begitu memahami materi pelajaran untuk memilihnya. Agar dapat melakukan fungsinya untuk mengecoh maka pengecoh harus dibuat homogen dengan kunci jawaban.

Pengecoh dikatakan berfungsi efektif apabila ada peserta didik yang terkecoh memilih. Pengecoh yang sama sekali tidak dipilih tidak dapat melakukan fungsinya sebagai pengecoh soal. Pengecoh yang berdasarkan hasil uji coba tidak efektif direkomendasikan untuk diganti dengan pengecoh yang lebih menarik (Purwanto, 2013: 108).

Wahidmurni, dkk (2010: 136) menyatakan, "Suatu pilihan jawaban (pengecoh) dapat dikatakan berfungsi apabila pengecoh paling tidak dipilih 5% peserta tes, serta lebih banyak dipilih oleh kelompok peserta didik yang belum paham materi".

Pada mata pelajaran matematika, aspek materi yang terdapat dalam SKL meliputi Aritmetika, Aljabar, Geometri dan Pengukuran, serta Statistika dan Peluang. Pada pelaksanaan ujian nasional, misalnya pelaksanaan tahun 2012/2013 materi Geometri dan Pengukuran meliputi 45% dari keseluruhan jumlah soal. Penguasaan

peserta didik terhadap materi geometri dan pengukuran sangat diperlukan agar berhasil dalam menempuh ujian nasional. Geometri dan pengukuran merupakan materi yang penting dan menjadi dasar dalam ilmu matematika, sehingga pada setiap tingkat pendidikan peserta didik selalu diajarkan tentang konsep geometri dan pengukuran. Oleh karena itu, diperlukan adanya pengembangan instrumen hasil belajar untuk menghasilkan butir-butir soal yang baik.

Penelitian ini bertujuan untuk mengembangkan instrumen tes pada aspek materi geometri dan pengukuran di jenjang SMP dan menetapkan karakteristik butir soal pada instrumen tes yang dikembangkan. Masalah dalam penelitian ini dibatasi pada pengembangan instrumen tes hasil belajar yang terdiri atas sejumlah butir soal. Penetapan penerimaan butir soal berdasarkan pada kriteria butir soal yang baik ditinjau dari tingkat kesukaran, daya beda, dan efektivitas pengecoh. Dari hasil penelitian ini diharapkan diperoleh sejumlah butir-butir soal yang telah terstandar, sehingga dapat dimanfaatkan oleh guru di sekolah untuk latihan, ulangan harian, ulangan tengah semester, ulangan akhir semester, dan ujian sekolah. Hasil penelitian ini juga dapat dijadikan referensi bagi guru atau calon guru, agar dapat memahami prosedur penyusunan dan penyajian instrumen tes yang tepat.

Metode Penelitian

Penelitian ini merupakan penelitian pengembangan menggunakan model pengembangan tes hasil belajar. Uji coba terhadap butir soal yang dikembangkan dilaksanakan di kelas IX SMP di Kota Banjarmasin pada 5 sekolah berbeda. Subjek uji coba penelitian adalah siswa kelas IX SMP di Kota Banjarmasin. Sekolah yang dijadikan sebagai tempat uji coba adalah SMP Negeri 7 Banjarmasin, SMP Negeri 9 Banjarmasin, SMP Negeri 10 Banjarmasin, SMP Negeri 13

Banjarmasin, dan SMP Negeri 26 Banjarmasin. Analisis butir dilakukan dengan program *Iteman*.

Prosedur pengembangan tes hasil belajar ini mengacu pada Mardapi (2012:110), yakni menyusun spesifikasi tes, menulis butir soal, menelaah butir soal, melakukan uji coba terhadap butir soal, menganalisis butir soal, memperbaiki tes, merakit tes, melaksanakan tes, dan menafsirkan hasil tes. Akan tetapi, karena tujuan dari penelitian pengembangan ini hanya untuk menghasilkan butir-butir soal yang baik ditinjau dari tingkat kesukaran, daya beda, dan efektivitas pengecoh, maka prosedur pengembangan tersebut hanya dilakukan sampai pada tahapan menganalisis butir soal yang telah diuji coba. Berikut tahapan-tahapan yang dilakukan sebagai berikut.

1. Menyusun spesifikasi tes

Pada tahapan ini, ditentukan tujuan diselenggarakannya tes. Tujuan yang ingin ditetapkan mengacu pada apa yang ingin diukur dari peserta didik. Setelah tujuan ditetapkan, dilakukan penyusunan kisi-kisi tes dan menentukan bentuk dan panjang tes.

2. Menulis butir soal

Berdasarkan kisi-kisi yang telah disusun, butir-butir soal dikembangkan sebanyak yang telah ditetapkan. Adapun banyaknya butir soal yang dibuat berjumlah 200 butir khusus pada aspek geometri dan pengukuran bangun datar.

3. Menelaah butir soal

Pada tahapan ini, butir-butir soal yang telah dibuat ditelaah dari segi kaidah penulisan soal bentuk pilihan ganda. Butir-butir soal yang telah ditelaah dan tidak sesuai dengan kaidah, maka perlu dilakukan revisi pada butir soal tersebut. Hal ini dilakukan guna menghasilkan butir soal yang baik sebelum diujicoba. Telaah mengacu pada 16 kaidah penulisan soal yang dikembangkan oleh Depdiknas. (2000), yang terdiri atas materi, konstruksi, dan bahasa.

Telaah butir soal dilakukan oleh 5 (lima) orang guru matematika. Lembar

telaah kaidah penulisan soal bentuk pilihan ganda dan perangkat tes diberikan kepada guru matematika untuk menilai kesesuaian butir soal yang diujikan berdasarkan kaidah penulisan soal bentuk pilihan ganda.

4. Melakukan uji coba terhadap butir soal

Pada tahapan uji coba, butir-butir soal yang telah ditelaah dari segi kaidah penulisan soal, diujicobakan kepada responden yaitu peserta didik sebagai testee. Butir-butir soal yang diujicobakan berjumlah 200 butir soal yang disusun kedalam 25 paket soal bentuk pilihan ganda.

Pada setiap paket terdiri atas 8 butir soal yang dirakit bersama dengan butir soal pada aspek lainnya sesuai SKL tingkat SMP. Dua puluh lima paket soal tersebut diujicobakan kepada siswa kelas IX SMP di lima sekolah berbeda secara bertahap, dengan masing-masing sekolah mendapatkan sejumlah paket soal untuk diujicobakan.

5. Menganalisis butir soal

Hasil uji coba kemudian dianalisis untuk mengetahui tingkat kesukaran, daya beda, dan efektivitas pengecoh dari setiap butir soal yang dibuat. Pada tahapan ini, akan ditentukan butir-butir soal yang layak diterima, tidak layak diterima, maupun yang harus diperbaiki.

Produk soal yang dihasilkan ditujukan untuk kepentingan pengujian secara lokal di kelas dalam bentuk soal latihan, ulangan harian, ulangan tengah semester, maupun ulangan akhir semester. Karena itu, analisis data yang digunakan mengacu pada analisis kuantitatif acuan norma (Kusaeri & Suprananto, 2012:173-177) Analisis dilakukan dengan menggunakan program *Iteman* untuk mengetahui tingkat kesukaran butir, daya beda butir, dan efektivitas pengecoh butir soal.

a. Tingkat kesukaran

Indeks tingkat kesukaran berkisar dari 0,00 sampai dengan 1,00. Klasifikasi

tingkat kesukaran tersebut dapat dilihat pada tabel berikut.

Tabel 1 Klasifikasi Tingkat Kesukaran

Indeks Tingkat Kesukaran	Klasifikasi
$0,00 \leq TK < 0,30$	Soal tergolong sukar
$0,30 \leq TK \leq 0,70$	Soal tergolong sedang
$0,70 < TK \leq 1,00$	Soal tergolong mudah

b. Daya beda

Indeks daya beda berkisar antara -1,00 sampai dengan +1,00. Semakin tinggi daya beda suatu butir soal, semakin baik soal yang dibuat. Adapun klasifikasi daya beda suatu butir soal dapat dilihat pada tabel berikut.

Tabel 2. Klasifikasi Daya Beda

Indeks Daya Beda	Klasifikasi
$DB \geq 0,30$	Butir soal diterima
$DB < 0,30$	Butir soal ditolak

c. Efektivitas pengecoh

Suatu pengecoh dikatakan berfungsi apabila dipilih paling tidak oleh 5% peserta tes.

Berdasarkan tingkat kesukaran butir soal, daya beda butir soal, dan efektivitas pengecoh dapat diklasifikasikan butir-butir soal yang diterima, diterima dengan perbaikan, dan ditolak mengacu pada kriteria sebagai berikut.

- 1) Butir soal diterima
 - a) Memiliki tingkat kesukaran $0,30 \leq TK \leq 0,70$
 - b) Memiliki daya beda $DB \geq 0,30$
 - c) Memiliki pengecoh yang berfungsi, yaitu dipilih oleh $\geq 5\%$ testee.
- 2) Butir soal diterima dengan perbaikan
 - a) Memiliki tingkat kesukaran $0,30 \leq TK \leq 0,70$
 - b) Memiliki daya beda $DB \geq 0,30$
 - c) Memiliki pengecoh yang berfungsi, yaitu dipilih oleh $< 5\%$ testee.

Butir-butir soal yang tidak memenuhi syarat sebagai butir soal diterima maupun butir soal yang diterima dengan perbaikan, maka butir soal tersebut ditolak.

Hasil dan Pembahasan

Hasil

Hasil penelitian diuraikan berdasarkan langkah-langkah pengembangan.

1. Menyusun spesifikasi tes

Pada tahapan ini disusun kisi-kisi penulisan soal berdasarkan SKL tingkat SMP. Adapun kisi-kisi yang disusun berdasarkan SKL tingkat SMP dapat dilihat pada Lampiran 1.

2. Menulis butir soal

Butir-butir soal yang dibuat berdasar pada indikator penulisan soal yang telah disusun dalam kisi-kisi. Butir soal yang dibuat untuk setiap indikator penulisan soal berjumlah 25 butir, sehingga dihasilkan 200 butir soal untuk 8 indikator penulisan soal.

3. Menelaah butir soal

Butir-butir soal yang dikembangkan, kemudian ditelaah oleh guru matematika untuk mengetahui kesesuaian butir soal terhadap kaidah penulisan bentuk soal pilihan ganda yang tepat. Berdasarkan hasil telaah yang dilakukan, diperoleh butir-butir soal yang dibuat telah memenuhi kaidah penulisan soal yang tepat dan layak untuk dilakukan uji coba.

Butir-butir soal yang masih belum memenuhi kriteria kaidah penulisan soal yang baik (hasil telaah yang belum mencapai 100%), dikarenakan adanya butir soal yang memiliki kunci jawaban benar lebih dari satu dan terdapat gambar, grafik, tabel, diagram dan sejenisnya pada soal harus jelas dan berfungsi atau kurangnya kesesuaian soal dengan gambar yang disajikan. Butir-butir soal tersebut kemudian diperbaiki terlebih dahulu sebelum dilakukan uji coba.

4. Melakukan uji coba

Butir-butir soal yang telah ditelaah, kemudian diujicoba di lima sekolah berbeda, yaitu SMP Negeri 7 Banjarmasin, SMP Negeri 9 Banjarmasin, SMP Negeri 10 Banjarmasin, SMP Negeri 13 Banjarmasin, dan SMP Negeri 26 Banjarmasin.

Uji coba di SMP Negeri 26 Banjarmasin dilakukan terhadap responden yang berjumlah 148 siswa yang terbagi dalam 6 kelas IX, yakni kelas IX A, IX B, IX C, IX D, IX E, dan IX F. Butir-butir soal yang diujicobakan di SMP Negeri 26 Banjarmasin untuk aspek geometri dan pengukuran khusus bangun datar berjumlah 40 butir soal yang terbagi kedalam 5 paket soal bentuk pilihan ganda. Setiap paket soal yang diujicobakan terdiri atas beberapa aspek atau materi yang terdapat dalam SKL tingkat SMP, sehingga untuk aspek geometri dan pengukuran khusus bangun datar terdiri atas 8 butir soal.

Uji coba di SMP Negeri 13 Banjarmasin dengan responden berjumlah 174 siswa yang tersebar dalam 6 kelas IX, yakni kelas IX A, IX B, IX C, IX D, IX E, dan IX F. Butir-butir soal yang diujicobakan di SMP Negeri 13 Banjarmasin berjumlah 48 butir yang terbagi kedalam 6 paket soal bentuk pilihan ganda. Butir-butir soal tersebut dirakit bersama dengan sejumlah butir soal selain aspek geometri dan pengukuran khusus bangun datar dalam SKL tingkat SMP.

Uji coba selanjutnya dilakukan di SMP Negeri 7 Banjarmasin. Uji coba di SMP Negeri 7 Banjarmasin diikuti oleh 191 siswa, dengan rata-rata jumlah siswa setiap kelas sebanyak 32 siswa. Adapun banyaknya butir soal yang diujicobakan untuk aspek geometri dan pengukuran khusus bangun datar berjumlah 48 butir yang terbagi kedalam 6 paket soal bentuk pilihan ganda.

Uji coba di SMP Negeri 10 Banjarmasin diikuti oleh 176 siswa yang terbagi dalam 6 kelas IX, yaitu kelas IX A, IX B, IX C, IX D, IX E, dan IX F. Adapun banyak butir soal yang diujicobakan

khusus geometri dan pengukuran pada bangun datar berjumlah 48 butir yang terbagi kedalam 6 paket soal bentuk pilihan ganda, serta tersusun bersama dengan butir-butir soal yang lain sesuai SKL tingkat SMP.

Uji coba yang terakhir dilakukan di SMP Negeri 9 Banjarmasin. Kelas yang dijadikan sebagai subjek uji coba adalah kelas IX D dan IX F dengan jumlah responden keseluruhan 54 siswa. Adapun banyaknya butir soal untuk aspek geometri dan pengukuran bangun datar yang diujicobakan berjumlah 16 butir yang terbagi dalam 2 paket soal bentuk pilihan ganda yang dirakit bersama dengan butir-butir soal sesuai SKL tingkat SMP.

5. Analisis butir soal

Respon terhadap butir-butir soal yang diujicoba kemudian dianalisis dengan program *Iteman*. Hasil analisis diperoleh bahwa dari 200 butir soal yang telah dibuat, terdapat butir soal yang belum memenuhi kriteria tingkat kesukaran, daya beda, dan efektivitas pengecoh yang tepat. Adapun butir-butir soal yang diklasifikasikan mudah, sedang, dan sukar dari 200 butir soal berdasarkan tingkat kesukaran dapat dilihat pada tabel berikut.

Tabel 3. Persentase Butir Soal Berdasarkan Tingkat Kesukaran

Klasifikasi	Jumlah butir	Persentase
Mudah	24 butir	12,0%
Sedang	89 butir	44,5%
Sukar	87 butir	43,5%
Total	200 butir	100,0%

Butir-butir soal yang termasuk tingkat kesukaran sedang lebih banyak dibandingkan dengan butir-butir soal yang mudah maupun sukar. Tampak bahwa persentase butir-butir soal dengan tingkat kesukaran sedang adalah 44,5%, sedangkan butir-butir soal yang mudah hanya 12,0% dan sukar sebesar 43,5%.

Analisis terhadap daya beda butir dapat dilihat pada Tabel 4. Berdasarkan tabel tersebut, butir-butir soal yang diterima dengan indeks daya beda soal $\geq 0,30$ adalah 59,5% atau 119 butir,

sedangkan sisanya butir-butir soal ditolak karena indeks daya beda $< 0,30$ adalah 81 butir atau 40,5%.

Tabel 4. Persentase Butir Soal Berdasarkan Analisis Daya Beda

Klasifikasi	Jumlah butir	Persentase
Butir soal diterima	119 butir	59,5%
Butir soal ditolak	81 butir	40,5%
Total	200 butir	100,0%

Dari 65 butir soal yang diterima terdapat 42 butir soal yang diterima dengan syarat revisi *option*, karena pengecoh dipilih kurang dari 5% peserta uji coba atau sebesar 21%.

Berdasarkan kriteria tingkat kesukaran soal, daya beda, dan efektivitas pengecoh disimpulkan bahwa banyak soal yang diterima adalah 23 butir, ditolak 135 butir, dan diterima dengan syarat perbaikan atau revisi *option* adalah 42 butir soal. Selengkapnya dapat dilihat pada Tabel 5 di bawah ini.

Tabel 5. Hasil Analisis Butir

Kategori	Jumlah butir	Persentase
Diterima	23 butir	11,5%
Ditolak	135 butir	67,5%
Revisi	42 butir	21,0%
Total	200 butir	100%

Pembahasan

Dari hasil deskripsi di atas menunjukkan bahwa untuk materi geometri dan pengukuran bangun datar, butir-butir soal yang disajikan termasuk kelompok soal yang sukar bagi responden, sehingga meskipun jumlah butir-butir soal dengan tingkat kesukaran sedang lebih banyak, tetapi mengamati persentase dari tingkat kesukaran soal tampak bahwa soal-soal yang disajikan termasuk kategori soal yang sukar bagi responden uji coba.

Beberapa faktor yang diduga menjadi penyebab, diantaranya:

- 1) tingkat penguasaan responden/peserta didik yang rendah terhadap aspek

materi geometri dan pengukuran bangun datar,

- 2) kesiapan responden/peserta didik pada saat uji coba berlangsung,
- 3) belum adanya pengulangan terhadap materi tersebut di sekolah.

Analisis terhadap indeks daya beda menunjukkan bahwa lebih dari setengah soal yang diujicobakan termasuk soal yang diterima, yaitu 59,5% atau 119 butir. Jumlah ini cukup besar jika dibandingkan dengan jumlah butir soal yang diterima yaitu 65 butir. Perbedaan jumlah ini juga tampak pada analisis tingkat kesukaran soal, yaitu pada soal dengan kategori sedang yaitu 89 butir. Hal ini disebabkan karena butir soal dengan klasifikasi tingkat kesukaran sedang, tetapi memiliki indeks daya beda $< 0,30$, atau butir soal yang memiliki daya beda $\geq 0,30$ tetapi tingkat kesukaran soal termasuk sukar atau mudah. Butir-butir soal yang ditolak bukan berarti tidak dapat digunakan. Butir-butir soal tersebut dapat direview kembali dan diperbaiki untuk selanjutnya diujicobakan pada kelompok responden yang lebih heterogen. Hal ini dapat dilakukan mengingat analisis dengan teori tes klasik sangat tergantung dengan kemampuan sampel uji coba yang dipilih (*dependent sampling*).

Simpulan

Simpulan

Dari hasil penelitian dan pembahasan di atas, dapat ditarik kesimpulan: (1) penelitian ini berhasil mengembangkan instrumen tes hasil belajar geometri dan pengukuran di jenjang SMP dengan memperhatikan prosedur-prosedur pengembangan yang berlaku yaitu, menyusun spesifikasi tes, menulis butir soal, menelaah butir soal, melakukan uji coba, dan menganalisis butir soal. Analisis butir dilakukan untuk mendapatkan informasi tentang tingkat kesukaran, daya beda, dan efektivitas pengecoh dari butir-butir soal yang telah

dibuat, (2) dari 200 butir soal yang dikembangkan berdasarkan hasil uji coba dan analisis butir diperoleh 23 butir soal diterima, 42 butir soal yang harus direvisi, dan 135 butir soal ditolak.

Saran

Disarankan untuk pembaca yang ingin melakukan penelitian sejenis dapat melakukan penelitian dengan subjek uji coba yang lebih besar dan memiliki kemampuan yang lebih heterogen. *Software* yang digunakan untuk menganalisis butir soal dalam penelitian ini adalah program *IteMan*, namun pembaca juga dapat menggunakan aplikasi komputer yang lain seperti *Excel*, *SPSS*, *Anates*, atau *software* lainnya yang kompatibel untuk menganalisis butir soal.

Daftar Pustaka

- Arifin, Zainal. 2012. *Evaluasi Pembelajaran*. Bandung: Remaja Rosdakarya.
- Depdiknas. 2000. *Pedoman pengembangan bank soal*. Jakarta: Balitbang, Pusat Penelitian dan Pengembangan Sistem Pengujian.
- Kusaeri dan Suprananto. 2012. *Pengukuran dan Penilaian Pendidikan*. Yogyakarta: Graha Ilmu.
- Mardapi, Djemari. 2008. *Teknik Penyusunan Instrumen Tes dan Nontes*. Yogyakarta: Mitra Cendikia.
- . 2012. *Pengukuran, Penilaian & Evaluasi Pendidikan*. Yogyakarta: Nuha Medika.
- Purwanto. 2013. *Evaluasi Hasil Belajar*. Yogyakarta: Pustaka Pelajar.
- Sukardi. 2012. *Evaluasi Pendidikan Prinsip dan Operasionalnya*. Jakarta: Bumi Aksara.
- Wahidmurni, dkk. 2010. *Evaluasi Pembelajaran Kompetensi dan Praktik*. Yogyakarta: Nuha Litera.