

URGENSI DAN ANALISIS BUTIR INSTRUMEN PENILAIAN TES

Ridwan Setiawan¹, Anis Fuad², Fitri Hilmiyati³
^{1,2,3} Pascasarjana UIN SMH Banten

Email: 232621108.ridwan@uinbanten.ac.id, 232621107.anis@uinbanten.ac.id,
fitri.hilmiyati@uinbanten.ac.id

Received : 08-12-2024

Revised : 10-12-2024

Accepted : 12-12-2024



This is Open Access article under the CC-BY-SA License (<http://creativecommons.org/licenses/by-sa/4.0/>).
Published by Pancasila and Citizenship Education Study Program, Faculty of Teacher Training and Education,
University of Palangka Raya.

Abstrak: Pendidikan merupakan usaha sadar untuk mewujudkan sesuatu pewarisan budaya dari satu generasi kepada generasi yang lainnya. Tujuan dari penulisan jurnal ini adalah untuk mengetahui bagaimana analisis butir instrumen penilaian tes. Metode yang digunakan dalam penulisan ini adalah metode penelitian kualitatif deskriptif kajian pustaka dimana dalam studi ini peneliti akan mencari sejumlah referensi baik dari buku, majalah, maupun internet untuk kemudian dikelompokkan mana data yang bersifat primer dan sekunder. Hasilnya yaitu yang pertama instrumen standar adalah alat evaluasi yang telah dikembangkan, diuji, dan terverifikasi secara sistematis untuk memastikan kualitas, keakuratan, serta keandalannya dalam mengukur kemampuan atau kompetensi peserta didik. Yang kedua Instrumen non standar dalam konteks pendidikan merujuk pada alat atau metode yang digunakan untuk mengukur atau menilai kemampuan siswa yang tidak mengikuti prosedur atau format yang sudah distandarisasi oleh lembaga atau badan tertentu. Yang ketiga analisis kualitatif butir instrumen bertujuan untuk memeriksa kualitas dan kesesuaian soal dalam mengukur kompetensi yang diinginkan. Yang keempat tingkat kesukaran dinyatakan dengan persentase siswa yang menjawab soal dengan benar. Yang kelima daya pembeda soal adalah kemampuan memisahkan siswa pandai dan siswa kurang pandai.

Kata Kunci: Pendidikan; Instrumen; Siswa

A. Pendahuluan

Pendidikan adalah salah satu patokan utama suatu negara dikatakan maju atau tidak. Pendidikan merupakan usaha sadar untuk mewujudkan sesuatu pewarisan budaya dari satu generasi kepada generasi yang lainnya (Rahman et al., 2022). Pendidikan melibatkan interaksi antara pendidik (guru atau instruktur) dan peserta didik (siswa atau mahasiswa) dalam rangka mencapai pembelajaran dan pengembangan pribadi. Pendidikan tidak hanya terjadi di lingkungan formal seperti sekolah dan perguruan tinggi, tetapi juga dapat terjadi di lingkungan informal seperti keluarga, masyarakat, dan lingkungan sekitar. Tujuan utama pendidikan adalah memberikan pengetahuan yang relevan, mengembangkan keterampilan, memfasilitasi pemahaman tentang nilai-nilai dan etika, serta membantu individu dalam mengembangkan potensi mereka secara optimal.

Pendidikan tidak hanya berfokus pada aspek akademik, tetapi juga melibatkan pengembangan sosial, emosional, dan fisik individu. Melalui pendidikan, individu dapat memperoleh pemahaman tentang dunia, mengembangkan kemampuan berpikir kritis, berkomunikasi dengan efektif, memecahkan masalah, dan berpartisipasi dalam kehidupan masyarakat. Selain itu, pendidikan juga berperan dalam membentuk sikap, moralitas, dan kepribadian individu. Pendidikan memiliki peran penting dalam membantu individu menjadi warga negara yang bertanggung jawab, menghormati perbedaan, dan berkontribusi positif terhadap masyarakat. Pendidikan di Indonesia sangatlah kompleks, dimulai dari rencana pembelajaran 1947 hingga kurikulum merdeka saat ini.

Dalam dunia pendidikan, evaluasi atau penilaian merupakan bagian yang sangat penting dalam mengukur pencapaian kompetensi peserta didik. Salah satu alat yang digunakan dalam proses evaluasi adalah instrumen tes, yang dapat berupa tes tertulis, ujian, atau penilaian lainnya. Instrumen tes ini dapat dibagi menjadi dua kategori, yaitu instrumen tes standar dan instrumen tes non-standar, yang masing-masing memiliki karakteristik dan cara penggunaan yang berbeda. Instrumen tes standar umumnya telah melalui serangkaian pengujian validitas dan reliabilitas, serta digunakan secara luas untuk tujuan evaluasi dalam skala besar. Sebaliknya, instrumen tes non-standar seringkali dibuat oleh guru atau penyusun tes untuk tujuan khusus atau dalam konteks yang lebih terbatas, seperti penilaian formatif dalam kelas.

Namun, apapun jenis instrumennya, kualitas butir-butir soal dalam tes sangat berpengaruh pada hasil evaluasi. Untuk memastikan bahwa instrumen tes yang digunakan dapat memberikan gambaran yang akurat mengenai kemampuan peserta didik, penting dilakukan analisis butir instrumen penilaian. Analisis butir instrumen tes berfungsi untuk mengevaluasi dan memastikan bahwa soal-soal dalam tes dapat

mengukur kompetensi dengan tepat, adil, dan sesuai dengan tujuan yang ingin dicapai.

B. Kajian Teori

1. Instrumen Standar dan Non Standar

a. Instrumen Standar

1) Pengertian Instrumen Standar

Instrumen standar adalah alat evaluasi yang telah dikembangkan, diuji, dan terverifikasi secara sistematis untuk memastikan kualitas, keakuratan, serta keandalannya dalam mengukur kemampuan atau kompetensi peserta didik. Instrumen ini sering digunakan dalam konteks yang lebih luas dan biasanya terstandarisasi, artinya memiliki prosedur yang baku dan konsisten dalam penerapannya.

2) Kegunaan Instrumen Standar

Menurut Suharsimi Arikunto kegunaan Instrumen Standar yaitu:

- a) Membandingkan hasil belajar dengan sifat atau karakteristik individu maupun kelompok.
- b) Membandingkan tingkat partisipasi siswa dalam keterampilan di berbagai mata pelajaran, baik untuk individu maupun kelompok.
- c) Membandingkan prestasi siswa berbagai sekolah atau kelas
- d) Menganalisis perkembangan siswa dalam jangka waktu atau periode tertentu.

3) Karakteristik Instrumen Standar

- a) Didasarkan atas bahan dan tujuan umum pendidikan di suatu negara
- b) Mencakup aspek yang luas meliputi pengetahuan dan keterampilan dengan butir soal yang sedikit
- c) Disusun dengan kelengkapan staf, professor, dan editor butir tes
- d) Menggunakan butir tes yang sudah dicobakan, dianalisis dan direvisi sebelum disajikan.
- e) Mempunyai reliabilitas yang tinggi

4) Contoh Instrumen Standar

a) Ujian Nasional

Ujian Nasional (UN) adalah ujian yang diadakan di Indonesia untuk menilai kompetensi dan pencapaian peserta didik pada akhir jenjang pendidikan tertentu, seperti di tingkat SMP/MTs (Sekolah Menengah Pertama), SMA/SMK/MA (Sekolah Menengah Atas), dan setingkatnya. Ujian ini menjadi salah satu indikator kelulusan siswa dan juga digunakan untuk mengukur keberhasilan proses pendidikan di tingkat nasional. UN diselenggarakan oleh Kementerian

Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia (Kemdikbudristek) atau Kementerian Agama (untuk jenjang pendidikan agama), dengan melibatkan berbagai lembaga pendidikan dan pengawasan secara ketat. Pada tahun 2020 UN resmi diganti oleh Penilaian Berbasis Asesmen Kompetensi yang diumumkan oleh Menteri Pendidikan dan Kebudayaan, Nadiem Makarim.

b) *TOEFL (Test of English as a Foreign Language)*

TOEFL adalah tes standar yang dirancang untuk mengukur kemampuan bahasa Inggris seseorang, terutama dalam konteks akademik. Tes ini banyak digunakan oleh universitas dan lembaga pendidikan di seluruh dunia sebagai syarat penerimaan mahasiswa internasional.

c) *TOAFL (Test of Arabic as a Foreign Language)*

TOAFL (Test of Arabic as a Foreign Language) adalah tes standar yang dirancang untuk mengukur kemampuan berbahasa Arab bagi penutur asing. Tes ini sering digunakan oleh universitas, lembaga pendidikan, dan organisasi yang membutuhkan penilaian terhadap kemampuan bahasa Arab peserta yang bukan penutur asli bahasa tersebut.

b. Instrumen Non Standar

1) Pengertian Instrumen Non Standar

Instrumen non standar dalam konteks pendidikan merujuk pada alat atau metode yang digunakan untuk mengukur atau menilai kemampuan siswa yang tidak mengikuti prosedur atau format yang sudah distandarisasi oleh lembaga atau badan tertentu. Berbeda dengan instrumen standar (seperti ujian nasional atau tes yang sudah memiliki pedoman baku), instrumen non-standar lebih fleksibel dan dapat disesuaikan dengan kebutuhan spesifik guru, sekolah, atau lembaga pendidikan lainnya. Instrumen non-standar sering digunakan dalam penilaian berbasis formatif atau alternatif, yang bertujuan untuk memberikan gambaran yang lebih menyeluruh tentang kemampuan atau perkembangan siswa dalam proses pembelajaran. Instrumen ini memungkinkan penilaian yang lebih individual dan dapat mencakup aspek-aspek yang tidak dapat diukur melalui tes standar.

2) Kegunaan Instrumen Non Standar

- a) Untuk menentukan seberapa baik siswa telah menguasai bahan Pelajaran yang diberikan dalam waktu tertentu
- b) Untuk menentukan posisi siswa dalam sebuah kelas atau kelompok.
- c) Untuk memperoleh suatu nilai

- 3) Karakteristik Instrumen Standar
 - a) Didasarkan atas bahan dan tujuan khusus yang dirumuskan guru untuk kelasnya sendiri
 - b) Dapat terjadi hanya mencakup pengetahuan atau keterampilan yang sempit dengan butir tes yang banya
 - c) Disusun sendiri oleh guru dengan sedikit bantuan atau tanpa bantuan orang lain
 - d) Jarang menggunakan butir-butir yang sudah diuji cobakan, dianalisis, dan direvisi
 - e) Mempunyai reliabilitas yang tinggi (Darwyan Syah, 2014).
- 4) Contoh Instrumen Standar
 - a) Tes Kinerja

Tes ini mengukur kemampuan siswa untuk melakukan tugas atau proyek yang menguji keterampilan mereka dalam konteks nyata. Dalam tes ini, siswa diharapkan untuk menyelesaikan tugas yang lebih praktis dan kompleks.
 - b) Tes Proyek

Tes proyek adalah jenis penilaian di mana siswa diberi tugas besar atau proyek yang berlangsung selama beberapa waktu. Tes ini memungkinkan siswa untuk menunjukkan pemahaman mendalam tentang suatu topik dengan menghasilkan produk atau hasil kerja yang relevan.
- 5) Tes Tertulis yang Tidak Baku

Tes tertulis non standar ini lebih fleksibel daripada tes pilihan ganda atau soal esai yang baku. Siswa dapat diminta untuk menjawab pertanyaan terbuka, menulis esai, atau menyelesaikan soal-soal yang mengharuskan mereka untuk berpikir kritis dan memberikan penjelasan yang lebih mendalam.

C. Metode Penelitian

Metode adalah suatu tata cara dalam mengetahui suatu hal dengan cara tersusun, sedangkan penelitian adalah pencarian awal dan kembali dalam suatu hal sampai menemukan sebuah hasil (Karmanis, 2020). Riset ini menggunakan metode penelitian kualitatif kajian pustaka, dimana peneliti akan mencari sumber data dari hasil wawancara dengan berbagai sumber dan menggali informasi berdasarkan studi dokumentasi, dimana dalam studi ini peneliti akan mencari sejumlah referensi baik dari buku, majalah, maupun internet untuk kemudian dikelompokkan mana data yang bersifat primer dan sekunder lalu diolah secara objektif (Anam et al., 2023).

D. Hasil Penelitian dan Pembahasan

1. Urgensi dan Analisis Kualitatif Butir Instrumen

Dalam konteks pengembangan instrumen penelitian atau evaluasi, urgensi kualitatif butir instrumen merujuk pada pentingnya meninjau kualitas setiap butir (item) dalam instrumen secara mendalam dari sudut pandang kualitatif. Pendekatan ini bertujuan untuk memastikan bahwa setiap butir instrumen benar-benar relevan, bermanfaat, dan mampu mengukur aspek yang dimaksudkan secara akurat dan bermakna. Untuk Analisis kualitatif butir instrumen bertujuan untuk memeriksa kualitas dan kesesuaian soal dalam mengukur kompetensi yang diinginkan. Dengan melakukan analisis kualitatif yang baik, kita dapat memastikan bahwa soal ujian atau tes yang disusun tidak hanya valid dan adil, tetapi juga efektif dalam mengukur pencapaian kompetensi siswa.

a. Keterkaitan dengan Tujuan Pembelajaran

Soal-soal ini berkaitan erat dengan tujuan pembelajaran Pendidikan Agama Islam (PAI), seperti memahami Rukun Islam, Rukun Iman, dan akhlak mulia dalam kehidupan sehari-hari. Semua soal dirancang untuk menguji sejauh mana siswa memahami dan mengamalkan ajaran Islam.

b. Kejelasan Bahasa

Bahasa dalam soal-soal ini jelas dan mudah dipahami oleh siswa. Soal pilihan ganda dan isian menggunakan kalimat yang langsung mengarah pada konsep yang ingin diuji. Pada soal esai, pertanyaan juga disusun secara jelas agar siswa tahu apa yang diminta dalam menjawab.

c. Tingkat Kesulitan

1) Soal Pilihan Ganda

Cukup mudah dan cocok untuk menguji pengetahuan dasar tentang Rukun Islam. Soal ini tidak terlalu rumit karena hanya membutuhkan pengenalan terhadap lima pokok ajaran Islam. Sedikit lebih sulit karena memerlukan siswa untuk mengingat dan menyebutkan seluruh Rukun Iman, yang mungkin tidak semua siswa hafal dengan mudah.

2) Soal Esai

Menilai kemampuan siswa untuk mengembangkan pemahaman dan memberikan penjelasan lebih mendalam tentang akhlak mulia. Ini soal yang lebih menantang karena mengharuskan siswa untuk berpikir secara kritis dan menyusun jawaban yang jelas.

d. Keadilan

Soal-soal ini adil karena menguji pengetahuan dasar yang seharusnya sudah diajarkan dalam pembelajaran PAI di kelas. Tidak ada soal yang

mengandung bias, dan soal-soal ini dapat dipahami oleh siswa dengan berbagai latar belakang.

e. Variasi Jenis Soal

Instrumen ini menggunakan variasi jenis soal yang cukup baik:

- 1) Pilihan Ganda untuk menguji pemahaman dasar.
- 2) Isian untuk menguji hafalan dan pemahaman konsep lebih mendalam.
- 3) Esai untuk mengukur kemampuan berpikir kritis dan penjelasan secara mendalam.

Variasi jenis soal ini memberikan gambaran yang lebih holistik mengenai pemahaman siswa terhadap materi PAI.

f. Konsistensi Soal

Soal-soal ini konsisten dalam hal fokus pada materi yang ingin diuji: Rukun Islam, Rukun Iman, dan Akhlak Mulia. Semua soal ini relevan dan mendukung satu sama lain untuk menilai aspek yang berbeda dari materi ajaran Islam.

g. Relevansi Waktu yang Diberikan

Soal Pilihan Ganda dan Isian kemungkinan bisa diselesaikan dalam waktu yang singkat, sedangkan soal Esai mungkin membutuhkan waktu lebih lama untuk memberikan penjelasan yang komprehensif. Waktu yang disediakan untuk masing-masing soal harus disesuaikan dengan tingkat kesulitan dan format soal.

2. Tingkat Kesukaran

Tingkat kesukaran soal diukur berdasarkan persentase siswa yang menjawab soal dengan benar. Semakin tinggi persentase siswa yang memberikan jawaban benar, semakin mudah soal tersebut. Sebaliknya, semakin rendah persentase siswa yang menjawab soal dengan benar, semakin sulit soal itu. Menurut Saifudin Azwar, tingkat kesukaran soal adalah rasio antara jumlah peserta tes yang menjawab soal dengan benar dan jumlah total peserta tes. Artinya, semakin banyak peserta tes yang menjawab soal dengan benar, semakin tinggi indeks kesukaran, yang menunjukkan bahwa soal tersebut lebih mudah. Sebaliknya, jika lebih sedikit peserta yang menjawab dengan benar, maka soal tersebut dianggap lebih sulit. Untuk menghitung indeks kesukaran digunakan rumus yaitu:

$$P=B:Js$$

Dimana:

P= Indeks kesukaran

B= Banyaknya siswa yang menjawab dengan benar

Js= Jumlah siswa seluruh peserta tes

Contoh:

Ada sepuluh orang dengan kode A s/d J yang mengikuti tes yang terdiri dari 10 soal. Jawaban tesnya dianalisis yaitu:

<i>Siswa</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>	<i>7</i>	<i>8</i>	<i>9</i>	<i>10</i>	<i>Skor Soal</i>
A	1	1	0	0	1	1	0	1	0	1	6
B	1	1	0	0	1	1	0	1	0	0	5
C	1	1	0	0	1	1	0	1	0	1	6
D	1	1	0	0	1	1	0	1	0	0	5
E	1	1	0	0	1	1	0	1	0	1	6
F	1	1	0	0	1	1	0	1	0	0	5
G	1	1	0	0	1	1	0	1	0	1	6
H	1	1	0	0	1	1	0	1	0	0	5
I	1	1	0	0	1	1	0	1	0	1	6
J	1	1	0	0	1	1	0	1	0	0	5
<i>Jumlah</i>	10	10	0	0	10	10	0	10	0	5	

Jawab soal nomor 1:

$$10:10=1$$

Berikut kategorinya yaitu:

1. 0,00-0,30= Sukar
2. 0,31-0,70= Sedang
3. 0,71-1,00= Mudah

3. Daya Pembeda

Fernandes mengatakan bahwa daya pembeda soal adalah kemampuan memisahkan siswa pandai dan siswa kurang pandai (Magdalena et al., 2021).

Untuk menghitung daya pembeda yaitu:

$$D=(U-L):N$$

Dimana:

D= Daya pembeda

U= Jumlah siswa dalam kelompok atas yang menjawab soal dengan benar

L= Jumlah siswa dalam kelompok bawah yang menjawab soal dengan benar

N= Jumlah siswa seluruh peserta tes

Contoh:

$$6-4=10$$

$$10:10=1$$

Berikut kategorinya yaitu:

1. Daya pembeda tinggi (>0.4): Soal sangat efektif dalam membedakan siswa yang menguasai materi dan yang tidak.
2. Daya pembeda sedang (0.3-0.39): Soal cukup baik dalam membedakan siswa, namun ada ruang untuk perbaikan.

3. Daya pembeda rendah (<0.2): Soal tidak efektif dalam membedakan siswa yang menguasai materi dengan yang tidak. Biasanya soal seperti ini perlu direvisi.
1 termasuk dalam daya pembeda tinggi karena lebih dari 0,4.

D. Simpulan

Pendidikan merupakan usaha sadar untuk mewujudkan sesuatu pewarisan budaya dari satu generasi kepada generasi yang lainnya. Untuk hasilnya yaitu yang pertama instrumen standar adalah alat evaluasi yang telah dikembangkan, diuji, dan diverifikasi secara sistematis untuk memastikan kualitas, keakuratan, serta keandalannya dalam mengukur kemampuan atau kompetensi peserta didik. Yang kedua Instrumen non standar dalam konteks pendidikan merujuk pada alat atau metode yang digunakan untuk mengukur atau menilai kemampuan siswa yang tidak mengikuti prosedur atau format yang sudah distandarisasi oleh lembaga atau badan tertentu. Yang ketiga analisis kualitatif butir instrumen bertujuan untuk memeriksa kualitas dan kesesuaian soal dalam mengukur kompetensi yang diinginkan. Yang keempat tingkat kesukaran dinyatakan dengan persentase siswa yang menjawab soal dengan benar. Yang kelima daya pembeda soal adalah kemampuan memisahkan siswa pandai dan siswa kurang pandai.

DAFTAR PUSTAKA

- Anam Syaiful, Ilzamuddin Ma'mur, Agus Gunawan, dkk. "Filsafat Pendidikan Sebagai Pondasi Dasar Lembaga Pendidikan Islam", *Jurnal Studi Islam*, Vol. 9, No. 1, Juni, 2023.
- Karmanis, *Buku Pedoman Belajar Metode Penelitian*, Semarang: CV Pilar Nusantara, 2020.
- Magdalena Ina, Indah Ayu Angraini, Siti Khoiriah, "Analisis Daya Pembeda, Dan Taraf Kesukaran Pada Soal Bilangan Romawi Kelas 4 SDN Tobat 1 Balaraja", *Jurnal Pendidikan dan Ilmu Sosial*, Vol. 3, No. 1, Maret, 2021.
- Rahman Abd BP, Sabhayati Asri Munandar, dkk. "Pengertian Pendidikan, Ilmu Pendidikan dan Unsur-Unsur Pendidikan", *Jurnal Al Urwatul Wutsqa*, Vol. 2, No. 1, Juni, 2022.
- Syah Darwyan, Supardi. *Evaluasi Pembelajaran Pendidikan Agama Islam*, (Jakarta: HAJA Mandiri, 2014.