

VALIDITAS DAN RELIABILITAS INSTRUMEN TES DIAGNOSTIK BERFORMAT *FOUR-TIER*  
UNTUK MATERI DINAMIKA ROTASI DAN KESETIMBANGAN BENDA TEGAR

Emita Miftakhul Jannah dan Frida U. Ermawati

Jurusan Fisika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Surabaya

Email: [emitajannah@mhs.unesa.ac.id](mailto:emitajannah@mhs.unesa.ac.id)

**Abstrak**

Tujuan dilakukan penelitian ini adalah untuk mengembangkan instrumen tes diagnostik empat tingkat (*four-tier*) untuk materi Dinamika Rotasi dan Keseimbangan Benda Tegar dan menentukan kelayakan instrumen yang telah dikembangkan tersebut ditinjau dari aspek validitas dan reliabilitas. Jenis penelitian ini adalah pengembangan dengan model R&D (*Research and Development*). Sampel Uji Coba dalam penelitian ini sebanyak 55 peserta didik yang terdiri dari dua kelas yaitu XI MIPA 2 dan XI MIPA 3. Dari hasil uji coba tersebut diketahui bahwa instrumen yang dikembangkan dinyatakan valid dan reliabel. Aspek validitas terdiri dari validitas internal dan validitas eksternal, dimana untuk validitas internal meliputi validitas isi, konstruk dan bahasa dengan hasil berturut-turut yaitu 97 %, 95 % dan 96 %. Sementara itu validitas eksternal meliputi validitas empiris isi dengan hasil *false positive* dan *false negative* sebesar 7,7 % dan 4,9 %, validitas empiris konstruk per butir diperoleh hasil bahwa 16 soal dinyatakan valid dan validitas empiris konstruk keseluruhan diperoleh  $r_{xy}$  eksperimen sebesar 0,608, dimana nilai tersebut lebih besar dari  $r_{tabel}$  0,266, sehingga dinyatakan valid. Selanjutnya adalah aspek reliabilitas dengan hasil  $r_{xy}$  eksperimen sebesar 0,646. Nilai  $r_{xy}$  eksperimen lebih besar dari  $r_{tabel}$  0,266, sehingga dinyatakan reliabel.

**Kata kunci:** Validitas, reliabilitas, tes diagnostik empat tingkat (*four-tier*)

**Abstract**

The purpose of this research is to develop a four-level diagnostic test instrument (*four-tier*) for the chapter of Dynamic Rotation and Rigid Body Equilibrium and determine the worthiness of the instrument in terms of the validity aspect and reliability aspect. This type of research is development with the R & D (Research and Development) model. The test sample in this research is 55 students of two classes is XI MIPA 2 and XI MIPA 3. From the results of the test know that the instruments developed were declared valid and reliable. The aspect of validity consists of internal validity and external validity, which for internal validity includes the validity of content, constructs and language with successive results is 97 %, 95 % and 96 %. Meanwhile, external validity includes empirical validity of contents with results of false positive and false negative is 7.7% and 4.9%, Empirical validity of construct by the item is 16 questions are declared valid and empirical validity of the overall construct is obtained  $r_{xy}$  experiment of 0,608, where the value is greater than  $r_{table}$  0,266, so it was declared valid. Next is the reliability aspect with the result  $r_{xy}$  experiment of 0.646. This value is greater than  $r_{table}$  0.266, so it is declared reliable.

**Keywords:** Validity, reliability, four-level diagnostic test instrument (*four-tier*)

**PENDAHULUAN**

Dalam kehidupan sehari-hari kita tidak terlepas dari suatu fenomena atau peristiwa Fisika yang ada di lingkungan sekitar. Sebagai contoh, ketika peserta didik (PD) diberikan fenomena yang merupakan penerapan dari konsep gerak rotasi yaitu pada materi Dinamika Rotasi dan Keseimbangan Benda Tegar. Berdasarkan pengalaman Penulis saat melakukan kegiatan Pra penelitian di SMA Negeri Jogoroto, Jombang, ketika PD diberikan sebuah sistem dimana gir sepeda bagian depan berukuran dua kali lebih besar dari ukuran gir belakang

dan keduanya dihubungkan dengan rantai yang sama, selanjutnya PD diminta untuk menentukan kecepatan linier dari kedua gir tersebut. Anggapan PD dari fenomena tersebut adalah kecepatan linier dari gir depan adalah lebih besar dari kecepatan linier gir belakang. Anggapan inilah yang disebut sebagai pemahaman awal atau *prakonsepsi* tentang konsep gerak rotasi. Sementara itu berdasarkan konsep Fisika, kecepatan linier dua gir yang serantai adalah sama. Apabila dari pemahaman awal yang dimiliki PD tersebut dengan konsep Fisika terdapat ketidaksesuaian, maka akan menimbulkan *miskonsepsi* atau salah konsep pada diri PD.

Miskonsepsi PD pada materi Dinamika Rotasi dan Kesetimbangan Benda Tegar perlu dideteksi sejak dini agar tidak mempengaruhi pemahaman PD terhadap konsep-konsep berikutnya (Anggraeni, 2017). Beberapa cara yang dapat dilakukan untuk mendeteksi miskonsepsi PD yaitu observasi, interview dan tes diagnostik (Rusilowati, 2015). Dari ketiga cara tersebut yang dianggap efektif dan mudah untuk mendeteksi miskonsepsi PD adalah dengan tes diagnostik. Rusilowati (2015) menyebutkan bentuk tes diagnostik pilihan ganda yaitu *one-tier* (satu tingkat yang menyajikan soal dan opsi pilihan jawaban, *two-tier* (dua tingkat) ditambahkan tingkat keyakinan PD dalam memilih opsi jawaban, *three-tier* (tiga tingkat) menyajikan opsi jawaban, tingkat keyakinan disertai alasan PD memilih jawaban tersebut, terakhir adalah *four-tier* (empat tingkat) dimana ditambahkan keyakinan dalam memilih alasan jawaban pada tiga tingkat. Dengan demikian dapat dikatakan bahwa tes pilihan ganda yang paling lengkap adalah *four-tier* atau empat tingkat.

Berdasarkan uraian latar belakang di atas, maka Penulis tergerak untuk mengembangkan instrumen tes diagnostik berformat *four-tier* untuk materi Dinamika Rotasi dan Kesetimbangan Benda Tegar. Adapun tujuan penelitian ini adalah menentukan kelayakan instrumen ditinjau dari aspek validitas dan reliabilitas instrumen tes diagnostik miskonsepsi berformat *four-tier*.

**METODE**

Sebelum melakukan uji coba di lapangan, terlebih dahulu dilakukan uji awal kepada 27 Mahasiswa Fisika Angkatan 2018 Universitas Negeri Surabaya, selanjutnya dari hasil tersebut dilakukan uji coba di lapangan kepada 55 PD kelas XI MIPA 2 dan XI MIPA 3 SMA Negeri Jogoroto, Jombang sebagai Uji Coba 1. Jenis penelitian ini merupakan penelitian pengembangan, dimana Penulis mengembangkan instrumen tes diagnostik *four-tier* pada materi Dinamika Rotasi dan Kesetimbangan Benda Tegar. Model pengembangan yang digunakan yaitu R&D (*Research and Development*) dengan tahapan : potensi dan masalah, desain awal produk, pengumpulan data, revisi desain, validasi produk, uji coba, analisis kelayakan (Adaptasi Sugiyono, 2015).

Analisis data yang dilakukan adalah kelayakan instrumen ditinjau dari aspek validitas dan reliabilitas. Validitas terdiri dari validitas internal dan validitas eksternal. Validitas internal dilakukan oleh dua orang dosen ahli Fisika untuk menentukan validitas isi, konstruk dan bahasa (Utari, 2018). Butir soal yang memenuhi indikator penilaian diberikan skor mulai dari 1 hingga 4. Selanjutnya total skor yang diperoleh dari dua orang validator tersebut diinterpretasikan dengan menggunakan ketentuan pada Tabel 1.

**Tabel 1.** Interpretasi Hasil Skor Validasi Internal (Riduwan & Akdon, 2013)

Persentase (%)	Kriteria Interpretasi Skor
0-20	Tidak valid
21-40	Kurang valid
42-60	Cukup valid
61-80	Valid
81-100	Sangat valid

Validitas eksternal ditinjau dari aspek validitas empiris isi dan validitas empiris konstruk (per butir soal dan keseluruhan). Analisis validitas empiris isi dilakukan dengan menghitung *false positive* (FP) dan *false negative* (FN). Persentase FP dan FN adalah kurang dari 10 % (Hestenes & Halloun, 1995). Validitas empiris konstruk ditentukan menggunakan persamaan *Pearson Product Moment*, dimana dari hasil perhitungan dengan persamaan tersebut dinyatakan valid apabila nilai  $r_{xy \text{ eksperimen}} > r_{tabel}$  (Suharsimi, 2014). Selanjutnya untuk aspek reliabilitas ditentukan dengan menggunakan persamaan *Alpha Cronbach*, dimana instrumen dinyatakan reliabel apabila nilai  $r_{xy \text{ eksperimen}} \geq r_{tabel}$ . Adapun  $r_{tabel}$  untuk penelitian ini dengan n=55 taraf kesukaran 5% adalah 0,266 (Sugiyono, 2015).

**HASIL DAN PEMBAHASAN**

Sebelum diuji cobakan di lapangan, instrumen yang dikembangkan terlebih dahulu ditentukan validitas internal dengan proses validasi yang dilakukan oleh dua orang dosen ahli Fisika. Indikator penilaian validitas internal meliputi: 1) validitas isi 2) validitas konstruk dan 3) validitas bahasa. Adapun hasil rekapitulasi hasil penilaian validitas isi oleh kedua dosen validator diberikan pada Tabel 2.

**Tabel 2.** Rekapitulasi Hasil Penilaian Validitas Empiris Isi Oleh Kedua Dosen Validator.

No	Aspek Yang Divalidasi	Validator		Persentase Validitas Isi (%)	Kriteria
		1	2		
1	Kesesuaian antara butir soal dengan materi Dinamika Rotasi dan Kesetimbangan Benda Tegar	4	4	97	Sangat Valid
2	Kesesuaian butir soal dengan indikator soal	4	4		
3	Kesesuaian antara butir soal dengan urutan materi	4	4		
4	Kejelasan batasan pertanyaan, jawaban, dan	4	3		

No	Aspek Yang Divalidasi	Validator		Persentase Validitas Isi (%)	Kriteria
		1	2		
	penjelasan alasan yang diharapkan				

Berdasarkan Tabel 2 diketahui bahwa instrumen tes diagnostik yang dikembangkan dalam penelitian ini tergolong sangat valid karena persentase validitas isi yaitu sebesar 97 %. Selanjutnya diberikan rekapitulasi penilaian validitas konstruk oleh kedua dosen validator yang disediakan pada Tabel 3.

**Tabel 3.** Rekapitulasi Hasil Penilaian Validitas Konstruk Oleh Dua Dosen Validator.

No	Aspek Yang Divalidasi	Validator		Persentase Validitas Konstruk (%)	Kriteria
		1	2		
1	Kejelasan petunjuk mengerjakan tes diagnostik	4	4	95	Sangat Valid
2	Kesesuaian antara kriteria butir soal dengan Taksonomi Bloom dan kompetensi dasar	4	4		
3	Butir tes diagnostik dapat mengidentifikasi miskonsepsi peserta didik	4	4		
4	Pengecoh pada pilihan alasan bersifat rasional dan homogen dengan jawaban tingkat pertama	4	3		
5	Tabel, grafik, gambar, dan sejenisnya bersesuaian dengan masalah yang disajikan	4	3		

Dari Tabel 3 diketahui bahwa validitas konstruk dinyatakan sangat valid karena memiliki skor sebanyak 95 %. Selain validitas isi dan validitas konstruk, aspek lainnya adalah validitas bahasa dengan hasil rekapitulasi data diberikan pada Tabel 4.

**Tabel 4.** Rekapitulasi Hasil Penilaian Validitas Bahasa Oleh Kedua Dosen Validator.

No	Aspek Yang Divalidasi	Validator		Persentase Validitas Bahasa (%)	Kriteria
		1	2		
1	Kalimat soal menggunakan Bahasa Indonesia yang baik dan benar	4	4	96	Sangat Valid
2	Ragam kalimat atau pertanyaan tidak menimbulkan penafsiran ganda	4	3		
3	Pertanyaan setiap butir tes dinyatakan dengan jelas dan komunikatif	4	4		

Berdasarkan Tabel 4 diketahui bahwa instrumen tes diagnostik yang dikembangkan dinyatakan sangat valid dari segi validitas bahasa karena persentase yang diperoleh sebesar 96 %. Dari ketiga validitas internal tersebut diketahui bahwa instrumen tes diagnostik miskonsepsi berformat four-tier yang dikembangkan dalam penelitian ini dinyatakan valid dan dapat digunakan untuk Uji Coba 1.

Validitas eksternal meliputi validitas empiris isi dan validitas empiris konstruk. Validitas empiris isi diperoleh melalui perhitungan FP dan FN yang disajikan pada Tabel 5 berikut ini.

**Tabel 5.** Jumlah dan Persentase FP dan FN dari Hasil Uji Coba 1.

No. Soal	FP	FN	No. Soal	FP	FN
1	1	3	12	3	4
2	4	6	13	6	4
3	4	2	14	6	1
4	5	0	15	11	1
5	5	1	16	2	2
6	2	4	17	5	3
7	2	6	18	3	5
8	4	0	19	1	4
9	5	1	20	10	1
10	2	5	21	5	0
11	0	5	22	7	2
Jumlah				93	60
Persentase (%)				7,7	4,9

Dari Tabel 5 dapat diketahui bahwa persentase FP dan FN adalah sebesar 7,7 % dan 4,9 %. Persentase keduanya adalah kurang dari 10 % (Hestenes & Halloun, 1995). Hasil tersebut juga sesuai dengan penelitian yang

dilakukan oleh Anggrayni dan Ermawati (2018) yang juga dihasilkan persentase Fpdan FN kurang dari 10 % yakni 5,6 % dan 3,4 %.

Selanjutnya untuk validitas empiris konstruk ditentukan dari validitas konstruk per butir soal dan validitas konstruk keseluruhan soal dengan menggunakan persamaan *Pearson Product Moment*. Validitas konstruk per butir soal dapat diketahui pada tabel berikut ini.

**Tabel 6.** Validitas Konstruk Per Butir Soal.

No.Soa	$r_{tabel}$	Koefisien Korelasi ( $r_{xy \text{ eksperimen}}$ )	Kriteria
1	0,266	0,337	Valid
2	0,266	0,044	Tidak valid
3	0,266	0,532	Valid
4	0,266	0,228	Tidak valid
5	0,266	0,134	Tidak valid
6	0,266	0,309	Valid
7	0,266	0,330	Valid
8	0,266	0,281	Valid
9	0,266	0,301	Valid
10	0,266	0,091	Tidak valid
11	0,266	0,403	Valid
12	0,266	0,378	Valid
13	0,266	0,341	Valid
14	0,266	0,147	Tidak Valid
15	0,266	0,253	Tidak Valid
16	0,266	0,419	Valid
17	0,266	0,478	Valid
18	0,266	0,434	Valid
19	0,266	0,332	Valid
20	0,266	0,478	Valid
21	0,266	0,532	Valid
22	0,266	0,343	Valid

Berdasarkan Tabel 6 diketahui bahwa terdapat enam soal yang dinyatakan tidak valid karena nilai  $r_{xy \text{ eksperimen}} < r_{tabel}$ . Berdasarkan hasil perhitungan validitas empiris konstruk per butir soal, hanya soal yang dinyatakan valid yang dapat digunakan untuk perhitungan reliabilitas dan kegiatan uji coba selanjutnya (Isfara, 2018). Sementara itu, untuk validitas konstruk keseluruhan soal di rekapitulasi pada Tabel 7 berikut ini.

**Tabel 7.** Validitas konstruk keseluruhan butir soal dalam penelitian ini.

Nomor	$r_{tabel}$	Koefisien Korelasi ( $r_{xy \text{ eksperimen}}$ )	Kriteria
1	0,266	0,608	Tinggi

Berdasarkan Tabel 7 diketahui bahwa validitas keseluruhan butir soal yang dikembangkan dalam penelitian ini termasuk kedalam kriteria tinggi karena nilai  $r_{xy \text{ eksperimen}} \geq r_{tabel}$ . Dari hasil tersebut dapat dikatakan bahwa instrumen yang dikembangkan dapat dinyatakan **valid** secara empiris isi. Kelayakan instrumen

selanjutnya adalah ditinjau dari aspek reliabilitas. Untuk aspek reliabilitas instrumen diperoleh dari hasil perhitungan dengan menggunakan persamaan *Alpha Cronbach*. Adapun hasil perhitungan tersebut disajikan pada Tabel 8 berikut ini.

**Tabel 8.** Reliabilitas instrumen Yang Diperoleh Dalam Penelitian Ini.

Nomor	$r_{tabel}$	Koefisien Korelasi ( $r_{xy \text{ eksperimen}}$ )	Kriteria
1	0,266	0,646	Tinggi

Berdasarkan Tabel 8, diketahui bahwa reliabilitas instrumen dalam penelitian ini tergolong tinggi karena nilai  $r_{xy \text{ eksperimen}} \geq r_{tabel}$ . dari hasil tersebut dapat dikatakan bahwa instrumen yang dikembangkan dikatakan **reliabel**.

## PENUTUP

### Simpulan

Berdasarkan analisis yang sudah dilakukan Penulis, dapat disimpulkan disimpulkan bahwa:

1. Instrumen tes diagnostik miskonsepsi berformat four-tier yang dikembangkan dalam penelitian ini dinyatakan layak karena valid dan reliabel.
2. Validitas internal meliputi validitas isi, konstruk dan bahasa secara berturut-turut dihasilkan persentase sebesar 97 %, 95 % dan 96 %.
3. Validitas eksternal meliputi validitas empiris isi diperoleh nilai FP dan FN sebesar 7,7 % dan 4,9 %. Validitas empiris konstruk per butir soal diperoleh 16 soal yang dinyatakan valid, semnetara itu untuk validitas empiris konstruk keseluruhan butir soal dihasilkan  $r_{xy \text{ eksperimen}}$  sebesar 0,608 yakni lebih besar dari  $r_{tabel}$  0,266, sehingga dinyatakan valid.
4. Reliabilitas instrumen yang dihasilkan dalam penelitian ini yaitu  $r_{xy \text{ eksperimen}}$  sebesar 0,6446 dimana hasil tersebut lebih besar dari  $r_{tabel}$  0,266, sehingga dinyatakan reliabel.

### Saran

Berdasarkan penelitian yang telah dilakukan, saran Penulis kepada Peneliti selanjutnya adalah mengembangkan butir soal yang memiliki potensi untuk dikembangkan lebih jauh lagi. Melakukan penelitian lanjutan dengan mengujicobakan insrtumen yang telah dikembangkan oleh Penulis kepada sampel yang berbeda dan lebih banyak lagi.

### DAFTAR PUSTAKA

- Anggraeni, D. M., & Suliyannah. (2017). *Diagnosis Miskonsepsi Siswa Pada Materi Momentum , Impuls , Dan Tumbukan Menggunakan Three-Tier*

Diagnostic Test. *Inovasi Pendidikan Fisika*, 06(03), 271–274.

Anggrayni, S., & Ermawati, F.U. (2019). The validity of Four-Tier 's misconception diagnostic test for Work and Energy concepts The validity of Four-Tier 's misconception diagnostic test for Work and Energy concepts. *Journal of Physics: Conf. Series*1171 (2019) 012037. <https://doi.org/10.1088/1742-6596/1171/1/012037>

Arikunto, Suharsimi. 2014. *Prosedur Penelitian Suatu Pendekatan Praktik (Edisi Revisi)*. Jakarta: Rineka Cipta.

Hestenes, D., & Halloun, I. (1995). Interpreting the Force Concept Inventory. *The Physics Teacher*, 33, 502–506. [https://doi.org/10.1007/978-1-4899-1695-2\\_7](https://doi.org/10.1007/978-1-4899-1695-2_7)

Isfara, L., & Ermawati, F. U. (2018). Validitas Instrumen *Four-Tier* Misconception Diagnostic Test untuk Materi Fluida Statis. *Inovasi Pendidikan Fisika*, 07(03), 429–433.

Riduwan, & Akdon. (2013). *Rumus dan Data dalam Analisis Statistika*. (Z. Arifin, Ed.) (Kelima). Bandung: ALFABETA.

Rusilowati, A. (2015). Pengembangan Tes Diagnostik sebagai Alat Evaluasi Kesulitan Belajar Fisika. In *Prosiding Seminar Nasional Fisika dan Pendidikan Fisika (SNFPF) Ke-6* (pp. 1–10). Semarang: UNNES.

Sugiyono. 2015. *Metode Penelitian Pendidikan Pendekatan Kuantitatif, Kualitatif, dan R&D*. Bandung: Alfabeta.

Utari, J. I., & Ermawati, F. U. (2018). Pengembangan Instrumen Tes Diagnostik Miskonsepsi Berformat *Four-Tier* untuk Materi Suhu, Kalor dan Perpindahannya. *Inovasi Pendidikan Fisika*, 07(03), 434–439.

Wiyono, F. M., Sugiyanto, & Yulianti, E. (2016). Identifikasi Hasil Analisis Miskonsepsi Gerak Menggunakan Instrumen Diagnostik *Three Tier* pada Siswa SMP. *Jurnal Penelitian Fisika Dan Aplikasinya*, 06(02), 61–69.

