

KONSTRUKSI DAN ANALISIS HASIL UJI-COBA TES HASIL BELAJAR BIOLOGI

I Gusti Ngurah Puger¹

E-mail: ngurah_puger@yahoo.co.id

Abstrak

Tujuan dari karya inovatif ini adalah untuk mengetahui: (1) langkah-langkah yang ditempuh di dalam mengonstruksi tes hasil belajar biologi dan (2) langkah-langkah yang ditempuh di dalam melakukan proses standarisasi tes hasil belajar biologi. Setelah tes hasil belajar biologi selesai dikonstruksi, selanjutnya dimintakan respon kepada dua orang *rater* yang menguasai bidang biologi, dan diujicobakan kepada 40 siswa kelas VIII SMP Negeri 1 Seririt. Data respon dua orang *rater* digunakan untuk menentukan koefisien validitas isi (VI) dan koefisien reliabilitas (r'') respon antar-*rater*. Sedangkan data respon 40 siswa SMP digunakan untuk menentukan tingkat kesulitan, menentukan daya pembeda, peran distraktor (pengecoh), menguji validitas butir, dan menghitung reliabilitas tes. Hasil kajian (*review*) pengembangan tes hasil belajar biologi dan analisis tes hasil belajar biologi diperoleh temuan sebagai berikut. (1) Langkah-langkah yang ditempuh dalam mengonstruksi tes hasil belajar biologi adalah: (a) analisis kurikulum, (b) tujuan pengembangan tes hasil belajar biologi, (c) pengembangan materi ajar, (d) menentukan standar kompetensi, kompetensi dasar, dan indikator pembelajaran, (e) membuat kisi-kisi tes hasil belajar biologi, dan (f) menulis butir soal tes hasil belajar biologi, dan (2) hasil analisis penilaian *rater* dan uji-coba pada sejumlah siswa adalah: (a) koefisien validitas isi (VI) dari tes hasil belajar biologi adalah 0,90, (b) hasil penghitungan reliabilitas respon antar-*rater* dengan menggunakan formula Hoyt, diperoleh koefisien r'' sebesar 0,89, (c) 16 butir soal termasuk kategori sedang; dan 4 butir soal termasuk kategori mudah, (d) sebanyak 4 butir yang menyusun tes hasil belajar biologi termasuk daya pembeda klasifikasi jelek (*poor*); sebanyak 10 butir termasuk daya pembeda klasifikasi cukup (*satisfactory*); dan sebanyak 6 butir termasuk daya pembeda kategori baik (*good*), (e) semua distraktor dari 20 butir yang menyusun tes hasil belajar biologi yang dikembangkan memiliki kategori baik, (f) butir tes hasil belajar biologi kategori *drop* (gugur), yakni butir tes nomor: 13 dan 19; dan selain butir tersebut termasuk kategori *valid* (sahih), dan (g) hasil penghitungan reliabilitas tes dengan menggunakan rumus K-R 20 diperoleh koefisien reliabilitas (r'') sebesar 0,84.

Kata kunci: Konstruksi, analisis, uji-coba, tes hasil belajar

¹I Gusti Ngurah Puger adalah staf edukatif pada Fakultas Keguruan dan Ilmu Pendidikan (FKIP), Universitas Panji Sakti, Singaraja.

CONSTRUCTION AND ANALYSIS OF TRIAL RESULTS BIOLOGY LEARNING TEST RESULTS

Abstract

The purpose of this innovative work was to find out: (1) the steps taken in constructing biology learning outcomes tests and (2) the steps taken in carrying out the standardization process of biology learning outcomes tests. After the biology learning test was completed, it was then asked to respond to two Rater who mastered the field of biology, and tested it on 40 students of the eighth graders SMP Negeri 1 Seririt. Two-rater response data was used to determine the content validity coefficient (VI) and reliability coefficient (r) of inter-rater responses. Whereas the response data of 40 junior high school students were used to determine the level of difficulty, to determine the distinguishing power, to determine distractor role, to define test item validity, and to calculate test reliability. The results of the review of the development of biology learning outcomes tests and analysis of biology learning outcomes tests were obtained as follows. (1) The steps taken in constructing biology learning outcomes tests are: (a) analyzing curriculum, (b) analyzing the purpose of developing biology learning outcomes tests, (c) developing teaching materials, (d) determining competency standards, basic competencies, and learning indicators, (e) creating a test of biology learning outcomes, and (f) writing biology learning outcomes test items; and (2) the results of the rater assessment analysis and trials on a number of students are: (a) content validity coefficient (VI) from the biology learning outcomes test was 0.90, (b) the results of the inter-rater response reliability calculation using the Hoyt formula, obtained the r coefficient was 0.89, (c) items numbered: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 13, 14, 16, 17, and 19 were in the medium category; and questions numbered: 12, 15, 18, and 20 belong to the easy category, (d) as many as 4 items compiling the biology learning outcomes test were categorized as weak or poor, namely the items number: 1, 15, 19, and 20; 10 items including satisfactory classification, namely items number: 2, 6, 7, 8, 9, 11, 12, 13, 14 and 16; and as many as 6 items were categorized as avergae/good, namely items number: 3, 4, 5, 10, 17, and 20, (e) all distractors from the 20 items compiling the biology learning outcomes test developed were categorized as good, (f) items dropped for biology namely test items number: 13 and 19; and the biology learning outcomes test which were categorized as valid, namely the test number: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 17, 18 and 20, and (g) the results of calculating the reliability of the test using the KR 20 formula obtained the reliability coefficient (r) of 0.84.

Keywords: Construction, analysis, trial, learning outcomes

PENDAHULUAN

Dalam melaksanakan kegiatan kesehariannya sebagai tenaga pendidik, seorang guru Sekolah Menengah Pertama (SMP) tidak bisa melepaskan diri dengan proses pembelajaran. Proses pembelajaran yang sudah dikomunikasikan di kelas harus diketahui efeknya pada siswa itu sendiri. Dalam artian, materi ajar yang sudah dikomunikasikan dalam proses pembelajaran tersebut apa sudah dipahami secara bermakna (meminjam istilah Ausubel) oleh siswa itu sendiri.

Pembelajaran merupakan upaya yang sengaja dan bertujuan yang berfokus kepada kepentingan, karakteristik, dan kondisi orang lain agar peserta didik dapat belajar dengan efektif dan efisien. Istilah ini merupakan paradigma baru yang menekankan pada prinsip keragaman peserta didik atau pembelajar (*learner*), dan menggantikan istilah *pengajaran* atau *mengajar* yang menekankan keseragaman. Istilah *pengajaran* lebih banyak berarti sebagai upaya penyampaian informasi kepada pihak lain. Latar belakang teoretisnya didasarkan pada teori psikologi behavioristik dan teori komunikasi searah. Sedangkan, konsep pembelajaran didasarkan pada teori psikologi konstruktivis dan teori komunikasi konvergensi. Konsep pembelajaran ini merupakan inti pada lapis pengalaman belajar, yaitu tempat peserta didik membangun diri sendiri berdasarkan pengetahuan dan pengalaman yang diperolehnya melalui interaksi dengan lingkungannya (Thobromi dan Mustofa, 2011).

Setiap proses pembelajaran yang dilakukan oleh seorang guru di SMP perlu dilakukan proses pengujian (*testing*) terhadap bermakna atau tidak-bermahnanya proses pembelajaran yang sudah dilakukan. Untuk menguji proses pembelajaran biologi di SMP dapat digunakan instrumen berupa tes hasil belajar biologi. Biasanya seorang guru biologi di dalam membuat tes hasil belajar biologi tidak memperhatikan persyaratan standarisasi tes. Guru-guru biologi di SMP sering memiliki asumsi yang keliru, yakni begitu tes hasil belajar biologi selesai dikonstruksi dari kisi-kisi tes langsung digunakan untuk menguji hasil belajar biologi.

Bahkan Widiyaningsih (2014) menyatakan dalam membuat tes hasil belajar dalam suatu bidang studi, banyak guru-guru di Sekolah Menengah tidak menggunakan dasar secara evaluatif di dalam menurunkan butir-butir yang menyusun tes hasil belajar

yang bersangkutan. Hal ini menyebabkan butir-butir yang menyusun tes hasil belajar tidak diturunkan berdasarkan sebaran materi yang tertuang dalam kurikulum bidang studi yang bersangkutan. Dengan kata lain, butir-butir tes hasil belajar diturunkan berbasiskan atas materi yang diminati oleh guru yang bersangkutan, atau tanpa menggunakan kisi-kisi tes hasil belajar.

Asumsi guru-guru biologi di SMP ini merupakan kekeliruan yang patut diberi penekanan dan perhatian yang serius. Hal ini sesuai dengan pernyataan Puger (2010), yang pada hakikatnya menyatakan seorang guru di SMP harus mengubah pandangan bahwa hasil konstruksi tes hasil belajarnya tidaklah langsung bisa digunakan untuk menguji hasil belajar siswa. Setiap tes hasil belajar, khususnya tes hasil belajar biologi yang sudah selesai dikonstruksi wajib dilakukan sekurang-kurangnya pengujian validitas isi, penghitungan reliabilitas respon antar-rater, pengujian validitas butir, dan penghitungan reliabilitas tes. Terkait dengan hal ini, Kasiram (1984) menambahkan bahwa selain menguji validitas butir dan menghitung reliabilitas tes, perlu juga dilakukan analisis tingkat kesukaran *item*, analisis daya beda, dan analisis peran pengecoh.

Berpijak atas kurang memahaminya mengenai cara mengonstruksi dan analisis hasil uji-coba tes hasil belajar, maka perlu dilakukan suatu kajian mengenai proses pengembangan dan analisis hasil uji-coba tes hasil belajar biologi di SMP.

Berdasarkan atas kenyataan-kenyataan yang sudah dikemukakan, dapat dirumuskan masalah sebagai berikut. (1) Bagaimanakah langkah-langkah yang ditempuh di dalam mengonstruksi tes hasil belajar biologi?, dan (2) langkah-langkah apakah yang ditempuh di dalam melakukan proses standarisasi tes hasil belajar biologi?

Materi ajar yang dikembangkan dalam karya inovatif ini adalah reproduksi generatif tumbuhan angiospermae. Materi ajar ini, pada hakikatnya terdiri atas tiga subtopik, yakni tata letak dan fungsi bagian-bagian bunga, penyerbukan (*polinasi*), dan pembuahan (*fertilisasi*). Dari ketiga subtopik ini, selanjutnya ditelusuri standar kompetensi, kompetensi dasar, dan indikatornya. *Item* yang menyusun tes hasil belajar biologi diturunkan dari indikator yang tercantum dalam kisi-kisi tes hasil belajar biologi.

Selanjutnya, tes hasil belajar biologi yang sudah selesai dikonstruksi dinilai oleh dua orang *rater*. Kecocokan *item* yang menyusun tes hasil belajar biologi dengan indikatornya pada kisi-kisi tes yang disiapkan oleh pengonstruksi tes sangat berkaitan dengan penguasaan materi ajar oleh penilai yang bersangkutan. Makin menguasai materi ajar yang dikonstruksi penulis tes oleh penilai ahli, maka makin akurat penilaiannya.

Dalam menulis *item* tes, pengonstruksi tes yang bersangkutan harus menyesuaikan tingkat kesulitan *item* tes dengan kemampuan siswa di mana tes yang bersangkutan akan diujicobakan atau digunakan. Bilamana kemampuan siswanya tergolong tinggi, maka komposisi *item* yang menyusun tes harus lebih banyak yang berkategori sukar, lalu diikuti kategori sedang, dan sedikit yang berkategori mudah.

Dalam menulis *item* tes hasil belajar biologi, perlu juga diperhatikan agar siswa terhindar dari proses *gambling* di dalam memberikan responnya pada alternatif jawaban yang tersedia. Suatu *item* yang diberikan respon melalui proses *gambling* akan tidak mampu membedakan kelompok siswa yang memiliki kemampuan tinggi (kelompok atas) dan memiliki kemampuan rendah (kelompok bawah). Ini berarti, *item* yang bersangkutan tidak memiliki daya beda yang baik.

Item yang baik juga bisa dilihat dari distraktornya yang sangat mirip dengan kunci jawabannya. Makin mirip distraktor *item* dengan kunci jawabannya akan menghindarkan siswa untuk menjawab *item* tersebut dengan cara menebak. Dari sini juga akan dapat diduga bahwa respon siswa terhadap kunci jawaban dan distraktornya akan menyebar. Dalam artian, siswa tidak menjawab pada kunci jawaban saja atau pada salah satu distraktor yang tersedia. Model *item* seperti ini dikatakan sebagai *item* yang mempunyai distraktor yang berfungsi secara efektif.

Langkah selanjutnya dalam analisis butir tes adalah menguji validitas butir. Tinggi-rendahnya butir tes hasil belajar yang mengalami *drop* (gugur), tergantung pada tiga hal pokok, yakni: tinggi-rendahnya koefisien validitas isi, tinggi-rendahnya hubungan antara skor butir dengan skor total, dan tinggi-rendahnya keterkaitan antara butir soal dengan kisi-kisi yang disusun sebelumnya. Seorang pengonstruksi tes hasil belajar dapat mengetahui atau meramalkan tinggi-rendahnya butir *item* yang menyusun

tes yang bersangkutan mengalami *drop* dengan mengkaji dari tiga hal *vital* yang sudah disebutkan.

Langkah terakhir di dalam analisis tes adalah menghitung reliabilitas tes. Tinggi-rendahnya koefisien r suatu tes, dapat diketahui dari empat faktor yang *vital*, yakni: tinggi-rendahnya nilai koefisien r respon antar-rater, tinggi-rendahnya koefisien r , pembuangan butir yang *drop*, dan panjang-pendeknya pertanyaan yang menyusun setiap butir tes hasil belajar.

METODE PENELITIAN

Metodologi Karya Inovatif

Karya ilmiah ini termasuk jenis karya inovatif (*innovative activity*). Menurut Aqib (2013) karya inovatif merupakan karya-karya yang dihasilkan oleh seorang guru yang merupakan kebaruan-kebaruan (*novelty*) atau mengandung nilai inovasi (*innovation*) dalam hal yang berkaitan dengan pembelajaran. Karya inovatif bagi guru dapat dibagi menjadi empat kategori, yakni: (1) menemukan teknologi tepat guna, (2) menemukan/menciptakan karya seni, (3) membuat/memodifikasi alat pelajaran/peraga/praktikum, dan (4) mengikuti pengembangan penyusunan standar, pedoman, soal dan sejenisnya. Karya inovatif ini, termasuk kategori pengembangan penyusunan standar, pedoman, soal, dan sejenisnya. Hal ini disebabkan oleh karya inovatif yang dikerjakan mengupas masalah konstruksi dan analisis hasil uji-coba tes hasil belajar biologi.

Sumber data karya inovatif ini adalah: draf konstruksi tes hasil belajar biologi, dua orang *rater*, dan 40 siswa kelas IX SMP Negeri 1 Seririt. Data dari draf konstruksi tes hasil belajar biologi digunakan untuk mengetahui langkah-langkah dalam konstruksi tes, data dari dua orang *rater* digunakan untuk menguji validitas isi dan menghitung reliabilitas respon antar-*rater*; dan data dari 40 siswa kelas IX SMP Negeri 1 Seririt digunakan untuk analisis tingkat kesukaran, menentukan daya beda, menentukan peran distraktor, menguji validitas butir, dan menghitung reliabilitas tes.

Untuk mengumpulkan data mengenai langkah-langkah di dalam mengonstruksi tes hasil belajar biologi digunakan pedoman pencatatan dokumen, untuk memperoleh

respon dua orang *rater* digunakan kisi-kisi tes hasil belajar biologi dan tes hasil belajar biologi, dan untuk memperoleh respon siswa digunakan tes hasil belajar biologi. Dua orang *rater* yang sudah menerima kisi-kisi tes hasil belajar biologi dan tes hasil belajar biologi selanjutnya memberikan responnya mengenai kecocokan antara suatu *item* dengan indikatornya. Respon yang diberikan terdiri atas 4 skala, yakni skala 1-2 termasuk kurang relevan, dan skala 3-4 termasuk sangat relevan. Untuk respon siswa, sebelum mengerjakan tes hasil belajar biologi yang diterima disarankan agar membaca terlebih dahulu petunjuk mengerjakan setiap butir yang menyusun tes hasil belajar biologi yang bersangkutan.

Data yang diperoleh dari pencatatan dokumen berupa draf konstruksi tes hasil belajar biologi dianalisis dengan analisis deskriptif kualitatif. Koefisien validitas isi (VI) diuji dengan formula tabulasi silang dari Gregory. Koefisien reliabilitas (r'') respon antar-*rater* dihitung dengan formula Anava HOYT sebagaimana dikutip oleh Sudijono (2001). Untuk menghitung tingkat kesukaran butir digunakan formula indeks kesukaran (P), sesuai dengan yang dikemukakan oleh Arikunto (1993). Untuk menghitung daya beda butir digunakan formula indeks diskriminasi (D). Indeks diskriminasi ini sesuai dengan yang dikemukakan oleh Daryanto (1999), untuk menentukan efektivitas peran pengecoh digunakan kriteria suatu distraktor dapat dikatakan berfungsi baik jika paling sedikit *dipilih oleh 5% pengikut tes* (Arikunto, 1993), untuk menguji validitas butir digunakan formula *korelasi point biserial* (r_{pbi}) (Djaali dan Muljono, 2008), dan untuk menghitung reliabilitas tes digunakan formula K-R 20 (Maba, 2002).

HASIL DAN PEMBAHASAN

Hasil Karya Inovatif

Berpijak atas hasil pencatatan draf konstruksi tes hasil belajar biologi dan selanjutnya dianalisis dengan analisis deskriptif kualitatif diperoleh hasil bahwa dalam konstruksi tes hasil belajar biologi ditempuh langkah-langkah sebagai berikut. (a) Analisis kurikulum, (b) tujuan pengembangan tes hasil belajar biologi, (c) pengembangan materi ajar, (d) menentukan standar kompetensi, kompetensi dasar, dan

indikator pembelajaran, (e) membuat kisi-kisi tes hasil belajar biologi, dan (f) menulis butir soal tes hasil belajar biologi.

Hasil pengujian validitas isi tes hasil belajar biologi dengan menggunakan formula tabulasi silang dari Gregory, diperoleh koefisien VI sebesar 0,90. Koefisien ini sudah memenuhi kriteria yang dikemukakan oleh Erwin, yang mensyaratkan sekurang-kurangnya koefisien VI sebesar 0,90. Hal ini berarti, dikaji dari koefisien validitas isinya, tes hasil belajar biologi yang dikembangkan bisa digunakan lebih lanjut.

Dalam menghitung reliabilitas respon antar-rater digunakan formula Anava HOYT. Hasil penghitungan reliabilitas respon antar-rater dengan menggunakan formula Anava HOYT, diperoleh koefisien r'' sebesar 0,89. Oleh karena koefisien reliabilitas hitung (r'' -hitung) lebih tinggi daripada koefisien reliabilitas yang dipersyaratkan oleh Fraenkel dan Wallen (1993) ($0,89 > 0,70$), maka tes hasil belajar biologi ini dapat diujicobakan lebih lanjut.

Untuk menghitung tingkat kesukaran butir tes hasil belajar biologi digunakan formula indeks kesukaran butir (P). Berdasarkan atas kategori indeks kesukaran butir (P), dapat dikatakan bahwa: butir soal nomor: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 13, 14, 16, 17, dan 19 termasuk kategori sedang; dan soal nomor: 12, 15, 18, dan 20 termasuk kategori mudah.

Dalam menentukan daya beda setiap butir yang menyusun tes hasil belajar biologi digunakan formula indeks diskriminasi (D) untuk menghitungnya. Berdasarkan atas hasil penghitungan indeks diskriminasi (D) diperoleh hasil sebagai berikut. (1) Sebanyak 4 butir yang menyusun tes hasil belajar biologi termasuk daya pembeda klasifikasi jelek (*poor*), yakni nomor butir: 1, 15, 19, dan 20; (2) sebanyak 10 butir termasuk daya pembeda klasifikasi cukup (*satisfactory*), yakni nomor butir: 2, 6, 7, 8, 9, 11, 12, 13, 14, dan 16; dan (3) sebanyak 6 butir termasuk daya pembeda kategori baik (*good*), yakni nomor butir: 3, 4, 5, 10, 17, dan 20.

Untuk mengetahui efektivitas peran pengecoh yang menyusun tes hasil belajar biologi yang dikembangkan, digunakan kriteria suatu distraktor dapat dikatakan berfungsi baik jika paling sedikit *dipilih oleh 5% pengikut tes*. Berdasarkan kriteria

tersebut, dapat ditarik suatu keputusan bahwa semua distraktor dari 20 butir yang menyusun tes hasil belajar biologi yang dikembangkan memiliki kategori baik. Hal ini disebabkan oleh semua distraktor dari 20 butir yang menyusun tes hasil belajar biologi dipilih oleh ≥ 2 siswa.

Langkah selanjutnya di dalam menganalisis butir tes hasil belajar adalah menguji validitas butir. Untuk menguji validitas butir tes hasil belajar digunakan formula *korelasi point biserial* (r_{pbi}). Setelah dilakukan pengujian validitas butir dengan menggunakan rumus *korelasi point biserial* (r_{pbi}) terdapat dua kategori butir tes hasil belajar biologi, yaitu: (1) butir tes hasil belajar biologi kategori *drop* (gugur), yakni butir tes nomor: 13 dan 19; dan (2) butir tes hasil belajar biologi kategori *valid* (sahih), yakni butir tes nomor: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 17, 18, dan 20.

Sebagai penutup di dalam analisis tes hasil belajar adalah menghitung reliabilitas tes. Untuk menghitung reliabilitas tes hasil belajar digunakan formula K-R 20. Hasil penghitungan reliabilitas tes dengan menggunakan rumus K-R 20, diperoleh koefisien reliabilitas (r'') sebesar 0,84. Bila dibandingkan dengan kriteria yang dipersyaratkan oleh Fraenkel dan Wallen (koefisien r'' sekurang-kurangnya 0,70), maka koefisien r'' tes hasil belajar biologi yang dikembangkan lebih besar. Hal ini berarti bahwa tes hasil belajar biologi yang dikembangkan dapat digunakan lebih lanjut.

Pembahasan

Hasil analisis deskriptif kualitatif terhadap draf konstruksi tes hasil belajar biologi menyatakan dalam konstruksi tes hasil belajar biologi ditempuh langkah-langkah sebagai berikut. (a) Analisis kurikulum, (b) tujuan pengembangan tes hasil belajar biologi, (c) pengembangan materi ajar, (d) menentukan standar kompetensi, kompetensi dasar, dan indikator pembelajaran, (e) membuat kisi-kisi tes hasil belajar biologi, dan (f) menulis butir soal tes hasil belajar biologi.

Hal ini sangat sesuai dengan langkah-langkah yang ditempuh di dalam pengembangan instrumen berupa tes hasil belajar. Langkah pertama yang harus

ditempuh adalah melakukan analisis kurikulum, yang dalam hal ini adalah kurikulum bidang studi biologi SMP kelas IX. Dari analisis kurikulum ini, seorang pengembang tes akan bisa mengetahui standar kompetensi, kompetensi dasar, dan cakupan materi tes yang akan dikembangkan. Tanpa melalui analisis kurikulum, maka tes hasil belajar biologi tidak akan bisa diwujudkan. Selanjutnya, pengembang tes harus merumuskan tujuan dari tes hasil belajar biologi yang dikembangkan, apa untuk tujuan formatif, sumatif, atau seleksi. Dari tujuan ini akan dapat dikembangkan materi ajar dari tes yang bersangkutan.

Dari standar kompetensi dan kompetensi dasar suatu materi ajar yang tercantum di dalam kurikulum bidang studi biologi, selanjutnya pengembang tes harus merumuskan indikator pembelajaran pada setiap subtopik. Berbasiskan atas materi ajar, standar kompetensi, kompetensi dasar, dan indikator pembelajaran inilah pengembang tes dapat menyusun kisi-kisi tes hasil belajar biologi. Sebetulnya, butir-butir tes hasil belajar biologi yang digunakan untuk mengukur hasil belajar biologi harus diturunkan dari kisi-kisi tes hasil belajar biologi. Dalam artian, butir-butir tes hasil belajar biologi yang berhasil dikembangkan harus sesuai dengan indikator pembelajaran yang tercantum dalam kisi-kisi tes hasil belajar biologi. Hal inilah yang digunakan oleh seorang penilai (*rater*) di dalam menilai butir yang menyusun tes hasil belajar biologi dengan kisi-kisi tes yang sudah disiapkan sebelumnya.

Bila dikaji dari hasil pengujian validitas isi dan penghitungan reliabilitas respon antar *rater*-nya, tes hasil belajar biologi yang dikembangkan dalam karya inovatif ini dapat diujicobakan lebih lanjut. Hal ini disebabkan oleh tes hasil belajar biologi tersebut sudah diturunkan dari materi ajar yang sesuai, yakni pada pokok bahasan reproduksi generatif tumbuhan angiospermae. Pada hakikatnya materi ajar reproduksi generatif tumbuhan angiospermae tersebut terdiri atas tiga sub-pokok bahasan, yakni tata letak dan fungsi bagian-bagian bunga, penyerbukan (*polinasi*), dan pembuahan (*fertilisasi*). Materi ajar tersebut diturunkan menjadi kisi-kisi tes hasil belajar biologi, yang di antaranya memuat indikator yang jelas. Setiap butir pertanyaan yang menyusun tes hasil belajar biologi tersebut diturunkan dari indikator yang bersesuaian.

Di samping diturunkan dari materi ajar yang bersesuaian dan kisi-kisi tes hasil belajar biologi, tes hasil belajar yang dikembangkan juga dinilai oleh dua orang *rater* yang menguasai dalam bidang biologi, terutama pada pokok bahasan reproduksi generatif tumbuhan angiospermae. Penguasaan materi yang menyusun tes hasil belajar biologi oleh masing-masing *rater* sangat berpengaruh terhadap pemberian responnya terhadap tes hasil belajar biologi yang dinilai.

Bila dikaji dari tingkat kesulitan butir yang menyusun tes hasil belajar biologi, dapat dikatakan bahwa butir soal nomor: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 13, 14, 16, 17, dan 19 termasuk kategori sedang; dan soal nomor: 12, 15, 18, dan 20 termasuk kategori mudah. Hal ini disebabkan karena butir soal yang dikembangkan peneliti dan melingkupi tes hasil belajar biologi memang disesuaikan dengan kondisi sekolah, di mana kemampuan siswa yang berada pada kisaran kriteria ketuntasan minimal (KKM) ke bawah. KKM bidang studi Biologi pada SMP Negeri 1 Seririt sebesar 65. Kalaupun ada siswa yang kemampuannya sedikit berada di atas 65, itu tidak berarti mampu mengerjakan butir tes yang berada pada kategori sulit.

Dari sudut kajian daya beda butir dapat dikatakan bahwa sebanyak 4 butir yang menyusun tes hasil belajar biologi termasuk daya pembeda klasifikasi jelek (*poor*), yakni nomor butir: 1, 17, 19, dan 20. Dalam artian, soal nomor: 1, 17, 19, dan 20 tidak bisa membedakan siswa yang memiliki kemampuan kelompok atas dengan siswa yang memiliki kemampuan kelompok bawah. Keempat nomor butir soal tersebut merupakan soal yang sebetulnya tergolong meragukan siswa, sehingga baik siswa yang memiliki kemampuan kelompok atas maupun kelompok bawah melakukan proses untung-untungan (*gambling*) di dalam menjawab. Proses *gambling* inilah yang menyebabkan keempat butir soal tersebut tidak memiliki daya beda yang baik.

Dari kajian peran distraktor dapat dikatakan bahwa semua distraktor dari 20 butir yang menyusun tes hasil belajar yang dikembangkan memiliki kategori baik. Hal ini disebabkan oleh semua distraktor dari 20 butir yang menyusun tes hasil belajar biologi dipilih oleh ≥ 2 siswa. Hal ini dapat disadari bahwa semua distraktor (pengecoh) yang menyusun setiap butir soal menggunakan istilah atau konsep yang mirip atau sesuai dengan kunci jawaban. Kemiripan konsep dan istilah yang digunakan

untuk menyusun distraktor dengan kunci jawaban menyebabkan jawaban soal tersebut tidak gampang ditebak. Karena tidak gampang ditebak, setiap butir soal harus dipikirkan oleh siswa yang akan memberikan respon terhadap tes hasil belajar biologi. Begitu siswa terkecoh dengan pilihan yang tersedia, maka siswa yang bersangkutan mudah mengalami *miskonsepsi*. *Miskonsepsi* inilah yang menyebabkan siswa memberikan respon pada pilihan jawaban yang termasuk kategori pengecoh.

Setelah dilakukan pengujian validitas butir dengan menggunakan rumus *korelasi point biserial* (r_{pbi}) terdapat dua kategori butir tes hasil belajar biologi, yaitu: (1) butir tes hasil belajar biologi kategori *drop* (gugur), yakni butir tes nomor: 13 dan 19; dan (2) butir tes hasil belajar biologi kategori *valid* (sahih), yakni butir tes nomor: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 17, 18, dan 20.

Rendahnya butir tes hasil belajar biologi yang mengalami *drop* setelah diuji dengan menggunakan analisis *korelasi point biserial* (r_{pbi}), yaitu sebanyak 2 butir sebetulnya dipengaruhi oleh tiga hal pokok, yaitu: (1) tingginya koefisien validitas isi, (2) tingginya hubungan antara skor butir dengan skor total, dan (3) tingginya keterkaitan antara butir soal dengan kisi-kisi yang disusun sebelumnya.

Tinggi-rendahnya nilai koefisien validitas isi (VI) suatu instrumen dapat digunakan untuk meramalkan banyak-sedikitnya butir yang menyusun instrumen tersebut yang mengalami *drop*. Sebagaimana sudah dikemukakan sebelumnya, bahwa nilai koefisien VI tes hasil belajar biologi sebesar 0,90. Nilai koefisien VI ini sama dengan kriteria Erwin (2001). Peneliti yang jeli, setelah menemukan nilai koefisien VI sebetulnya sudah bisa berasumsi bahwa setelah tes hasil belajar biologi tersebut disebarkan kepada siswa dan skor butirnya diuji dengan analisis *korelasi point biserial* pasti akan sedikit yang mengalami *drop*.

Sebetulnya yang paling berperan mengenai *drop* atau *valid*-nya butir yang menyusun tes hasil belajar biologi adalah adanya korelasi antara skor butir dengan skor total. Hal ini disebabkan oleh pada esensinya yang dikorelasikan dalam uji validitas butir tes hasil belajar biologi adalah skor butir setiap responden dengan skor total setiap responden. Makin tinggi keterkaitan antara skor butir dengan skor total setiap responden, maka makin tinggi nilai koefisien r . Fenomena ini sering dikenal dengan

istilah *part-whole correlation*. Atau, skor setiap butir tes hasil belajar biologi berperan sebagai penyusun skor total. Berarti bila mencari korelasi antara skor butir dengan skor totalnya, secara otomatis skor butir berkorelasi dengan skor dirinya sendiri. Dalam dunia kalibrasi instrumen, kejadian ini dikenal dengan istilah efek *spurious overlaving*.

Tingginya koefisien r dalam uji *part-whole correlation* sangat berkaitan dengan tinggi-rendahnya keterkaitan antara butir soal yang menyusun tes hasil belajar biologi dengan kisi-kisinya. Dari sini dapat dibuktikan bahwa makin tinggi koefisien VI akan dapat digunakan sebagai salah satu indikator untuk menentukan tinggi-rendahnya butir tes hasil belajar yang mengalami *drop*.

Hasil penghitungan reliabilitas tes dengan menggunakan rumus K-R 20, diperoleh koefisien reliabilitas (r'') sebesar 0,84. Tingginya reliabilitas respon tes hasil belajar biologi ini disebabkan oleh empat faktor, yaitu: (1) tingginya nilai koefisien r'' respon antar-rater, (2) tinggi-rendahnya koefisien r, (3) pembuangan butir yang *drop*, dan (4) panjang-pendeknya pertanyaan yang menyusun setiap butir tes.

Tinggi-rendahnya nilai koefisien r'' respon antar-rater tes hasil belajar biologi dapat digunakan sebagai salah satu indikator untuk menentukan tinggi-rendahnya nilai koefisien r'' tes hasil belajar biologi. Makin tinggi nilai koefisien r'' respon antar-rater, maka makin tinggi juga nilai koefisien r'' tes hasil belajar biologi.

Faktor lain yang mempengaruhi nilai koefisien r'' tes hasil belajar biologi adalah tinggi-rendahnya koefisien r setiap butir yang menyusun tes hasil belajar. Makin tinggi nilai koefisien r setiap butir yang menyusun tes hasil belajar, maka makin tinggi nilai koefisien r'' tes hasil belajar yang bersangkutan.

Butir tes yang *drop*, yakni butir tes yang nilai koefisien r-hitungnya berada di bawah nilai r-tabel tidak diikutkan dalam penghitungan koefisien r'' tes hasil belajar dengan menggunakan formula K-R 20. Hal ini akan berpengaruh terhadap nilai k (jumlah butir yang *valid*) yang menyusun formula K-R 20. Penyebut dari komponen k pada formula K-R 20 adalah k-1. Oleh karena butir yang *valid* pada tes hasil belajar biologi sebanyak 18 butir, maka penyebutnya sebesar $18-1 = 17$, sehingga $18/17 = 1,059$. Nilai ini bila dikalikan dengan $\left(1 - \frac{\sum p_i q_i}{S_i^2}\right)$ akan menghasilkan nilai r'' yang termasuk kategori tinggi.

Selain ketiga faktor yang sudah disebutkan, faktor panjang pendeknya pertanyaan yang menyusun setiap butir tes hasil belajar juga mempengaruhi tinggi rendahnya nilai koefisien r suatu instrumen. Hal ini disebabkan oleh pertanyaan yang terlalu panjang akan dapat mengaburkan makna inti dari pertanyaan yang bersangkutan. Kekaburan makna inti dari pertanyaan suatu butir yang menyusun tes hasil belajar dapat menimbulkan respon yang keliru dari penjawab tes. Kekeliruan memberikan respon sebagai akibat panjangnya pertanyaan menyebabkan rendahnya skor butir. Skor butir yang rendah menyebabkan rendahnya nilai koefisien r yang diperoleh melalui penghitungan dengan formula K-R 20. Dengan kata lain, rendahnya nilai koefisien r sampai berada di bawah kriteria Fraenkel dan Wallen (1993), maka tes hasil belajar tersebut tidak bisa digunakan lebih lanjut.

Tindak lanjut dari tes hasil belajar yang nilai koefisien r berada di bawah kriteria Fraenkel dan Wallen adalah butir tes yang menyusun tes hasil belajar yang bersangkutan harus direvisi sehingga betul-betul merupakan turunan dari kisi-kisi tes hasil belajar. Setelah itu, harus dimintakan ulang mengenai respon kepada *rater*, dan diuji-cobakan ulang kepada sejumlah siswa.

SIMPULAN

Berdasarkan atas kajian (*review*) pada bagian pengembangan tes hasil belajar, dan cakupan analisis tes hasil belajar dapat disimpulkan hal-hal sebagai berikut.

1. Dalam melakukan konstruksi tes hasil belajar biologi harus ditempuh langkah-langkah sebagai berikut. (a) Analisis kurikulum, (b) tujuan pengembangan tes hasil belajar biologi, (c) pengembangan materi ajar, (d) menentukan standar kompetensi, kompetensi dasar, dan indikator pembelajaran, (e) membuat kisi-kisi tes hasil belajar biologi, dan (f) menulis butir soal tes hasil belajar biologi.
2. Sebelum tes hasil belajar digunakan sebagai instrumen dalam pembelajaran, maka tes hasil belajar biologi tersebut harus dinilai oleh *rater* dan diuji-cobakan kepada sejumlah siswa. Hasil penilaian *rater* dan uji-coba tes hasil belajar selanjutnya dianalisis. Adapun hasil analisis data penilaian *rater* dan uji-coba pada 40 siswa adalah: (a) koefisien validitas isi (VI) dari tes hasil belajar biologi

adalah 0,90, (b) hasil penghitungan reliabilitas respon antar-rater dengan menggunakan formula Anava HOYT, diperoleh koefisien r'' sebesar 0,89, (c) butir soal nomor: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 13, 14, 16, 17, dan 19 termasuk kategori sedang; dan soal nomor: 12, 15, 18, dan 20 termasuk kategori mudah, (d) sebanyak 4 butir yang menyusun tes hasil belajar biologi termasuk daya pembeda klasifikasi jelek (*poor*), yakni nomor butir: 1, 15, 19, dan 20; sebanyak 10 butir termasuk daya pembeda klasifikasi cukup (*satisfactory*), yakni nomor butir: 2, 6, 7, 8, 9, 11, 12, 13, 14, dan 16; dan sebanyak 6 butir termasuk daya pembeda kategori baik (*good*), yakni nomor butir: 3, 4, 5, 10, 17, dan 20, (e) semua distraktor dari 20 butir yang menyusun tes hasil belajar biologi yang dikembangkan memiliki kategori baik, (f) butir tes hasil belajar biologi kategori *drop* (gugur), yakni butir tes nomor: 13 dan 19; dan butir tes hasil belajar biologi kategori *valid* (sahih), yakni butir tes nomor: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 17, 18, dan 20, dan (g) hasil penghitungan reliabilitas tes dengan menggunakan rumus K-R 20 diperoleh koefisien reliabilitas (r'') sebesar 0,84.

DAFTAR PUSTAKA

- Aqib, Zainal. 2013. *Pengembangan Keprofesional Berkelanjutan bagi Guru*. Bandung: Yrama Widya.
- Arikunto, Suharsimi. 1993. *Dasar-Dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.
- Daryanto, H. 1999. *Evaluasi Pendidikan*. Jakarta: PT Rineka Cipta.
- Djaali, H. dan Pudji Muljono. 2008. *Pengukuran dalam Bidang Pendidikan*. Jakarta: PT Gramedia Widiasarana Indonesia.
- Gregory, Robert J. 2000. *Psychological Testing: History, Principles, and Application*. Boston: Allyn and Bacon.
- Kasiram, Moh. 1984. *Teknik Analisa Item Tes Hasil Belajar dan Cara-Cara Menghitung Validitas dan Reliabilitas*. Surabaya: Usaha Nasional.
- Maba, Wayan. 2002. "Evaluasi Pembelajaran." *Makalah yang Disajikan dalam Penataran PBM Dosen Kopertis Wilayah VIII, Tanggal 27-30 Oktober 2002*.

- Puger, I Gusti Ngurah. 2010. *Standarisasi Tes Hasil Belajar Biologi*. Laporan Hasil Penelitian FKIP Universitas Panji Sakti Singaraja.
- Sudijono, Anas. 2001. *Pengantar Evaluasi Pendidikan*. Jakarta: Rajawali Pers.
- Thobromi, Muhammad dan Arif Mustofa. 2011. *Belajar & Pembelajaran: Pengembangan Wacana dan Praktik Pembelajaran dalam Pembangunan Nasional*. Yogyakarta: Ar-Ruzz Media.
- Widiyaningsih, Ni Nyoman. 2014. *Tes Hasil Belajar Sebagai Instrumen Pembelajaran dan Penelitian*. Surabaya: Usaha Nasional.