

## Pengembangan dan Validasi Instrumen Tes untuk Mengukur Keterampilan Menyelesaikan Masalah Peserta Didik SMA pada Pelajaran Fisika

R Putri Z\*, R Sari, Jumadi, Ariswan

Program Studi Pendidikan Fisika Program Pascasarjana Universitas Negeri Yogyakarta, Jl. Colombo No 1, Karang Malang, Catur Tunggal, Sleman, Yogyakarta

\*E-mail: rahmiputriz09@gmail.com

*Received: 28 Juli 2019. Accepted: 11 Maret 2020. Published: 16 Maret 2020*

**Abstrak.** Keterampilan menyelesaikan masalah adalah salah satu keterampilan penting yang perlu dimiliki oleh peserta didik untuk dapat melanjutkan ke pendidikan yang lebih tinggi dan kebutuhan dimasa mendatang. Adapun instrumen penilaian yang baik diperlukan guru untuk mengetahui tingkat pengetahuan dan keterampilan yang dicapai peserta didik. Adapun tujuan dari penelitian ini adalah untuk menghasilkan instrumen penilaian yang valid digunakan untuk mengukur keterampilan menyelesaikan masalah peserta didik fisika berdasarkan para ahli dan uji empiris. Jenis penelitian merupakan penelitian R & D dengan pendekatan 4D. Penelitian ini menggunakan 7 orang validator dan subjek uji empiris terdiri atas 200 orang peserta didik kelas XI SMA. Data dikumpulkan dengan lembar validasi dan instrumen penilaian. Data hasil validasi ahli di analisis menggunakan V'Aiken, dan data uji empiris di analisis menggunakan teori respon butir. Adapun hasil penelitian menjelaskan bahwa instrumen penilaian yang dikembangkan telah valid menurut para ahli dengan nilai validasi antara 0,93 sampai 1. Adapun berdasarkan uji empiris didapatkan dari 15 butir pertanyaan tes terdapat 2 butir pertanyaan yang tidak valid dan harus dibuang. Nilai reliabilitas berdasarkan uji empiris adalah 0,83 berada pada kategori tinggi yang menunjukkan instrumen dapat digunakan.

**Kata Kunci:** Pengembangan dan Validasi, Instrumen Penilaian, Keterampilan Menyelesaikan Masalah

**Abstract.** Problem solving skills are one of the important skills that students need to have to be able to continue to higher education and needs in the future. The good assessment instrument is needed by the teacher to determine the level of knowledge and skills achieved by the students. The purpose of this study is to produce a valid assessment instrument used to measure problem solving skills of physics students based on experts and empirical tests. This type of research is an R & D study with a 4D approach. This study uses 7 validators and empirical test subjects consisting of 200 students of class XI senior high school. Data is collected with validation sheets and assessment instruments. The data from expert validation were analyzed using V'Aiken, and empirical test data was analyzed using item response theory. The results of the study explain that the assessment instruments developed have been valid according to experts with validation values between 0.93 to 1. As for the empirical tests obtained from 15 test questions there are 2 items that are invalid and must be discarded. The reliability value based on the empirical test is 0.83 in the high category which shows the instrument can be used.

**Keywords:** Developing and Validation, Instrument Test, Problem Solving Skills.

## 1. PENDAHULUAN

Adanya pelaksanaan proses belajar menuntut adanya pengukuran untuk mengevaluasi apakah indikator dan tujuan belajar telah tercapai atau tidak. Oleh sebab itu perlu dilakukannya penilaian. Penilaian dilakukan dengan pendekatan otentik [1]. Guru dapat memberikan tes maupun non test kepada peserta didik untuk dilakukan penilaian. Berbagai macam jenis tes yang digunakan dalam penilaian antara lain adalah tes pilihan ganda, tes jawaban singkat, tes uraian [2]. Ketika guru akan memberikan tes, guru harus menyusun tes yang dapat mengukur indikator dan tujuan pembelajaran dengan benar. [3] Tes dimaksudkan untuk mengukur kompetensi peserta didik yang telah ditetapkan dalam Standar Kompetensi Lulusan. Selain itu, [4] dijelaskan bahwa tes pencapaian sangat berguna dalam menganalisis kegagalan dalam pembelajaran, mengevaluasi hasil belajar, evaluasi kurikulum, dan memberikan motivasi kepada peserta didik untuk belajar. Oleh sebab itu, hasil tes yang diperoleh diharapkan benar-benar mampu menggambarkan kemampuan peserta didik, sehingga guru dapat membedakan mana peserta didik yang telah memenuhi dan yang belum memenuhi standar yang ada pada standar lulusan yang ditetapkan.

Pelaksanaan tes bertujuan pula membantu guru dalam memperbaiki pelaksanaan pembelajaran selanjutnya. Akan tetapi [5] mengadakan perbaikan tidak hanya untuk pembelajaran selanjutnya. Melainkan berkaitan pula kepada model tes dan sistem penskoran hasil belajar yang digunakan. Adanya model tes dan sistem penskoran yang tepat diharapkan dapat meningkatkan kualitas hasil belajar. Adapun peningkatan kualitas hasil belajar dari suatu lembaga pendidikan dapat dimulai dari peningkatan kualitas program pembelajaran dan peningkatan kualitas penilaian yang dilakukan di dalam kelas [6]. Hal ini juga dijelaskan oleh peneliti lain, [7] bahwa kualitas dari pembelajaran dapat diperbaiki dengan melakukan perbaikan asesmen yang akan digunakan dalam mengukur hasil belajar peserta didik. Oleh sebab itu, guru harus bisa memilih bentuk tes yang tepat untuk mengukur kemampuan dan keterampilan peserta didik. Guru harus memiliki pertimbangan-pertimbangan sehingga tes yang digunakan benar-benar dapat menunjukkan kemampuan peserta didik.

[8] Teknik tes dilakukan untuk melakukan penilaian terhadap pengetahuan. Penilaian dapat dilakukan secara tertulis atau lisan. Adapun secara tertulis, penilaian dapat dilakukan dengan memberikan tes berupa tes dengan pemilihan jawaban yakni berupa soal tes pilihan ganda, jawaban benar-salah, menjodohkan, dan sebab akibat, dan mengisi jawaban yakni berupa soal tes isian singkat, dan uraian. Pada pelajaran fisika, penilaian dilakukan terhadap ranah kognitif, afektif dan psikomotor. Adapun Tes uraian sangat baik digunakan untuk mengetahui pengetahuan peserta didik secara lebih mendalam. [9] Dijelaskan bahwa dengan tes uraian diharapkan peserta didik dapat mengembangkan kemampuan penalaran dalam menyelesaikan berbagai jenis masalah. Selain itu, tes uraian sering digunakan guru untuk melihat proses berfikir dari peserta didik [10].

Salah satu keterampilan yang penting pada abad 21 adalah keterampilan menyelesaikan masalah [11]. [12] Dijelaskan bahwa memiliki kemampuan untuk menyelesaikan masalah adalah suatu hal penting dari pembelajaran Fisika, dan merupakan hal yang sangat penting dalam pemahaman prinsip dan konsep. Adapun indikator keterampilan menyelesaikan masalah [13] yakni; memahami masalah, mampu menyusun rencana penyelesaian, melaksanakan rencana, dan mengecek kembali hasil. Oleh sebab itu, dalam pembuatan instrumen tes dan penilaian perlu disesuaikan dengan indikator keterampilan menyelesaikan masalah sehingga guru dapat mengidentifikasi kelebihan dan kelemahan penyelesaian masalah peserta didik dengan tepat.

Adapun berdasarkan observasi penilaian hasil belajar di sekolah diketahui bahwa dalam mengukur keterampilan menyelesaikan masalah peserta didik, ternyata guru masih menggunakan sistem penilaian, benar 1 dan salah 0 untuk soal pilihan ganda dan penilaian soal uraian dengan penilaian yang kurang jelas rubriknya. Hal ini menunjukkan bahwa terdapat kekurangan dalam pelaksanaan penilaian tes, yakni belum dimasukkannya indikator-indikator keterampilan menyelesaikan masalah yang mengharuskan peserta didik untuk menjawab pertanyaan sesuai dengan tahapan yang disebut dalam indikator penyelesaian masalah.

Berdasarkan masalah tersebut, perlunya dikembangkan instrumen tes yang tepat, yakni memasukkan indikator keterampilan yang akan diukur ke dalam instrumen tes dan menggunakan penilaian yang tepat untuk menganalisis hasil tes. Selain itu, instrumen yang digunakan haruslah telah valid dan reliabel. Ada beberapa hal yang perlu dipertimbangkan dalam mengembangkan soal tes uraian antara lain; 1) Menggunakan pertanyaan uraian untuk mengukur hasil pembelajaran yang tidak dapat diukur dengan tes pilihan ganda, 2) Membuat pertanyaan sesuai dengan keterampilan yang akan

diukur, 3) Menyusun pertanyaan dengan baik sehingga peserta didik dapat menentukan masalah yang perlu diselesaikan, 4) Menunjukkan perkiraan waktu untuk pertanyaan [14]. [4] Adapun standar dari tes terdiri atas intruksi yang detail tentang bagaimana tes harus dikerjakan-yakni cara yang tepat dalam mengerjakan tes, batas waktu mengerjakan (jika ada), dan cara guru dalam mengatasi beberapa masalah yang muncul atau disebut dengan manual tes.

Setelah melakukan pengembang instrumen tes, maka perlu dilakukan analisis terhadap kualitas dan kelayakan instrumen melalui analisis instrumen. Salah satu tujuan dilakukan analisis instrumen adalah untuk mengetahui nilai validitas dan reliabilitas instrumen. [15] Validitas dan reliabilitas adalah hal yang sangat penting dalam mengembangkan tes. Oleh sebab itu, perlu dilakukan uji validasi ahli dan uji empiris terhadap instrumen tes yang dikembangkan sehingga dapat menghasilkan instrumen yang baik. Salah satu cara analisis adalah menggunakan teori respon butir. [16] Adapun penelitian menggunakan teori respon butir memiliki kelebihan dibanding dengan teori tes klasik yakni probabilitas peserta didik atau subjek dalam menjawab butir soal dengan benar tergantung pada keterampilan peserta didik atau subjek dan karakteristik butir, [17] melalui model item respon butir dapat menterjemahkan nilai tes dan pelaporan hasil tes. Salah satu jenis analisis respon butir adalah *partial credit model*. Adapun dengan menggunakan *partial credit model* cocok untuk dikenakan pada tes prestasi, termasuk soal fisika yang membutuhkan tahap identifikasi permasalahan hingga solusi akhir [18]. [19] Selain itu, dijelaskan pula bahwa model PCM cocok untuk dikenakan pada tes uraian pelajaran fisika, dimana peserta didik diberi nilai sesuai dengan tahap-tahap jawaban yang diberikan. Misalnya, jika menjawab kurang sempurna maka diberi nilai 1, sempurna 2, dan tidak menjawab 0. Agar penilaian menjadi lebih objektif, penyusunan instrumen perlu membuat suatu rubrik pedoman penskoran. Ketika kita mengasumsikan bahwa sebuah butir mengikuti pola kredit parsial maka trait/abilitas individu lebih tinggi diharapkan memiliki skor yang lebih tinggi daripada individu yang memiliki trait/abilitas rendah.

Beberapa penelitian berkaitan dengan pengembangan instrumen tes telah dilakukan. [20] Penelitian tentang pengembangan dan analisis terhadap instrumen untuk mengukur pemahaman peserta didik pada konsep dasar sebelum penugasan biochemistry dengan instrumen tes memenuhi syarat validitas isi yang diukur dengan Confirmatory Factor Analysis dan reliabilitas yang merupakan kekonsistenan diukur dengan koefisien alfa Cronbach. Penelitian [21] menggunakan pendekatan 4D dalam pengembangan instrumen penilaian HOTS berbasis kurikulum 2013 terhadap sikap disiplin dan menghasilkan instrumen penilaian yang memiliki efektifitas tinggi dalam mengukur sikap disiplin. Penelitian [22] dengan judul *how do student understand energy in biology, chemistry, and physics? Development and validation of an assessment instrument* menghasilkan instrumen tes yang valid dan reliabel digunakan untuk menguji kemampuan peserta didik materi energi yang terdapat pada antardisiplin ilmu sains (biologi, kimia, dan fisika) dengan salah satu statistik analisis menggunakan rash model. Penelitian [23] tentang pengembangan instrumen evaluasi dengan menggunakan CIPP model dalam pengimplementasian penilaian proyek pada topik optik mengembangkan instrumen dengan menggunakan pendekatan 4D dan menggunakan V'Aiken untuk analisis hasil validasi dan ICC untuk menganalisis reliabilitas instrumen. Hasil penelitian menunjukkan bahwa instrumen evaluasi yang dikembangkan valid dan reliabel untuk digunakan.

Berdasarkan hal di atas, diketahui bahwa instrumen tes yang valid dan reliabel sangat diperlukan dalam pembelajaran sehingga dapat mengukur keterampilan dan kemampuan peserta didik dengan baik. Adapun tujuan dari penelitian ini adalah untuk menghasilkan instrumen penilaian yang valid digunakan untuk mengukur keterampilan menyelesaikan masalah peserta didik fisika berdasarkan para ahli dan uji empiris pada materi usaha dan energi.

## 2. Metode

### 2.1 Jenis Penelitian

Adapun penelitian ini merupakan jenis research and development (R & D) dengan pendekatan 4D yakni *Define, Design, Develop, Disseminate* dengan pembatasan sampai pada tahapan *Develop*. Adapun tahapan-tahapan pengembangan instrumen tes sebagai berikut:

### 2.1.1 Define

Pada tahapan define dilakukan pendefinisian terhadap instrumen yang akan dikembangkan berdasarkan kepada analisis silabus yakni analisis terhadap kompetensi dasar yang dikembangkan, analisis materi yakni menuliskan indikator dan tujuan pembelajaran yang akan dimasukkan dalam instrumen tes

### 2.1.2 Design

Pada tahapan design dihasilkan draft awal pengembangan berupa lembar kisi-kisi instrumen tes yang dikembangkan, draft awal dari instrumen tes yang terdiri dari butir soal tes, dan lembar validasi instrumen tes.

### 2.1.3 Develop

Pada tahapan develop dilakukan validasi dari draft awal yang telah dihasilkan dan kemudian dilakukan revisi. Hasil instrumen tes yang telah valid kemudian dilakukan uji empiris di sekolah. Hasil dari validasi dan uji empiris kemudian dianalisis.

## 2.2 Subjek Penelitian

Adapun subjek penelitian digunakan untuk uji empiris di sekolah. Adapun subjek penelitian terdiri dari 200 peserta didik yang telah mempelajari materi usaha dan energi. 105 peserta didik SMAN 1 Sungai Penuh, dan 95 peserta didik SMAN 2 Sungai Penuh.

## 2.3 Teknik Pengumpulan dan Analisis Data

Adapun data dikumpulkan dengan menggunakan angket lembar validasi untuk memperoleh data validasi isi. Pada uji empiris digunakan lembar instrumen tes yang telah valid menurut para ahli untuk mengumpulkan data validasi berdasarkan uji empiris. Adapun data yang dikumpulkan dianalisis dengan menggunakan persamaan V'aiken untuk menghitung nilai validasi isi dari ahli dan menggunakan teori respon butir untuk menghitung nilai validasi berdasarkan uji empiris.

### 2.3.1 Validasi Isi Instrumen Tes

Validasi yang digunakan berupa validitas isi. Data yang diperoleh dari validasi instrumen penilaian oleh validator, di analisis menggunakan persamaan V'Aiken [24] yakni:

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

s = r - Io

Io = Angka Penilaian validitas yang terendah (dalam hal ini 1)

c = Angka penilaian validitas yang tertinggi (dalam hal ini 3)

r = angka yang diberikan oleh seorang penilai

n = jumlah validator

Adapun nilai minimal validasi dengan 3 kriteria pilihan disajikan dalam tabel 1.

**Tabel 1.** Nilai Minimal Validasi

Jumlah Validator	Nilai Minimal
5	1.00
6	0.92
7	0.86
8	0.81

Setelah mendapatkan hasil dari persamaan, kemudian nilai dikonversi ke dalam tiga kategori dengan rentang indek aiken V dari 0-1, seperti yang disajikan di tabel 2 dibawah ini:

**Tabel 2.** Kriteria Validasi

Hasil Validasi	Kriteria Validasi
$0,8 < V \leq 1$	Tinggi
$0,4 < V \leq 0,8$	Sedang
$V \leq 0,4$	Rendah

Sumber: [25]

### 2.3.2 Validasi Instrumen Tes berdasarkan Uji Empiris

Validasi berdasarkan uji empiris dilakukan dengan analisis respon butir dan didapat nilai validitas dan reliabilitas instrumen tes. Adapun penilaian validitas butir instrumen tes dilihat pada *item fit statistics*. Sedangkan Nilai reliabilitas tes dapat dilihat pada *Summary of item Estimates*.

### 3. Hasil dan Pembahasan

Instrumen tes dikembangkan dengan mengarahkan kepada penilaian keterampilan menyelesaikan masalah sesuai dengan indikator yang disampaikan oleh Polya [13]. Adapun konsep yang dikembangkan yakni usaha dan energi dengan rubrik penilaian dapat dilihat pada Tabel 3.

**Tabel 3.** Indikator dan Rubrik Penilaian Hasil Tes

Indikator penyelesaian masalah matematis	Aspek Penilaian	Nilai
Memahami masalah	Peserta didik menuliskan apa yang diketahui dan ditanyakan dari permasalahan	1
Menyusun rencana	Peserta didik menggunakan pengetahuan yang telah mereka pelajari untuk merencanakan penyelesaian masalah dengan menuliskan tahapan penyelesaian dan persamaan yang akan digunakan	1
Melaksanakan rencana	Peserta didik melakukan penyelesaian masalah berdasarkan rencana yang telah disusun	1
Melihat/mengecek kembali	Peserta didik mengecek kembali apa yang telah ditulis pada tahapan memahami masalah, menyusun rencana, dan melaksanakan rencana dengan cara merangkumnya dalam bentuk tulisan.	1

Berdasarkan indikator dan rubrik di atas dijelaskan bahwa jika peserta didik hanya dapat melakukan tahapan memahami masalah, maka peserta didik mendapat nilai 1. Jika dapat melakukan tahapan memahami masalah, dan menyusun rencana, maka peserta didik mendapat nilai 2, dan begitu seterusnya. Adapun jika peserta didik tidak dapat melakukan semua indikator penyelesaian masalah, maka diberi nilai 0.

Setelah instrumen tes dikembangkan, kemudian instrumen diberikan kepada 7 validator ahli untuk dilakukan validasi isi. Adapun berdasarkan hasil validasi isi dapat dilihat pada Tabel 4.

**Tabel 4.** Hasil Validasi Instrumen Tes Keterampilan Menyelesaikan Masalah

Validator (Rater)	Nomor butir pernyataan														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Dosen 1	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
Dosen 2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
Guru 1	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
Guru 2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
Teman 1	3	3	3	2	3	3	2	3	3	3	3	2	3	3	3
Teman 2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
Teman 3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
$V = \sum S/n(c-1)$	1	1	1	0,93	1	1	0,93	1	1	1	1	0,93	1	1	1

Berdasarkan Tabel 4, didapatkan bahwa setiap butir memiliki nilai validasi pada kategori tinggi, sehingga instrumen tes dapat digunakan. Adapun beberapa revisi yang dilakukan terhadap butir instrumen sebagai berikut:

- Besaran gaya yang dilakukan pada soal nomor 1 dan 3 diperbaiki dengan besaran yang sesuai

- b) Soal nomor 4 perlu ditambahkan kata “tentang”
- c) Soal nomor 5 keterangan gambar perlu diperbaiki lebih mendetail
- d) Soal nomor 11 struktur kalimat perlu diperbaiki sesuai struktur bahasa yang benar.

Adapun karakteristik dari instrumen tes yang dikembangkan dan telah valid menurut ahli antara lain: 1) Butir tes disusun berdasarkan indikator dari materi usaha dan energi dengan tujuan menguji keterampilan menyelesaikan masalah peserta didik pada materi usaha dan energi; 2) pada lembar tes diberi petunjuk pengerjaan tes; 3) lembar jawaban disusun untuk menilai keterampilan menyelesaikan masalah peserta didik dengan setiap nomor butir memenuhi indikator untuk menjawab pertanyaan berupa memahami masalah, merencanakan penyelesaian masalah, melaksanakan rencana penyelesaian masalah dan mengecek kembali hasil.

Setelah dilakukan validasi isi, kemudian dilakukan uji empiris di sekolah untuk mengetahui nilai validasi empiris. Hasil uji empiris dilakukan analisis respon butir dengan melihat keluaran pada INFIT t. Nilai infit t ini juga menandakan tingkat kesukaran dari setiap butir. Butir dikatakan valid, jika INFIT t bernilai -2 sampai 2. [15] Tingkat kesulitan bervariasi antara -2,0 sampai dengan 2,0. Item dengan tingkat kesulitan -2 menandakan item yang sangat mudah, sedangkan tingkat kesulitan 2,0 berarti item sangat sulit.

Adapun hasil validasi berdasarkan uji empiris dapat dilihat pada Tabel 5.

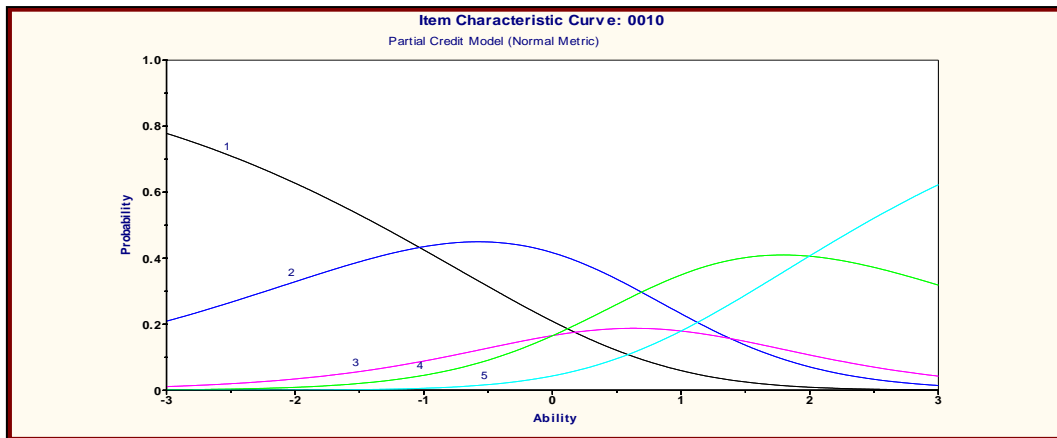
**Tabel 5.** Hasil Uji Klasik Tes Keterampilan Menyelesaikan Masalah

No. Butir Soal	INFIT MNSQ	INFIT t	interpretasi	Keterangan
1	1,21	2	Fit	Digunakan
2	1,35	2,9	Unfit	Tidak dapat digunakan
3	1,00	0,1	Fit	Digunakan
4	0,97	-0,1	Fit	Digunakan
5	0,87	-0,7	Fit	Digunakan
6	1,06	0,7	Fit	Digunakan
7	1,37	3,9	Unfit	Tidak dapat digunakan
8	0,95	-0,4	Fit	Digunakan
9	0,89	-0,8	Fit	Digunakan
10	0,86	-1,6	Fit	Digunakan
11	0,98	-0,1	Fit	Digunakan
12	0,87	-1,1	Fit	Digunakan
13	0,87	-1,0	Fit	Digunakan
14	0,82	-1,0	Fit	Digunakan
15	0,90	-0,6	Fit	Digunakan

Berdasarkan pada uji respon butir yang telah dilakukan, maka diketahui bahwa butir tes pada nomor 2 dan 7 tidak memenuhi syarat validiras sehingga tidak dapat digunakan atau harus dibuang.

Nilai reliabilitas diperlukan untuk mendukung validitas. Adapun nilai reliabilitas instrumen tes dilihat dari *reliability of estimate*. Berdasarkan *reliability of estimate* diketahui nilai reliabilitas instrumen tes sebesar 0,83 dan berada pada kategori tinggi.

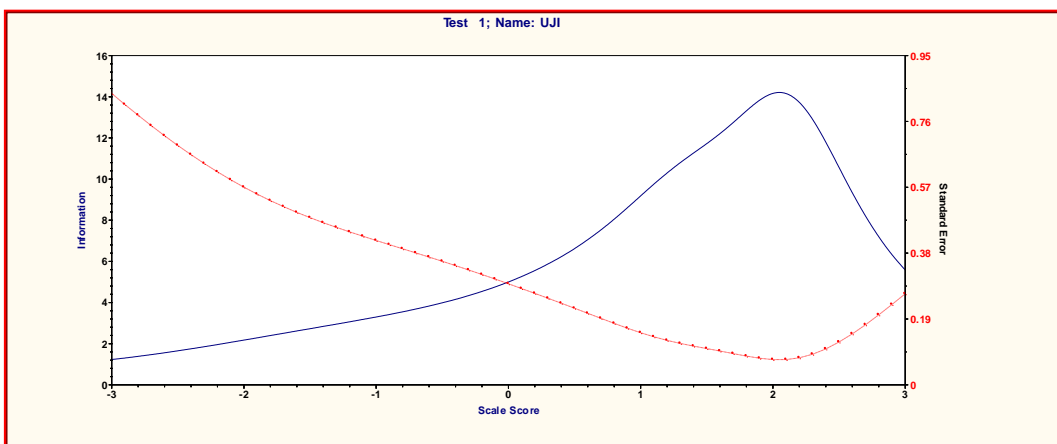
Adapun tingkat kemampuan peserta didik untuk menjawab setiap butir tes dapat dilihat pada tampilan plot ICC. Adapun contohnya yakni butir nomor 10 sebagai berikut:



**Gambar 1** Grafik Hasil Plot ICC

Adapun berdasarkan Gambar 1 dapat dijelaskan sebagai berikut: 1) Katagori 1 mewakili nilai 0, yakni menjelaskan bahwa untuk soal nomor 10, peserta didik dapat menjawab dengan skor 0 dengan kemampuan  $> -3$ ; 2) Kategori 2 mewakili nilai 1, yakni menjelaskan bahwa untuk nomor 10, peserta didik dapat menjawab dengan skor 1 dengan kemampuan  $> -0,5$ ; 3) Kategori 3 mewakili nilai 2, yakni menjelaskan bahwa untuk nomor 10, peserta didik dapat menjawab dengan skor 2 dengan kemampuan  $> 0,7$ ; 4) Kategori 4 mewakili nilai 3, yakni menjelaskan bahwa untuk nomor 10, peserta didik dapat menjawab dengan skor 3 dengan kemampuan  $> 1,9$ ; 5) Kategori 5 mewakili nilai 4, yakni menjelaskan bahwa untuk nomor 10, peserta didik dapat menjawab dengan skor 4 dengan kemampuan  $> 3$ .

Selain nilai validitas, reliabilitas, dan ICC, pada analisis respon butir dapat pula dilihat tingkat kemampuan yang diperlukan untuk menjawab seluruh butir tes. Adapun informasi ini dapat dilihat pada grafik Total Info.



**Gambar 2.** Grafik Total Info

Berdasarkan Gambar 2 menunjukkan bahwa secara keseluruhan butir soal keterampilan menyelesaikan masalah dapat diselesaikan oleh peserta didik dengan kemampuan antara 0 sampai  $>3$ .

Berdasarkan hasil-hasil yang telah di dapat, maka diketahui bahwa instrumen tes uraian yang disusun telah valid dan reliabel sehingga dapat digunakan untuk mengukur keterampilan menyelesaikan masalah peserta didik SMA pada materi usaha dan energi. Tes uraian baik digunakan, karena secara empirik dan simulasi, tes yang disajikan dalam bentuk uraian cenderung memiliki nilai fungsi informasi yang tinggi [26]. Nilai fungsi informasi yang tinggi akan sangat membantu guru dalam mengetahui karakteristik dan tingkat pengetahuan peserta didik terhadap materi atau konsep yang diajarkan. Adapun pada analisis data untuk validasi ahli berupa hasil validasi isi menggunakan V'Aiken dan analisis data uji empiris menggunakan teori respon butir. Pemilihan analisis yang tepat sangat diperlukan, agar dapat menghasilkan instrumen tes yang dapat mengukur keterampilan menyelesaikan masalah peserta didik dengan tepat. Selain itu, melakukan validasi isi memiliki peranan penting dalam pengembangan dan validasi instrumen [18]. Adapun berdasarkan penelitian

terdahulu [4] menjelaskan bahwa instrumen tes yang baik harus memenuhi kriteria validitas dan reliabilitas. Adapun Validitas dan reliabilitas merupakan sifat psikometrik penting dalam literatur pengukuran dan evaluasi pendidikan dengan yang sebelumnya lebih penting karena reliabilitas merupakan perkiraan. Hubungan antara validitas dan reliabilitas adalah bahwa suatu tes mungkin dapat diandalkan tetapi tidak valid. Tetapi tes mungkin tidak valid kecuali itu dapat diandalkan. Oleh karena itu, pengujian harus memiliki keduanya agar dianggap memuaskan untuk tujuan pengujian.

Adapun validitas isi digunakan untuk menunjukkan kualitas konten dari instrumen yang dikembangkan. [20] Validitas isi sangat penting dalam pengembangan semua jenis instrumen. Validitas isi tidak bergantung kepada nilai peserta didik, akan tetapi ditentukan oleh para ahli yang ahli dibidangnya, yakni dari para ahli diambil keputusan untuk melakukan revisi terhadap konten instrumen untuk kejelasan, kebenaran, dan relevansi dan memutuskan sejauh mana item mencerminkan domain konten. Adapun dari validasi isi diketahui apakah item/butir tes telah mencerminkan konsep yang sedang di ukur atau belum. Jika belum, maka konten dari item/butir harus dimodifikasi atau direvisi.

#### 4. Simpulan

Lembar validasi isi digunakan untuk menguji validitas berdasarkan validator. Adapun validator terdiri dari 7 orang. Hasil validasi isi menyatakan instrumen tes valid digunakan dengan nilai antara 0,93 sampai 1. Adapun penilaian berdasarkan uji empiris nilai validasi tes dilihat dari nilai infit T. berdasarkan infit T nilai validasi menunjukkan bahwa dari 15 butir soal terdapat 2 soal yang tidak memenuhi kategori valid dan harus dibuang. Nilai reliabilitas ditunjukkan instrumen oleh reliability of estimate dari uji teori respon butir. Berdasarkan reliability of estimate diketahui nilai reliabilitas instrumen tes sebesar 0,83 dan berada pada kategori tinggi. Nilai validitas dan reliabilitas ini menunjukkan bahwa instrumen tes dapat digunakan dengan membuang butir yang tidak memenuhi syara infit T.

Adapun pelaksanaan pengembangan instrumen tes dapat dikembangkan dalam berbagai bentuk evaluasi pembelajaran. Sebelum melakukan pengembangan instrumen, perlu dilakukan analisis terhadap indikator pembelajaran yang ingin di buat dalam butir tes dan mempertimbangkan jumlah setiap butir tes untuk setiap indikator, sehingga jika ada butir yang tidak valid dalam satu indikator maka dapat digantikan oleh butir tes lainnya. Selain itu, jumlah validator yang semakin besar akan menghasilkan instrumen tes yang lebih baik. Adapun analisis yang digunakan dalam menentukan validitas dan reliabilitas juga perlu dipertimbangkan, sehingga instrumen yang dihasilkan memiliki kualitas yang baik.

#### Ucapan Terima Kasih

Tim penulis mengucapkan terima kasih kepada Direktorat Riset dan Pengabdian Masyarakat, Direktorat Jendral Penguatan Riset dan Pengembangan, Kementerian Riset, Teknologi, dan Pendidikan Tinggi yang telah menyediakan fasilitas dan dana pendukung sehingga penelitian dapat dilakukan dengan baik.

#### Daftar Pustaka

- [1] Peraturan Menteri Pendidikan dan Kebudayaan 2016 *Standar proses pendidikan dasar dan menengah No 22 Tahun 2016*
- [2] B Subali 2016 *Pengembangan Tes beserta Penyelidikan Validitas dan Reliabilitas secara Empiris* (Yogyakarta: UNY Press)
- [3] D Rosana dan Sukardiyono 2015 *Penilaian jurnal Kependidikan* **45** 2 pp 130-141
- [4] A A Adeleke and E O Joshua 2015 *Journal of Education and Practice* **6** 7 pp 28-42
- [5] E Istiyono 2013 *Pengembangan instrumen untuk mengukur kemampuan berpikir tingkat tinggi dalam mata pelajaran fisika di SMA* (Laporan Penelitian Disertasi Doktor)
- [6] D Mardapi, Kumaidi and B Kartowagiran 2011 *Jurnal Penelitian dan Evaluasi Pendidikan* **15** 2 pp 326-341
- [7] B Bashooir dan Supahar 2018 *Jurnal Penelitian dan Evaluasi Pendidikan* **22** 2 pp 219-230



- [8] N D L Dewi dan Z K Prasetyo 2016 *Jurnal Inovasi Pendidikan IPA* **2** 2 pp 213-222
- [9] S Munadi 2018 *Prosiding Seminar Nasional Matematika dan Pendidikan Matematika: Integrasi Budaya, Psikologi, dan Teknologi dalam Membangun Pendidikan Karakter Melalui Matematika dan Pembelajarannya* pp 559-571
- [10] A R Nova, Parno, S Koes 2016 *Jurnal Pendidikan: Teori, Penelitian, dan Pengembangan* **1** 6 pp 1197-1203
- [11] The Partnership for 21st Century Skills 2019 *P21 Framework Definitions*
- [12] D C Giancoli 2014 *Physics: Principles with Applications (7<sup>th</sup> Edition)* (Pearson: New York)
- [13] G Polya 2004 *How to Solve It. A New Aspect of Mathematical Method with a New Foreward by John H Conway* (United State of America: Princeton Science Library)
- [14] M D Miller, R L Linn and N E Gronlund 2009 *Measurement and assessment in teaching*, (United State of America: Pearson)
- [15] T M Haladyna and M C Rodgiguez 2013 *Developing and validating test items* (New York: Routledge)
- [16] H Retnawati 2014 *Teori respons butir dan penerapannya* (Yogyakarta: Nuha Medika)
- [17] R K Hamblen, H Swaminathan and Rogers 1991 *Fundamental of item response theory* (USA: Sage)
- [18] E Istiyono, D Mardapi dan Suparno 2014 *Prosiding Konferensi Ilmiah Tahunan Himpunan Evaluasi Pendidikan Indonesia (HEPI)*
- [19] W Widhiarso 2010 *Model Politomi dalam Teori Respon Butir* (Yogyakarta: Fakultas Psikologi UGM)
- [20] S M Villafane, C P Bailey, J Loertscher, V Minderhout, and J E Lewis 2010 *Biochemistry And Molecular Biology Education* **39** 2 pp 102–109
- [21] U Pratiwi and E F Fasham 2015 *Jurnal Penelitian dan Pembelajaran IPA* **1** 1 pp 123-142
- [22] S T Opitz, K Neumann, S Bernholt and U Harms 2017 *EURASIA Journal of Mathematics Science and Technology Education* **13** 7 pp 3019-3042
- [23] J A Asfaroh, D Rosana and Supahar 2017 *AIP Conference Proceedings* **1868** 080006 pp 1-9
- [24] L R Aiken 1985 *Educational and psychological measurement* **45** 1 pp 131-142
- [25] H Retnawati 2016 *Validitas Reliabilitas & Karakteristik Butir: Panduan untuk Peneliti, Mahasiswa, dan Psikometrian* (Yogyakarta: Parama Publishing)
- [26] P Susangko 2010 *Jurnal Penelitian dan Evaluasi Pendidikan* **14** 2 pp 269-288