

PENGEMBANGAN INSTRUMEN PENILAIAN PRAKTIKUM GENETIKA PADA SISWA SEKOLAH MENENGAH PERTAMA (SMP)

Yulinda Asnita

Universitas Negeri Jakarta

Awaluddin Tjalla

Universitas Negeri Jakarta

ABSTRACT

The objective of the study is to develop the assessment instrument for genetic practicum for junior high school students. The sample of the study is computed by using proportionate stratified random sampling method. Conceptually, the construction of genetic practicum measures the three dimensions of abilities: cognitive, psychomotor, and affective. According to the experts and panelists, the results of the assessment items show that there are 51 items which are valid and with interrater of over 0.7 both for the exact indicators and the use of the language. In the first and second steps, the instruments have been tried out to 255 9th grade students. The result of the empiric test in the first and the second steps based on the analysis test for confirmation factors shows that some indicators have loading factor of over 0.5 with T- score of > 1,96. According to the computation of multi dimension for reliability coefficient in the first and second try outs, the score of the construction for reliability coefficient is more than 0.9 and the score of Average Variance Extracted (AVE) is above 0.4 which means that it proves the high level of the validity and reliability of the assessment instruments of genetic practicum for the Junior High School students.

Keywords:

Assessment, the developing instruments, genetic practical, the reliability construction

ABSTRAK

Tujuan dari penelitian ini adalah mengembangkan instrumen penilaian praktikum genetika pada siswa SMP. Sampel penelitian ini menggunakan metode *proportionate stratified random sampling*. Secara konseptual konstruk praktikum genetika mengukur tiga dimensi kemampuan, yaitu: pengetahuan (kognitif), keterampilan (psikomotorik), dan sikap (afektif). Pengembangan butir pernyataan komponen penilaian berdasarkan telaah pakar dan panelis menghasilkan 51 butir yang valid dan reliabilitas interrater yang cukup tinggi di atas 0,7 baik untuk ketepatan indikator dan penggunaan bahasa. Pada tahap pertama dan kedua instrumen diujicobakan pada 255 siswa kelas IX SMP. Hasil uji empiris pertama dan kedua melalui pengujian analisis faktor konfirmatori didapatkan beberapa indikator memiliki nilai *loading factor* di atas 0,5 dengan nilai $T_{hitung} > 1,96$. Dari perhitungan nilai koefisien reliabilitas multidimensi pada uji coba pertama dan kedua menunjukkan bahwa nilai koefisien reliabilitas konstruk lebih dari 0,9 dan nilai *Average Variance Extracted* (AVE) lebih dari 0,4 yang berarti tingkat validitas dan reliabilitas instrumen penilaian praktikum genetika pada siswa SMP.

Kata Kunci

Penilaian, pengembangan instrumen, praktikum genetika, reliabilitas konstruk

Alamat Korespondensi

-
e-mail:
yulindabio@gmail.com

I. Pendahuluan

Tantangan yang dihadapi oleh sistem pendidikan nasional dewasa ini menjadi kian berat dengan berbagai dinamika dan perubahan yang terjadi sebagai dampak dari proses globalisasi. Di era globalisasi pendidikan yang baik diharapkan memberikan pengalaman bagi siswa untuk

menghasilkan lulusan yang berkualitas (H.A.R Tilaar, 2003) dan mampu berkompetisi dengan masyarakat global.

Program for International Student Assessment (PISA) 2012 mencatat kecakapan anak-anak Indonesia dalam usia sekolah menengah untuk mengimplementasikan pengetahuan yang dimiliki

untuk menyelesaikan masalah-masalah di dunia nyata dalam bidang membaca, matematika, dan sains masih di bawah rata-rata (Mailizar, 2013) serupa dengan hasil survey internasional lainnya, yaitu: hasil studi *Trends in International Mathematics and Science Study* (TIMSS) untuk penguasaan IPA tahun 2011 masih di bawah rata-rata (Elin Drian, 2012). Hasil ini menunjukkan bahwa hasil belajar siswa Indonesia khususnya pada mata pelajaran IPA masih terpaku pada penguasaan materi yang hanya menghafal dan mengukur kognitif pada tingkat yang dasar. Siswa tidak menguasai soal-soal aplikasi maupun keterampilan yang diperlukan dalam menghadapi perkembangan zaman.

Sejatinya fungsi dari pendidikan merupakan sebuah proses yang menggali dan memaksimalkan semua aspek potensi pada siswa yang diperlukan dalam kehidupan bermasyarakat serta memiliki peran pengembangan diri dalam kodrat manusia (John Dewey, 2004). Siswa dapat mengembangkan kemampuannya secara maksimal sesuai dengan kompetensi yang dimiliki dan menerapkan kemampuannya itu dalam kehidupan bermasyarakat. Pendidikan tidak hanya berfungsi sebagai transfer pengetahuan akan tetapi memberikan pengalaman belajar kepada siswa untuk kehidupannya. Sejalan dengan prinsip penerapan kurikulum 2013 yang berkaitan dengan penilaian pembelajaran yang berbasis pada kompetensi, pengajaran berdasarkan kompetensi merupakan suatu sistem di mana siswa dikategorikan menyelesaikan pelajaran jika telah melaksanakan tugas yang dipelajari dan dilakukan sesuai dengan kemampuan yang dimilikinya (Suhaenah Suparno, 2001).

Kompetensi adalah suatu pengetahuan, keterampilan, kemampuan atau kapabilitas yang dimiliki oleh seseorang yang telah menjadi bagian dari dirinya sehingga mewarnai perilaku kognitif, afektif, dan psikomotoriknya (Wina Sanjaya, 2008). Berdasarkan pendapat tersebut kompetensi harus didukung oleh pengetahuan, sikap, dan apresiasi. Ada empat alasan mengenai pentingnya praktikum IPA, yaitu: (1) praktikum membangkitkan motivasi belajar IPA, (2)

praktikum mengembangkan keterampilan dasar melakukan eksperimen, (3) praktikum menjadi wahana belajar pendekatan ilmiah, dan (4) praktikum menunjang materi pelajaran (Tuti Alawiyah, 2014). Dalam kegiatan praktikum, pengembangan keterampilan proses sains memungkinkan siswa untuk membangun pengetahuan, menyelesaikan masalah, berpikir kritis, dan menemukan jawaban untuk memuaskan rasa ingin tahu tidak hanya sekedar mengingat konsep-konsep yang sudah jadi. IPA merupakan suatu disiplin ilmu yang mengkaji kehidupan dan alam sekitar, membahas aspek keilmuan dan proses IPA, berorientasi pada kegiatan laboratorium atau percobaan atau pengamatan, dan menggunakan keterampilan fisik, kemampuan imajinasi, dan kreativitas (Ministry of Education New Zealand, 1994).

Pengembangan instrumen penilaian dilakukan secara holistik pada setiap pembelajaran untuk mengukur berbagai ranah (kognitif, psikomotorik, dan afektif) dan mendorong siswa untuk lebih memahami hakikat sains secara komprehensif untuk itu perlu model pengembangan yang dapat mengukur semua aspek tersebut khususnya di sekolah menengah pertama khususnya untuk rubrik penilaian dalam kegiatan praktikum perbandingan genetika sehingga guru dapat mengukur ketiga aspek tersebut dalam kegiatan di laboratorium dengan instrumen yang objektif, valid, dan reliabel. Hal ini juga memudahkan guru dalam melakukan evaluasi kinerja siswa pada saat kegiatan praktikum.

2. Metode Penelitian

Metode penelitian ini adalah pengembangan instrumen. Pada penelitian ini diawali dengan mengkonstruksi teori menjadi 3 dimensi yang diukur untuk praktikum genetika meliputi ranah pengetahuan (kognitif) terdiri dari: 1) pengetahuan istilah pewarisan sifat, 2) pemahaman menghitung peluang genotip dan fenotip dari persilangan, dan 3) menerapkan prinsip persilangan dalam kasus penyakit hereditas. Pada ranah keterampilan (psikomotor) aspek yang dinilai meliputi: 1) mempersiapkan

praktikum, 2) melaksanakan praktikum, dan 3) kegiatan akhir praktikum. Sedangkan ranah sikap (afektif) menilai sikap: 1) ketaqwaan, 2) disiplin, 3) jujur, 4) tanggung jawab, dan 5) gotong-royong. Pakar melakukan telaah secara konstruktif dan memberikan masukan untuk perangkat instrumen serta dilanjutkan dengan uji panelis untuk ketepatan indikator dan penggunaan bahasa dihasilkan 22 indikator dengan 51 butir pernyataan. Dilanjutkan dengan uji empiris tahap satu dan dua dengan masing-masing jumlah sampel sebanyak 255 peserta yang berasal dari empat sekolah menengah pertama dan dilanjutkan dengan analisis hasil uji empiris.

3. Hasil Penelitian dan Pembahasan

Instrumen yang telah disusun dilanjutkan dengan validasi secara teori oleh pakar dan diberikan masukan pada butir pernyataan untuk ketepatan mengukur indikator dan penggunaan bahasanya. Dilanjutkan dengan uji panelis dengan teknik analisis validitas Aiken. Butir pernyataan akan valid jika nilainya di atas 0,2 (Dali Santun Naga, 2013). Dilanjutkan dengan menghitung reliabilitas antar rater menggunakan rumus Hoyt. Dari 51 butir dihasilkan semua butir valid dan memiliki reliabilitas antar rater di atas 0,7 sehingga semua butir valid dan reliabel untuk digunakan dalam uji empiris tahap pertama.

Tabel 1. Rekapitulasi Reliabilitas Panelis Butir Instrumen Penilaian Praktikum Genetika

Dimensi	Kriteria Penilaian		Status
	Ketepatan Butir Dengan Indikator reliabilitas antar rater r_{11}	Ketepatan Penggunaan Bahasa reliabilitas antar rater r_{11}	
Kognitif	0,79	0,95	Reliabel
Psikomotorik	0,97	0,97	Reliabel
Afektif	0,98	0,81	Reliabel

Uji empiris tahap pertama menunjukkan terdapat sembilan butir pernyataan yang tidak valid dan harus didrop karena mempunyai nilai t_{hitung} lebih kecil dari 1,7. Butir tersebut adalah butir nomor 12, 16, 17, 31, 36, 37, 38, 39, dan 47. Butir-butir yang drop dibuang dan tidak dihitung dengan CFA dan terdapat satu indikator pada aspek afektif atau sikap yang hilang, yaitu: ketaqwaan pada butir 37 dan 38. Butir-butir yang menghasilkan nilai $t_{hitung} < 1,7$ dalam penilaiannya adalah skor yang memiliki nilai yang umumnya homogen atau semua siswa melakukan prosedur praktikum yang sama dalam kelompoknya ketika pemantauan penilaian psikomotorik. Hilangnya indikator ketaqwaan karena kedua butir, yaitu: 37 dan 38 dilakukan hampir oleh seluruh siswa ketika memulai pembelajaran dengan berdoa dan

mengucapkan salam merupakan kebiasaan awal untuk memulai pelajaran sehingga hasil $t_{hitung} < 1,7$ dan indikator tersebut tidak valid karena sudah menjadi pembiasaan dalam pembinaan sikap spiritual.

Penelitian dilanjutkan dengan uji reliabilitas untuk menunjukkan konsistensi suatu alat pengukur di dalam mengukur gejala yang sama dengan koefisien Alpha Cronbach (*Cronbach's alpha*) dengan hasil di atas 0,70. Tahap berikutnya dilakukan uji kecocokan model dengan *Confirmatory Factor Analysis* (CFA) terhadap 22 indikator dan 42 butir menggunakan *Maximum Likelihood* dengan menguji kesesuaian (fit) model. Terdapat lima ukuran GOF yang menunjukkan hasil kecocokan yang kurang, yaitu: *Chi Square*, GFI, RMSEA, RMR, dan AGFI.

Tabel 2. Nilai Ukuran Derajat Kecocokan Model Secara Keseluruhan pada Model Uji Coba I

Ukuran Goodness of Fit	Nilai	Keterangan
A. Absolute Fit Measures		
1. Chi Square	847,53	kurang fit
2. GFI	0,77	< 0,80 → kurang fit
3. RMSEA	0,11	> 0,08 → kurang fit
4. RMR	0,10	> 0,05 → kurang fit
B. Incremental Fit Measures		
1. AGFI	0,71	< 0,80 → kurang fit
2. NFI	0,88	< 0,90 → marginal fit
3. TLI / NNFI	0,89	≤ 0,90 → marginal fit
4. CFI	0,90	≥ 0,90 → good fit
5. IFI	0,90	≥ 0,90 → good fit
6. RFI	0,86	≤ 0,90 → marginal fit
C. Parsimonious Fit Measures		
1. AIC	941,53	> 506,00 (Saturated) dan < 7465,21 (Independence Model) → good fit
2. CAIC	1154,97	< 1654,94 (Saturated) dan < 7565,12 (Independence Model) → good fit
3. ECVI	3,71	> 1,99 (Saturated) dan < 29,39 (Independence Model) → good fit
4. PGFI	0,62	> 0,60 → good fit

Tiga ukuran GOF yang menunjukkan hasil kecocokan yang marginal, yaitu: NFI, TLI/NNFI, dan RFI, dan enam ukuran GOF yang menunjukkan kecocokan yang baik, yaitu: CFI, IFI, AIC, CAIC, ECVI dan PGFI. Dengan demikian terdapat beberapa ukuran GOF yang menunjukkan kecocokan yang kurang baik, namun sebagian besar ukuran GOF menunjukkan kecocokan yang baik sehingga kecocokan keseluruhan model adalah baik (*model fit*).

Nilai *loading factor* tiap-tiap indikator dari masing-masing dimensi dapat digunakan untuk menentukan validitas konvergen. Diperoleh 5 indikator memiliki nilai *loading factor* di bawah 0,3 namun empat diantaranya menunjukkan nilai *t-value* > 1,96 sehingga seluruh butir instrumen dapat dinyatakan valid dan dapat digunakan untuk menjangkau data dari responden, kecuali pada X_{18} yang memiliki nilai *loading factor* dan *t-value* sangat kecil, maka indikator ini tidak digunakan pada pengujian selanjutnya (tidak valid). Indikator yang hilang dari X_{18} adalah penggunaan waktu menyelesaikan praktikum dan mengumpulkan laporan tepat waktu. Hasil *factor loading* menunjukkan nilai yang sangat rendah (negatif) pada indikator ini disebabkan seluruh siswa sudah

menunjukkan sikap taat pada waktu yang disepakati untuk penyelesaian praktikum begitu juga dengan pengumpulan laporannya, hal ini menyebabkan data yang diperoleh umumnya sama (homogen).

Setelah dilakukan uji validasi konstruk, maka langkah selanjutnya adalah melakukan uji reliabilitas konstruk atau *Construct Reliability* (CR) dan *Average Variance Extracted* (AVE), keduanya (CR dan AVE) menentukan reliabel atau tidaknya suatu konstruk dan model pengukuran. Nilai *Construct Reliability* (CR) sebesar 0,92 dan nilai *Average Variance Extracted* (AVE) sebesar 0,4 menurut Igbaria, et. al. sebuah konstruk mempunyai reliabilitas tinggi CR ≥ 0,70 namun nilai AVE sedikit di bawah 0,50 masih dapat dipertahankan (Magib Igbaria, 1995). Dengan demikian model pengukuran ini dapat dipercaya dan mempunyai nilai konsistensi yang baik. Jadi model pengukuran untuk mengukur kompetensi praktikum genetika siswa SMP memiliki reliabilitas yang baik atau reliabel. Model fit secara keseluruhan dengan 21 indikator dan 41 butir.

Tabel 3. Nilai *Loading Factor* dan t_{hitung} Masing-masing Indikator pada Tiap-tiap Dimensi pada Model Uji Coba I

Dimensi	Indikator	Loading Factor	t_{hitung}	Ket
Kognitif	X ₁	0,81		Valid
	X ₂	0,86	14,75	Valid
	X ₃	0,66	10,91	Valid
	X ₄	0,88		Valid
	X ₅	0,70	12,83	Valid
	X ₆	0,24	3,77	Valid
	X ₇	0,77	14,88	Valid
	X ₈	0,14	2,11	Valid
Psikomotorik	X ₉	0,26	4,03	Valid
	X ₁₀	0,67	11,97	Valid
	X ₁₁	0,66	11,73	Valid
	X ₁₂	0,66	11,84	Valid
	X ₁₃	0,81	15,99	Valid
	X ₁₄	0,81	16,05	Valid
	X ₁₅	0,77	14,90	Valid
	X ₁₆	0,21	3,24	Valid
	X ₁₇	0,74	13,97	Valid
	X ₁₈	-0,18	-2,77	Tidak Valid
Afektif	X ₁₉	0,61		Valid
	X ₂₀	0,38	4,71	Valid
	X ₂₁	0,38	4,68	Valid
	X ₂₂	0,52	6,95	Valid

Pada uji empiris tahap dua data ditabulasi dan dianalisa dengan CFA menggunakan software Lisrel 8.80. Tahap awal analisa adalah dengan melakukan uji model CFA pada seluruh skor komposit pada instrumen terhadap 21 indikator penelitian yang mengandung 41 butir dengan menggunakan metode *Maximum Likelihood*. Terdapat lima ukuran GOF yang menunjukkan hasil kecocokan

yang kurang, yaitu: *Chi Square*, GFI, RMSEA, RMR, dan AGFI. Tiga ukuran GOF yang menunjukkan hasil kecocokan yang marginal, yaitu: NFI, TLI/NNFI, dan RFI, dan enam ukuran GOF yang menunjukkan kecocokan yang baik, yaitu: CFI, IFI, AIC, CAIC, ECVI, dan PGFI disajikan pada Tabel 4.

Tabel 4. Nilai Ukuran Derajat Kecocokan Model Secara Keseluruhan pada Model Uji Coba II

Ukuran Goodness of Fit	Nilai	Keterangan
A. Absolute Fit Measures		
1. <i>Chi Square</i>	791.31	kurang fit
2. GFI	0,77	< 0,80 → kurang fit
3. RMSEA	0,11	> 0,08 → kurang fit
4. RMR	0,10	> 0,05 → kurang fit
B. Incremental Fit Measures		
1. AGFI	0,72	< 0,80 → kurang fit
2. NFI	0,88	< 0,90 → marginal fit
3. TLI / NNFI	0,89	≤ 0,90 → marginal fit
4. CFI	0,90	≥ 0,90 → good fit
5. IFI	0,90	≥ 0,90 → good it
6. RFI	0,86	≤ 0,90 → marginal fit

Ukuran Goodness of Fit	Nilai	Keterangan
C. Parsimonious Fit Measures		
1. AIC	881,31	> 462,00 (Saturated) dan < 7389,13 (Independence Model) → good fit
2. CAIC	1085,67	< 1511,03 (Saturated) dan < 7484,50 (Independence Model) → good fit
3. ECVI	3,47	> 1,82(Saturated) dan < 29,09(Independence Model) → good fit
4. PGFI	0,62	> 0,60 → good fit

Hasil ini menunjukkan meskipun terdapat ukuran GOF dengan kecocokan yang kurang baik, namun sebagian besar ukuran GOF menunjukkan kecocokan yang baik sehingga dapat disimpulkan bahwa kecocokan keseluruhan model adalah baik

(model fit). Dari hasil analisis *loading factor* dari uji empiris tahap dua dihasilkan *Construct Reliability* (CR) sebesar 0,95 dan nilai *Average Variance Extracted* (AVE) sebesar 0,43.

Tabel 5. Nilai *Loading Factor* dan t_{hitung} Masing-masing Indikator pada Tiap-tiap Dimensi pada Model Uji Coba II

Dimensi	Indikator	Loading Factor	t_{hitung}	Ket
Kognitif	X ₁	0,81		Valid
	X ₂	0,86	14,77	Valid
	X ₃	0,66	10,92	Valid
Psikomotorik	X ₄	0,88		Valid
	X ₅	0,70	12,84	Valid
	X ₆	0,24	3,82	Valid
	X ₇	0,77	14,83	Valid
	X ₈	0,13	2,09	Valid
	X ₉	0,26	4,03	Valid
	X ₁₀	0,66	11,94	Valid
	X ₁₁	0,66	11,76	Valid
	X ₁₂	0,66	11,84	Valid
	X ₁₃	0,81	15,93	Valid
	X ₁₄	0,81	16,05	Valid
	X ₁₅	0,77	14,86	Valid
	X ₁₆	0,21	3,22	Valid
	X ₁₇	0,74	14,03	Valid
	X ₁₈	0,79	8,27	Valid
Afektif	X ₁₉	0,50		Valid
	X ₂₀	0,40	3,95	Valid
	X ₂₁	0,55	4,58	Valid

Dengan demikian dapat dinyatakan bahwa CR dan AVE memenuhi kriteria, dengan nilai CR > 0,7 dan AVE mendekati 0,5 serta dihasilkan 21 indikator dengan 41 butir sebagai instrumen

untuk mengukur penilaian praktikum genetika siswa kelas IX SMP.

4. Kesimpulan

Hasil pengembangan instrumen penilaian praktikum genetika pada siswa sekolah menengah pertama yang telah teruji validitas dan reliabilitasnya diharapkan dapat memberi manfaat bagi pembaca, guru, dosen yang mengajarkan materi praktikum genetika. Instrumen penilaian praktikum genetika pada siswa sekolah menengah pertama ini diharapkan dapat dijadikan acuan oleh guru dalam melakukan penilaian kemampuan siswa baik ranah kognitif (pengetahuan), psikomotor (keterampilan), dan afektif (sikap) dalam proses pembelajaran praktikum khususnya.

5. Daftar Pustaka

- Alawiyah, Tuti. (2014). "Pengembangan Instrumen Penilaian Praktikum di Laboratorium Kimia SMA pada kurikulum 2013". Jakarta: Tesis, PPs UNJ.
- Dali S Naga. (2010). *Teori Skor pada Pengukuran Mental*. Jakarta: PPs UNJ.
- Dewey, John. (2004). *Experience and Education: Pendidikan Berbasis Pengalaman*. Jakarta: Teraju.
- Elin Drian. *Gawat Darurat Pendidikan*. Jakarta: Education Forum, 14 Desember 2012. <http://nasional.kompas.com/read/2012/12/14>, diunduh 30 Maret 2016.
- Igbaaria, Magib et. Al. (1995). Testing The Determinants of Microcomputer Usage via a Structural Equation Model. *Journal of Management Information Systems*, Vol. 11(4).
- Mailizar. *PISA 2012: Pembelajaran untuk Indonesia*. Southampton: The Jakarta Post.com, 6 Desember 2013. <http://acdipindonesia.wordpress.com/2013/12/09/pisa-2012>. diunduh 30 Maret 2016.
- Tilaar, H.A.R. (2003). *Kekuasaan dan Pendidikan: Suatu Tinjauan dari Perspektif Studi Kultural*. Magelang: Indonesiatera.
- Suparno, Suhaenah. (2001). *Membangun Kompetensi Belajar*. Jakarta: Direktorat Jenderal Pendidikan Tinggi Departemen Pendidikan Nasional.
- Sanjaya, Wina. (2008). *Pembelajaran dalam Implementasi Kurikulum Berbasis Kompetensi*. Jakarta: Kencana.
- Azwar, Saifuddin. (2012). *Penyusunan Skala Psikologi*. Yogyakarta: Pustaka Pelajar.
- Ministry of Education. (1994). *Science in The New Zealand Curriculum*. Wellington: Learning Media.