



ANALISIS TINGKAT KESUKARAN DALAM PENGEMBANGAN EVALUASI PEMBELAJARAN PENDIDIKAN AGAMA ISLAM (PAI)

Rukayah^{1*}, Dina Hermina²

^{1,2} Mahasiswa Pascasarjana, PAI, UIN Antasari Banjarmasin

*Penulis Korespondensi: rukayahtariwin@gmail.com

Abstract. *Learning evaluation is a crucial component for measuring the achievement of students' competencies. In the context of Islamic Education (PAI), evaluation instruments often face obstacles related to an imbalanced proportion of difficulty levels and item quality bias. This article aims to deeply examine the basic concepts, classification, implementation, and follow-up actions of item difficulty level analysis in PAI subjects. Utilizing a literature review approach (conceptual/literature review), this article analyzes psychometric theories relevant to the broad characteristics of PAI material, spanning from textual cognitive aspects to Higher Order Thinking Skills (HOTS). The results of the study indicate that difficulty level analysis serves a diagnostic function to assess exam fairness and the effectiveness of teacher pedagogy. The implementation of an ideal difficulty index (P) moves toward a balanced proportion of easy, medium, and difficult questions. These findings recommend the need for a Standard Operating Procedure (SOP) to follow up on evaluation results, such as multicultural-based editorial improvements and storing high-quality items in a question bank to realize a continuous quality improvement cycle for PAI evaluations.*

Keywords: *Difficulty Level, Learning Evaluation, Islamic Religious Education, Item Analysis.*

Abstrak. Evaluasi pembelajaran merupakan komponen krusial untuk mengukur ketercapaian kompetensi peserta didik. Dalam konteks Pendidikan Agama Islam (PAI), instrumen evaluasi sering kali menghadapi kendala ketidakseimbangan proporsi tingkat kesukaran dan bias kualitas butir soal. Artikel ini bertujuan untuk mengkaji secara mendalam konsep dasar, klasifikasi, implementasi, serta tindak lanjut hasil analisis tingkat kesukaran pada mata pelajaran PAI. Melalui pendekatan kajian pustaka (conceptual/literature review), artikel ini menganalisis teori psikometrika yang relevan dengan karakteristik materi PAI yang luas, meliputi aspek kognitif tekstual hingga berpikir tingkat tinggi (Higher Order Thinking Skills/HOTS). Hasil kajian menunjukkan bahwa analisis tingkat kesukaran berfungsi sebagai fungsi diagnostik untuk menilai keadilan ujian dan efektivitas pedagogi guru. Implementasi indeks kesukaran ((P)) yang ideal bergerak pada proporsi seimbang antara soal mudah, sedang, dan sukar. Temuan ini merekomendasikan perlunya Standard Operating Procedure (SOP) tindak lanjut hasil evaluasi, seperti perbaikan redaksi berbasis multikultural dan penyimpanan butir berkualitas ke dalam bank soal demi mewujudkan siklus peningkatan mutu evaluasi PAI yang berkelanjutan.

Kata Kunci: Tingkat Kesukaran, Evaluasi Pembelajaran, Pendidikan Agama Islam, Analisis Butir Soal.

1. LATAR BELAKANG

Naskah Dalam evaluasi pembelajaran Pendidikan Agama Islam (PAI), tingkat kesukaran merupakan sebuah instrumen analisis utama untuk menentukan seberapa mudah atau sulitnya suatu butir soal bagi peserta didik. Tujuan esensial dari penentuan tingkat kesukaran bukan untuk menyulitkan siswa, melainkan untuk memastikan bahwa alat ukur yang digunakan mampu membedakan kemampuan setiap peserta didik secara

adil, objektif, dan akurat¹. Namun, realitas di lapangan menunjukkan adanya kesenjangan yang signifikan antara kondisi ideal psikometrika dengan praktik evaluasi yang dilakukan oleh para pendidik.

Terdapat empat permasalahan utama yang menjadi kunci kesenjangan dalam evaluasi PAI di sekolah atau madrasah, di antaranya adalah ketidakseimbangan proporsi tingkat kesukaran, rendahnya validitas alat ukur pada materi HOTS (*Higher Order Thinking Skills*), evaluasi yang hanya bersifat formalitas, dan karakteristik materi PAI yang luas serta kompleks². Masalah yang paling sering muncul adalah ujian PAI didominasi oleh soal yang "terlalu mudah" (hanya bersifat hafalan tekstual) atau "terlalu sulit" tanpa dasar analisis yang jelas³. Akibatnya, hasil ujian tidak mampu memetakan kemampuan siswa yang sebenarnya karena soal tidak dapat membedakan siswa yang benar-benar belajar dengan yang hanya beruntung saja⁴.

Pemerintah mendorong penggunaan soal berbasis HOTS, akan tetapi dalam PAI hal tersebut menjadi tantangan besar. Guru sering kali kesulitan menentukan apakah suatu soal menjadi "sulit" karena menuntut berpikir kritis (seperti analisis hukum/dalil) atau menjadi "sulit" hanya karena kalimatnya yang membingungkan serta penggunaan istilah Arab yang tidak umum. Hal ini berdampak pada kegagalan siswa menjawab bukan karena tidak paham esensi agama, melainkan karena kendala bahasa atau redaksi soal. Di sisi lain, analisis tingkat kesukaran sering kali hanya berhenti di atas kertas sebagai pemenuhan administrasi guru, di mana pendidik melakukan evaluasi hanya untuk mengambil nilai tanpa melakukan analisis butir soal setelah ujian berakhir. Dampaknya, soal-soal yang kualitasnya buruk atau tidak proporsional terus digunakan kembali di tahun-tahun berikutnya tanpa adanya perbaikan (revisi).

Karakteristik cakupan materi PAI sangat luas dan kompleks, meliputi Al-Qur'an Hadits, Aqidah Akhlak, Fiqih, dan Sejarah Kebudayaan Islam (SKI). Setiap sub-materi tersebut memiliki tingkat kesulitan yang berbeda, namun sering mendapatkan perlakuan yang sama dalam pembuatan soal, yang berdampak pada terjadinya ketimpangan nilai siswa pada aspek tertentu. Oleh karena itu, membahas topik tingkat kesukaran ini sangat

1
2
3
4

krusial karena menyentuh akar dari kualitas pendidikan agama itu sendiri, agar setiap siswa mendapatkan penilaian yang adil di mana soal yang disajikan sesuai dengan porsi kemampuan rata-rata kelompoknya.

2. METODE PENELITIAN

Artikel ini disusun menggunakan metode kajian pustaka (*literature review*) dengan pendekatan konseptual non-penelitian, fokus utama metode ini adalah melakukan analisis teoritis dan konseptual mengenai tingkat kesukaran pada instrumen evaluasi PAI. Sumber data sekunder diperoleh dari buku-buku dasar evaluasi pendidikan, regulasi penilaian, serta sintesis metodologi yang menghubungkan kedalaman materi agama dengan indeks kesukaran empiris. Data yang telah dikumpulkan kemudian dianalisis secara deskriptif-analitis guna menghasilkan draf narasi ilmiah yang dapat menjadi rujukan akademik bagi perbaikan mutu ujian sekolah.

3. HASIL DAN PEMBAHASAN

Bagian Secara umum, tingkat kesukaran adalah sebuah kategori atau derajat yang menunjukkan seberapa sulit atau mudahnya suatu butir soal bagi peserta didik dalam evaluasi PAI. Hal ini dinyatakan melalui sebuah angka bernama Indeks Kesukaran yang mencerminkan persentase siswa yang mampu menjawab soal tersebut dengan benar. Secara teknis, tingkat kesukaran adalah analisis terhadap respons siswa terhadap soal ujian, di mana jika sebuah soal dijawab benar oleh hampir seluruh siswa maka dianggap mudah, dan jika hanya sedikit siswa yang menjawab benar maka dianggap sukar. Soal yang ideal biasanya berada pada kategori sedang karena mampu memberikan gambaran distribusi kemampuan siswa secara akurat.

Menurut Suharsimi Arikunto dalam bukunya "*Dasar-Dasar Evaluasi Pendidikan*", soal yang baik ialah soal yang tidak terlalu mudah ataupun sukar. Soal yang terlalu mudah tidak merangsang siswa untuk mempertinggi usaha memecahkannya, sedangkan soal yang terlalu sukar akan menyebabkan siswa menjadi putus asa dan tidak mempunyai semangat untuk mencoba lagi karena di luar jangkauannya. Dalam mata pelajaran PAI, tingkat kesukaran berkaitan dengan kedalaman kompetensi yang diuji, yang dibedakan menjadi:

1. Aspek hafalan (LIT/LOTS): Soal dengan tingkat kesukaran rendah (mudah) biasanya berupa hafalan ayat, nama malaikat, atau urutan rukun iman
2. Aspek pemahaman & analisis (HIT/HOTS): Soal dengan tingkat kesukaran tinggi biasanya menuntut siswa melakukan ijtihad sederhana, analisis perbandingan mazhab (fiqih), atau kontekstualisasi sejarah Islam dengan kehidupan modern.

Sejalan dengan hal tersebut, Nana Sudjana mendefinisikan tingkat kesukaran sebagai derajat kesulitan atau proporsi siswa yang menjawab benar dari sejumlah peserta didik yang mengikuti tes⁵. Hal ini menunjukkan kualitas soal tidak hanya ditentukan oleh materi, tetapi juga oleh kemampuan kelompok peserta didik yang diuji⁶. Lebih lanjut, Anas Sudijono menjelaskan bahwa tingkat kesukaran adalah angka yang menunjukkan imbangan (perbandingan) antara jumlah peserta didik yang dapat menjawab dengan benar terhadap seluruh jumlah peserta tes⁷.

Konsep dan Formulasi Matematis Indeks Kesukaran (P)

Secara matematis, karakteristik indeks kesukaran dilambangkan dengan huruf (P) dan dipahami melalui rentang nilai 0,00 sampai 1,00. Nilai $(P = 0,00)$ menunjukkan soal sangat sukar (tidak ada satupun siswa yang benar), sedangkan $(P = 1,00)$ menunjukkan bahwa soal tersebut sangat mudah (seluruh siswa menjawab benar). Rumus yang digunakan untuk menghitung tingkat kesukaran adalah sebagai berikut⁸.

$$P = \frac{B}{J S}$$

Ket:

P= Indeks Kesukaran

B= Jumlah Siswa yang Menjawab Benar

JS= Jumlah Seluruh Siswa)

Para ahli evaluasi membagi kriteria indeks kesukaran ke dalam tiga klasifikasi utama, yaitu:

⁵ Arikunto, S. (2019). *Dasar-Dasar Evaluasi Pendidikan* (Ed. 3). Bumi Aksara.

⁶ Sudjana, N. (2017). *Penilaian Hasil Proses Belajar Mengajar*. Remaja Rosdakarya.

⁷ Sudijono, Anas. 2015. *Pengantar Evaluasi Pendidikan*. Jakarta: Raja Grafindo Persada.

⁸ Nana Sudjana, *op.cit.*, hlm. 135

1. Soal Sukar (Difficult/Hard): Nilai rentang $(P < 0,20)$ (atau $(P < 0,30)$) (pp. 9, 13). Soal yang hanya dapat dijawab benar oleh sebagian kecil peserta didik atau menuntut pemikiran mendalam.
2. Soal Sedang (Moderate/Medium): Nilai rentang $(0,20 - 0,80)$ (atau $(P = 0,31 - 0,70)$) (pp. 9, 14). Soal yang dijawab benar oleh sebagian peserta didik dan menunjukkan keseimbangan kualitas.
3. Soal Mudah (Easy): Nilai rentang $(P > 0,80)$ (atau $(P > 0,70)$) (pp. 9, 13). Soal yang sebagian besar atau hampir seluruh peserta didik dapat menjawab dengan benar⁹.

SOP Tindak Lanjut Analisis Tingkat Kesukaran Soal PAI

Setelah mendapatkan nilai Indeks Kesukaran ((P)), guru harus melaksanakan *Standard Operating Procedure* (SOP) tindak lanjut demi menjaga mutu soal secara konsisten.

1. Klasifikasi dan Identifikasi: Memasukkan setiap butir soal ke dalam tabel distribusi (Mudah, Sedang, Sukar) untuk melihat kesesuaian komposisi dengan kisi-kisi (misalnya proporsi ideal menurut Suharsimi Arikunto).
2. Evaluasi Soal Kategori "Sukar" ($(P < 0,30)$): Guru harus melakukan bedah soal secara kualitatif untuk menganalisis apakah soal sulit karena menuntut nalar tinggi (HOTS) atau sulit karena bahasa yang membingungkan/materi belum diajarkan¹⁰. Jika soal berbobot (HOTS), simpan untuk ujian pengayaan; jika cacat konstruksi (bahasa sulit), lakukan revisi redaksi; dan jika materi belum tuntas, lakukan *Remedial Teaching*.
3. Evaluasi Soal Kategori "Mudah" ($(P > 0,70)$): Guru menganalisis apakah pengecoh (*distractor*) tidak berfungsi atau jawaban terlalu mencolok. Keputusannya adalah memperbaiki pilihan jawaban agar lebih homogen atau menaikkan level kognitif soal¹¹.

⁹ Arifin, Zainal. 2016. *Evaluasi Pembelajaran: Prinsip, Teknik, Prosedur*. Bandung: Remaja Rosdakarya.

¹⁰ Zainal Arifin, *op.cit.*, hlm. 270, hlm.13

¹¹ Suharsimi Arikunto, *op.cit.*, hlm. 228

4. Seleksi untuk "Bank Soal": Soal-soal dengan kategori sedang ($P = 0,31 - 0,70$) dimasukkan ke dalam Bank Soal sekolah agar dapat digunakan kembali pada ujian tahun berikutnya dengan tingkat kepercayaan yang tinggi¹².
5. Refleksi Metode Pembelajaran: Jika satu kelas gagal menjawab soal kategori "Sedang", analisis ini berfungsi sebagai fungsi diagnostik bagi pendidik untuk mengevaluasi efektivitas metode mengajarnya di kelas¹³.

Cara Mengimplementasikannya dalam Pembelajaran

Implementasi tingkat kesukaran dalam pembelajaran PAI dilaksanakan secara sistematis melalui empat tahapan utama, yaitu:

1. Tahap Perencanaan (Penyusunan Kisi-kisi): Sebelum membuat soal, guru menentukan proporsi tingkat kesukaran berdasarkan kedalaman materi, misalnya membagi soal menjadi 25% sukar, 50% sedang, dan 25% mudah¹⁴.
2. Tahap Konstruksi Butir Soal: Guru menggunakan Kata Kerja Operasional (KKO) yang sesuai dengan level kognitif Taksonomi Bloom (C1-C2 untuk mudah/sedang, C4-C6 untuk sukar/HOTS) serta memastikan kesukaran berasal dari kompleksitas materi, bukan bahasa yang ambigu¹⁵.
3. Tahap Analisis Empiris (Uji Coba): Setelah ujian dilaksanakan, guru menghitung statistik indeks kesukaran untuk memetakan soal mana yang tidak terjawab oleh kelompok pintar atau soal mana yang benar semua¹⁶.
4. Tahap Tindak Lanjut (*Feed Forward*): Hasil analisis digunakan langsung untuk memperbaiki Proses Belajar Mengajar (PBM). Jika banyak soal "sukar" pada materi tertentu (misalnya Perhitungan Waris), guru harus melakukan *Remedial Teaching* pada pertemuan berikutnya¹⁷.

4. KESIMPULAN

¹² Nana Sudjana, *op.cit.*, hlm. 140

¹³ Purwanto, N. (2013). *Prinsip-Prinsip dan Teknik Evaluasi Pengajaran*. Remaja Rosdakarya.

¹⁴ Suharsimi Arikunto, *op.cit.*, hlm. 224

¹⁵ Daryanto. (2014). *Evaluasi Pendidikan*. Rineka Cipta.

¹⁶ Anas Sudijono, *op.cit.*, hlm. 375

¹⁷ Zainal Arifin, *op.cit.*, hlm. 272

Tingkat kesukaran dalam evaluasi PAI adalah instrumen ukur yang menganalisis respons siswa terhadap butir soal untuk mengetahui keseimbangan proporsi soal (mudah, sedang, sukar) agar selaras dengan kemampuan rata-rata siswa. Karakteristik indeks kesukaran ((P)) bergerak secara matematis antara nilai 0,00 hingga 1,00, di mana nilai yang paling ideal berada pada kategori "sedang" karena mampu mendistribusikan gambaran kemampuan belajar siswa secara akurat dan adil. Implementasi analisis ini tidak boleh dilakukan hanya setelah ujian (*post-test*), melainkan harus direncanakan sejak awal penyusunan kisi-kisi dan konstruksi butir soal. Apabila terdapat kegagalan massal pada soal dengan materi penting, tingkat kesukaran yang tinggi menjadi indikator fungsi diagnostik perlunya evaluasi menyeluruh terhadap metode pengajaran guru maupun redaksi konstruksi instrumen ujian.

DAFTAR REFERENSI

- Arifin, Zainal. (2016). *Evaluasi Pembelajaran: Prinsip, Teknik, Prosedur*. Bandung: Remaja Rosdakarya.
- Arikunto, Suharsimi. (2019). *Dasar-Dasar Evaluasi Pendidikan* (Ed. 3). Jakarta: Bumi Aksara.
- Daryanto. (2014). *Evaluasi Pendidikan*. Jakarta: Rineka Cipta.
- Purwanto, Ngalim. (2013). *Prinsip-Prinsip dan Teknik Evaluasi Pengajaran*. Bandung: Remaja Rosdakarya.
- Sudjana, Nana. (2017). *Penilaian Hasil Proses Belajar Mengajar*. Bandung: Remaja Rosdakarya.
- Sudijono, Anas. (2015). *Pengantar Evaluasi Pendidikan*. Jakarta: Raja Grafindo Persada.
- Thoha, M. Chabib. (2016). *Teknik Evaluasi Pendidikan*. Jakarta: Raja Grafindo Persada.