



Analisis Butir Soal Sejarah Kebudayaan Islam Kelas VIII di SMP Islam Al Musthofa Berdasarkan Level Kognitif Taksonomi Bloom Revisi

Millati Qanitatin¹

Universitas Islam Negeri K.H. Abdurrahman Wahid Pekalongan, Indonesia

Email: millati.qanitatin24191@mhs.uingusdur.ac.id

Jihan Nayla Atikah²

Universitas Islam Negeri K.H. Abdurrahman Wahid Pekalongan, Indonesia

Email: jihan.nayla.atikah24189@mhs.uingusdur.ac.id

Rara Anis Sabila³

Universitas Islam Negeri K.H. Abdurrahman Wahid Pekalongan, Indonesia

Email: rara.anis.sabila24192@mhs.uingusdur.ac.id

*Korespondensi: email: millati.qanitatin24191@mhs.uingusdur.ac.id

Abstrak

History Artikel:
Diterima 1 Juni 2026
Direvisi 15 Juni 2026
Diterima 20 Juni 2026
Tersedia online 26 Juni 2026

Learning evaluation is an essential component of the educational process to measure students' learning achievement. The quality of evaluation instruments is strongly influenced by the alignment between the cognitive level of test items and the learning objectives formulated according to the Revised Bloom's Taxonomy. This study aims to analyze and map the distribution of cognitive levels in the multiple-choice test items of *Islamic Cultural History* (Sejarah Kebudayaan Islam/SKI) on the topic of the Umayyad Dynasty for eighth-grade students at SMP Islam Al Musthofa based on the Revised Bloom's Taxonomy. This research employed a descriptive qualitative approach, with 20 multiple-choice test items as the research object, which were administered on a limited basis to 20 eighth-grade students. Data were collected through document analysis and students' test results. The data were analyzed by classifying each test item into the cognitive levels of C1 (remembering), C2 (understanding), and C3 (applying). The findings reveal that the evaluation instrument is still dominated by Lower-Order Thinking Skills (LOTS), particularly at the remembering and understanding levels, while Higher-Order Thinking Skills (HOTS) are not yet adequately represented. Although the instrument includes a range of cognitive levels, it requires further improvement by incorporating more HOTS-oriented items to promote students' critical and analytical thinking skills. The empirical trial showed that students performed very well on remembering-level questions but experienced a significant decline in conceptual understanding. At the application level (C3), a construction bias was identified, as all students (100%) answered the questions correctly due to distractors that were insufficiently homogeneous and overly predictable. Consequently, the correct answers could be identified using general logic rather than deeper reasoning. These findings are expected to serve as a reference for prospective teachers in enhancing their professional competence in developing evaluation instruments with diverse cognitive levels and designing more effective distractors that align with the demands of 21st-century learning.

Kata kunci:

item analysis, Revised Bloom's Taxonomy, cognitive levels, learning evaluation, Islamic Cultural History.

Pendahuluan

Pendidikan adalah aspek yang sangat krusial untuk pengembangan potensi manusia. Melalui pendidikan, diharapkan individu memiliki kemampuan serta keterampilan yang cukup sebagai manusia utuh yang siap bersaing menghadapi tantangan zaman. Pendidikan di sekolah sangat terkait dengan perancangan pembelajaran berupa kurikulum yang bertujuan untuk mengembangkan sumber daya manusia dalam aspek kognitif, psikomotor, dan afektif (Arifin, 2018). Proses pembelajaran kemudian ditutup dengan evaluasi atau penilaian untuk mengukur hasil belajar siswa. Dalam menentukan nilai secara objektif, diperlukan suatu alat ukur atau kriteria yang jelas. Perolehan hasil belajar ini erat kaitannya dengan kemampuan siswa dalam mengolah informasi pada aspek kognitif yang telah dipelajari. Oleh sebab itu, ahli pendidikan kemudian menciptakan metode pengklasifikasian tujuan pendidikan yang dikenal dengan istilah taksonomi. Salah satu konsep pengklasifikasian kemampuan berpikir yang paling luas digunakan dalam dunia pendidikan adalah Taksonomi Bloom.

Pada tahun 2001, Lorin Anderson bersama Krathwohl melakukan rekonstruksi terhadap ranah kognitif tersebut agar tetap relevan dengan tuntutan pendidikan kontemporer di abad ke-21 yang berfokus pada pembelajaran aktif dan berbasis kompetensi (Zainuddin, 2019). Perubahan pokok pada versi revisi ini meliputi penggantian istilah dari kata benda menjadi kata kerja aktif, serta penambahan dimensi pengetahuan (Retno Utari dkk, 2011). Melalui revisi ini, ranah kognitif dikategorikan secara hierarkis dari tingkat dasar menuju tingkat yang lebih kompleks, meliputi mengingat (C1) sebagai usaha mengambil kembali informasi dari memori, dan memahami (C2) untuk menciptakan pemahaman dari berbagai sumber melalui kegiatan mengklasifikasi atau membandingkan. Selanjutnya, tingkatan bergerak pada kemampuan menerapkan (C3) yang mengacu pada penggunaan prosedur untuk mengatasi masalah, menganalisis (C4) sebagai proses membagi komponen dan mencari keterkaitannya, mengevaluasi (C5) yang memberikan penilaian berdasarkan kriteria standar, hingga puncaknya yaitu mencipta (C6) untuk menghasilkan satu kesatuan baru yang harmonis (Retno, 2012).

Penerapan tingkatan level kognitif yang variatif dan seimbang ini sangat krusial, khususnya dalam mata pelajaran Sejarah Kebudayaan Islam (SKI). Sebagai disiplin ilmu, SKI mendasari pembelajaran sistematis terkait pertumbuhan, rekonstruksi peradaban, dan hasil karya umat Islam di masa lalu dalam berbagai bidang kehidupan agar siswa dapat merefleksikan peristiwa bersejarah tersebut (Arlina dkk, 2023). Namun, realita di lapangan menunjukkan bahwa pembelajaran SKI termasuk materi Sejarah Bani Umayyah di kelas VIII seringkali masih bersifat *teacher-centered* (berpusat pada guru), yang berdampak negatif pada pasifnya partisipasi siswa di kelas serta munculnya kejenuhan. Kondisi ini diperparah dengan adanya persepsi keliru bahwa SKI hanyalah materi yang menuntut hafalan fakta kronologis, nama tokoh, atau tahun peristiwa belaka. Padahal, instrumen evaluasi SKI yang baik idealnya harus mampu menstimulus siswa hingga ke taraf berpikir tingkat tinggi atau *Higher Order Thinking Skills* (HOTS) berdasarkan Taksonomi Bloom Revisi, seperti menganalisis sebab-akibat kebijakan masa lampau (C4) atau mengevaluasi runtuhnya sebuah dinasti (C5) untuk diambil ibrahnya (*lesson learned*). Kesenjangan antara tuntutan ideal ini dengan realita kompetensi calon pendidik dalam merancang soal bervariasi menjadi tantangan tersendiri yang perlu dikaji lebih dalam.

Berdasarkan kesenjangan antara tuntutan ideal evaluasi dan realita di lapangan tersebut, diperlukan sebuah upaya refleksi akademis yang nyata dari calon pendidik dalam merancang instrumen penilaian yang bermutu. Oleh karena itu, penelitian ini dilakukan untuk menganalisis dan memetakan kualitas distribusi level kognitif berdasarkan Taksonomi Bloom Revisi pada instrumen soal pilihan ganda Sejarah Kebudayaan Islam materi Sejarah Bani Umayyah kelas VIII yang telah dikembangkan oleh peneliti. Untuk melihat keterbacaan soal, kesesuaian tingkat kesukaran, serta respons kognitif nyata dari peserta didik, instrumen evaluasi ini telah diujikan secara empiris terbatas kepada 20 siswa kelas VIII di SMP Islam Al

Musthofa. Melalui penelitian deskriptif kualitatif ini, diharapkan hasil pemetaan dan analisis butir soal yang diperoleh dapat menjadi bahan evaluasi mandiri sekaligus refleksi kritis bagi calon pendidik dalam meningkatkan kompetensi profesionalitas penyusunan instrumen penilaian pembelajaran sejarah di masa depan.

Metode Penelitian

Penelitian ini menggunakan pendekatan deskriptif kualitatif dengan metode analisis isi (content analysis) atau analisis dokumen. Pendekatan ini dipilih karena penelitian ini bertujuan untuk mendeskripsikan, membedah, dan memetakan distribusi level kognitif butir soal Sejarah Kebudayaan Islam (SKI) kelas VIII materi Sejarah Bani Umayyah berdasarkan kerangka teori Taksonomi Bloom Revisi. Penelitian ini dilaksanakan di SMP Islam Al Musthofa sebagai lokus uji coba empiris instrumen.

Sumber data utama dalam penelitian ini adalah dokumen instrumen evaluasi berupa butir soal pilihan ganda materi Sejarah Bani Umayyah kelas VIII yang dirancang dan dikembangkan sendiri oleh peneliti selaku calon pendidik. Sebagai data pendukung untuk mengonfirmasi keterbacaan soal, kesesuaian tingkat kesukaran, dan respons kognitif riil peserta didik, instrumen ini diujikan kepada 20 siswa kelas VIII di SMP Islam Al Musthofa. Teknik pengumpulan data dalam penelitian ini meliputi: Dokumentasi, yaitu menghimpun dokumen berupa kisi-kisi soal, lembar butir soal pilihan ganda yang telah dirancang, serta lembar jawaban hasil uji coba dari 20 siswa dan studi Pustaka, yaitu mengkaji berbagai literatur, jurnal ilmiah, dan buku panduan penulisan soal yang relevan dengan Taksonomi Bloom Revisi.

Uji keabsahan data kualitatif dalam penelitian ini menggunakan triangulasi teori dan diskusi sejawat. Triangulasi teori dilakukan dengan menyandingkan teks stimulus dan pokok soal secara silang menggunakan berbagai referensi standar Taksonomi Bloom Revisi. Sementara itu, diskusi sejawat dilakukan melalui analisis kolaboratif antaranggota tim peneliti guna mereduksi subjektivitas dalam pengklasifikasian level kognitif butir soal.

Teknik analisis data dilakukan melalui empat tahap, yaitu: (1) mengumpulkan seluruh butir soal beserta kisi-kisinya, (2) mereduksi dan memilah soal berdasarkan indikator materi Bani Umayyah, (3) menganalisis dan mengodekan setiap butir soal ke dalam enam level kognitif (C1 sampai C6) berdasarkan kesesuaian Kata Kerja Operasional (KKO) yang terkandung pada soal, dan (4) menarik kesimpulan berupa persentase distribusi level kognitif instrumen secara utuh.

Hasil dan Pembahasan

A. Distribusi Level Kognitif Butir Soal SKI

Bagian ini menyajikan hasil pemetaan dan deskripsi kuantitatif mengenai sebaran level kognitif pada instrumen evaluasi Sejarah Kebudayaan Islam (SKI) materi Sejarah Bani Umayyah kelas VIII yang telah dirancang oleh peneliti. Pemetaan level kognitif ini didasarkan pada kerangka teori Taksonomi Bloom Revisi, yang mengklasifikasikan ranah kognitif ke dalam enam tingkatan, mulai dari mengingat (C1) hingga mencipta (C6). Proses identifikasi dilakukan dengan menganalisis kesesuaian antara stimulus soal, pokok soal (stem), serta indikator ketercapaian materi dengan Kata Kerja Operasional (KKO) yang digunakan pada setiap butir. Adapun data utuh mengenai pemetaan ragam level kognitif dari keseluruhan butir soal pilihan ganda yang telah dikembangkan dipaparkan secara rinci pada Tabel 1 berikut:

Analisis Butir Soal Sejarah Kebudayaan Islam Kelas VIII di SMP Islam Al Musthofa Berdasarkan Level Kognitif Taksonomi Bloom Revisi | 626

No	Indikator Materi	Kata Kerja Operasional (KKO)	Level Kognitif
1	Tokoh pendiri Bani Umayyah	Menyebutkan	C1
2	Mengidentifikasi nama lembaga administrasi pos Bani Umayyah	Mengidentifikasi	C1
3	Menjelaskan tujuan pemindahan ibu kota ke Damaskus oleh Muawiyah	Menjelaskan	C2
4	Menunjukkan bukti nyata sikap zuhud khalifah Umar bin Abdul Aziz	Mengidentifikasi	C1
5	Mengidentifikasi silsilah kakek buyut Bani Umayyah	Menyebutkan	C1
6	Mengidentifikasi pusat kota pemerintah daulah Bani Umayyah	Menyebutkan	C1
7	Mengidentifikasi nama panglima perang penakluk Andalusia	Menyebutkan	C1
8	Membedakan sistem kepemimpinan Bani Umayyah dengan Khulafaurrasyidin	Membedakan	C2
9	Mengidentifikasi kebijakan perluasan wilayah oleh khalifah Walid bin Abdul Malik	Mengidentifikasi	C1
10	Menjelaskan cara memperkuat nilai persatuan pada sejarah Bani Umayyah	Menjelaskan	C2
11	Mengklasifikasikan beberapa kebijakan khalifah Umar bin Abdul Aziz	Mengategorikan	C2
12	Mengidentifikasi keahlian tokoh cendekiawan muslim (Zaid bin Tsabit)	Mengidentifikasi	C1
13	Menyeleksi tokoh yang bukan termasuk cendekiawan Bani Umayyah	Memilih/memilah	C2
14	Mengidentifikasi tahun berdirinya dinasti Bani Umayyah	Menyebutkan	C1

15	Menjelaskan perubahan sistem pemerintahan pasca masa khulafaurrasyidin	Menjelaskan	C2
16	Menginterpretasi nilai islami dalam pengelolaan baitul mal	Menafsirkan	C2
17	Menjelaskan nilai dibalik pembentukan lembaga diwan	Menjelaskan	C2
18	Mengontekstualisasikan perilaku etos kerja umat islam dalam kehidupan	Menerapkan	C3
19	Menentukan teladan toleransi dari kondisi masyarakat multikultural	Menentukan	C3
20	Mengontekstualisasikan keteladanan menuntut ilmu dalam peran pelajar	Menerapkan	C3

Berdasarkan data yang tersaji pada Tabel 1, ditemukan bahwa dari total 20 butir soal pilihan ganda yang dikembangkan, sebaran level kognitif masih menunjukkan adanya pemusatan pada tingkat berpikir tingkat rendah (Lower Order Thinking Skills). Secara akumulatif, instrumen evaluasi ini didominasi oleh kelompok LOTS yang meliputi level C1 (Mengingat) sebanyak 9 butir soal atau sebesar 45%, dan level C2 (Memahami) sebanyak 8 butir soal atau sebesar 40%. Sementara itu, kelompok soal Middle Order Thinking Skills (MOTS) yang diwakili oleh level C3 (Menerapkan) memiliki porsi sebesar 15% (3 butir soal). Di sisi lain, untuk tingkatan berpikir tingkat tinggi atau Higher Order Thinking Skills (HOTS) yang meliputi level C4 hingga C6 sama sekali belum terakomodasi (0%).

Dominasi tingkat berpikir LOTS dan MOTS pada instrumen ini sejalan dengan kerangka rekonstruksi teori Anderson dan Krathwohl (2001) yang telah dipaparkan pada bagian pendahuluan, di mana penguasaan kemampuan mengingat fakta (C1) dan memahami konsep (C2) merupakan fondasi awal yang wajib dilewati siswa sebelum mampu memecahkan masalah atau berpikir analitis. Namun, jika ditinjau dari teori evaluasi modern menurut Arifin (2018), ketiadaan porsi soal berskala HOTS mengonfirmasi adanya bias dalam konstruksi instrumen evaluasi sejarah di sekolah. Calon pendidik cenderung terjebak pada pemenuhan target konten substantif-faktual yang bersifat linier (seperti silsilah, nama tokoh, dan penamaan lembaga) demi mengejar keterbacaan dasar siswa, sehingga belum optimal dalam menstimulus daya nalar kritis-evaluatif peserta didik sebagaimana yang dicita-citakan dalam tuntutan pendidikan abad ke-21.

B. Analisis Kualitatif dan Refleksi Butir Soal

Pada bagian ini, peneliti melakukan analisis isi *content analysis* secara mendalam terhadap karakteristik stimulus, pokok soal (stem), dan pilihan jawaban pada butir soal yang telah dikembangkan. Analisis didasarkan pada kesesuaian tingkat kedalaman berpikir siswa dengan kerangka Taksonomi Bloom Revisi, serta dikonfirmasi melalui data empiris hasil jawaban 20 siswa di SMP Islam Al Musthofa.

1. Refleksi Butir Soal Kategori Level C1(mengingat)

Berdasarkan hasil pemetaan pada Tabel 1, level C1 merupakan porsi terbesar dalam instrumen evaluasi yang dirancang oleh peneliti, yaitu mencapai 45% (9 butir soal). Kategori level ini pada dasarnya menguji kemampuan dasar peserta didik dalam mengambil kembali

(retrieving) pengetahuan faktual yang relevan dari memori jangka panjang mereka. Sebagai representasi, analisis dilakukan secara mendalam pada butir soal nomor 14. Pada soal nomor 14, terdapat pertanyaan: “Tahun berdirinya Dinasti Bani Umayyah adalah...”. Butir soal ini murni berada pada ranah kognitif C1 karena menggunakan Kata Kerja Operasional (KKO) Menyebutkan/Mengidentifikasi. Karakteristik utama dari soal ini adalah menguji ingatan jangka pendek siswa terkait data faktual kronologis sejarah yang sifatnya mutlak tanpa adanya tuntutan penalaran sebab-akibat atau kontekstualisasi masalah. Berdasarkan data uji coba empiris pada 20 siswa di SMP Islam Al Musthofa, tingkat kebenaran pada soal nomor 14 ini sangat tinggi, yaitu sebanyak 17 siswa (85%) mampu menjawab dengan benar, sedangkan hanya 3 siswa (15%) yang keliru.

Hasil empiris ini mengonfirmasi pendapat Anderson dan Krathwohl, bahwa soal level C1 memiliki tingkat kesukaran yang relatif rendah bagi siswa kelas VIII karena pola penalarannya bersifat linier dan langsung (Anderson dan Krathwohl, 2010). Selaras dengan penelitian sebelumnya, kemudahan siswa dalam melewati level ini membuktikan bahwa instrumen yang didominasi aspek ingatan dapat dengan mudah diselesaikan siswa, namun belum mampu memfasilitasi pengukuran kemampuan berpikir kritis dan kreatif siswa secara menyeluruh (Syahriani, 2021). Oleh karena itu, tinggi-rendahnya kelulusan siswa pada nomor ini murni bergantung pada kuatnya daya hafal peserta didik terhadap angka tahun sejarah, bukan pada ketajaman analisis kritis mereka.

2. Refleksi Butir Soal Kategori Level C2 (Memahami)

Berdasarkan hasil pemetaan pada Tabel 1, level C2 merupakan salah satu porsi penting dalam instrumen evaluasi yang dirancang oleh peneliti, yaitu mencapai 40% (8 butir soal). Kategori soal pada level ini bertujuan untuk menguji sejauh mana peserta didik mampu membangun makna dan memahami pesan instruksional yang diberikan. Sebagai representasi, analisis dilakukan secara mendalam pada butir soal nomor 15. Pada soal nomor 15, terdapat pertanyaan: “Salah satu perubahan sistem pemerintahan pada masa Dinasti Umayyah adalah...”. Butir soal ini berada pada ranah kognitif C2 karena menuntut kemampuan tingkat Kata Kerja Operasional (KKO) Mengidentifikasi/Mengklasifikasikan. Karakteristik utama dari soal ini adalah peserta didik tidak sekadar diminta mengingat fakta mentah (C1), melainkan harus memahami esensi makna di balik fenomena perubahan sistem tersebut. Berdasarkan data uji coba empiris pada 20 siswa di SMP Islam Al Musthofa, tingkat kebenaran pada soal nomor 15 ini menunjukkan hasil yang kurang optimal, yaitu hanya sebanyak 9 siswa (45%) yang mampu menjawab dengan benar, sedangkan 11 siswa (55%) lainnya masih keliru.

Hasil empiris ini mengonfirmasi landasan teori Taksonomi Bloom Versi Revisi dari Anderson dan Krathwohl (2001) mengenai Pengetahuan Konseptual (Conceptual Knowledge), di mana siswa dinilai berhasil jika mampu mengklasifikasikan atau mengelompokkan suatu fenomena sejarah ke dalam skema kategori atau prinsip keilmuan yang lebih luas. Pendapat ini diperkuat oleh Teori Belajar Kognitif yang menyatakan bahwa pembelajaran berfokus pada bagaimana proses mental internal, seperti memori, perhatian, dan pemecahan masalah mempengaruhi perolehan pengetahuan (Marta, 2025). Berdasarkan teori tersebut, siswa diposisikan sebagai pengolah informasi aktif yang mengelola proses mental internalnya sendiri, bukan sekadar penerima stimulus pasif. Secara keseluruhan, teori kognitif ini mendukung pendekatan aktif yang mendorong pemahaman mendalam dan keterampilan berpikir kritis (Magdalena, 2020). Oleh karena itu, tingginya angka kekeliruan peserta didik yang mencapai 55% pada nomor ini menunjukkan bahwa sebagian besar siswa masih terjebak pada pola belajar hafalan linier (C1) dan belum sepenuhnya mampu melakukan proses mental internal untuk menangkap esensi, karakteristik, serta klasifikasi dari perubahan struktural pemerintahan Islam pada masa tersebut. Temuan ini menjadi refleksi penting bagi peneliti bahwa penguatan pemahaman konseptual pada materi sejarah kebudayaan Islam masih perlu ditingkatkan agar siswa dapat mencapai tingkat pemahaman yang lebih mendalam.

3. Refleksi Butir Soal Kategori C3 (menerapkan)

Berdasarkan hasil pemetaan pada Tabel 1, level C3 (Menerapkan) merupakan kategori yang mengukur kemampuan peserta didik dalam menggunakan pengetahuan yang telah dipelajari untuk menyelesaikan situasi atau permasalahan tertentu. Pada level ini, peserta didik tidak hanya memahami informasi yang diterima, tetapi juga mampu mengimplementasikan konsep, nilai, atau prinsip ke dalam konteks kehidupan nyata. Sebagai representasi, analisis dilakukan pada butir soal nomor 20. Pada soal nomor 20, terdapat pertanyaan: “Khalifah Umar bin Abdul Aziz dikenal sebagai pemimpin yang gemar menuntut ilmu dan menghabiskan waktunya untuk belajar. Sebagai seorang pelajar, tindakan yang dapat dilakukan untuk meneladani sikap beliau adalah...”. Soal ini secara teoretis termasuk dalam ranah kognitif C3 (Menerapkan) karena menuntut peserta didik untuk menggunakan pemahaman tentang keteladanan tokoh sejarah dan mengaplikasikannya ke dalam situasi konkret yang relevan dengan kehidupan sehari-hari melalui Pilihan B (“Memanfaatkan waktu luang untuk membaca buku”).

Menurut Anderson dan Krathwohl, kategori *applying* (menerapkan) adalah kemampuan menggunakan prosedur, konsep, prinsip, atau pengetahuan yang telah dipelajari dalam situasi tertentu (Anderson dan Krathwohl, 2001) Melalui kacamata teori konstruktivisme, peserta didik dituntut menghubungkan pengetahuan tentang karakter tokoh sejarah dengan pengalaman dan konteks kehidupannya sendiri sebagai dasar pengambilan keputusan dalam memilih tindakan yang paling tepat (Wardani & Rumiati, 2011). Namun, temuan empiris yang sangat unik diperoleh dari hasil uji coba terbatas terhadap 20 siswa di SMP Islam Al Musthofa. Pada butir soal nomor 20 ini, tingkat kebenaran siswa mencapai angka mutlak, yaitu sebesar 100% (20 siswa menjawab benar). Melalui perspektif evaluasi pendidikan, fenomena kelulusan mutlak pada tingkat penerapan (C3) ini mengindikasikan adanya kelemahan pada konstruksi pilihan jawaban (distraktor). Opsi pengecoh pada pilihan A, C, dan D dirancang kurang homogen dan cenderung menampilkan bias perilaku yang terlalu ekstrem negatif (seperti hanya memajang buku, malas membaca jika tidak diperintah, dan belajar hanya saat ujian). Akibatnya, fungsi pengecoh tidak berjalan secara optimal karena siswa dapat dengan mudah mengeliminasi jawaban yang salah menggunakan logika umum (*common sense*) tanpa harus mendalami substansi materi sejarahnya secara mendalam. Hasil refleksi kualitatif ini menjadi catatan evaluatif yang sangat berharga bagi peneliti bahwa dalam merancang soal C3 ke depan, tingkat homogenitas pilihan pengecoh harus dibuat lebih kompetitif dan setara agar mampu mengukur ranah aplikasi kognitif peserta didik secara objektif di lapangan.

Kesimpulan

Berdasarkan hasil analisis isi dan uji coba empiris instrumen soal pilihan ganda, sejarah kebudayaan Islam materi sejarah Bani Umayyah kelas 8 di SMP Islam Al Musthofa, dapat disimpulkan bahwa distribusi level kognitif masih didominasi oleh kelompok dengan kemampuan berpikir tingkat rendah yang meliputi aspek mengingat dan memahami. Tingkat kemampuan berpikir menengah pada aspek menerapkan hanya terakomodasi dalam porsi yang sangat kecil dan untuk tingkat berpikir tinggi sama sekali belum terpenuhi di dalam instrumen evaluasi yang dikembangkan. Dengan demikian hasil uji empiris menunjukkan bahwa peserta didik dapat melewati soal kategori meningkat dengan sangat baik namun, mengalami penurunan drastis pada soal kategori pemahaman konseptual. Disisi lain, Pada kategori penerapan ditemukan adanya bias konstruksi di mana seluruh siswa mampu menjawab dengan benar akibat opsi pengecoh atau distraktor yang kurang homogen dan terlalu mudah, sehingga jawaban dapat diidentifikasi menggunakan logika umum, siswa tidak perlu berpikir mendalam untuk menjawabnya.

Sebagai bentuk refleksi akademis untuk meningkatkan profesionalisme dalam menyusun Instrumen penilaian di masa depan, seorang calon pendidik disarankan dalam

pembuatan soal yang semula bersifat substantif faktual menjadi menstimulus daya nalar kritis dan evaluatif siswa. Selain itu, pendidik juga harus lebih cermat dalam merancang tingkat homogenitas opsi pengecoh atau distraktor pada soal pilihan ganda agar dibuat lebih kompetitif dan logis.

Ucapan Terima Kasih

Ucapan Terima kasih ini dengan penuh rasa syukur kepada Tuhan yang Maha Esa, kami panjatkan terima kasih atas selesainya artikel ini yang tidak terlepas dari dukungan berbagai pihak. Ucapan terimakasih yang sebesar besarnya kami sampaikan kepada Dosen Mata Kuliah Evaluasi Pendidikan Ibu Irfa Ma'alina Li'iliyyina atas arahan, ilmu, dan kesabarannya selama proses penyusunan artikel ini. Ucapan yang sama pula kami sampaikan kepada pihak SMP Islam Al Musthofa yang telah menerima kami untuk melakukan penelitian ini.

Referensi

- Anderson, L.W., Krathwohl, D.R. 2001. Taksonomi untuk Pembelajaran, Pengajaran, dan Penilaian: Revisi Taksonomi Tujuan Pendidikan Bloom. New York: Addison Wesley Longman, Inc.
- Arifin, Z. 2018. Pengembangan Instrumen Analisis Kesalahan untuk Mengukur Critical Thinking Skill Berdasarkan Kompetensi Inti Kurikulum 2013. EduMa. Volume 7. Nomor 1. ISSN 2086-3918.
- Arlina Nurjannah Tumanggor, Astriana Hidayah, Hikmah Bayani Situmorang, Mitha Shaskila Sinaga (2023), implementasi Strategi Pembelajaran Inquirypada Mata Pelajaran SKI. Journal of Educational Management and Strategy(JEMAST).
- Hidayat, A., & Sholihah, M. (2022). Dilema Evaluasi Pembelajaran Sejarah: Analisis Soal Berbasis LOTS pada Tingkat Pendidikan Menengah. Jurnal Pendidikan Sejarah dan Evaluasi, 4(2), 115-128.
- Retno Utari, Widyaiswara Madya, KNPk Pusdiklat. 2011. Taksonomi bloom. Jurnal: Pusdiklat KNPk 766 (1), 1-7, 2011
- Retno, U.2012. Taksonomi Bloom: Apa dan bagaimana menggunakannya? Pusdiklat KNPk. Tersedia pada www.bppk.depkeu.go.id.
- Wardhani, & Rumiati, 2011. _Instrumen Penilaian Hasil Belajar Matematika SMP: Belajar dari PISA dan TIMSS_. Yogyakarta: PPPPTK Matematika.