



Analisis Butir Soal *Dynamic Assessment* untuk Mengukur Pemahaman Konsep Klasifikasi Tumbuhan pada Mahasiswa

Tutik Wulandari, Murni Ramli*, Muzzazinah

Program Studi Magister Pendidikan Biologi, Universitas Sebelas Maret, Surakarta, Indonesia

*Email: mramlim@staff.uns.ac.id

DOI: 10.24815/jpsi.v10i1.22082

Article History:

Received: August 7, 2021
Accepted: January 11, 2022

Revised: December 5, 2021
Published: January 19, 2022

Abstract. Based on the study of literature and documents, the material analysis of plant classification is considered difficult. The assessments that have been used on plant classification materials have not been equipped with feedback. This study aims to determine the validity, reliability, level of difficulty, discriminatory power, detection of bias and unidimensionality of dynamic assessment instruments on plant classification materials. Development research method with the ADDIE model. The research sample was 121 biology education students who were selected by purposive random sampling. The results showed the validity of the questions in the form of 12 fit questions with the Rasch modeling having an MNSQ infit of 0.80-1.22. The reliability of the items in the 0.97 category is very good and the respondent's reliability is 0.31 in the weak category. While the level of difficulty of the questions is unknown, there are four categories, namely 3 very easy questions, 4 easy questions, 2 difficult questions, and 3 very difficult questions. The discriminatory power of questions with a separation index of 6.71 and an index of separation of questions of 0.67. Detection bias shows that gender does not discriminate against item items and unidimensionality of item items with 36.5% raw variance with sufficient category. The dynamic assessment instrument is suitable to be used to measure the understanding of the concept of plant classification in students.

Keywords: analysis of items, dynamic assessment, multiple-choice, plant classification

Pendahuluan

Ilmu biologi sebagian besar mencakup konsep abstrak, peristiwa, topik, dan fakta yang menyebabkan sulit untuk dipahami (Çimer, 2012). Selain itu, kegiatan belajar biologi seringkali dihadapkan pada materi abstrak dan di luar pengalaman sehari-hari siswa sehingga materi sulit dipahami (Jayawardana, 2017). Konsep biologi yang diajarkan tidak hanya sekedar hafalan, melainkan harus dipahami (Wulandari, dkk., 2021). Oleh karena itu, diperlukan pembelajaran yang bermakna dalam memfasilitasi mahasiswa untuk memahami konsep dengan baik. Selain itu, pemahaman konsep dapat dikembangkan dari kegiatan belajar dan pengalaman untuk meningkatkan kemampuan berpikir dalam memecahkan masalah (Amanda, dkk., 2021).

Pemahaman konsep yang baik melalui pembelajaran yang bermakna dapat membantu mahasiswa memiliki struktur kognitif lebih kompleks dan dapat menghubungkan konsep satu dengan konsep yang lain dengan mudah (Demirbas & Ertugul, 2014; Özarslan & Çetin, 2018). Salah satu konsep dasar yang perlu dikuasai oleh mahasiswa pendidikan biologi yaitu, klasifikasi dan identifikasi taksonomi tumbuhan.

Seiring perkembangan teknologi, metode yang efektif dalam proses klasifikasi tumbuhan dengan menggunakan pendekatan filogenetik berdasarkan data molekuler dan genetik (Maskour, dkk., 2019).

Berdasarkan hasil penelitian mengenai pemahaman konsep klasifikasi tanaman pada mahasiswa masih rendah dan terdapat miskonsepsi (Yangin, dkk., 2014). Hasil penelitian lain menunjukkan bahwa sebanyak 42,9% mahasiswa menganggap klasifikasi tanaman baru hanya berdasarkan fosil yang ada dan sebanyak 45,7% mahasiswa menganggap proses identifikasi tanaman hanya melihat karakteristik tanaman secara langsung (Maskour, dkk., 2016). Selain itu, rendahnya pemahaman konsep dapat dipengaruhi oleh beberapa factor, yaitu kurang fokus dalam memperhatikan pembelajaran dan penggunaan metode pembelajaran yang monoton (Vellayati, dkk., 2020). Berdasarkan beberapa hasil penelitian mengenai pemahaman konsep mahasiswa terkait klasifikasi tanaman dapat dikategorikan masih rendah sehingga perlu adanya upaya untuk meningkatkan pemahaman konsep mahasiswa. Jika mengharapkan pemahaman konsep yang baik maka pembelajaran harus ditekankan pada proses berpikir untuk dapat menyelesaikan masalah dengan baik. Selain itu, profil pemahaman konsep merupakan salah satu alat untuk menganalisis proses berpikir (El-Hani, dkk., 2015). Oleh karena itu, untuk mengetahui pemahaman konsep mahasiswa diperlukan alat ukur yang valid (Bisson, dkk., 2016).

Alat ukur yang dapat digunakan dalam mengukur pemahaman konsep, yaitu melalui penilaian. Penilaian memiliki peranan penting dalam pembelajaran, salah satu peran yang signifikan dapat digunakan untuk memotivasi tujuan belajar siswa (Derakhshan, dkk., 2015). Oleh karena itu, penilaian digunakan sebagai alat untuk membuat keputusan mengenai pencapaian siswa melalui hasil tes atau prosedur penilaian lainnya (Alemi, 2015). Selain itu, proses penilaian tidak hanya fokus terhadap hasil akhir tetapi harus berorientasi pada proses belajar. Penilaian yang berorientasi pada proses didukung oleh teori Vygotsky zona pengembangan proksimal (ZPD), siswa memerlukan *feedback* atau petunjuk langsung dalam menyelesaikan masalah (Konzulin, dkk., 2003). Oleh karena itu, *feedback* dari pendidik diperlukan sebelum siswa dapat menyelesaikan permasalahan secara mandiri.

Salah satu penilaian yang disertai dengan *feedback*, yaitu *dynamic assessment*. Penilaian tersebut digunakan untuk memperoleh informasi mengenai tingkat kemampuan siswa dan proses mencapai dalam mencapai kemampuan yang dipengaruhi oleh intervensi pendidikan tertentu (Dörfler, dkk., 2009). Tujuan *dynamic assessment* sebagai perantara dan memberikan petunjuk kepada siswa selama proses penilaian (Lantolf, dkk., 2011). Selain itu, *dynamic assessment* terfokus dalam membantu meningkatkan keterampilan siswa, menekankan pada proses penilaian, dan pemberian *feedback* sehingga proses penilaian memiliki peran utama untuk memecahkan permasalahan selama kegiatan pembelajaran (Mauludin, 2018).

Petunjuk atau bimbingan dalam *dynamic assessment* disebut dengan mediasi, mediasi merujuk pada proses mahasiswa ketika terlibat aktif dalam mencapai tujuan pembelajaran (Shrestha, dkk., 2012). Selain itu, mediasi memungkinkan adanya kolaborasi dan interaksi antara pengajar dan pembelajar dalam menyelesaikan penilaian (Mauludin, 2018). Peran mediasi dan intervensi dalam *dynamic assessment* diprioritaskan untuk meningkatkan pemahaman pembelajar sehingga diperlukan mediasi yang dapat membantu menyelesaikan permasalahan dan mendorong perkembangan kemampuan mereka (Shakki, 2016).

Dynamic assessment dianggap sebagai sarana yang efektif dalam menjembatani kesenjangan antara pembelajaran dengan penilaian dan membantu mengurangi sifat pesimis mahasiswa (Amiri, dkk., 2016). Selain itu, *dynamic assessment* digunakan untuk mengukur kemampuan mahasiswa, memperoleh informasi yang berguna memberikan intruksi yang efektif, dan membantu pengajar dalam meningkatkan potensi mahasiswanya (Konzulin, dkk., 2003).

Hasil analisis dokumen terhadap instrumen soal evaluasi pada mata kuliah klasifikasi tumbuhan yang telah digunakan di Pendidikan Biologi Universitas Sebelas Maret, yaitu instrumen soal evaluasi berupa tes tertulis dalam bentuk uraian, soal evaluasi belum dilengkapi dengan umpan balik atau *feedback*. Selain itu, soal evaluasi yang digunakan berupa kuesioner dan pertanyaan open-ended (Maskour, dkk., 2019; Yangin, dkk., 2014). Soal evaluasi yang telah digunakan belum dilengkapi dengan umpan balik atau *feedback* sehingga perlu adanya pengembangan instrumen yang mampu memberdayakan pemahaman konsep mahasiswa. Instrumen *dynamic assessment* yang dipergunakan untuk memberdayakan pemahaman konsep pada materi klasifikasi tumbuhan, sehingga diharapkan mampu membantu mahasiswa lebih memahami konsep-konsep dasar klasifikasi tumbuhan.

Tujuan penelitian ini adalah untuk mengetahui kualitas butir soal instrumen. Selanjutnya hasil analisis butir soal digunakan sebagai salah satu dasar perbaikan instrumen soal sebelum dilakukan uji coba skala luas.

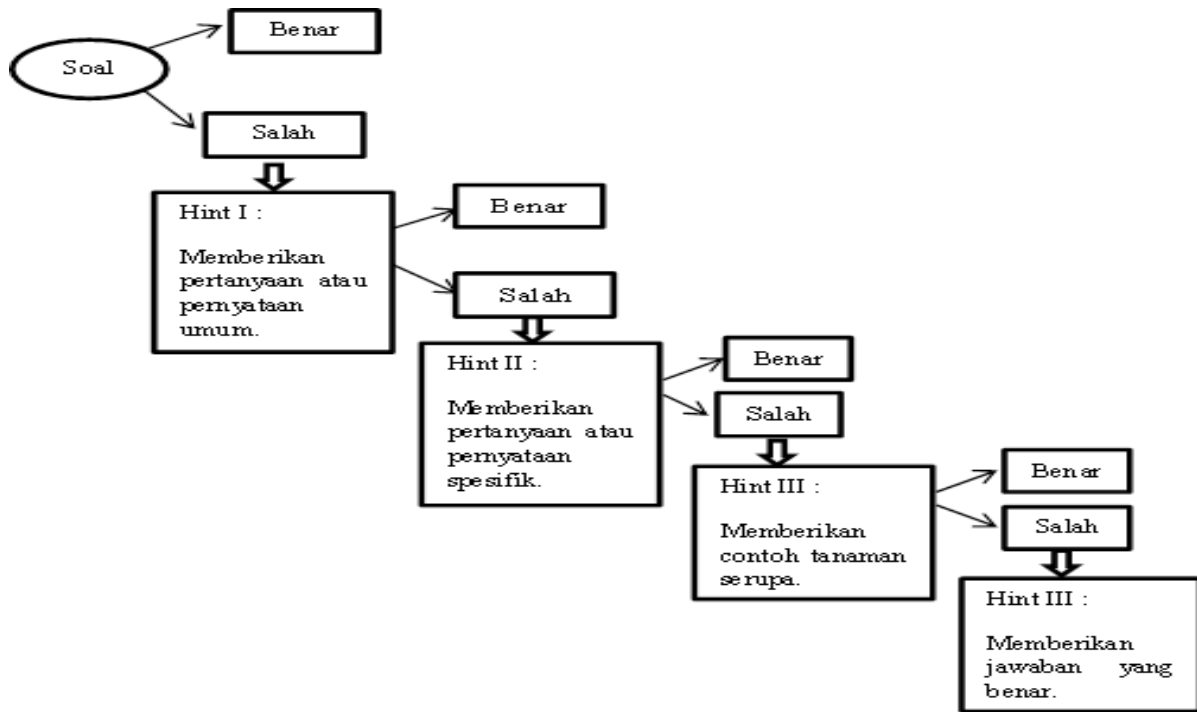
Metode

Penelitian ini merupakan penelitian deskriptif dan kuantitatif untuk mengecek kualitas instrumen. Data yang diperoleh berasal dari jawaban mahasiswa kemudian dianalisis menggunakan program *winstep*. Tahapan penelitian yang dilakukan, yaitu : 1) menyusun instrumen soal berupa multiple-choice disertai dengan umpan balik/*feedback* untuk mengetahui pemahaman konsep mahasiswa, 2) melakukan validasi terkait content materi dengan validator ahli, 3) memilih universitas yang digunakan untuk melakukan uji coba soal skala terbatas instrumen, 4) melaksanakan pengujian instrumen kepada 121 mahasiswa Pendidikan Biologi sebagai responden, 5) menganalisis hasil pengujian instrumen dengan menggunakan program *winstep*.

Sampel penelitian, yaitu mahasiswa Pendidikan Biologi Universitas Sebelas Maret semester 3 dan 5 pada tahun ajaran 2019 yang telah menempuh mata kuliah Klasifikasi dan Keanekaragaman Phanerogamae. Sampel yang digunakan 121 mahasiswa terdiri atas 67 mahasiswa semester 3 dan 54 mahasiswa semester 5. Prosedur pengumpulan data dilakukan secara *purposive random sampling*, yaitu mahasiswa yang telah menempuh mata kuliah klasifikasi tumbuhan tingkat tinggi. Teknik pengumpulan data melalui analisis dokumen penilaian yang telah digunakan, lembar observasi, dan wawancara. Teknik analisis data menggunakan permodelan rasch dengan menggunakan program *winstep*. Analisis data berkaitan dengan validasi butir soal, reliabilitas soal, daya pembeda, tingkat kesukaran soal, deteksi bias soal dan unidimensionalitas soal.

Instrumen soal yang digunakan terdiri atas 12 butir soal berupa pilihan ganda disertai dengan umpan balik atau *feedback*. Tujuan pemberian umpan balik atau *feedback* untuk membantu mahasiswa dalam membangun konsep klasifikasi tumbuhan dengan benar. Setiap soal terdiri dari empat umpan balik atau *feedback* sehingga mahasiswa memiliki empat kali kesempatan menjawab soal dengan benar. Umpan balik atau *feedback* yang diberikan berupa pertanyaan dan pernyataan terkait dasar klasifikasi tumbuhan. Konsep yang diukur pada instrumen, yaitu 1) langkah pengelompokan tumbuhan dengan benar, 2) dasar klasifikasi filogenetik berdasarkan kekerabatan takson, 3) dasar pengelompokan tumbuhan Angiospermae berdasarkan sistem APG (*Angiospermae Phylogenetic Group*), 4) dasar perbedaan tumbuhan Angiospermae dan Gymnospermae, 5) dasar klasifikasi tumbuhan Spermatophyta, 6) dasar klasifikasi dan karakteristik tumbuhan eudicot dan monokotil, 7) dasar klasifikasi dan ciri khusus pengelompokan tumbuhan pada tingkatan takson famili, 8) dasar klasifikasi pengelompokan tumbuhan tingkat genus. Tipe soal yang dikembangkan dalam penelitian ini, yaitu mahasiswa diberikan pernyataan dan pertanyaan mengenai tahapan pengelompokan hingga dasar klasifikasi tumbuhan pada tingkat genus.

Konsep yang diberikan terdiri atas konsep yang benar dan konsep yang salah kemudian mahasiswa memilih jawaban dengan benar. Selain itu, tipe soal yang diberikan berupa mahasiswa mengamati gambar beberapa tumbuhan kemudian diminta untuk memilih jawaban yang benar terkait pengelompokkan tumbuhan tersebut. Konsep pemberian *feedback* pada setiap butir soal dapat dilihat pada Gambar 1.



Gambar 1. Alur Pemberian *Feedback*

Hasil dan Pembahasan

Hasil analisis data pada penelitian ini meliputi: validitas, reliabilitas, tingkat kesukaran soal, daya pembeda, deteksi bias dan unidimensionalitas instrumen.

Validitas Butir Soal

Validitas merupakan kesesuaian suatu tes dengan apa yang seharusnya diukur, dengan kata lain instrumen tes yang digunakan akurat untuk mengukur variabel yang telah ditentukan (Ghazali, 2016). Soal memiliki validitas yang tinggi, jika dapat mengukur kompetensi yang diharapkan. Tetapi, jika soal tidak dapat mengukur kompetensi yang diharapkan maka soal memiliki validitas rendah. Soal dengan validitas tinggi mampu untuk mengukur kemampuan mahasiswa. Sebaliknya, butir soal perlu dilakukan perbaikan apabila memiliki validitas rendah (Mitha, dkk., 2018). Kesesuaian item soal dengan Model Rasch dapat dilihat dari nilai infit dan outfit MNSQ dan ZSTD. Item soal diterima jika nilai infit dan outfit MNSQ berada pada rentang 0,70-1,30 sedangkan nilai infit dan outfit ZSTD berada pada rentang -2,0 sampai +2,0 (Bond & Fox, 2007). Keseluruhan item soal yang dianalisis memiliki nilai infit MNSQ, yaitu 0,80-1,14 dan infit ZSTD, yaitu -0,68 sampai 1,22 sehingga dapat dikatakan bahwa semua item soal fit atau valid. Hasil analisis menunjukkan bahwa seluruh item soal fit atau memenuhi kriteria outfit MNSQ dan infit ZSTD sehingga

dapat disimpulkan bahwa butir soal sesuai untuk mengukur pemahaman konsep mahasiswa.

Uji validitas butir soal digunakan untuk memetakan butir soal yang dapat menyebabkan soal memiliki kriteria rendah. Butir soal dapat dikatakan memiliki validitas tinggi apabila nilai butir soal sejajar dengan nilai total (Borualogo, dkk., 2016; Mufida Nofiana, dkk., 2016). Faktor yang mempengaruhi hasil validitas butir soal, yaitu jawaban responden, penskoran, instrumen dan administrasi evaluasi (Arifin, 2011). Faktor jawaban responden diantaranya kecenderungan responden menjawab spontan, selain itu responden kesulitan mengenali bahasa ilmiah yang terdapat dalam soal.

Reliabilitas Butir Soal

Reliabilitas diartikan sebagai konsistensi dari suatu instrumen. Tes dikatakan reliabel apabila menunjukkan hasil yang sama jika diujikan pada kelompok yang sama dalam waktu berbeda. Reliabilitas soal ditentukan dari nilai *item reliability* dan *person reliability* menggunakan model rasch (Sumintono & Widhiarso, 2015). Kriteia reliabilitas pada permodelan rasch disajikan pada Tabel 1. Reliabilitas item soal dan responden disajikan pada Tabel 2. Reliabilitas item soal, yaitu 0,97 dikategorikan sangat baik sehingga dapat disimpulkan bahwa kualitas butir-butir soal dalam instrumen memiliki konsistensi yang bagus. Nilai reliabilitas responden, yaitu 0,31 dikategorikan rendah sehingga dapat disimpulkan bahwa konsistensi dalam menjawab soal mengenai konsep klasifikasi tumbuhan termasuk sangat lemah atau tidak konsisten.

Tabel 1. Kriteria reliabilitas pada permodelan rasch

Nilai Reliability (person/item)	Interpretasi
>0,94	Istimewa
0,91-0,94	Bagus sekali
0,81-0,90	Bagus
0,67-0,80	Cukup
<0,67	Lemah

Tabel 2. Reliabilitas item soal dan responden

	Rata-rata logit (SD)	Reliabilitas
Responden	0,04 (0,49)	0,31
Item soal	0,00 (1,53)	0,97

Tingkat Kesukaran Soal

Tingkat kesukaran soal digunakan untuk membedakan kemampuan peserta tes yang mampu menjawab soal dan yang tidak mampu menjawab soal dengan benar. Tingkat kesukaran soal dalam model rasch menggunakan skala logit atau nilai yang muncul bergantung dari pola respon yang diberikan sehingga mampu digunakan untuk mengelompokkan kemampuan peserta tes. Hasil distribusi tingkat kesukaran butir soal disajikan pada Tabel 3. Berdasarkan hasil distribusi tingkat kesukaran butir soal dikategorikan menjadi empat, yaitu 3 soal sangat mudah; 4 soal mudah; 2 soal sulit; dan 3 soal sangat sulit. Butir soal sebaiknya memiliki proporsi tingkat kesukaran berimbang, yaitu butir soal dengan kategori mudah sebanyak 25%; kategori sedang 50%, dan kategori sukar 25% (Amalia & Widayati, 2016). Selain itu, butir soal yang memiliki kualitas baik adalah soal yang tidak terlalu mudah dan tidak terlalu sukar. Hal ini dikarenakan soal yang terlalu mudah dapat merangsang peningkatan usaha menjawab soal dan soal yang terlalu sukar dapat menyebabkan putus asa karena diluar kemampuan peserta tes (Iskandar & Rizal, 2017).

Tabel 3. Hasil distribusi tingkat kesukaran butir soal

Nilai logit	Kategori Butir Soal	Jumlah butir soal
$Measure\ logit < -1,53$	Sangat mudah	3
$-1,53 \leq Measure\ logit \leq 0,00$	Mudah	4
$0,00 \leq Measure\ logit \leq 1,53$	Sulit	2
$Measure\ logit > 1,53$	Sangat sulit	3

Daya Pembeda Soal

Daya pembeda butir soal digunakan untuk mengetahui kemampuan soal membedakan peserta tes yang memiliki tingkat kemampuan tinggi dan rendah. Daya pembeda butir dalam analisis model rasch berdasarkan indeks separasi item soal. Berdasarkan indeks separation item soal, yaitu 6,15 atau dapat dikatakan bahwa item soal dapat dikategorikan dalam kelompok sangat sulit, sulit, mudah, dan sangat mudah. Indeks separation item responden tes, yaitu 0,67 atau dapat dikatakan bahwa kemampuan responden masih didominasi oleh satu kemampuan saja dan belum dapat membedakan kemampuan responden dengan baik. Berdasarkan rata-rata logit pada responden, yaitu 0,04 atau dapat dikatakan bahwa kemampuan responden cenderung lebih rendah dibanding dengan tingkat kesukaran soal. Semakin besar nilai separasi item maka kualitas instrumen dan butir soal semakin bagus karena mampu mengidentifikasi kelompok responden dan kelompok butir soal (Sumintono & Widhiarso, 2015)

Deteksi Bias Responden

Analisis deteksi bias dilakukan untuk apakah item soal mempunyai bias dalam kategori responden tertentu. Deteksi bias yang dianalisis dalam instrumen ini, yaitu data demografi jenis kelamin. Analisis deteksi bias pada model rasch menggunakan keberfungsian butir diferensial (DIF/*differential item functioning*). Penggunaan jenis kelamin untuk mendeteksi bias dikarenakan terdapat hubungan antara jenis kelamin dan kognitif seseorang (Amedu, 2015; Chung & Chang, 2017; Halpern, dkk., 2007).

Analisis bias responden dengan model rasch menggunakan nilai probabilitas, jika nilai probabilitas dibawah 0,05 menunjukkan adanya bias terhadap kelompok tertentu (Sumintono & Widhiarso, 2015). Hasil analisis bias responden dapat dilihat pada Tabel 4, nilai probabilitas di atas 0,05 sehingga jenis kelamin tidak mendiskriminatif terhadap butir soal.

Tabel 4. Hasil keberfungsian butir diferensial (DIF)

SUMMARY DIF CHI-SQUARED	PROB.	BETWEEN-CLASS/GROUP		ITEM NUMBER
		MNSQ	ZTSD	
2.7175	.0993	2.9411	1.39	1
.2607	.6096	.2700	-.28	2
.9595	.3273	1.0167	.48	3
.0040	.9497	.0040	-1.31	4
.6727	.4121	.7321	.26	5
.0471	.8281	.4335	-.04	6
1.1174	.2905	1.1588	.58	7
.1527	.6960	.1558	-.51	8
.0078	.9295	.0087	-1.21	9
.6943	.4047	.7182	.25	10
.2645	.6071	.2744	-.27	11
.0967	.7558	.0992	-.67	12

Unidimensionalitas Butir Soal

Unidimensionalitas instrumen digunakan untuk mengevaluasi apakah instrumen yang digunakan sesuai dengan apa yang seharusnya diukur. Model rasch menggunakan analisis komponen utama (*principal component analysis*) dari residual. Unidimensionalitas instrumen terpenuhi jika nilai minimal *raw varians* 20% dengan kategori cukup, lebih besar dari 40% kategori bagus, dan lebih besar dari 60% kategori istimewa.

Hasil analisis nilai *raw varians* item soal, yaitu 35,6% dengan kategori cukup sehingga instrumen sesuai untuk mengukur pemahaman konsep klasifikasi tumbuhan pada mahasiswa. Hasil analisis varians yang tidak dapat dijelaskan oleh instrumen sebesar 9,8% juga memenuhi kriteria, yaitu tidak melebihi 15%. Selain itu, untuk mengetahui ada atau tidaknya butir soal yang tidak sesuai untuk mengukur pemahaman konsep dapat dilihat nilai *eigenvalue* dan *observed* dalam *unexplained variance 1st contrast* (<3.0) sedangkan nilai *eigenvalue* item soal 1,8 sehingga dapat dikatakan tidak ada butir soal yang bermasalah atau sudah sesuai untuk mengukur pemahaman konsep. Identifikasi dimensionalitas digunakan untuk mengoptimalkan pengukuran sehingga informasi yang diberikan lebih difokuskan pada apa yang diukur (Sumintono & Widhiarso, 2015). Hasil unidimensionalitas disajikan pada Tabel 5.

Tabel 5. Hasil unidimensionalitas butir soal

	Eigenvalue	Observed	Expected
Total raw variance in observations	18.6374	100.0%	100.0%
Raw variance explained by measures	6.6374	35.6%	35.7%
Raw variance explained by persons	1.6897	9.1%	9.1%
Raw variance explained by items	4.9477	26.5%	26.6%
Raw unexplained variance (total)	12.0000	64.4%	64.3%
Unexplnd variance in 1st contrast	1.8229	9.8%	15.2%
Unexplnd variance in 2nd contrast	1.7056	9.2%	14.2%
Unexplnd variance in 3rd contrast	1.4352	7.7%	12.0%
Unexplnd variance in 4th contrast	1.1446	6.1%	9.5%
Unexplnd variance in 5th contrast	1.1102	6.0%	9.3%

Instrumen yang dikembangkan sudah memenuhi kriteria *dynamic assessment* yang menekankan pada pemberian *feedback* atau umpan balik. Tujuan pemberian umpan balik membantu mahasiswa dalam menyelesaikan permasalahan terkait klasifikasi tumbuhan dan membantu mendorong mahasiswa dalam meningkatkan kemampuan berpikirnya sehingga dapat meningkatkan perubahan pemahaman konsep dengan benar (Amiri & Saberi, 2016; Shakki, 2016). Pemberian umpan balik dalam penelitian ini memungkinkan responden untuk melakukan refleksi diri, meningkatkan motivasi belajar, dan memberikan informasi kepada pengajar untuk merancangan pembelajaran selanjutnya (Marriott, 2009).

Kriteria *dynamic assessment* untuk memberdayakan pemahaman konsep pada materi klasifikasi tumbuhan, yaitu menekankan pada pemberian umpan balik dan materi yang digunakan berdasarkan pada konsep-konsep yang sequencing nya berurutan dari konsep dasar ke konsep kompleks, dan berhubungan dengan perkembangan klasifikasi tumbuhan secara modern. Petunjuk atau bimbingan dalam *dynamic assessment* disebut dengan mediasi, yang merujuk pada proses keterlibatan aktif mahasiswa dalam mencapai tujuan pembelajaran (Shrestha & Coffin, 2012).

Proses penilaian *dynamic assessment* berbeda dengan penilaian sumatif, yaitu penilaian *dynamic assessment* lebih menekankan pada proses peningkatan keterampilan dalam menyelesaikan permasalahan dan memberikan kesempatan pada pendidik untuk melihat perkembangan pembelajar melakukan proses mediasi atau pemberian umpan balik (Mehri & Amerian, 2015; Mohamadi, 2018). Hal ini sesuai dengan kriteria instrumen yang

dikembangkan, yaitu menekankan pada membantu proses konstruksi konsep klasifikasi tumbuhan dengan benar. Selain itu, penilaian *dynamic assessment* memiliki tujuan untuk meningkatkan kemampuan dengan memberikan pembelajaran melalui pemberian umpan balik mendorong responden dapat melakukan penilaian diri serta belajar lebih aktif (Allal & Ducrey, 2000; Moore-Brown, dkk., 2006; Wang, 2010).

Dynamic assessment terbukti mampu mengatasi permasalahan dalam proses penilaian yang sudah terstandar dan mampu memberikan peningkatan pembelajaran dalam waktu jangka panjang (Aghaebrahimian, dkk., 2014). Selain itu, dapat meningkatkan kemampuan mahasiswa secara otomatis, memungkinkan pendidik untuk dapat mempersiapkan penilaian yang lebih baik melalui adanya umpan balik dan proses mediasi atau pemberian umpan balik dapat mengembangkan potensi mahasiswa (Mauludin, 2018; Miao & Mian, 2013). Oleh karena itu, memberikan informasi ilmiah yang akurat dalam proses pengembangan potensi siswa dapat digunakan mengatasi hambatan belajar melalui kegiatan mediasi, intervensi, kolaborasi dan keterlibatan aktif mahasiswa dalam pembelajaran (Mauludin, 2018).

Kesimpulan

Instrumen pemahaman konsep klasifikasi tumbuhan telah dikembangkan dan dilakukan uji coba kemudian dilakukan analisis menggunakan permodelan rasch dikategorikan memiliki kualitas baik dari segi validitas yaitu, semua butir soal *fit* atau valid berada pada 0,80-1,14. Analisis reliabilitas butir soal menunjukkan 0,97 dikategorikan sangat bagus sedangkan reliabilitas responden 0,31 dikategorikan lemah. Tingkat kesukaran soal dikategorikan menjadi empat, yaitu 3 soal sangat mudah; 4 soal mudah; 2 soal sulit dan 3 soal sangat sulit. Daya pembeda soal dengan indeks separation 6,15 dan indeks separation responden 0,67. Analisis deteksi bias menunjukkan jenis kelamin tidak mendiskriminatif item soal dengan nilai probabilitas diatas 0,05. Unidimensionalitas item soal dengan kategori cukup, yaitu 36,5% sehingga instrumen yang dikembangkan sesuai untuk mengukur pemahaman konsep klasifikasi tumbuhan.

Daftar Pustaka

- Aghaebrahimian, A., Rahimirad, M., Ahmadi, A., & Alamdari, J.K. 2014. Dynamic assessment of writing skill in advanced EFL Iranian Learners. *Procedia - Social and Behavioral Sciences*, 98: 60–67. <https://doi.org/10.1016/j.sbspro.2014.03.389>.
- Alemi, M. 2015. The impact of dynamic assessment on iranian efl s tudents ' writing self-assessment. *Tell*, 9(1):145–169.
- Allal, L. & Ducrey, G.P. 2000. Assessment of or in the zone of proximal development. *Learning and Instruction*, 10(2):137–152. [https://doi.org/10.1016/s0959-4752\(99\)00025](https://doi.org/10.1016/s0959-4752(99)00025).
- Amanda, F.F., Sumitro, S.B., Lestari, S.R., & Ibrohim 2021. Analysis of the relationship between concept mastery and problem-solving skills of pre-service biology teachers in human physiology courses. *Jurnal Pendidikan Sains Indonesia*, 9(3):421-431.
- Amedu, O.I. 2015. The Effect of gender on the achievement of students in biology using the Jigsaw method. *Journal of Education and Practice*, 6(17):176–179.

- Amiri, F. & Saberi, L. 2016. Dynamic assessment: the effect of mediated learning experience on Iranian EFL learners ' writing skills. *International Academic Journal of Humanities*, 3(2):1-9.
- Arifin, Z. 2011. *Evaluasi Pembelajaran*. Bandung: Remaja Rosdakarya.
- Bisson, M.J., Gilmore, C., Inglis, M., & Jones, I. 2016. Measuring conceptual understanding using comparative Judgement. *International Journal of Research in Undergraduate Mathematics Education*, 2(2):141-164. <https://doi.org/10.1007/s40753-016-0024-3>.
- Bond, T.G. & Fox, C.M. 2007. *Applying the rasch model: Fundamental measurement in the human sciences: Second edition*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Borualogo, I.S., Kusdiyati, Sulisworo, Susandari, & Sirodj, D.A.N. 2016. Analisis item soal UTS pedologi semester ganjil 2015-2016. *Schema: Journal of Psychological Research*, 3(1):46-57.
- Chung, L.Y. & Chang, R.C. 2017. The effect of gender on motivation and student achievement in digital game-based learning: A case study of a contented-based classroom. *Eurasia Journal of Mathematics, Science and Technology Education*, 13(6): 2309-2327. <https://doi.org/10.12973/EURASIA.2017.01227A>.
- Çimer, A. 2012. What makes biology learning difficult and effective: Students' Views. *Educational Research and Reviews*, 7(3):61-71. <https://doi.org/10.5897/ERR11.205>.
- Demirbas, M. & Ertugul, N. 2014. A study on preschoolers ' conceptual perceptions of states of matter : a case study of Turkish students. *South African Journal of Education*, 34(3): 1-13.
- Derakhshan, A. & Kordjazi, M. 2015. Implications of dynamic assessment in second/foreign language contexts. *English Linguistics Research*, 4(1):41-48. <https://doi.org/10.5430/elr.v4n1p41>.
- Dörfler, T., Golke, S., & Artelt, C. 2009. Dynamic assessment and its potential for the assessment of reading competence. *Studies in Educational Evaluation*, 35(2-3): 77-82. <https://doi.org/10.1016/j.stueduc.2009.10.005>.
- El-Hani, C.N., Amaral, E.M.R., do Sepulveda, C., & Mortimer, E.F. 2015. Conceptual profiles: theoretical-methodological grounds and empirical studies. *Procedia - Social and Behavioral Sciences*, 167:15-22. <https://doi.org/10.1016/j.sbspro.2014.12.636>.
- Halpern, D.F., Benbow, C.P., Geary, D.C., Gur, R.C., Hyde, J.S., & Gernsbacher, M.A. 2007. The science of sex differences in science and mathematics. *Psychological Science in the Public Interest, Supplement*, 8(1):1-51. <https://doi.org/10.1111/j.1529-1006.2007.00032>.
- Iskandar, A. & Rizal, M. 2017. Analisis kualitas soal di perguruan tinggi berbasis aplikasi TAP. *Jurnal Penelitian dan Evaluasi Pendidikan*, 21(2):12-23.
- Jayawardana, H.B.A. 2017. Paradigma pembelajaran biologi di era digital. *Jurnal Bioedukatika*, 5(1):12-18. <https://doi.org/10.26555/bioedukatika.v5i1.5628>.

- Konzulin, A., Gindis, B., Ageyev, V.S., & Suzanne M.M. 2003. *Vygotsky's Educational Theory in Cultural Context*. Cambridge: Cambridge University Press.
- Lantolf, J.P. & Poehner, M.E. 2011. Dynamic assessment in the classroom: Vygotskian praxis for second language development. *Language Teaching Research*, 15(1):11–33. <https://doi.org/10.1177/1362168810383328>.
- Marriott, P. 2009. Students' evaluation of the use of online summative assessment on an undergraduate financial accounting module. *British Journal of Educational Technology*, 40(2):237–254. <https://doi.org/10.1111/j.1467-8535.2008.00924>.
- Maskour, L., Alami, A., & Agrraki, M. 2016. Study of some learning difficulties in plant classification among university students. *The Eurasia Proceedings of Educational & Social Sciences (EPESS)*, 5:294–297.
- Maskour, L., Alami, A., Zaki, M., & Agorram, B. 2019. Plant classification knowledge and misconceptions among university students in morocco. *Education Sciences*, 9(1):1–21. <https://doi.org/10.3390/educsci9010048>.
- Mauludin, L.A. 2018. Dynamic assessment to improve students' summary writing skill in an ESP class. *Southern African Linguistics and Applied Language Studies*, 36(4):355–364. <https://doi.org/10.2989/16073614.2018.1548296>.
- Md Ghazali, N.H. 2016. A reliability and validity of an instrument to evaluate the School-Based Assessment System: A Pilot Study. *International Journal of Evaluation and Research in Education (IJERE)*, 5(2):148-157 <https://doi.org/10.11591/ijere.v5i2.4533>.
- Mehri, E. & Amerian, M. 2015. Challenges to dynamic assessment in second language learning. *Theory and Practice in Language Studies*, 5(7):1458-1466 <https://doi.org/10.17507/tpls.0507.19>.
- Miao, T. & Mian, L. 2013. Dynamic assessment in ESL writing classroom. *International Conference on Education Technology and Management Science (ICETMS 2013)*, (Icetms):676–679. <https://doi.org/10.2991/icetms.2013.1>.
- Mitha, P., Setyaningrum, P., Ramli, M., & Rinanto, Y. 2018. Analisis kualitas butir soal instrumen Assessment Diagnostic untuk mendeteksi miskonsepsi siswa SMA pada materi virus. *Jurnal Bioedukatika*, 6(2):91–101.
- Mohamadi, Z. 2018. Comparative effect of online summative and formative assessment on EFL student writing ability. *Studies in Educational Evaluation*, 59:29–40. <https://doi.org/10.1016/j.stueduc.2018.02.003>.
- Moore-Brown, B., Huerta, M., Uranga-Hernandez, Y., & Peña, E.D. 2006. Using dynamic assessment to evaluate children with suspected learning disabilities. *Intervention in School and Clinic*, 41(4):209–217. <https://doi.org/10.1177/10534512060410040301>.
- Mufida, Nofiana, & Sajidan. 2016. Pengembangan instrumen evaluasi higher order thinking skills pada materi Kingdom Plantae. *Jurnal Pedagogi Hayati*, 1(1):46–53.

- Özarslan, M. & Çetin, G. 2018. Biology Students' Cognitive Structures about Basic Components of Living Organisms. *Science Education International*, 29(2):62–74.
- Shakki, F. 2016. The interplay between language skills and dynamic assessment. *International Journal of Linguistics*, 8(2):141-157 <https://doi.org/10.5296/ijl.v8i2.9221>
- Shrestha, P. & Coffin, C. 2012. Dynamic assessment, tutor mediation and academic writing development. *Assessing Writing*, 17(1):55–70. <https://doi.org/10.1016/j.asw.2011.11.003>.
- Sumintono, B. & Widhiarso, W. 2015. *Aplikasi Permodelan Rasch pada Assessment Pendidikan*. Cimahi: Trim Komunikata.
- Vellayati, S., Nurmaliah, C., Sulastri, Yusrizal, & Saidi, N. 2020. Identifikasi Tingkat Pemahaman Konsep Siswa Menggunakan Tes Diagnostik Three-Tier Multiple Choice Pada Materi Hidrokarbon. *Jurnal Pendidikan Sains Indonesia*, 8(1):128-240.
- Wang, T.H. 2010. Web-based dynamic assessment: taking assessment as teaching and learning strategy for improving students' e-Learning effectiveness. *Computers and Education*, 54(4):1157–1166. <https://doi.org/10.1016/j.compedu.2009.11.001>.
- Wulandari, S., Gusmalini, A., & Zulfarina. 2021. Analisis miskonsepsi mahasiswa pada konsep genetika menggunakan instrumen four tier diagnostic test. *Jurnal Pendidikan Sains Indonesia*, 9(4):642-654.
- Yangin, S., Sidekli, S., & Gokbulut, Y. 2014. Prospective teachers' misconceptions about classification of plants and changes in their misconceptions during pre-service education. *Journal of Baltic Science Education*, 13(1):105–117.