

PENGEMBANGAN INSTRUMEN PENILAIAN UNTUK MENGUKUR KETERAMPILAN PROSES SAINS SISWA SMA

DEVELOPMENT OF ASSESSMENT INSTRUMENTS TO MEASURE THE SCIENCE PROCESS SKILLS OF HIGH SCHOOL STUDENTS

Etty Nurmala Fadillah

FKIP Universitas Muhammadiyah Palembang

Jalan Ahmad Yani, 13 Ulu Palembang, Sumatera Selatan

ettynurmalafadillah@gmail.com; ettynurmalafadillah@um-palembang.ac.id

Diterima: September 2017; Disetujui: November 2017; Diterbitkan: November 2017

Abstrak

Penelitian pengembangan ini bertujuan untuk mengetahui: (1) kelayakan instrumen penilaian untuk mengukur keterampilan proses sains siswa SMA ditinjau dari karakteristik standar tes dan (2) seberapa besar tingkat penguasaan instrumen penilaian untuk mengukur keterampilan proses sains siswa SMA. Prosedur pengembangan dalam penelitian ini dengan model 4-D yang dikemukakan oleh Thiagarajan yaitu *define* (pendefinisian), *design* (perancangan), *develop* (pengembangan), dan *disseminate* (penyebarluasan). Instrumen pengumpulan data yang digunakan berupa angket, wawancara, dan soal pilihan ganda beralasan. Teknik analisis data menggunakan analisis kualitatif dengan menggunakan statistik Aiken's V dan analisis kuantitatif dengan menggunakan program QUEST. Hasil penelitian adalah sebagai berikut. (1) Instrumen penilaian untuk mengukur keterampilan proses sains siswa SMA dikatakan layak ditinjau dari karakteristik standar tes yaitu kecocokkan butir soal dengan INFIT MNSQ 0,86 sampai 1,29, Reliabilitas sebesar 0,80, dan tingkat kesukaran dengan jangkauan -1,47 sampai 1,59 terdiri dari 3 kategori: 11 butir soal mudah, 7 butir soal sedang, dan 5 butir soal Sukar. (2) Persentase tingkat penguasaan keterampilan proses sains siswa pada sekolah dengan kategori tinggi 76,64%, kategori sedang 73,71% dan rendah 70,12%.

Kata kunci: instrumen penilaian, keterampilan proses sains

Abstract

This development research aimed to determine: (1) the feasibility of assessment instrument to measure the science process skills of high school students in terms of test standard characteristics and (2) how the level of assessment instrument mastery to measure the science process skills of high school students. The development procedure used in this study was by using 4-D model that was found by Thiagarajan; define, design, develop, and disseminate. The instruments used in collecting the data were questionnaires, interview, and multiple choice questions reasoned. The technique for analyzing the data used in this study was qualitative analysis by using Aiken's V statistics and quantitative analysis by using QUEST program. The results of this study were (1) assessment instrument to measure the process skills of high school students where could be categorized as effective from test standard characteristic namely the matching of question items with INFIT MNSQ was 0.86 until 1.29. The reliability was 0.80 and the difficulty level was -1.47 until 1.59 consisted of 3 categories: 11 items of easy question, 7 items of medium questions, and 5 questions of difficult question. (2) The percentage of assessment instrument mastery level science process skills of high school students in high level was 76.64%, medium category was 73.71%, and poor level was 70.12%.

Keywords: *assessment instrument, science process skills*

©Didaktika Biologi: Jurnal Penelitian Pendidikan Biologi

p-ISSN 2549-5267

e-ISSN 2579-7352

Pendahuluan

Pembangunan suatu negara merupakan salah satu aspek penting tujuan dari adanya pendidikan. Berdasarkan hasil dari *Program for International Student Assessment (PISA)* di tahun 2012 dan hasil *The Learning Curve*

Pearson 2014 yang menggambarkan indeks global kemampuan kognitif, Indonesia mendapatkan peringkat 64 dari 65 negara anggota PISA dengan skor literasi sains sebesar 382 dan berada pada posisi terendah yaitu peringkat 40 dari keseluruhan negara

anggotanya pada *The Learning Curve Pearson 2014*. Hal tersebut menunjukkan bahwa kualitas pendidikan di Indonesia masih rendah jika dibandingkan dengan negara-negara lainnya. Salah satu yang mempengaruhi kualitas pendidikan di Indonesia adalah guru sebagai pendidik.

Peranan guru sebagai pendidik terdapat pada Undang-Undang No. 20 Tahun 2003 tentang Sistem Pendidikan Nasional pasal 1 ayat 20 yang menyatakan bahwa pembelajaran adalah proses interaksi peserta didik dengan pendidik dan sumber belajar pada suatu lingkungan belajar. Pembelajaran merupakan usaha sengaja, terarah dan bertujuan agar orang lain dapat memperoleh pengalaman yang bermakna. Hamalik (2013) menyatakan bahwa pembelajaran merupakan suatu proses dari suatu kegiatan dan bukan hanya sekedar hasil atau tujuan. Hamalik melanjutkan bahwa belajar bukan hanya mengingat, akan tetapi lebih luas dari itu, yakni perubahan tingkah laku. William Burton (Hamalik, 2013) menambahkan “*a good learning situation consist of a rich and varied series of learning experiences unified around a vitorous purpose and carried on in interaction with a rich, varied and propocative environment.*” Berdasarkan beberapa definisi tersebut, pembelajaran dapat didefinisikan sebagai proses interaksi yang disengaja oleh peserta didik, pendidik, sumber belajar serta lingkungan belajar untuk memperoleh serangkaian pengalaman belajar yang bermakna.

Pembelajaran biologi pada hakikatnya merupakan suatu proses untuk menghantarkan siswa ke tujuan belajarnya dan biologi itu sendiri berperan sebagai alat untuk mencapai tujuan tersebut. Biologi sebagai ilmu dapat diidentifikasi melalui objek, benda alam, persoalan/gejala yang ditunjukkan oleh alam, serta proses keilmuan dalam menemukan konsep-konsep biologi. Proses pembelajaran biologi merupakan penciptaan situasi dan kondisi yang kondusif sehingga terjadi interaksi antara subjek didik dengan objek belajarnya yang berupa makhluk hidup dan segala aspek kehidupannya. Melalui interaksi antara subjek didik dengan objek belajar dapat menyebabkan perkembangan proses mental dan sensori motorik yang optimal pada diri siswa. Berdasarkan KTSP (BSNP, 2006),

mata pelajaran biologi dikembangkan melalui kemampuan berpikir analitis, induktif dan deduktif untuk menyelesaikan masalah yang berkaitan dengan peristiwa alam sekitar dan penyelesaian masalah bersifat kualitatif dan kuantitatif dilakukan dengan menggunakan pemahaman dalam bidang lainnya. Salah satu indikator untuk melihat keberhasilan pembelajaran biologi adalah dengan hasil penilaian.

Istilah penilaian merupakan alih bahasa dari istilah *assessment*, bukan dari istilah *evaluation*. Depdikbud (Arifin, 2014) mengemukakan penilaian adalah suatu kegiatan untuk memberikan berbagai informasi secara berkesinambungan dan menyeluruh tentang proses dan hasil yang telah dicapai siswa. Kata “menyeluruh” mengandung arti bahwa penilaian mencakup aspek pengetahuan, keterampilan, sikap, dan nilai-nilai. Selanjutnya, Gronlund (Arifin, 2014) mengartikan penilaian sebagai suatu proses yang sistematis dari pengumpulan, analisis, dan interpretasi informasi/data untuk menentukan sejauh mana peserta didik telah mencapai tujuan pembelajaran. Dengan demikian dapat disimpulkan bahwa penilaian merupakan suatu kegiatan untuk memberikan informasi yang sistematis dan berkesinambungan tentang proses dan hasil belajar peserta didik dalam rangka membuat keputusan-keputusan berdasarkan kriteria dan pertimbangan tertentu.

Griffin dan Nix (Kurniawan & Mutaqimah, 2009) mengatakan bahwa penilaian adalah suatu pernyataan yang didasarkan pada sejumlah fakta untuk menjelaskan karakteristik seseorang atau sesuatu. Menurut Reynolds, Livingston & Willson (2010), penilaian adalah beberapa prosedur yang bersifat sistematis untuk mengumpulkan informasi dan digunakan dalam membuat kesimpulan mengenai karakteristik seseorang atau objek, sedangkan menurut Sudjiono (2011), penilaian berarti menilai sesuatu dan menilai berarti mengambil keputusan atas suatu hal yang berdasar akan baik atau buruk, pandai atau bodoh, dan lain sebagainya. Berdasarkan beberapa definisi tersebut, penilaian dapat diartikan sebagai semua aktifitas yang dilakukan oleh pendidik dan peserta didik untuk menilai diri mereka sendiri, yang memberikan informasi sebagai umpan balik

untuk memodifikasi aktifitas belajar dan mengajar yang lebih baik.

Hasil penilaian tidak identik dengan hasil kognitif saja, namun juga keterampilan-keterampilan yang menunjang pembelajaran dari siswa. Salah satu keterampilan yang dapat dinilai dari aktifitas siswa adalah keterampilan proses. Keterampilan proses sains sangat penting dikuasai siswa karena dibandingkan dengan keterampilan lainnya, keterampilan proses sains dapat lebih melihat perkembangan kemampuan peserta didik dalam melakukan suatu aktivitas untuk menemukan suatu konsep. Dengan menemukan sendiri konsep tersebut, siswa dapat lebih lama mengingatnya dibandingkan dengan menghafal.

Chiappetta & Koballa (2010) menyatakan bahwa keterampilan proses sains (*science process skill*) dibedakan menjadi dua bagian, yaitu keterampilan proses sains dasar (*basic science process skill*) dan keterampilan proses sains yang terintegrasi (*integrated science process skill*). Keterampilan proses sains dasar mencakup: (1) mengobservasi (*observing*); (2) mengklasifikasikan (*classifying*); (3) hubungan ruang/waktu (*space/time relation*); (4) menggunakan bilangan (*using number*); (5) melakukan pengukuran (*measuring*); (6) menginferensi (*inferring*); dan (7) memprediksi (*predicting*).

Adapun keterampilan proses sains terintegrasi menurut Bundu (2006) meliputi (1) mengidentifikasi variabel, (2) menyusun tabel data, (3) menyusun grafik (4) menggambarkan hubungan antar variabel, (5) memperoleh dan memproses data, (6) menganalisis investigasi, (7) menyusun hipotesis, (8) merumuskan variabel secara operasional, (9) merancang investigasi, dan (10) melakukan eksperimen

Pembelajaran biologi menekankan adanya interaksi antara subjek dan objek yang dipelajari. Pembelajaran biologi di Sekolah Menengah Atas (SMA) diharapkan dapat menjadi wahana bagi peserta didik untuk mempelajari diri sendiri dan alam sekitar, serta prospek pengembangan lebih lanjut dalam kehidupan sehari-hari. Salah satu materi yang terdapat pada SMA kelas X semester genap adalah materi ekosistem. Materi ekosistem dipilih karena pada materi ini terdapat banyak permasalahan terkait

dengan lingkungan. Permasalahan-permasalahan ini sangat bagus untuk mengembangkan keterampilan proses sains siswa. Keterampilan proses sains dapat diperoleh dari aktivitas siswa mulai dari pengamatan sampai membuat kesimpulan.

Berdasarkan studi awal yang telah dilakukan pada beberapa guru biologi di provinsi Daerah Istimewa Yogyakarta (DIY), guru mengalami kesulitan menilai keterampilan proses sains dari siswa. *Performance test* yang diketahui dan guru gunakan sangat sulit untuk diterapkan dengan kondisi jumlah siswa yang cukup banyak dalam satu kelas, sehingga aktivitas siswa tidak dapat diamati satu persatu. Tes penilaian tertulis merupakan salah satu alternatif permasalahan ini. Penggunaan instrumen ini akan membantu guru dalam menilai keterampilan proses tanpa melakukan observasi atau metode diskusi yang selama ini digunakan. Hasil studi awal yang lainnya adalah belum tersedianya instrumen penilaian biologi untuk mengukur keterampilan proses dalam bentuk tes tertulis.

Agar hasil dari penilaian tersebut benar-benar menilai keterampilan siswa, hasil penilaian haruslah mengungkapkan informasi secara lengkap dan sesuai dengan data yang diperlukan. Untuk mendapatkan hasil penilaian yang sesuai, seharusnya juga digunakan instrumen penilaian yang tepat. Menurut Arifin (2014), salah satu karakteristik instrumen yang baik adalah bersifat relevan. Relevan berarti instrumen yang digunakan haruslah sesuai antara materi yang diajarkan dengan konteks penilaian hasil belajar. Kegiatan penilaian yang dilakukan oleh guru dapat dilakukan dengan berbagai cara, salah satunya dalam bentuk *paper and pencil test*. Namun, selama ini *paper and pencil test* yang digunakan guru hanya mengukur kemampuan kognitif saja dan kurang memperhatikan pada penilaian keterampilan proses siswa.

Kegiatan penilaian tidak dapat dipisahkan dengan kegiatan pembelajaran. Menurut Sudjana (2005), tujuan dari adanya penilaian adalah (a) mendeskripsikan kecakapan belajar para siswa, (b) mengetahui keberhasilan proses pendidikan dan pengajaran di sekolah, (c) menentukan tindak lanjut hasil penilaian, (d) memberikan pertanggungjawaban dari pihak sekolah

kepada pihak-pihak yang berkepentingan. Tujuan penilaian inilah yang harus dicapai dalam menilai keterampilan-keterampilan dalam pembelajaran biologi. Untuk itu perlu adanya instrumen penilaian yang digunakan oleh guru untuk menilai keterampilan siswa. Siswa SMA yang mendapatkan pembelajaran biologi sudah seharusnya menggunakan instrumen penilaian yang bukan hanya mengukur keterampilan kognitif saja, namun meliputi keterampilan proses sains. Oleh karena itu, penelitian ini mengembangkan instrumen penilaian dalam bentuk tes tertulis pada mata pelajaran biologi untuk mengukur keterampilan proses sains siswa SMA kelas X.

Indikator keterampilan proses sains yang diukur di dalam penelitian ini merupakan sintesis dari indikator keterampilan proses sains dasar yang disampaikan oleh Chiappetta & Koballa (2010) dan keterampilan proses sains terintegrasi yang disampaikan oleh Bundu (2006), sehingga indikator keterampilan proses sains yang diukur meliputi: (1) mengamati, (2) mengelompokkan, (3) menafsirkan, (4) meramalkan, (5) merumuskan hipotesis, (6) merencanakan percobaan, (7) berkomunikasi.

Adapun rumusan masalah dari penelitian ini adalah (1) apakah instrumen penilaian mata pelajaran biologi yang dikembangkan layak atau berkualitas menurut aspek-aspek penilaiannya? dan (2) seberapa besar tingkat keterpakaian instrumen penilaian untuk mengukur keterampilan proses sains siswa SMA kelas X pada mata pelajaran biologi materi ekosistem?

Berdasarkan perumusan masalah, tujuan penelitian yang akan dicapai dalam penelitian ini adalah mengetahui: (1) kelayakan instrumen penilaian untuk mengukur keterampilan proses sains siswa SMA kelas X pada mata pelajaran biologi materi ekosistem; dan (2) seberapa besar tingkat keterpakaian instrumen penilaian untuk mengukur keterampilan proses sains siswa SMA kelas X pada mata pelajaran biologi materi ekosistem.

Metode Penelitian

Jenis Penelitian

Penelitian ini merupakan *Research and Development (R&D)*. Tujuan dari penelitian ini yaitu mengembangkan instrumen penilaian biologi untuk mengukur keterampilan proses siswa SMA kelas X pada materi ekosistem. Model pengembangan instrumen yang digunakan dalam penelitian ini adalah model pengembangan 4-D yang dikemukakan oleh Thiagarajan (1974), yaitu *define, design, develop, dan disseminate*.

Target/Subjek Penelitian

Subjek dari penelitian ini adalah siswa SMA kelas X semester genap dari empat SMA Negeri di Kabupaten Sleman. Pengambilan SMA Negeri ini menggunakan teknik *purposive sampling* yaitu sekolah yang telah menjalankan proses pembelajaran menggunakan Kurikulum 2013 dan dalam pembelajaran biologi telah menerapkan keterampilan proses sains. Pemilihan sekolah berdasarkan kriteria kelulusan Ujian Nasional tahun 2014 yaitu pada tingkat tinggi, sedang, dan rendah.

Uji coba terbatas dilakukan di SMA Negeri 2 Ngaglik yang terdiri atas empat kelas X MIA dengan jumlah total subjek 121 orang peserta didik, sedangkan uji coba luas dilakukan di tiga SMA yaitu SMA Negeri 1 Godean, SMA Negeri 1 Prambanan, dan SMA Negeri 1 Seyegan. Distribusi subjek penelitian uji coba luas yang digunakan dapat dilihat dalam Tabel 1.

Tabel 1. Distribusi Subjek Penelitian Uji Coba Luas

No.	SMA	Nilai UN Biologi	Sampel
1	SMAN 1 Godean	7,40	63
2	SMAN 1 Prambanan	6,36	60
3	SMAN 1 Seyegan	5,55	64
Jumlah peserta didik			187

Prosedur

Penelitian ini merupakan *Research and Development (R&D)* yang menggunakan model pengembangan 4-D. Tujuan dari penelitian ini adalah menghasilkan instrumen penilaian biologi yang dapat mengukur keterampilan proses sains siswa SMA kelas X pada materi ekosistem. Adapun prosedur pengembangan instrumen penilaian biologi

ini meliputi empat tahap, yaitu: *Define, Design, Develop, dan Disseminate*.

1. *Define*

Tahap ini bertujuan untuk menetapkan dan mendefinisikan kebutuhan dalam penilaian keterampilan siswa. Tahapan ini dilakukan dengan studi wawancara kepada guru biologi SMA yang mengajar di kelas X. Hasil yang diperoleh menggambarkan bahwa teknik observasi yang dilakukan guru selama ini untuk menilai keterampilan siswa dinilai cukup sulit oleh guru karena guru tidak mampu mengamati aktivitas siswa satu persatu. Oleh karena itu perlu dikembangkan suatu instrumen penilaian biologi yang dapat mengukur kognitif siswa sekaligus keterampilan proses sains siswa.

2. *Design*

Tahap ini berisi perancangan model dan prosedur pengembangan. Pada tahapan ini dilakukan perancangan instrumen penilaian biologi. Tahapan ini berupa menentukan KI, KD, materi, menjabarkan indikator dari keterampilan proses sains, merancang kisi-kisi penyusunan soal, serta naskah soal dan rubrik penilaiannya.

3. *Develop*

Tahap ini merupakan tahap pengembangan instrumen sebagai produk. Produk yang dikembangkan berupa kisi-kisi penyusunan soal, naskah soal pilihan ganda yang disertai alasan memberi jawaban, dan rubrik penilaian. Pada tahapan ini dilakukan validasi produk oleh validator, yang meliputi *peer reviewer* (teman sejawat), dosen ahli yang terdiri dari dosen ahli evaluasi terutama di bidang pengukuran dan evaluasi pendidikan, serta dosen ahli materi lingkungan, serta guru sebagai pelaksana pembelajaran. Setelah dilakukan validasi oleh *peer reviewer*, produk direvisi berdasarkan masukan *peer reviewer* (revisi I), kemudian divalidasi oleh dosen ahli evaluasi pendidikan dan ahli materi lingkungan, revisi produk berdasarkan hasil validasi (revisi II), selanjutnya divalidasi oleh guru, revisi produk berdasarkan hasil validasi guru (revisi III), dilanjutkan uji coba terbatas, revisi produk berdasarkan uji coba terbatas (revisi IV), serta uji coba luas.

4. *Disseminate*

Tahapan *disseminate* merupakan tahapan penyebarluasan produk hasil akhir pengembangan ke seluruh populasi, yaitu setelah melewati tahap *Define, Design, dan Develop*.

Data, Instrumen dan Teknik Pengumpulan Data

Teknik pengumpulan data berupa tes dan non tes. Teknik non tes dengan melalui wawancara dan angket, sedangkan teknik tes dengan menggunakan produk instrumen penilaian.

Angket yang digunakan yakni angket penilaian soal oleh dosen ahli evaluasi, dosen ahli materi, dan guru biologi kelas X. Angket ini digunakan untuk mendapatkan data tentang kelayakan instrumen penilaian. Angket untuk dosen ahli evaluasi, dosen ahli materi dan guru biologi kelas X merupakan angket tertutup.

Wawancara digunakan untuk mengetahui informasi cara penilaian guru terhadap keterampilan proses sains siswa yang dilakukan selama ini. Tujuan dari wawancara adalah untuk mengetahui masalah guru dalam mengukur keterampilan proses sains siswa. Hasil wawancara ini dijadikan *need assessment* sebagai langkah awal penelitian.

Produk instrumen penilaian yang dikembangkan adalah naskah soal dalam bentuk soal pilihan ganda beralasan beserta rubrik penilaian dan kisi-kisinya. Produk ini divalidasi secara teori terlebih dahulu oleh enam validator, kemudian divalidasi secara empiris dalam uji coba terbatas. Tahapan selanjutnya produk soal ini digunakan untuk mengukur ketercapaian keterampilan proses sains siswa SMA.

Teknik Analisis Data

Draf awal yang divalidasi menghasilkan data kualitatif. Data tersebut berupa masukan dan saran pada instrumen penilaian yang berasal dari validator. Masukan dan saran dari validator dipadukan dan digunakan untuk merevisi draf awal yang sudah diberikan. Hasil penilaian dari ahli materi, ahli evaluasi, dan guru biologi dianalisis dengan menggunakan statistik Aiken's V. Tujuan dari analisis ini adalah

untuk mengetahui kelayakan isi instrumen penilaian.

Data berupa skor dianalisis dengan Statistik Aiken's V yang dirumuskan sebagai berikut (Azwar, 2014):

$$V = \frac{\sum s}{[n(c-1)]}$$

keterangan:

V = validitas;

s = r - lo;

lo = angka penilaian validitas yang terendah;

c = angka penilaian validitas yang tertinggi;

r = angka yang diberikan oleh seorang penilai.

Nilai V tersebut akan diinterpretasikan dalam rentang antara 0,00 sampai dengan 1,00 sebagai koefisien validitas isi yang baik ataupun tidak baik serta nilai tersebut sebagai ukuran mendukung atau tidaknya validitas isi secara keseluruhan (Aiken, 1985). Untuk menginterpretasikan nilai validitas isi yang diperoleh dari perhitungan di atas, maka digunakan pengklasifikasian validitas seperti yang ditunjukkan pada Tabel 2.

Tabel 2. Kriteria Validitas

No.	Hasil Validitas	Kriteria Validitas
1	0,80 < V < 1,00	Sangat tinggi
2	0,60 < V < 0,80	Tinggi
3	0,40 < V < 0,60	Cukup
4	0,20 < V < 0,40	Rendah
5	0,00 < V < 0,20	Sangat rendah

Penguasaan keterampilan proses sains dibagi menjadi tiga kategori, yaitu tinggi, sedang, dan rendah. Kategori ini diadaptasi dari Azwar (2014) seperti yang ditunjukkan pada Tabel 3.

Tabel 3. Kategori Penguasaan Keterampilan Proses Sains

No.	Kategori	Persentase Rerata Skor (%)
1	Tinggi	$66,67\% \leq X$
2	Sedang	$33,33\% \leq X < 66,67\%$
3	Rendah	$X < 33,33\%$

Persentase tersebut didapatkan dengan membagi rerata skor pada keterampilan proses sains yang dijawab benar oleh siswa dengan jumlah butir soal. Jika dituliskan secara matematis adalah sebagai berikut.

$$\text{Persentase Penguasaan} = \frac{x}{n} \times 100\%$$

Keterangan:

x = rerata skor dari keterampilan proses sains yang dijawab oleh siswa

n = jumlah butir soal keterampilan proses sains

Analisis kuantitatif atau sering disebut validitas empiris (*empirical validity*) adalah penelaahan butir soal berdasarkan karakteristik internal tes melalui data yang diperoleh secara empiris. Karakteristik internal yang dimaksud meliputi kecocokan butir instrumen, reliabilitas soal, dan tingkat kesukaran. Analisis kuantitatif ini dilakukan setelah soal diujikan dengan menggunakan program QUEST.

Kecocokan Butir Instrumen

Pengujian *goodness of fit* (kecocokan butir instrumen) untuk mengetahui *fit* item dan testi terhadap model. Keeves & Alagumalai (Subali & Suyata, 2012) menyatakan bahwa *Item Characteristic Curve* (ICC) akan membentuk kurva yang mendatar (*flat*) bila besar INFIT MNSQ untuk item lebih besar dari satuan logit 1,30 atau lebih kecil dari satuan 0,77 dengan rata-rata 1,0. Oleh karena itu dalam program QUEST, Adam & Khoo (Subali & Suyata, 2012) menyatakan bahwa suatu item atau *testi/case/person* dinyatakan *fit* dengan model dengan batasan kisaran INFIT MNSQ dari 0,77 sampai 1,30. Dengan demikian, suatu item menjadi tidak *fit* menurut model Rasch bila memiliki nilai <-2,0 atau >+2,0 (probability atau peluang <0,05).

Reliabilitas

Hasil analisis program QUEST juga menampilkan estimasi reliabilitas tes berdasarkan item. Subali & Suyata (2012) menyatakan bahwa dari hasil analisis program QUEST, nilai estimasi reliabilitas menurut IRT jika dihitung berdasarkan item disebut indeks sparasi item. Apabila berdasarkan testi (*case/person*) disebut dengan *indeks sparasi person*. Semakin tinggi estimasi indeks sparasi item, maka semakin tepat keseluruhan item dianalisis menurut model yang digunakan dalam hal ini sesuai dengan Rasch Model. Semakin tinggi indeks sparasi person semakin konsisten setiap item pengukur digunakan untuk mengukur testi bersangkutan. Estimasi reliabilitas berdasarkan testi yakni reliabilitas

alpha Cronbach untuk data politimus. Menurut Wright & Master (Subali & Suyata, 2012), indeks sparasi item (*item separation index*) disebut dengan istilah reliabilitas sampel, sedangkan *indeks sparasi person* disebut reliabilitas tes.”

Tingkat Kesukaran

Karakteristik yang kedua adalah tingkat kesukaran atau indeks kesukaran. Indeks kesukaran atau tingkat kesukaran diperoleh dengan menggunakan program QUEST. Subali & Suyata (2012) menyatakan skala politimus memiliki skor x sebesar 0, 1, 2, 3, ..., m_i . Butir dikatakan baik jika indeks kesukaran lebih dari $-2,00$ atau kurang dari $2,00$. *Difficulty* merupakan *mean* (rerata) indeks kesukaran pada skor 0, skor 1, skor 2, dan skor 3 pada setiap butir.

Hasil dan Pembahasan

Hasil Pengembangan

Hasil pengembangan merupakan pencapaian yang telah didapatkan setelah menyelesaikan prosedur penelitian. Produk yang dikembangkan berupa instrumen penilaian biologi yang bertujuan untuk mengukur keterampilan proses sains siswa

SMA kelas X pada materi ekosistem. Uraian dari hasil pengembangan instrumen penilaian sebagai berikut.

1. Define

Produk instrumen penilaian yang dikembangkan berupa kisi-kisi soal, naskah soal pilihan ganda beralasan, serta rubrik penilaian. Instrumen penilaian dalam penelitian ini merupakan alat yang digunakan untuk menilai capaian pembelajaran siswa, baik aspek pengetahuan yaitu kognitif maupun keterampilan proses sains. Penilaian dilakukan dengan menggunakan tes dalam bentuk tertulis. Tes yang dikembangkan berupa pilihan ganda beralasan dengan jumlah item soal sebanyak 35 butir.

2. Design

Tahap ini merupakan tahap merancang instrumen penilaian. Pada tahap ini dihasilkan kisi-kisi penyusunan soal, naskah soal awal, dan rubrik penilaian. Kisi-kisi penyusunan soal dirancang dengan adanya identitas, KI, KD, indikator pembelajaran, indikator keterampilan proses sains, nomor soal dan ranah kognitif. Hal ini dapat dilihat pada Tabel 4.

Tabel 4. Kisi-kisi Penyusunan Soal KPS

No.	Indikator Pembelajaran	Indikator KPS	Nomor Soal	Ranah Kognitif
1	Mengamati dampak yang ditimbulkan oleh faktor abiotik terhadap faktor biotik	Mengamati	1 14	C2 C2
2	Mengelompokkan interaksi makhluk hidup serta faktor biotik berdasarkan hasil pengamatan	Mengelompokkan	2 15	C4 C4
3	Mengelompokkan jenis-jenis ekosistem berdasarkan ciri-cirinya	Mengelompokkan	16	C2
4	Mengelompokkan jenis organisme berdasarkan peranannya	Mengelompokkan	3 17	C2 C2
5	Menafsirkan hasil pengamatan tentang ekosistem dalam bentuk grafik, diagram, atau skema	Menafsirkan	4 18	C5 C5
6	Menafsirkan konsep eutrofikasi pada ekosistem perairan	Menafsirkan	5 19	C4 C4
7	Meramalkan hasil pengamatan dan memprediksi kemungkinan yang terjadi dari hasil pengamatan	Meramalkan	6 20	C5 C5
8	Mengkomunikasikan konsep aliran energi, rantai makanan, jaring-jaring makanan, piramida ekologi, daur biogeokimia	Berkomunikasi	7 21	C4 C4
9	Merumuskan hipotesis suatu permasalahan terkait dengan piramida makanan dan jaring-jaring makanan	Merumuskan hipotesis	8 22	C4 C4

Lanjutan Tabel 4

No.	Indikator Pembelajaran	Indikator KPS	Nomor Soal	Ranah Kognitif
10	Merumuskan hipotesis hasil penelitian berdasarkan grafik atau tabel	Merumuskan hipotesis	9 23	C2 C2
11	Merencanakan percobaan terkait dengan ekosistem	Merencanakan percobaan Merencanakan percobaan Merencanakan percobaan Merencanakan percobaan	10 11 12 13	C1 C2 C1 C2

Naskah soal awal terdiri dari petunjuk soal, contoh soal, dan soal sebanyak 35 item soal, sedangkan rubrik penilaian terdiri dari identitas, nomor, nomor soal, ide pokok, skor, dan kriteria. Pada skor diberikan skor 0, 1, 2, dan 3. Pada skor 0 apabila tidak menjawab, skor 1 apabila jawaban pilihan ganda salah dan alasan salah, skor 2 apabila jawaban pilihan ganda benar dan alasan salah atau apabila jawaban pilihan ganda salah dan alasan benar dengan memunculkan semua ide pokok, dan skor 3 apabila jawaban pilihan ganda benar dan alasan benar dengan memunculkan semua ide pokok.

3. *Develop*

Tahap pengembangan merupakan tahap pengembangan produk instrumen penilaian yang telah direncanakan sebelumnya disusun untuk menjadi produk yang siap untuk diujicobakan. Langkah-langkah yang dilakukan dalam tahap pengembangan instrumen penilaian biologi ini antara lain validasi oleh dosen ahli dan guru biologi sebagai pelaksana pembelajaran.

Validasi Dosen Ahli

Berdasarkan penelitian yang dilakukan oleh Akani (2015), draf instrumen yang digunakannya divalidasi oleh lima orang ahli, satu dari ahli pengukuran dan evaluasi, satu dari ahli biologi, satu dari ahli kimia, satu dari ahli fisika, dan satu lagi dari ahli ipa terintegrasi.

Produk instrumen penilaian biologi dikembangkan berdasarkan kisi-kisi yang mengacu pada KI, KD dan indikator yang ada, divalidasi oleh dosen ahli untuk mengetahui kelayakan produk instrumen

penilaian biologi sebelum divalidasi oleh guru sebagai pelaksana pembelajaran. Instrumen penilaian biologi ini berbentuk soal pilihan ganda beralasan. Soal ini terdiri dari 35 soal. Validasi instrumen penilaian biologi ini dilakukan oleh dosen ahli evaluasi dan dosen ahli materi.

Hasil validasi dibagi menjadi tiga yaitu valid tanpa revisi, valid dengan revisi, dan tidak valid. Saran ketika terjadi revisi dibagi menjadi 21 macam, 6 untuk ahli materi dan 15 untuk ahli evaluasi. Selain pada lembar validasi, masukan dan saran juga didapatkan dari validator secara operasional melalui lembar soal. Soal terdiri dari 35 butir yang mewakili keterampilan proses sains.

Hasil dari validasi instrumen penilaian dari dosen ahli menunjukkan bahwa terdapat beberapa komponen yang memerlukan perbaikan lebih lanjut. Setelah itu dianalisis menggunakan statistik Aiken's V dengan bantuan program *Microsoft Excel*. Dari analisis Aiken's V diperoleh nilai rata-rata 0,86 yang berarti valid.

Pada rubrik penilaian ada beberapa aspek yang dinilai yaitu substansi, konstruksi, dan bahasa. Berdasarkan lembar validasi yang telah dinilai oleh dosen ahli menunjukkan hasil validasi rubrik penilaian untuk pengembangan keterampilan proses sains siswa pada materi ekosistem dinyatakan valid.

Validasi Guru

Validasi selanjutnya dilakukan oleh guru sebagai pelaksana pembelajaran. Validasi dilakukan oleh guru A dari SMAN 1 Prambanan, guru B dari SMAN 1 Godean, guru C dari SMAN 1 Seyegan, dan guru D dari SMAN 2 Ngaglik sebagai guru kelas X

di SMAN Kabupaten Sleman. Berdasarkan lembar validasi, hasil validasi dibagi menjadi tiga yaitu valid tanpa revisi, valid dengan revisi, dan tidak valid. Soal terdiri dari 35 butir yang mewakili keterampilan proses sains. Hasil dari validasi instrumen penilaian dari guru menunjukkan bahwa terdapat beberapa komponen yang memerlukan perbaikan lebih lanjut. Hal yang harus diperhatikan yaitu pada kategori layak dengan revisi.

Berdasarkan validasi soal dari guru diperoleh saran berupa pengurangan butir soal dari yang semula 35 butir soal menjadi 23 butir soal. Masukan dan saran dari guru tersebut digunakan untuk merevisi produk awal instrumen penilaian, sehingga diperoleh instrumen penilaian yang siap diujicobakan. Selanjutnya validasi instrumen penilaian menggunakan statistik Aiken's V dengan bantuan program *Microsoft Excel*. Dari analisis Aiken's V diperoleh nilai rata-rata 0,94 yang berarti valid.

Dari hasil analisis validasi instrumen penilaian oleh guru dapat diketahui bahwa sebanyak 17 item soal dengan kategori sangat tinggi, 5 item soal dengan kategori tinggi, dan 1 soal dengan kategori cukup.

Uji Coba Terbatas

Tindak lanjut hasil revisi produk oleh dosen ahli dan guru adalah uji coba terbatas. Uji coba terbatas dilaksanakan di SMAN 2 Ngaglik dengan sampel penelitian 121 orang siswa. Uji coba terbatas ini dilakukan untuk mengetahui kecocokan butir instrumen (*Goodness of fit*), reliabilitas, dan tingkat kesukaran dari instrumen penilaian yang dikembangkan.

Informasi tentang karakteristik soal diperoleh dari analisis menggunakan program komputer QUEST. Uji coba yang dilakukan memanfaatkan waktu pelajaran biologi yaitu 2 jam pelajaran biologi dengan total waktu yang diperlukan yaitu 90 menit. Penskoran dilakukan dengan diskor politomus.

Goodness of fit pada Uji Coba

Pengujian *fit* tes secara keseluruhan maupun tiap butir dengan menggunakan program QUEST. Pengujian *fit* tes keseluruhan dikembangkan oleh Adam & Khoo (Subali & Suyata, 2012) berdasarkan

nilai rerata INFIT *Mean of Square (Mean INFIT MNSQ)* beserta simpangan bakunya atau mengamati nilai rata-rata INFIT *t (Mean INFIT t)* beserta simpangan bakunya. Jika rerata INFIT MNSQ sekitar 1,00 dan simpangan bakunya 0,00 atau rerata INFIT *t* mendekati 0,00 dan simpangan bakunya 1,00, maka keseluruhan tes *fit* dengan model PCM 1 PL. Berdasarkan perhitungan, nilai rerata INFIT MNSQ 1,00 (sekitar 1,00) dan simpangan baku 0,12 (sekitar 0,00), maka keseluruhan tes *fit* dengan PCM 1 PL.

Pengujian penetapan *fit* setiap butir terhadap model mengikuti kaidah Adam & Khoo (Subali & Suyata, 2012), yakni suatu butir *fit* terhadap model jika nilai INFIT MNSQ antara 0,77 sampai dengan 1,30. Dengan batas penerimaan butir menggunakan INFIT MNSQ atau *fit* menurut model (antara 0,77 sampai dengan 1,30) dan menggunakan INFIT *t* dengan batas -2,00 sampai 2,00, maka diperoleh butir-butir yang cocok memenuhi *goodness of fit*. Nilai INFIT MNSQ memiliki jangkauan dari 0,86 sampai dengan 1,29.

Berdasarkan batas penerimaan butir menggunakan INFIT MNSQ atau *fit* menurut model, maka semua butir sebanyak 23 butir *fit* semua.

Reliabilitas Tes pada Uji Coba

Selain untuk menguji kecocokan, output program QUEST juga menampilkan estimasi reliabilitas perangkat instrumen tes. Berdasarkan hasil analisis dengan program QUEST, reliabilitas tes berdasarkan item (indeks aparasi item sebesar 0,76 yang tergolong kategori cukup tinggi, sedangkan reliabilitas tes berdasarkan testi (indeks sparasi person) sebesar 0,80 yang tergolong kategori sangat tinggi (Sumintono & Widhiarso, 2014).

Tingkat Kesukaran (Difficulties Index)

Penelitian yang dilakukan oleh Sukarno, Permanasari, dan Hamidah (2013) dengan menggunakan soal tes keterampilan sains dasar yang dikembangkan oleh The International Energy Management Courses (IEC) dan The University of Queensland. Item yang dikembangkan meliputi keterampilan menyimpulkan, mengobservasi, memprediksi, mengukur, dan mengklasifikasikan. Data yang telah

dikumpulkan kemudian dilakukan kategorisasi skor menjadi tiga yaitu mudah, sedang, dan tinggi.

Berdasarkan hasil analisis dengan QUEST, tingkat kesukaran (*difficulty*) butir-butir terletak dalam jangkauan dari -1,47 sampai dengan 1,59 dengan rata-rata 0,00 dan simpangan baku 0,87. Butir dikatakan baik jika nilai *difficulty* lebih dari -2,00 atau kurang dari 2,00. Nilai *difficulty* rata-rata perangkat tes keterampilan proses sains yang dipakai pada tahap uji coba sebesar $0,00 \pm 0,87$ dengan kategori sedang. Jadi, berdasarkan nilai *difficulty* dari 23 butir soal, terdapat 11 soal dengan kriteria mudah, 7 soal dengan kriteria sedang dan 5 soal dengan kriteria sukar. Distribusi tingkat kesukaran butir soal dapat dilihat pada Tabel 4.

Tabel 4. Distribusi Tingkat Kesukaran Butir Soal

No.	Kriteria	Jumlah
1	Mudah	11
2	Sedang	7
3	Sulit	5
Jumlah Soal		23

Berdasarkan data pada Tabel 4, maka dapat diperoleh produk instrumen penilaian yang diterima dari uji coba terbatas sebanyak 23 butir soal, artinya semua butir soal diterima. Soal yang diterima digunakan sebagai produk akhir pada instrumen penilaian sebelum digunakan pada uji coba luas.

Uji Coba Luas

Setelah produk instrumen penilaian keterampilan proses sains yang telah diuji coba secara terbatas direvisi dan diperoleh soal pilihan ganda beralasan pada uji terbatas sebanyak 23 butir soal, berarti tidak ada butir soal yang ditolak atau gugur melainkan butir soal bisa diuji lapangan seluruhnya.

Uji coba luas melibatkan 3 SMA Negeri dengan kategori sekolah berbeda di Kabupaten Sleman yaitu SMA Negeri 1 Godean (kategori sekolah tinggi), SMA Negeri 1 Prambanan (kategori sekolah sedang), dan SMA Negeri 1 Seyegan (kategori sekolah rendah). Kategori ini berdasarkan peringkat Ujian Nasional (UN) tahun ajaran 2013/2014.

Berdasarkan hasil uji coba luas di lapangan, analisis penguasaan keterampilan

proses sains dilakukan dengan mencari rerata nilai persentase jawaban benar pada masing-masing jenis keterampilan proses sains yang muncul. Hasil analisis ketercapaian keterampilan proses sains pada uji coba luas tersaji pada Tabel 5. Tabel 5 menunjukkan tingkat penguasaan keterampilan proses sains pada uji coba luas. Hasil tersebut diperoleh dari siswa di tiga sekolah di SMA Negeri Sleman yang dijadikan subjek uji coba.

Tabel 5. Penguasaan Keterampilan Proses Sains pada Uji Coba Luas

No.	Subjek Penelitian	Rata-rata Penguasaan (%)
1	SMAN 1 Godean	76,64
2	SMAN 1 Prambanan	73,71
3	SMAN 1 Seyegan	70,12

Dari Tabel 5 dapat diketahui untuk penguasaan keterampilan proses sains paling tinggi dari semua indikator adalah siswa SMAN 1 Godean dengan indikator mengelompokkan, menafsirkan, meramalkan merumuskan hipotesis, dan merencanakan percobaan lebih tinggi persentase penguasaannya dibandingkan dengan sekolah lainnya.

Revisi Produk

1. Revisi Berdasarkan Hasil Validasi

Produk instrumen penilaian yang dikembangkan melalui revisi 3 validator yaitu satu orang dosen ahli evaluasi, satu orang dosen ahli materi, dan empat orang guru sebagai pelaksana pembelajaran kelas X. Hasil validasi instrumen penilaian menunjukkan instrumen penilaian valid. Menurut Azwar (2014) estimasi validitas tidak dapat dituntut suatu koefisien sangat tinggi sebagaimana halnya dalam penilaian terhadap koefisien reliabilitas. Selanjutnya Azwar mengungkapkan apabila koefisien validitas itu lebih dari 0,33 dianggap valid dan jika kurang dari 0,33 biasanya dianggap tidak memadai atau tidak valid. Jadi keseluruhan penilaian dari validator tersebut dapat disimpulkan valid atau dengan kata lain layak.

Validator ahli evaluasi memutuskan 23 butir soal valid tanpa revisi dan 12 soal valid dengan revisi, sedangkan validator ahli materi 30 soal valid tanpa revisi dan 5 soal valid dengan revisi. Validator guru

memutuskan 33 butir soal valid tanpa revisi dan 2 soal valid dengan revisi. Semua butir soal yang valid tanpa revisi dapat digunakan langsung pada uji coba terbatas. Butir soal yang valid dengan revisi sebelum digunakan pada uji coba terbatas, diperbaiki terlebih dahulu. Revisi dari ketiga validator ahli tersebut kemudian diperbaiki dengan memperhatikan masukan dan saran pada lembar validasi dan lembar soal.

Semua butir soal yang telah direvisi berdasarkan masukan dan saran validator selanjutnya disusun kembali sehingga diperoleh produk instrumen penilaian yang siap diujicobakan secara terbatas. Produk instrumen penilaian yang berjumlah 35 butir soal direduksi menjadi 23 butir soal.

Pada rubrik penilaian saran dan masukan dari validator adalah skor penilaian. Awalnya skor penilaian adalah Skor 4 jika jawaban benar dan alasan tepat, Skor 3 jika jawaban benar dan alasan kurang tepat. Skor 2 jika jawaban benar dan alasan tidak tepat, skor 1 jika jawaban benar dan tidak ada alasan dan jika jawaban salah dan alasan tidak tepat, dan skor 0 jika jawaban salah dan tidak ada alasan. Saran dan masukan validator adalah skor 3 apabila jawaban benar dan alasan benar, skor 2 apabila jawaban benar dan alasan salah dan apabila jawaban salah dan alasan benar, skor 1 apabila jawaban salah dan alasan salah, skor 0 apabila tidak menjawab.

2. Revisi Berdasarkan Uji Coba Terbatas

Berdasarkan uji coba terbatas, semua butir soal yang sebelumnya sudah divalidasi oleh dosen ahli dan guru ketika dilaksanakan ujicoba terbatas sebanyak 23 butir soal hasilnya memenuhi kriteria standar tes dan berada pada kategori diterima. Semua butir soal yang diterima langsung dimasukkan produk instrumen penilaian, sedangkan yang berada pada kategori diperbaiki, dilakukan perbaikan sehingga menjadi produk instrumen penilaian yang siap dipakai dalam uji coba luas (pengukuran). Semua perbaikan dilakukan sesuai dengan hasil validitas empiris, maka semua butir soal disusun dan dirakit kembali. Butir soal yang sudah dirakit menjadi produk instrumen penilaian kemudian digunakan dalam uji coba luas.

Simpulan

Dari hasil analisis dan pembahasan dapat dikemukakan beberapa simpulan sebagai berikut: (1) Instrumen penilaian biologi untuk mengembangkan keterampilan proses sains siswa SMA kelas X pada materi ekosistem dikatakan layak ditinjau dari karakteristik standar tes. Penilaian dari dosen ahli diperoleh validitas isi Aiken's V sebesar 0,86 dan dari guru diperoleh validitas isi Aiken's V sebesar 0,92. Berdasarkan hasil uji coba secara empiris diperoleh sebanyak 23 butir soal *fit* dengan nilai *goodness of fit* memiliki jangkauan dari 0,86 sampai dengan 1,41. Reliabilitas soal diperoleh nilai 0,80 dengan kategori bagus dan tingkat kesukaran dengan rentang antara -1,47 sampai dengan 1,59; dan (2) Persentase tingkat penguasaan keterampilan proses sains sekolah dengan kategori tinggi adalah sebesar 76,64%; sekolah dengan kategori sedang sebesar 73,71%; dan sekolah dengan kategori rendah sebesar 70,12%.

Daftar Pustaka

- Aiken, L.R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement*. 45, 132-142.
- Akani, Omiko. (2015). Levels of Possession of Science Process Skills by Final Year Students of Colleges of Education in South-Eastern States of Nigeria. *Journal of Education and Practice*, 6(27), 94-101.
- Arifin, Z. (2014). *Evaluasi pembelajaran*. Bandung: PT Remaja Rosdakarya.
- Azwar, S. (2014). *Penyusunan skala psikologi*. Yogyakarta: Pustaka Pelajar.
- BSNP. (2006). *Panduan penyusunan kurikulum tingkat satuan pendidikan jenjang pendidikan dasar dan menengah*. Jakarta: BSNP.
- Bundu, P. (2006). *Penilaian keterampilan proses dan sikap ilmiah dalam pembelajaran sains SD*. Jakarta: Depdiknas Dirjen Dikti.
- Chiappetta, E.L. & Koballa, T.R. (2010). *Science instruction in the middle and secondary schools (7th ed.)*. Boston: Pearson Education Inc.

- Hamalik, O. (2013). *Proses belajar mengajar*. Jakarta: Bumi Aksara.
- Kurniawan, E & Mutaqimah, E. (2009). *Penilaian*. Jakarta: Depdiknas.
- Republik Indonesia. (2008). *Undang-undang RI Nomor 20, Tahun 2003, tentang Sistem Pendidikan Nasional*.
- Reynolds, C.R., Livingston, R.B., & Willson, V. (2010). *Measurement and assessment in education (2nd ed.)*. Upper Saddle River: Pearson.
- Subali, B. & Suyata, P. (2012). *Pengembangan item tes konvergen dan divergen dan penyelidikannya secara empiris*. Yogyakarta: Diandra.
- Sudjana, N. (2005). *Penilaian hasil proses belajar mengajar*. Bandung: PT Remaja Rosdakarya.
- Sudjiono, A. (2011). *Pengantar evaluasi pendidikan*. Jakarta: Rajawali Pers.
- Sumintono, B. & Widhiarso, W. (2014). *Aplikasi model RASCH untuk penelitian ilmu-ilmu sosial*. Cimahi: Trim Komunikata Publishing House.
- Sukarno, Permanasari, A, & Hamidah, I. (2013). The Profile of Science Process Skill (SPS) Student at Secondary High School (Case Study in Jambi). *International Journal of Scientific Engineering and Research (USER)*, 1(1), 79-83.
- Thiagarajan, S., Semmel, D.S. & Semmel, M.I. (1974). *Instructional development for training teachers of exceptional children*. Minneapolis: University Of Minneasota.