

Pengembangan Instrumen Tes Standar Kognitif pada Mata Pelajaran IPA Kelas 7 SMP Di Kabupaten Banjar

The Development of Standardized Test Instrument Natural Science Subjects Seventh Grade Junior High School in Kabupaten Banjar

Dyah Febria Wardhani*, Aminuddin Prahatama Putra

Universitas Lambung Mangkurat, Jl. Brigjend H.Hasan Basrie, Banjarmasin, Indonesia

*corresponding author: aminuddinpatra@unlam.ac.id

Abstract: The purpose of this study was to develop a standardized test instrument cognitive science subjects 7 junior class in accordance with steps of developing tests based on lattice UN year 2015/2016. Study Cognitive tests include the dimensions of factual knowledge, conceptual and procedural accordance with the revised Bloom cognitive levels. This type of research is the development of research. The data collection was done by using the test and non test. Source of data obtained from the results of the test answer sheets study, questionnaires and interviews with teachers and students in grade 7 junior in Banjar district. The sampling technique used purposive sampling and data analysis done qualitatively covering material, construction and language. Quantitative data analysis were performed using the MICROCAT Bigsteps 2:30 version. Results of the research is a description of the procedural process and product development are developed. The steps used to develop test instruments is to determine the purpose test, preparing test specifications, penulisan soal, a qualitative analysis by 9 experts (3 experts material, 3 experts construct, and 3 linguists), the revised phase I trials practicality in 4 teachers, a revised phase II, simulated use of the product in 3 students, the revision phase III, trial products on a single class of students from each school samples, analysis of items, and the revision stage IV. Learning achievement test instruments developed have excellent content validity based on the review of experts and high reliability butir soal the 0.95. The end product development test research results in the form of 29 questions multiple choice test that can be used as a question bank and ready to be implemented.

Keywords: Development of test instrument, Natural sciences, Junior high school, Kabupaten Banjar

1. PENDAHULUAN

Kurikulum, proses pembelajaran dan penilaian merupakan tiga dimensi dari sekian banyak dimensi yang sangat penting dalam pendidikan. Ketiga dimensi tersebut saling berkaitan antara yang satu dengan yang lainnya. Oleh sebab itu, disamping kurikulum yang cocok dan proses pembelajaran yang benar perlu ada sistem penilaian yang baik dan terencana (Surapranata, 2004).

Penilaian hasil belajar peserta didik jenjang pendidikan dasar dan menengah dilaksanakan berdasarkan Standar Kompetensi Lulusan (PERMENDIKNAS RI No. 23 Tahun 2006 dan PERMENDIKBUD RI No. 54 Tahun 2013) dan Standar Penilaian Pendidikan (PERMENDIKNAS RI No. 20 Tahun 2007 dan PERMENDIKBUD RI No. 66 Tahun 2013) yang berlaku secara nasional.

Standar kompetensi lulusan mengharuskan peserta didik lulusan SMP/MTs/SMPLB/Paket B memiliki dimensi pengetahuan yaitu pengetahuan factual, konseptual, dan prosedural dalam ilmu pengetahuan, teknologi, seni dan budaya dengan wawasan kemanusiaan, kebangsaan, kenegaraan, dan peradaban terkait fenomena dan kejadian yang tampak mata (Warso, 2014).

Warso (2014) dan Sudijono (2015) menjelaskan bahwa penilaian dilakukan secara menyeluruh meliputi aspek kognitif, afektif dan psikomotor untuk

setiap jenjang pendidikan. Keterampilan berpikir dan hasil belajar kognitif penting bagi siswa dalam mempelajari berbagai mata pelajaran, termasuk ilmu pengetahuan alam (IPA).

Biologi merupakan salah satu mata pelajaran di SMP yang terintegrasi dalam mata pelajaran IPA. Pada aspek Biologi, IPA mengkaji berbagai persoalan yang terkait dengan berbagai fenomena yang terjadi pada makhluk hidup dan interaksinya dengan berbagai faktor lingkungan (Mamu, 2014).

Standar kompetensi lulusan mata pelajaran IPA SMP/MTs khususnya aspek Biologi adalah 1) melakukan pengamatan dengan peralatan yang sesuai, melaksanakan percobaan sesuai prosedur, mencatat hasil pengamatan dan pengukuran dalam tabel dan grafik yang sesuai, membuat kesimpulan dan mengkomunikasikannya secara lisan dan tulisan sesuai dengan bukti yang diperoleh, 2) Memahami keanekaragaman hayati, klasifikasi keragamannya berdasarkan ciri, cara-cara pelestariannya, serta saling ketergantungan antar makhluk hidup di dalam ekosistem dan 3) Memahami sistem organ pada manusia dan kelangsungan makhluk hidup.

Standar penilaian pendidikan menyebutkan bahwa prinsip-prinsip yang mendasari penilaian hasil belajar yaitu sah, objektif, adi, terpadu, terbuka, menyeluruh dan berkesinambungan, sistematis, beracuan kriteria, ekonomis, akuntabel, dan edukatif.

Prinsip-prinsip tersebut menghendaki soal yang tes yang standar (Mendiknas, 2007; Mendikbud, 2013).

Tes standar adalah tes dimana soal-soalnya sudah mengalami proses analisis baik secara kualitatif maupun secara kuantitatif (Puspendik Balitbang Depdiknas, 2007).

Sempat saat ini banyak instrumen hasil belajar yang belum memenuhi persyaratan sebagai tes yang baik (Mardapi, 2011). Salah satu hal yang mungkin menjadi penyebabnya adalah kemampuan guru dalam membuat tes yang masih rendah sehingga pengukuran menjadi tidak akurat (Suwanto, 2009).

Hasil analisis butir soal ulangan umum kenaikan kelas tahun pelajaran 2014/2015 yang dilakukan guru di Kabupaten Banjar diperoleh hasil nilai reliabilitas sebesar 0,67; soal yang tuntas dijawab oleh siswa sebesar 25,40%, jumlah siswa yang tidak tuntas sebanyak 81,70%; persentase daya serap sebesar 50,10%, soal yang memiliki kualitas pengecoh baik sebanyak 26,25%; soal yang memiliki daya beda baik sebanyak 26,88%; tingkat kesukaran soal 14,50% sangat mudah, 7,50% mudah, 33,50% sedang, 20,75% sukar dan 23,75% sangat sukar. Analisis butir soal melalui proportion correct dan koefisien Point Biserial diperoleh hasil 27% soal yang diterima, 43% soal yang harus direvisi dan 30% soal yang ditolak.

Analisis butir soal tersebut menunjukkan bahwa instrumen tes yang telah digunakan selama ini belum standar. Analisis yang dilakukan guru masih belum menunjukkan independensi dan nirbias dari soal yang telah diujikan. Oleh karena itu, diperlukan suatu instrumen tes kognitif standar yang dapat dihasilkan melalui tahapan pengembangan tes.

Standar penilaian pendidikan menyebutkan bahwa instrumen penilaian yang digunakan harus memenuhi persyaratan substansi (isi), konstruksi dan bahasa serta memiliki bukti validitas empiric serta menghasilkan skor yang dapat diperbandingkan antar sekolah, antardaerah dan antar nasional (Mendiknas, 2007). Instrumen penilaian yang standar harus valid, reliabel dan bebas dari adanya unsur bias (Mardapi, 2011).

Teori respon butir adalah teori statistik umum mengenai soal ujian dan karakteristik tes serta bagaimana karakteristik tersebut dihubungkan dengan kemampuan yang diukur oleh soal-soal dalam tes. Respon butir dapat diukur secara dikotomus dan politomus (Hambleton, 1996). Jika data yang digunakan diskor secara dikotomus, penggunaan analisis kuantitatif tes menjadi sangat sesuai sehingga tes ini lebih reliable dan valid dibandingkan dengan teori tes klasik (Scheiblechner, 2009).

Penelitian yang dilakukan oleh Rahayu (2012) menyebutkan bahwa pengembangan instrumen penilaian kognitif sangat penting untuk diuji kelayakannya. Mardapi(2011) menjelaskan bahwa pengembangan instrumen tes dengan mengikuti langkah-langkah yang sistematis dapat memberikan hasil yang optimal. Aplikasi Rasch model dalam pengembangan soal tes dapat menjadi alat yang kinerjanya sangat baik dalam evaluasi dan pemilihan soal yang baik sehingga hasilnya bagus dan valid (Sukor, 2013). Melalui pengembangan tes dan analisis

butir soal yang digunakan, soal-soal yang telah dikembangkan benar-benar bisa mengukur dan mengetahui seberapa jauh kemampuan siswa (Widyantoro, 2009).

Berdasarkan uraian sebelumnya, tujuan penelitian ini adalah unuk mengembangkan instrumen tes standar kognitif mata pelajaran IPA Kelas 7 SMP sesuai dengan langkah-langkah pengembangan tes berdasarkan kisi-kisi UN tahun 2015/2016. Jenjang kognitif tes meliputi dimensi pengetahuan factual, pengetahuan konseptual dan procedural sesuai dengan tingkatan kognitif Bloom revisi.

Hasil dari penelitian pengembangan ini adalah diperolehnya instrumen tes kognitif standar mata pelajaran IPA khususnya Biologi berupa soal-soal pilihan ganda yang terdiri atas 4 pilihan jawaban yang disusun sesuai dengan kisi-kisi UN IPA SMP/MTs tahun pelajaran 2015/2016 untuk tingkatan materi kelas 7 SMP di Kabupaten Banjar yang dapat dijadikan Bank Soal yang berkualitas baik.

2. METODE PENELITIAN

Penelitian ini dilakukan di kelas 7 SMP di Kabupaten Banjar yang menggunakan Kurikulum KTSP 2006 sebanyak 4 sekolah yaitu SMPN 3 Martapura, SMPN 1 Pengaron, SMPN 2 Sungai Pinang dan SMP Darussalam.

Subjek penelitian ini yaitu (1) Siswa sebagai subjek untuk simulasi dan uji coba produk (uji kelas), (2) Pakar sebanyak 9 orang untuk validasi instrumen tes yang dibuat. Tim ahli sebagai validator yang terdiri atas 3 orang ahli substansi (bidang studi Biologi), 3 orang ahli konstruksi (evaluasi) dan 3 orang ahli bahasa dan (3) Guru sebanyak 4 orang untuk uji keterbacaan instrument soal.

Penelitian ini merupakan jenis penelitian pengembangan. Model penelitian yang dijadikan acuan adalah tahapan pengembangan dari Model R&D oleh Sugiyono (2015). Tahapan tersebut terdiri dari desain produk, validasi desain, revisi desain, uji coba produk, revisi produk, uji coba pemakaian, revisi produk dan produksi massal. Penelitian yang dilakukan hanya sampai pada revisi produk setelah uji coba produk, sehingga tiga tahapan terakhir akan dilakukan pada tahapan penelitian selanjutnya.

Instrumen tes yang dikembangkan dalam penelitian ini adalah soal-soal pilihan ganda yang terdiri atas 4 pilihan jawaban sesuai dengan kisi-kisi UN SMP/MTs mata pelajaran IPA khususnya Biologi Tahun Pelajaran 2015/2016.

Teknik pengumpulan data yang dilakukan dalam penelitian ini adalah (1) Teknik Tes, dilakukan kepada siswa menggunakan lembar soal instrumen tes kognitif standard, dan (2) Teknik Non Tes yang dilakukan kepada siswa dan guru menggunakan angket dan wawancara.

Teknik *sampling* menggunakan *purposive sampling* berdasarkan letak (*region*) daerah sekolah yaitu sekolah yang terletak di daerah perkotaan, pinggiran dan desa terpencil serta sekolah swasta.

Teknik analisis data dilakukan secara kualitatif dan kuantitatif. Analisis secara kualitatif



menggunakan hasil validasi desain (substansi, konstruksi dan bahasa) dan kritik saran dari pakar serta hasil penelitian secara keseluruhan. Analisis data secara kuantitatif menggunakan hasil analisis program *MicroCAT Bigsteps* Versi 2.30.

Analisis kualitatif data berupa nilai validitas isi instrument tes kognitif standard didasarkan pada kriteria dari Rahayu (2012) seperti pada table 1 berikut.

Table 1. Kriteria Interpretasi Validitas Butir Tes

Validitas	Interpretasi Validitas Item
0,80-1,00	Sangat tinggi
0,60-0,80	Tinggi
0,40-0,60	Cukup
0,20-0,40	Rendah
0,00-0,20	Sangat rendah

Sumber : Rahayu (2012)

Menurut Prilantari (2007), Dalam pendekatan teori respons butir model logistik satu parameter, kualitas butir soal dikelompokkan baik, cukup baik dan tidak baik dengan ketentuan sebagai berikut:

- Baik, apabila butir cocok dengan model dan parameter tingkat kesukaran $-2,00 \leq b_i \leq 2,00$
- Cukup baik, apabila butir cocok dengan model dan parameter tingkat kesukaran $\leq -2,00$ atau $\geq 2,00$.
- Tidak baik apabila butir tidak cocok dengan model.

Berdasarkan hasil (*output*) dari analisis model Rasch dengan program *MicroCAT BIGSTEPS* versi 2.30, diperoleh informasi estimasi nilai taraf sukar butir tes yang selanjutnya dikategorikan menurut rentang nilai dengan pengkategorian skala logits. Butir soal dengan nilai taraf sukar butir pada skala logits $\delta < -2,00$ dikategorikan sebagai butir yang sangat mudah, butir dengan $-2,00 \leq \delta \leq -1,00$ dikategorikan sebagai butir yang mudah, butir dengan $-1,00 \leq \delta \leq +1,00$ dikategorikan sebagai butir yang sedang, butir dengan $+1,00 \leq \delta \leq +2,00$ dikategorikan sebagai butir yang sukar dan $\geq +2,00$ dikategorikan sebagai butir yang sangat sukar (Rogayah, 2010)

3. HASIL DAN PEMBAHASAN

3.1. Proses Pengembangan Instrumen Tes Standar Kognitif IPA Kelas 7 SMP di Kabupaten Banjar

Instrumen tes kognitif standard yang di kembangkan berupa soal-soal pilihan ganda dengan empat pilihan jawaban berdasarkan kisi-kisi UN IPA SMP/MTs khususnya Biologi kelas 7 SMP di Kabupaten Banjar. Proses pengembangan instrument tes standar kognitif yang dilakukan sesuai dengan tahapan pengembangan dari penelitian R&D Sugiyono (2015) yaitu:

3.1.1 Desain Produk

Desain produk ini berkaitan dengan desain instrument tes kognitif yang baru. Tahapan ini dilakukan sesuai dengan tahapan pengembangan tes. Langkah awal dalam pengembangan tes adalah penyusunan spesifikasi instrument tes dan perakitan soal.

Penyusunan spesifikasi instrument tes terdiri atas:

3.1.1.1 Menentukan Tujuan Tes

Tujuan tes yang dikembangkan adalah mengukur kemampuan kognitif siswa pada level kognitif pengetahuan, aplikasi dan penalaran siswa pada materi makhluk hidup dan lingkungannya serta struktur dan fungsi makhluk hidup.

3.1.1.2 Menyusun Kisi-Kisi

Menyusun kisi-kisi dilakukan dengan mengisi komponen kisi-kisi yang telah ditentukan dan memasukkannya ke dalam suatu format kisi-kisi berupa matriks. Kisi-kisi soal yang digunakan dalam penelitian ini adalah kisi-kisi soal UN tahun pelajaran 2015/2016 yang telah diterbitkan oleh BSNP.

Setiap standar kompetensi lulusan dikembangkan menjadi 3 butir soal yang terdiri atas dimensi pengetahuan factual, konseptual dan prosedural.

3.1.1.3 Memilih Bentuk Instrumen Tes

Bentuk instrument tes kognitif standard yang dikembangkan adalah tes objektif berbentuk pilihan ganda (*multiple choice*) dengan empat pilihan jawaban.

3.1.1.4 Menentukan Panjang Instrumen Tes

Panjang instrument tes yang dikembangkan adalah 36 butir soal dengan alokasi waktu 72 menit.

Setelah penyusunan spesifikasi instrument tes dilanjutkan dengan perakitan soal. Setiap butir soal ditulis berdasarkan rumusan indicator yang sudah disusun dalam kisi-kisi dengan memperhatikan kaidah penulisan soal pilihan ganda yang baik dalam bentuk kartu soal. Pengembangan instrument tes kognitif pada tahap ini disebut produk awal (prototype 1).

3.1.2 Validasi Desain

Validasi desain dilakukan untuk mengetahui sejauh mana tingkat kelayakan produk instrument tes kognitif yang dikembangkan berdasarkan data yang diperoleh oleh ahli. Pakar yang ditunjuk meliputi 7 orang dosen (pakar substansi, konstruksi dan bahasa) dan 2 orang guru (pakar substansi). Hasil validasi ini kemudian dianalisis dan digunakan sebagai bahan pertimbangan dalam merevisi prototype 1. Soal yang valid dilanjutkan dan soal yang tidak valid masuk pada tahapan berikutnya.

Hasil validasi desain dari aspek substansi, konstruksi dan bahasa dapat dilihat pada table berikut.

Tabel 2. Nilai Validitas Isi Instrumen Tes Kognitif Mata P

elajaran IPA Kelas 7 SMP di Kabupaten Banjar

No Soal	Nilai Validitas Isi			Rata-rata	Kriteria
	Substansi	Konstruksi	Bahasa		
1.	84,00	100,00	85,33	89,78	sangat tinggi
2.	86,67	100,00	82,67	89,78	sangat tinggi
3.	93,33	100,00	86,67	93,33	sangat tinggi
4.	90,67	93,33	92,00	92,00	sangat tinggi
5.	84,00	94,67	93,33	90,67	sangat tinggi
6.	93,33	94,67	92,00	93,33	sangat tinggi
7.	86,67	93,33	93,33	91,11	sangat tinggi
8.	86,67	88,00	93,33	89,33	sangat tinggi
9.	86,67	88,00	84,00	86,22	sangat tinggi
10.	88,00	88,00	92,00	89,33	sangat tinggi
11.	81,33	85,33	90,67	85,78	sangat tinggi
12.	81,33	85,33	89,33	85,33	sangat tinggi
13.	86,67	86,67	92,00	88,45	sangat tinggi
14.	85,33	85,33	93,33	88,00	sangat tinggi
15.	93,33	86,67	93,33	91,11	sangat tinggi
16.	86,67	93,33	90,67	90,22	sangat tinggi
17.	73,33	100,00	92,00	88,44	sangat tinggi
18.	73,33	86,67	90,67	83,56	sangat tinggi
19.	86,67	93,33	92,00	90,67	sangat tinggi
20.	85,33	90,67	92,00	89,33	sangat tinggi
21.	86,67	86,67	92,00	88,45	sangat tinggi
22.	86,67	100,00	93,33	93,33	sangat tinggi
23.	86,67	100,00	93,33	93,33	sangat tinggi
24.	86,67	90,67	92,00	89,78	sangat tinggi
25.	85,33	86,67	90,67	87,56	sangat tinggi
26.	86,67	86,67	92,00	88,45	sangat tinggi
27.	93,33	93,33	86,67	91,11	sangat tinggi
28.	81,33	92,00	93,33	88,89	sangat tinggi
29.	93,33	93,33	88,00	91,55	sangat tinggi

No Soal	Nilai Validitas Isi			Rata-rata	Kriteria
	Substansi	Konstruksi	Bahasa		
30.	100,00	93,33	81,33	91,55	sangat tinggi
31.	86,67	100,00	93,33	93,33	sangat tinggi
32.	78,67	92,00	90,67	87,11	sangat tinggi
33.	78,67	93,33	92,00	88,00	sangat tinggi
34.	93,33	100,00	93,33	95,55	sangat tinggi
35.	85,33	93,33	85,33	88,00	sangat tinggi
36.	86,67	93,33	90,67	90,22	sangat tinggi
TE S	86,37	92,44	90,52	89,78	sangat tinggi

Tabel 2 menunjukkan bahwa analisis kualitatif soal dari aspek substansi, konstruksi dan bahasa memiliki nilai validitas isi yang sangat valid sesuai dengan criteria yang ditentukan oleh Azizah (2015).

3.1.3 Revisi Desain

Revisi dilakukan berdasarkan analisis hasil validasi desain serta kritik dan saran dari pakar. Hasil revisi ini disebut dengan prototipe 2.

Revisi berdasarkan kritik dan saran dari pakar baik ahli substansi (bidang studi biologi), ahli konstruksi dan ahli bahasa. Adapun revisi yang dilakukan pada masing-masing aspek tersebut dapat dilihat pada table berikut.

Tabel 3. Kritik dan Saran Ahli Substansi, Konstruksi dan Bahasa dalam Tahap Revisi Desain

No	Kritik dan Saran Ahli	Perbaikan
1.	Aspek Substansi	
	<ul style="list-style-type: none"> Sesuaikan antara rumusan butir soal dengan pembahasan. Pilihan jawaban sebaiknya sepadan Pada rumusan butir soal, setiap gambar diberikan keterangan 	<ul style="list-style-type: none"> Menyesuaikan antara rumusan butir soal dengan pembahasan. Menyetarakan pilihan jawaban Memberikan keterangan pada setiap gambar pada rumusan butir soal,
2.	Aspek Konstruksi	
	<ul style="list-style-type: none"> Gambar kurang jelas karena hitam putih Keterangan soal kurang jelas Pokok soal tidak jelas sehingga soal mengarah kepada jawaban 	<ul style="list-style-type: none"> Memperjelas gambar dengan gambar yang berwarna Memperjelas keterangan soal Memperjelas pokok soal sehingga soal tidak mengarah kepada jawaban



No	Kritik dan Saran Ahli	Perbaikan
3.	Aspek Bahasa <ul style="list-style-type: none"> Soal-soal yang EYD-nya belum baik perlu diperbaiki. 	Melakukan perbaikan terhadap soal sesuai dengan EYD yang baik dan benar.

Tabel 3 merupakan dasar yang dijadikan acuan dalam proses revisi butir soal yang akan digunakan pada tahapan selanjutnya.

3.1.4 Uji Coba Produk

Tahapan ini dimulai dengan uji keterbacaan kepada guru masing-masing sekolah sampel, simulasi penggunaan instrumen tes kognitif dan uji kelompok terbatas (uji kelas).

3.1.4.1 Uji Keterbacaan

Uji keterbacaan dilakukan kepada masing-masing guru sebanyak 1 orang dari 4 sekolah yang menjadi sampel penelitian. Hasil uji keterbacaan yang diberikan oleh guru dapat dilihat pada table berikut.

Tabel 4. Hasil Uji Keterbacaan Instrumen Tes Kognitif Standar dari guru.

No	Indikator	Rata-Rata	Kriteria
1.	Petunjuk dalam menjawab soal diberikan dengan runut dan jelas.	5,00	Sangat baik
2.	Setiap butir soal sesuai dengan standar kompetensi lulusan.	4,75	Sangat baik
3.	Penggunaan bahasa sesuai dengan EYD.	4,75	Sangat baik
4.	Kesederhanaan struktur kalimat soal.	4,75	Sangat baik
5.	Kejelasan gambar, tabel, grafik dan skema pada soal.	5,00	Sangat baik
6.	Kebenaran kunci jawaban	5,00	Sangat baik
RATA-RATA		4,875	Sangat baik

Tabel 4 menunjukkan bahwa rata-rata hasil uji keterbacaan guru sangat baik sehingga dapat dilanjutkan untuk simulasi kepada siswa.

Saran yang diberikan oleh guru terdapat pada soal nomor 6 yaitu gambar rangkaian percobaan yang kurang jelas, sehingga harus diberikan anak panah mengenai arah jalannya larutan eosin dan member warna merah pada larutan eosin.

3.1.4.2 Simulasi Uji Coba Produk

Simulasi uji coba produk ini dilakukan kepada 3 orang siswa dari masing-masing sekolah yang menjadi sampel penelitian dengan kemampuan penguasaan

mata pelajaran IPA yang berbeda yaitu siswa dengan kemampuan tertinggi, kemampuan sedang dan kemampuan terendah. Hasil simulasi uji coba produk kepada siswa dapat dilihat pada table berikut.

Tabel 5. Hasil Simulasi Uji Coba Produk pada Siswa

No	Nama Sekolah	Siswa	Jawaban Benar
1.	SMPN 3 Martapura	Atas	14
		Sedang	11
		bawah	10
2.	SMPN 1 Pengaron	Atas	26
		Sedang	21
		bawah	15
3.	SMPN 1 Paramasan	Atas	18
		Sedang	12
		bawah	9
4.	SMP Darussalam	Atas	25
		Sedang	17
		bawah	11

Tabel 5 menunjukkan bahwa hasil simulasi uji coba produk instrumen tes kognitif standar pada tiga orang siswa dengan kemampuan yang berbeda dapat menunjukkan kemampuan yang sebenarnya dilihat dari jumlah jawaban yang dapat diberikan dengan benar.

Pada tahapan ini terdapat satu sekolah yang bukan sampel karena SMPN 2 Sungai Pinang hanya memiliki satu rombongan belajar kelas 7 sehingga melakukan simulasi pada sekolah yang memiliki karakteristik yang sama yaitu sekolah sangat terpencil di Kabupaten Banjar yaitu SMPN 1 Paramasan. Jika simulasi dilakukan di sekolah SMPN 2 Sungai Pinang, maka tidak akan bias melakukan uji coba produk untuk satu kelas siswa.

3.1.4.3 Uji Coba Produk

Uji coba produk dilakukan untuk mengetahui analisis butir soal secara kuantitatif. Uji ini dilakukan pada siswa sebanyak satu kelas pada masing-masing sekolah yang menjadi sampel penelitian. Analisis butir secara kuantitatif ini dilakukan dengan menggunakan program *MicroCAT Bigsteps* Versi 2.30 yang menggunakan model Rasch pada data dikotomis.

Hasil analisis kualitatif menggunakan program *Bigsteps* pada instrumen tes kognitif standar yang dibuat dapat dilihat pada table 6.

Tabel 6. Output Tabel 3.1 tahap ketiga Analisis Instrumen Tes Kognitif IPA Kelas 7 SMP di Kabupaten Banjar.

SUMMARY OF 92 MEASURED (NON-EXTREME) PERSONS							
SCORE	COUNT	MEASURE	ERROR	MNSQ	INFIT	MNSQ	OUTFIT
MEAN	13,3	29,0	-.21	.43	1,00	.0	1,01
S.D.	3,9	.0	.72	.03	.16	.9	.27
RMSE	.43	ADJ.S.D.	.58	PERSON SEP	1,33	PERSON SEP REL.	.64

SUMMARY OF 29 MEASURED (NON-EXTREME) ITEMS							
SCORE	COUNT	MEASURE	ERROR	MNSQ	INFIT	MNSQ	OUTFIT
MEAN	42,1	92,0	.00	.25	1,00	.0	1,01
S.D.	19,1	.0	1,07	.03	.06	.7	.12
RMSE	.25	ADJ.S.D.	1,04	ITEM SEP	4,20	ITEM SEP REL.	.95

Tabel 5 menunjukkan bahwa terdapat 92 orang peserta uji coba dengan 36 item soal. Adapun yang dianalisa terdapat 67 orang dan 7 item dihilangkan (*drop*) karena tidak cocok dengan model pada tahap 1-3. Tabel *summary person*, rata-rata kemampuan peserta adalah -0,59 logit dengan rata-rata eror sebesar 0,45. Tabel *summary item* menunjukkan bahwa koefisien reliabilitasnya adalah 0,95 sehingga tes tersebut termasuk sangat andal dengan rata-rata kesalahan pengukuran 0,25. Standar eror ini lebih kecil dibandingkan dengan standar eror pada peserta tes (*person*).

3.2. Produk Pengembangan Instrumen Tes Standar Kognitif Mata Pelajaran Ipa Kelas 7 SMP Di Kabupaten Banjar

3.2.1 Instrumen Tes Kognitif Tes Standar Hasil Pengembangan

Produk yang dihasilkan dari penelitian pengembangan ini adalah soal-soal pilihan ganda dengan empat pilihan jawaban sebanyak 36 butir yang terdiri dari 12 Indikator Standar kompetensi Lulusan yang tercantum dalam kisi-kisi ujian nasional tahun pelajaran 2015/2016.

Masing-masing indikator tersebut dibagi menjadi 3 dimensi pengetahuan yaitu pengetahuan factual, konseptual dan procedural. Berdasarkan dimensi pengetahuan, instrumen tes standar yang telah dibuat terdiri atas 12 buah soal pengetahuan faktual, 12 buah soal pengetahuan konseptual dan 12 buah soal pengetahuan procedural, sedangkan berdasarkan dimensi kognitif Bloom revisi terdapat 9 buah soal C1, 3 buah soal C2, 15 buah soal C3, 3 buah soal C4, 3 buah soal C5 dan 3 buah soal C6.

Menurut kisi-kisi UN terdapat soal yang termasuk pemahaman sebanyak 12 buah, soal aplikasi sebanyak 15 buah dan soal penalaran sebanyak 9 buah.

3.2.2 Hasil Analisis Bigsteps Instrumen Tes Kognitif Mata Pelajaran IPA Kelas 7 SMP di Kabupaten Banjar

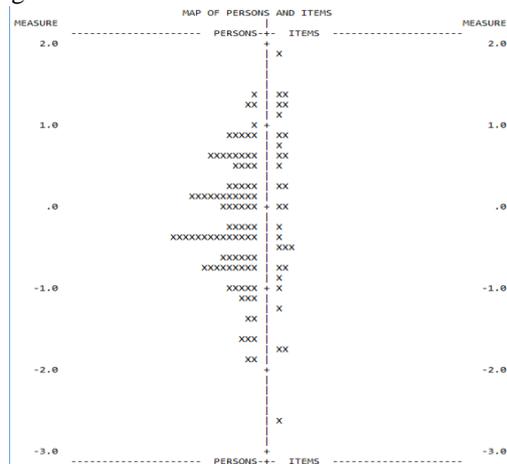
Berdasarkan hasil *output* tabel 6.1 dan 10.1 pada analisis Bigsteps pada tahap ketiga, diperoleh statistic masing-masing sebagaimana pada tabel 6 berikut.

Tabel 6. Ringkasan Analisis RUN III Instrume Tes Kognitif Standar Menggunakan Analisis Bigsteps.

Karakteristik	Statistik			
	min	maks	mean	SD
Skor mentah	6	22	13,30	3,90
Measure	-1,91	1,41	-0,21	0,72
SEM	0,41	0,54	0,43	0,03
Infit	-2,50	2,00	0,00	0,90
Outfit	-1,80	2,50	0,10	0,90

Berdasarkan hasil analisis RUN 1 sampai dengan RUN 3, masih terdapat peserta yang menjawab dengan tebakan sebanyak 2 orang yaitu peserta tes nomor 5 dan 45. Tidak terdapat item yang tidak konsisten sehingga analisis dapat diberhentikan sampai pada tahap 3. Terdapat 29 butir soal yang memenuhi criteria Rasch model untuk dijadikan soal yang baik dalam mengukur kemampuan siswa dari 36 butir soal yang diujikan. Dengan demikian terdapat 7 butir soal yang tidak baik yaitu butir soal nomor 34, 10, 9, 13, 17, 15 dan 31. Ditemukan juga peserta yang menjawab dengan tebakan sebanyak 3 orang dari 93 orang peserta yang ikut ujian. Indeks keandalan tes secara keseluruhan (30 item) pada mulanya sebesar 0,94 tetapi setelah dilakukan penghapusan butir yang tidak baik, maka indeks keandalan tes (29 item) menjadi 0,95. Standar eror pengukuran pada analisis dengan Bigsteps sebesar 4,20.

Output tabel 1.1 menunjukkan pola sebaran taraf sukar butir tes dengan sebaran kemampuan peserta tes. Berikut adalah hasil *output* tabel 1.1 Program Bigsteps pada instrumen tes standar mata pelajaran IPA Kelas 7 SMP di Kabupaten Banjar yang dapat dilihat pada gambar 1 berikut.



Gambar 1. Peta Persebaran Peserta Tes dan Item

Gambar di atas menunjukkan bahwa taraf sukar butir tes memiliki nilai sedikit lebih tinggi daripada



tingkat kemampuan peserta tes. Terdapat satu item yang lebih tinggi daripada kemampuan peserta tes tertinggi. Item tersebut adalah butir soal nomor 30 yang dapat dilihat pada *Map of Item* pada *Output table 12.1* yang merupakan butir soal tersulit dan butir soal termudah adalah butir soal nomor 4. Berdasarkan *Map Of Person* pada *Output table 16.1* menunjukkan bahwa peserta yang memiliki kemampuan tertinggi sebanyak 2 orang yaitu peserta nomor 37 dan peserta yang memiliki kemampuan terendah yaitu peserta nomor 23 dan 40.

Gambar 1 juga menunjukkan bahwa tidak ada butir soal yang memiliki nilai measure di atas 2 dan hanya satu soal yang memiliki nilai measure kurang dari -2 yaitu soal nomor 4. Hal ini menunjukkan bahwa butir soal nomor 4 termasuk soal yang sangat mudah dan banyak dipahami oleh peserta tes.

Output tabel 13.2 memberikan informasi tentang sebaran jawaban yang dipilih oleh peserta tes. Pengecoh yang tidak berfungsi dengan baik yakni yang dipilih kurang dari 5% dari peserta tes.

Tabel 8. *Output* tabel 13.2 Analisis Bigsteps Instrumen Tes Kognitif Mata Pelajaran IPA kelas 7 SMP di Kabupaten Banjar.

TABLE 13.2 IPA KTSP KELAS 7 Jun 29 15:20 2016 INPUT: 93 PERSONS 36 ITEMS ANALYZED: 92 PERSONS 29 ITEMS 2 CATEGORIES "BIGSTEPS" RASCH ANALYSIS VER. 2.30

ITEMS	OPTION	FREQUENCIES	MEASURE	ORDER									
NUM	COUNT	MISSING	% SCR	D % SCR	A % SCR	B % SCR	C % SCR						
30	92	0	0	**	22	23	0	18	19	0	12	13	1
28	92	40	43	0	37	40	0	22	23	0	16	17	0
6	92	17	18	1	34	36	0	20	21	0	20	21	0
11	92	0	0	**	39	42	0	19	20	0	14	15	0
12	92	20	21	1	48	52	0	7	7	0	17	18	0
36	92	0	0	**	13	14	0	43	46	0	22	23	1
22	92	0	0	**	19	20	0	16	17	0	32	34	0
27	92	25	27	1	28	30	0	26	28	1	26	28	0
18	92	12	13	0	17	18	0	11	11	0	37	40	0
32	92	27	29	1	16	17	0	13	14	0	29	31	1
33	92	0	0	**	15	16	0	40	43	0	29	31	1
8	92	0	0	**	13	14	0	33	35	1	11	11	0
26	92	35	38	0	27	29	0	36	39	1	22	23	0
7	92	7	7	0	31	33	0	38	41	1	9	9	0
2	92	14	15	0	18	19	0	22	23	0	41	44	1
35	92	11	11	0	16	17	0	42	45	1	14	15	0
21	92	20	21	0	48	52	1	28	30	0	10	10	0
24	92	6	6	0	20	21	0	49	53	1	10	10	0
14	92	0	0	**	15	16	0	51	55	1	11	11	0
16	92	15	16	0	52	56	1	13	14	0	19	20	0
		8	8	0									

TABLE CONTINUED ON NEXT SHEET
 CONTINUED FROM PREVIOUS SHEET

ITEMS	OPTION	FREQUENCIES	MEASURE	ORDER									
NUM	COUNT	MISSING	% SCR	D % SCR	A % SCR	B % SCR	C % SCR						
25	92	0	0	**	13	14	0	13	14	0	13	14	0
3	92	53	57	1	15	16	0	10	10	0	11	11	0
1	92	56	60	1	19	20	0	8	8	0	58	63	1
20	92	7	7	0	14	15	0	12	13	0	59	64	1
29	92	7	7	0	12	13	0	61	66	1	14	15	0
5	92	5	5	0	67	72	1	2	2	0	9	9	0
19	92	14	15	0	12	13	0	1	1	0	74	80	1
23	92	5	5	0	8	8	0	74	80	1	5	5	0
4	92	5	5	0	2	2	0	4	4	0	83	90	1
		3	3	0									

diketahui bahwa pengecoh yang tidak baik terdapat pada butir soal

nomor 5 (B), 19(B) dan 4 (A, B, dan D). dari uraian tersebut terdapat 3 butir soal yang harus direvisi.

Data yang tidak fit dari hasil analisis tes standar kognitif mata pelajaran IPA kelas 7 SMP di Kabupaten Banjar menggunakan program Bigsteps, diperoleh data seperti pada tabel 8 berikut.

Tabel 9. Kecocokan Butir Soal dengan Model Rasch

Statistik	Nomor Butir Soal	Jumlah
Fit	1, 2, 3, 4, 5, 6, 7, 8, 11, 12, 14, 16, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 32, 33, 35, 36	29
Tidak fit	34, 10, 9, 13, 17, 15, 31	7
Jumlah soal seluruhnya		36

Tabel 9 menunjukkan bahwa terdapat 7 butir soal yang tidak fit dengan model. Soal fit dengan model yaitu soal yang mempunyai nilai outfit $-2,00 \leq \text{fit} \leq 2,00$. Goodness of fit instrumen tes kognitif standar IPA kelas 7 SMP di Kabupaten Banjar adalah 19,44% yang tidak fit dengan model dan 80,56% yang fit dengan model.

Data hasil estimasi nilai taraf sukar butir soal sesuai dengan kategori taraf sukar butir menurut Rogayah (2010) dapat dilihat pada tabel 10 berikut.

Tabel 10. Hasil Estimasi Nilai Taraf Sukar Butir

Statistik	Nomor Soal	Jlh	%	Kategori
$\delta_i < -2,00$	4	1	2,78%	Sangat mudah
$-2,00 < \delta_j < -1,00$	23, 19, 5	3	8,33%	Mudah
$-1,00 < \delta_j < +1,00$	29, 20, 1, 3, 25, 16, 14, 24, 21, 35, 2, 7, 26, 8, 33, 32, 18, 27, 22	19	52,78%	Sedang
$+1,00 < \delta_j < +2,00$	36, 12, 11, 6, 28, 30	6	16,67%	Sukar
$\delta_j > +2,00$	-	-	-	Sangat sukar

Tabel 10 menunjukkan bahwa dari hasil estimasi nilai taraf sukar butir soal terdapat 77, 78% butir soal yang baik dari 36 butir soal yang diujikan yaitu yang cocok dengan model dan memiliki taraf sukar yang nilainya berkisar antara $-2,00 \leq \text{fit} \leq 2,00$ (Rogayah, 2010).

Berdasarkan ketentuan dari Prilantari (2007) tentang kualitas butir soal dalam pendekatan teori respon butir model logistik satu parameter, hasil analisis menunjukkan bahwa terdapat 28 butir soal (23, 19, 5, 29, 20, 1, 3, 25, 16, 14, 24, 21, 35, 2, 7, 26, 8, 33, 32, 18, 27, 22, 36, 12, 11, 6, 28, dan 30) yang memiliki kualitas baik, 1 butir soal (4) yang memiliki kualitas cukup baik, dan 7 butir soal (34, 10, 9, 13, 17, 15, dan 31) yang memiliki kualitas tidak baik.



4. KESIMPULAN DAN SARAN

Proses pengembangan tes yang dilakukan adalah desain produk, validasi desain oleh 9 orang pakar yang memiliki nilai validitas sangat valid, dan uji coba produk yang sangat baik bagi guru maupun siswa. Terdapat 7 butir soal yang tidak cocok dengan model melalui analisis kuantitatif dengan program Bigsteps dari 36 butir soal yang diujikan yaitu soal nomor 34, 10, 9, 13, 17, 15 dan 31. Terdapat 3 butir soal yang harus direvisi dalam hal pilihan jawaban yaitu butir soal nomor 4, 5 dan 19. Ditemukan 3 orang peserta yang menjawab dengan menebak dan terdapat 77,78% atau 28 butir soal yang memiliki kualitas baik untuk mengukur kemampuan siswa.

5. UCAPAN TERIMA KASIH

Ucapan terima kasih disampaikan kepada Dosen pembimbing, Program Studi Magister Pendidikan Biologi Universitas Lambung Mangkurat, Dinas Pendidikan Kabupaten Banjar Kalimantan Selatan, Kepala Sekolah, Guru dan Siswa yang sangat membantu dalam pelaksanaan penelitian ini.

6. DAFTAR PUSTAKA

- Hambleton, R.K and R.W. Jones. 1996. *Comparison of Classical Test Theory and Item Response Theory and Their Applications to Test Development*. ITEMS (Instructional Topics in Educational Measurement).
- Mamu, H.D. 2014. Pengaruh Strategi Pembelajaran, Kemampuan Akademik dan Interaksinya terhadap Keterampilan Berpikir Kritis dan Hasil Belajar Kognitif IPA Biologi. *Jurnal Pendidikan Sains* 2 (1): 1-11.
- Mardapi, D; Kumaidi; Kartowagiran,B. 2011. Pengembangan Instrumen Pengukur Hasil Belajar Nirbias dan Terskala Baku. *Jurnal Penelitian dan Evaluasi Pendidikan*15(2): 326-341.
- Mendikbud. 2013. PERMENDIKBUD RI No. 54 tentang Standar Kompetensi Lulusan.
- Mendikbud. 2013. PERMENDIKBUD RI No. 66 tentang Standar Penilaian Pendidikan.
- Mendiknas. 2006. PERMENDIKNAS RI No. 20 tentang Standar Kompetensi Lulusan.
- Mendiknas. 2007. PERMENDIKNAS RI No. 20 tentang Standar Penilaian Pendidikan.
- Pusat Penilaian Pendidikan Balitbang-Depdiknas. 2007. *Panduan Penulisan Soal Pilihan Ganda*. Depdiknas. Jakarta.
- Prilantari, H. 2007. Pengembangan Bank Soal Keterampilan Proses IPA (Biologi) Kelas X Tingkat Madrasah Aliyah. *Jurnal Penelitian dan Evaluasi Pendidikan* 2(10):223-234.
- Rahayu, D. dan U. Azizah. 2012. Pengembangan Instrumen Penilaian Kognitif Berbasis Komputer dengan Kombinasi Permainan “Who Wants To Be A Chemist” pada Materi Pokok Struktur Atom untuk Kelas X SMA RSBI. *Prosiding Seminar Nasional Kimia UNESA*. Hal.41-50.

- Rogayah, S. Dan Ekaria. 2010. Evaluasi Taraf Sukar Butir Tes Matematika USM PMB STIS Tahun 2007/2008 dan Tahun 2008/2009 dengan Model Rasch. *Jurnal Aplikasi Statistika dan Komputasi Statistik, UPPM-STIS* 2(1):76-91
- Scheiblechner, H.H. 2009. Rasch and pseudo-Rasch Models: Suitableness for Practical Test and Applications. *Psychology Science Quarterly* 51 (2): 181-194.
- Sudijono, A. 2015. *Pengantar Evaluasi Pendidikan*. Rajawali Pers. Jakarta.
- Sugiyono. 2015. *Metode Penelitian Pendidikan (Pendekatan Kuantitatif, Kualitatif dan R&D)*. Alfabeta. Bandung.
- Sukor, N.S., K.Osman, T.M.T. Soh. 2013. Chemistry Test Item Development: Assessing Conceptual Understanding among Malaysian Students. *Asian Social Science* 9 (16): 126-132.
- Surapranata, S. 2004. *Panduan Penilaian Tes Tertulis: Implementasi Kurikulum 2004*. PT. Remaja Rosdakarya Offset. Bandung.
- Suwarto. 2009. Pengembangan Tes dan Analisis Hasi Tes yang Terintegrasi dalam Program Komputer. *Jurnal Penelitian dan Evaluasi Pendidikan* 12 (1): 40-56.
- Warso, A.W.D.D.W. 2014. *Standar Kompetensi Lulusan (SKL) dan Struktur Kurikulum SMP/Mts*. Graha Cendikia. Yogyakarta.
- Widyantoro, D; Boenasir; Karsono. 2009. Pengembangan Soal Tes Pilihan Ganda Kompetensi Sistem Starter dan Pengisian Program Keahlian Teknik Mekanik Otomotif Kelas XII. *Jurnal PTM* 9 (1): 14-21.

Penanya: Marheny Lukitasari (IKIP PGRI Madiun)

Pertanyaan:

1. Apakah instrument tes yang dibuat akan valid lagi jika digunakan pada tahun berikutnya?
2. Apakah soal yang dikembangkan bersifat subjektif?
3. Apakah C1-C6 digunakan dalam penelitian ini?

Jawaban:

1. Penelitian yang dilakukan tidak sampai pada tahap uji pemakaian, sehingga valid yang dimaksud dalam penelitian ini adalah valid secara teoritik dari para ahli (substansi, konstruksi, dan Bahasa). Instrumen tes ini akan standar jika dilakukan tahapan penelitian lebih lanjut yaitu uji coba pemakaian yang lebih luas dan diujikan juga pada sekolah yang menggunakan kurikulum 2013.
2. Soal yang diujikan berupa soal pilihan ganda (objektif) dengan 4 pilihan jawaban. Siswa tidak diminta untuk memberikan alasan pemilihan jawaban. Meskipun begitu, melalui analisis Bigsteps siswa yang menjawab dengan menebakakan teridentifikasi.
3. Instrumen yang dibuat menggunakan taksonomi Bloom revisi ulang yang terdiri atas dimensi pengetahuan (12 soal factual, 12 soal konseptual, dan 12 soal procedural) dan dimensi kognitif C1-C6 (9 soal C1, 3 Soal C2, 15 soal C4, 3 soal C5, dan 3 soal C6)

