

PENGEMBANGAN INSTRUMEN HOTS FIKIH BERBASIS SOCIO-SCIENTIFIC ISSUES UNTUK PENGUATAN PENALARAN ETIS

Dilla Muckti Jihansyah

Universitas Islam Negeri Sayyid Ali Rahmatullah Tulungagung

dillamuckti@gmail.com

ABSTRACT

This study aims to develop and examine the feasibility of a Higher Order Thinking Skills (HOTS) test instrument for the subject of Fiqh on the topic of Ikhlasul Amal (slaughtering, sacrificial offering/qurban, and aqiqah), designed based on a socio-scientific issues (SSI) approach to strengthen students' ethical reasoning. The instrument consists of 35 multiple-choice items and 5 essay items constructed around contextual cases such as ostentation (riya') in performing qurban, unequal distribution of sacrificial meat, and social pressure in religious practice, requiring students to analyze, evaluate, and create solutions (cognitive levels C4-C6) rather than merely recall textual fiqh rulings. This Research and Development (R&D) study involved a limited trial with 25 ninth-grade students at MTsN 5 Magetan. The instrument's feasibility was examined through content validity (Aiken's V) by five expert validators, empirical item validity (product moment correlation), reliability (Cronbach's Alpha), normality testing (Shapiro-Wilk), and classical item analysis using Iteman 4.3. Results show that all 35 items obtained Aiken's V indices between 0.80 and 1.00 (very valid category), all items were empirically valid ($r\text{-count } 0.463\text{--}0.782 > r\text{-table } 0.396$), the instrument was highly reliable ($\alpha = 0.959$), and student score data were normally distributed (Shapiro-Wilk Sig. = 0.077). Nonetheless, classical item analysis revealed eleven items with low to negative discrimination indices (R_{pbis}), indicating the need to revise certain distractors. Overall, the instrument is considered feasible for measuring students' higher order thinking skills and ethical reasoning on the topic of Ikhlasul Amal, with recommended revisions to several items prior to wider-scale use.

Keywords: *HOTS instrument, socio-scientific issues, ethical reasoning, fiqh, Aiken's V validity*

ABSTRAK

Penelitian ini bertujuan untuk mengembangkan dan menguji kelayakan instrumen soal Higher Order Thinking Skills (HOTS) mata pelajaran Fiqh pada materi Ikhlasul Amal (penyembelihan, kurban, dan akikah) yang dirancang berbasis pendekatan socio-scientific issues (SSI) guna memperkuat penalaran etis peserta didik. Instrumen disusun dalam bentuk 35 butir soal pilihan ganda dan 5 butir soal uraian yang mengangkat kasus-kasus kontekstual seperti riya' dalam berkurban, kesenjangan distribusi daging kurban, dan tekanan sosial dalam pelaksanaan ibadah, sehingga peserta didik dituntut menganalisis, mengevaluasi, dan mencipta solusi (level kognitif C4-C6), bukan sekadar mengingat ketentuan fiqh secara

tekstual. Penelitian menggunakan pendekatan Research and Development (R&D) dengan uji coba terbatas pada 25 peserta didik kelas IX MTsN 5 Magetan. Kelayakan instrumen diuji melalui validitas isi (Aiken's V) oleh lima validator ahli, validitas butir empiris (korelasi product moment), reliabilitas (Cronbach's Alpha), uji normalitas (Shapiro-Wilk), serta analisis butir klasik menggunakan program Iteman 4.3. Hasil penelitian menunjukkan seluruh 35 butir memperoleh indeks Aiken's V pada rentang 0,80-1,00 (kategori sangat valid), seluruh butir dinyatakan valid secara empiris (r hitung 0,463-0,782 > r tabel 0,396), instrumen sangat reliabel ($\alpha = 0,959$), dan data skor peserta didik berdistribusi normal (Sig. Shapiro-Wilk = 0,077). Meskipun demikian, analisis butir klasik menunjukkan sebelas butir memiliki daya beda (R_{pbis}) rendah hingga negatif, yang mengindikasikan perlunya revisi pada distraktor tertentu. Instrumen ini secara umum dinyatakan layak digunakan sebagai alat ukur kemampuan berpikir tingkat tinggi dan penalaran etis peserta didik pada materi Ikhlasul Amal, dengan catatan perbaikan pada sejumlah butir sebelum digunakan pada skala yang lebih luas.

Kata Kunci: instrumen HOTS, socio-scientific issues, penalaran etis, fikih, validitas Aiken's V

A. Pendahuluan

Kurikulum pendidikan agama Islam di madrasah, termasuk mata pelajaran Fikih, menempatkan penguasaan Higher Order Thinking Skills (HOTS) sebagai salah satu capaian penting agar peserta didik tidak sekadar menghafal ketentuan syariat, tetapi mampu menganalisis, mengevaluasi, dan menyusun solusi atas persoalan keagamaan yang mereka temui dalam kehidupan sehari-hari (Anderson & Krathwohl, 2001). Materi Ikhlasul Amal yang membahas penyembelihan, kurban, dan akikah merupakan materi yang sarat dengan dimensi sosial: praktik kurban di masyarakat kerap diwarnai persoalan riya', kesenjangan distribusi daging, tekanan sosial untuk ikut

kurban kolektif, hingga penyembelihan yang tidak memperhatikan kesejahteraan hewan.

Kondisi tersebut menjadikan materi Ikhlasul Amal sangat relevan dikembangkan melalui pendekatan socio-scientific issues (SSI), yaitu pendekatan pembelajaran yang menghadapkan peserta didik pada isu-isu sosial kontroversial atau dilematis yang memiliki keterkaitan dengan konsep keilmuan tertentu, sehingga mendorong peserta didik berargumentasi dan bernalar secara etis, bukan hanya kognitif semata (Sadler & Zeidler, 2005; Zeidler et al., 2005). Dalam konteks pembelajaran Fikih, SSI dapat diwujudkan melalui

soal-soal berbasis kasus yang menuntut peserta didik menilai keabsahan suatu praktik, mengevaluasi fenomena sosial di lingkungan sekitar, dan merumuskan sikap yang selaras dengan nilai keikhlasan beramal.

Namun demikian, penyusunan soal HOTS tidak dapat dilakukan tanpa proses pengujian kelayakan yang memadai. Soal yang tampak menuntut penalaran tinggi belum tentu valid dan reliabel apabila tidak diuji secara empiris. Oleh karena itu, penelitian pengembangan ini disusun untuk (1) mengembangkan instrumen soal HOTS Fikih berbasis SSI pada materi Ikhlasul Amal, dan (2) menguji kelayakan instrumen tersebut melalui uji validitas isi, validitas butir empiris, reliabilitas, normalitas data, dan analisis butir secara klasik, sehingga diperoleh instrumen yang benar-benar layak digunakan untuk mengukur kemampuan berpikir tingkat tinggi sekaligus penalaran etis peserta didik.

B. Metode Penelitian

Penelitian ini menggunakan pendekatan Research and Development (R&D) dengan fokus pada tahap pengembangan produk dan uji coba terbatas, mengikuti

prosedur pengembangan instrumen tes yang mencakup penyusunan kisi-kisi, penulisan butir soal, telaah/validasi ahli, uji coba lapangan terbatas, dan analisis data hasil uji coba (Sugiyono, 2019).

Produk yang Dikembangkan

Produk yang dikembangkan berupa perangkat soal HOTS Fikih kelas IX semester ganjil pada Bab Ikhlasul Amal, terdiri atas 35 butir soal pilihan ganda dan 5 butir soal uraian, dengan alokasi waktu 90 menit. Penyusunan soal diturunkan dari Capaian Pembelajaran (CP), Tujuan Pembelajaran (TP), dan Alur Tujuan Pembelajaran (ATP) yang menekankan kemampuan menganalisis ketentuan penyembelihan, mengevaluasi praktik kurban di masyarakat, menelaah hikmah akikah, menyusun solusi permasalahan fikih, serta menunjukkan sikap kritis terhadap ibadah yang berorientasi formalitas semata. Kisi-kisi soal disajikan pada Tabel 1.

Tabel 1. Kisi-Kisi Instrumen Soal

N o	Materi Pokok	Lev el Kog nitif	Indikat or Soal	Ben tuk Soa l	B uti r
1	Penyem belihan	C4	Mengan alisis keabsah an penyem belihan berdasa rkan kasus	Pilih an Gan da	1- 8
2	Penyem belihan	C5	Mengev aluasi praktik penyem belihan modern	Pilih an Gan da	9- 12
3	Kurban	C4	Menela ah syarat hewan kurban	Pilih an Gan da	13 - 18
4	Kurban	C5	Mengev aluasi pelaksa naan kurban di masyar akat	Pilih an Gan da	19 - 24
5	Akikah	C4	Mengan alisis ketentu an akikah	Pilih an Gan da	25 - 29

N o	Materi Pokok	Lev el Kog nitif	Indikat or Soal	Ben tuk Soa l	B uti r
6	Akikah	C6	Mengan alisis ketentu an akikah (lanjuta n)	Pilih an Gan da	30 - 35
7	Seluruh Materi	C4- C6	Menyele saikan studi kasus berbasis fenome na	Urai an	1- 5

Ciri khas SSI pada instrumen ini tampak pada rumusan stimulus soal yang mengangkat kasus kontekstual, misalnya panitia kurban yang membagikan daging tidak merata, praktik pamer jumlah hewan kurban di media sosial, keluarga yang berutang demi kurban mewah, hingga anak yang minder karena keluarganya belum mampu berakikah. Melalui kasus-kasus tersebut peserta didik tidak hanya diminta mengenali hukum fikih secara tekstual, tetapi juga menimbang aspek keikhlasan, kepedulian sosial, dan tanggung

jawab sebagaimana ditekankan dalam capaian pembelajaran ikhlasul amal.

Ciri khas SSI pada instrumen ini tampak pada rumusan stimulus soal yang mengangkat kasus kontekstual, misalnya panitia kurban yang membagikan daging tidak merata, praktik pamer jumlah hewan kurban di media sosial, keluarga yang berutang demi kurban mewah, hingga anak yang minder karena keluarganya belum mampu berakikah. Melalui kasus-kasus tersebut peserta didik tidak hanya diminta mengenali hukum fikih secara tekstual, tetapi juga menimbang aspek keikhlasan, kepedulian sosial, dan tanggung jawab sebagaimana ditekankan dalam capaian pembelajaran ikhlasul amal.

Subjek dan Teknik Analisis Data

Uji coba terbatas dilakukan terhadap 25 peserta didik. Validitas isi diuji dengan teknik Aiken's V berdasarkan penilaian lima validator ahli menggunakan skala 1-5, dengan kriteria butir dinyatakan valid apabila $V \geq 0,80$ (Aiken, 1985). Validitas butir empiris dihitung menggunakan korelasi product moment antara skor butir dan skor total, dibandingkan dengan r tabel pada taraf signifikansi 5% untuk $n=25$ (r tabel = 0,396).

Reliabilitas instrumen dihitung dengan koefisien Cronbach's Alpha, dengan kriteria reliabel apabila $\alpha > 0,70$ (Ebel & Frisbie, 1991). Uji normalitas data skor menggunakan Shapiro-Wilk karena jumlah sampel kurang dari 50. Selain itu, dilakukan analisis butir secara klasik (tingkat kesukaran/P dan daya beda/point-biserial) menggunakan perangkat lunak Iteaman 4.3 untuk mengidentifikasi butir yang memerlukan revisi.

C.Hasil Penelitian dan Pembahasan

Hasil Uji Validitas Isi (Aiken's V)

Penilaian lima validator ahli terhadap 35 butir soal menunjukkan bahwa seluruh butir memperoleh indeks Aiken's V pada rentang 0,80 hingga 1,00, dengan rata-rata sekitar 0,90. Ringkasan hasil disajikan pada Tabel 2.

Tabel 2. Ringkasan Hasil Uji Validitas Isi (Aiken's V)

Statistik	Nilai	Rentang	Kategori
Jumlah butir dinilai	35 butir	-	-
Jumlah validator ahli	5 orang	-	-
Indeks Aiken's V terendah	0,80	Butir S05	Sangat Valid
Indeks Aiken's V tertinggi	1,00	Butir S03, S25	Sangat Valid
Rata-rata indeks V	$\pm 0,90$	0,80 - 1,00	Sangat Valid
Butir tidak valid ($V < 0,80$)	0 butir (0%)	-	-

Nilai tertinggi (1,00) diperoleh pada butir S03 dan S25, sedangkan nilai terendah (0,80) terdapat pada butir S05, namun tetap berada di atas ambang batas 0,80 sehingga seluruh butir termasuk kategori sangat valid.

Temuan ini mengindikasikan bahwa dari sisi konten, rumusan stimulus SSI pada setiap butir telah dinilai relevan dan sesuai dengan tujuan pengukuran oleh para ahli, sehingga instrumen layak dilanjutkan ke tahap uji coba lapangan tanpa revisi butir yang bersifat mendasar.

Hasil Uji Validitas Empiris, Reliabilitas, dan Normalitas

Hasil uji coba lapangan terhadap 25 responden menunjukkan bahwa seluruh 35 butir dinyatakan valid secara empiris, dengan nilai r hitung berkisar antara 0,463 (butir P31) hingga 0,782 (butir P32), seluruhnya berada di atas r tabel (0,396) pada taraf signifikansi 5%. Uji reliabilitas menghasilkan koefisien Cronbach's Alpha sebesar 0,959, jauh melampaui ambang minimal 0,70, yang menunjukkan konsistensi internal instrumen yang sangat tinggi. Uji normalitas dengan Shapiro-Wilk memperoleh nilai signifikansi 0,077 ($> 0,05$), sehingga data skor peserta didik dinyatakan berdistribusi normal. Ringkasan hasil disajikan pada Tabel 3.

Tabel 3. Ringkasan Uji Validitas Empiris, Reliabilitas, dan Normalitas

Uji Statistik	Hasil	Keterangan
Validitas butir (r hitung vs r tabel 0,396; $n=25$)	r hitung 0,463 - 0,782	Seluruh 35 butir valid (sig. $< 0,05$)
Reliabilitas (Cronbach's Alpha)	0,959	Sangat reliabel ($> 0,70$)
Normalitas (Shapiro-Wilk)	Sig. = 0,077	Data berdistribusi normal ($> 0,05$)
Rata-rata skor total (Iteman)	24,96 dari 35 (SD 4,04)	P rata-rata 0,713
Rata-rata daya beda (Rpbis)	0,153	Tergolong rendah pada uji coba terbatas

Ketiga hasil pengujian tersebut secara bersama-sama menunjukkan bahwa instrumen memenuhi syarat psikometrik dasar sebagai alat ukur: valid pada setiap butirnya, reliabel sebagai satu kesatuan alat ukur, dan data yang dihasilkan layak dianalisis lebih lanjut dengan statistik parametrik.

Hasil Analisis Butir Secara Klasik

Analisis butir menggunakan Iteman 4.3 menunjukkan rata-rata skor peserta didik sebesar 24,96 dari 35 butir (SD = 4,036), dengan tingkat kesukaran (P) rata-rata 0,713 yang tergolong sedang menuju mudah, serta daya beda (Rpbis) rata-rata hanya 0,153. Sebelas butir mendapat tanda (flag) karena tingkat kesukaran yang tergolong ekstrem dan/atau daya beda yang rendah hingga negatif, sebagaimana disajikan pada Tabel 4.

Tabel 4. Butir dengan Catatan Daya Beda Rendah/Negatif (Hasil Iteman)

Butir	P	Rpbis	Tanda	Catatan
Butir 2	0,800	0,073	K	Tingkat kesukaran tinggi, daya beda lemah
Butir 3	0,880	0,040	K	Tergolong mudah, daya beda sangat lemah
Butir 5	0,600	0,143	K	Daya beda lemah
Butir 8	0,800	-0,108	K, LR	Daya beda negatif, distraktor perlu revisi
Butir 12	0,800	0,180	K	Daya beda lemah
Butir 13	0,680	0,005	K	Daya beda sangat lemah
Butir 15	0,640	-0,024	K, LR	Daya beda negatif, distraktor perlu revisi
Butir 22	0,720	0,039	K	Daya beda sangat lemah
Butir 25	0,680	-0,105	K, LR	Daya beda negatif, distraktor perlu revisi
Butir 28	0,640	0,019	K	Daya beda sangat lemah
Butir 29	0,760	-0,163	K, LR	Daya beda negatif, distraktor perlu revisi

Empat di antara sebelas butir tersebut (butir 8, 15, 25, dan 29) bahkan menunjukkan daya beda negatif, yang berarti peserta didik berskor rendah justru cenderung lebih sering menjawab benar dibanding peserta didik berskor tinggi. Kondisi ini lazim dijumpai pada uji coba terbatas dengan jumlah responden yang relatif kecil ($n=25$), sehingga estimasi daya beda menjadi kurang stabil. Meskipun butir-butir tersebut tetap lolos uji validitas isi dan validitas total-item karena dihitung dengan pendekatan yang berbeda, temuan pada analisis klasik ini penting sebagai bahan revisi, khususnya pada opsi pengecoh (distraktor) yang mungkin mengandung ambiguitas makna dalam narasi kasus SSI yang digunakan.

Kontribusi terhadap Penguatan Penalaran Etis

Secara substansi, instrumen yang dikembangkan berhasil menghadirkan dimensi penalaran etis melalui narasi-narasi SSI, seperti dilema panitia kurban yang mendahulukan kerabat dibanding warga miskin, fenomena pameran kurban di media sosial, dan sikap anak yang minder karena keterbatasan ekonomi keluarga dalam

berakikah. Butir-butir semacam ini tidak dapat dijawab hanya dengan mengingat hukum fikih secara tekstual, melainkan menuntut peserta didik menimbang nilai keikhlasan, kepedulian sosial, dan keadilan distribusi, sejalan dengan karakteristik penalaran etis yang menjadi salah satu capaian pendekatan SSI (Sadler & Zeidler, 2005).

Di sisi lain, hasil analisis butir klasik menunjukkan bahwa kualitas psikometrik tidak selalu berbanding lurus dengan kompleksitas narasi SSI yang dimuat dalam soal. Butir dengan cerita kasus yang kaya konteks sosial namun opsi jawabannya kurang tegas perbedaannya cenderung menghasilkan daya beda yang lemah. Hal ini menegaskan pentingnya keseimbangan antara kekayaan konteks isu sosial-saintifik dan kejelasan konstruksi opsi jawaban agar instrumen tetap mampu membedakan peserta didik berkemampuan tinggi dan rendah secara akurat.

D. Kesimpulan

Instrumen soal HOTS Fikih berbasis socio-scientific issues pada materi Ikhlasul Amal (penyembelihan, kurban, dan akikah) yang

dikembangkan dalam penelitian ini dinyatakan valid secara isi (seluruh 35 butir memperoleh indeks Aiken's V 0,80-1,00, kategori sangat valid), valid secara empiris (seluruh butir r hitung > r tabel 0,396), sangat reliabel (Cronbach's Alpha = 0,959), dan diujicobakan pada data yang berdistribusi normal (Shapiro-Wilk Sig. = 0,077). Instrumen ini secara umum layak digunakan sebagai alat ukur kemampuan berpikir tingkat tinggi sekaligus penalaran etis peserta didik. Namun demikian, hasil analisis butir klasik mengidentifikasi sebelas butir dengan daya beda rendah hingga negatif yang perlu direvisi, terutama pada konstruksi distraktor, sebelum instrumen digunakan pada uji coba dengan skala responden yang lebih luas.

DAFTAR PUSTAKA

Artikel in Press :

- Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement*, 45(1), 131-142.
- Anderson, L. W., & Krathwohl, D. R. (2001). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. New York: Longman.
- Ebel, R. L., & Frisbie, D. A. (1991). *Essentials of Educational Measurement* (5th ed.). Englewood Cliffs, NJ: Prentice Hall.
- Sadler, T. D., & Zeidler, D. L. (2005). Patterns of informal reasoning in the context of socioscientific decision making. *Journal of Research in Science Teaching*, 42(1), 112-138.
- Sugiyono. (2019). *Metode Penelitian Pendidikan: Pendekatan Kuantitatif, Kualitatif, dan R&D*. Bandung: Alfabeta.
- Zeidler, D. L., Sadler, T. D., Simmons, M. L., & Howes, E. V. (2005). Beyond STS: A research-based framework for socioscientific issues education. *Science Education*, 89(3), 357-377.