

# ANALISIS TINGKAT KESULITAN (*DIFFICULTY LEVEL*) SOAL PADA BUKU SEJARAH KEBUDAYAAN ISLAM KURIKULUM 2013

R. Ahmad Nur Kholis

MTs Nahdlatul Ulama Kepuharjo Karangploso Malang &

SMP Islam Abu Ghonaim Bumiaji Batu

e-mail: [kholis3186@gmail.com](mailto:kholis3186@gmail.com)

---

DOI : 10.14421/jpai.2017.142-08

---

## **Abstract**

*This study aims to describe the proportion of problem difficulty levels presented in chapters 1 and 2 of the Teachers' Book of Curriculum 2013 of Islamic Culture History Subject (SKI) published by the Ministry of Religious Affairs in 2015. The data analyzed in this study is a matter of discussion on: (1) Traces of Abbasid Abbasid Civilizations; and (2) The Brilliance of Muslim Scientists of the Abbasid Dynasty. Analytical techniques used in this study there are 2 (two), namely: (1) Descriptive statistical analysis or also called social statistics; (2) The technique of proportional analysis responds correctly in the analysis of difficulty level (difficulty level). The result of the analysis shows that the problems in the two sections of the book are presented in proportion: (1) 25% of the questions in the easy category; (2) 35% matter in medium category; and (3) 40% problem in difficult category (difficult). Thus, it can be concluded that these problems can be said to be ideal as a matter of practice.*

**Keywords :** *Difficulty level of question, History and Culture of Islam subject.*

## **Abstrak**

Penelitian ini bertujuan untuk mendeskripsikan proporsi tingkat kesulitan soal yang disajikan dalam bab 1 dan 2 Buku Guru Kurikulum 2013 Mata Pelajaran Sejarah Kebudayaan Islam (SKI) yang diterbitkan oleh Kementerian Agama tahun 2015. Data yang dianalisis dalam penelitian ini adalah soal pada pembahasan tentang: (1) Jejak Peradaban Dinasti Abbasiyah; dan (2) Kecemerlangan Ilmuwan Muslim Dinasti Abbasiyah. Teknik analisis yang digunakan dalam penelitian ini ada 2 (dua) yaitu: (1) Analisis statistik deskriptif atau yang disebut juga dengan statistik sosial; (2) Teknik analisis proporsi menjawab benar dalam analisis tingkat kesukaran (*difficulty level*). Hasil analisis menunjukkan bahwa soal pada kedua pembahasan dalam buku SKI tersebut disajikan dengan proporsi: (1) 25% soal dalam kategori mudah; (2) 35% soal dalam kategori sedang; dan (3) 40% soal dalam kategori sulit (*sukar*). Dengan demikian, maka dapat disimpulkan bahwa soal-soal tersebut dapat dikatakan ideal sebagai soal latihan.

**Kata Kunci:** Tingkat Kesulitan Soal, Mata Pelajaran SKI.

## **Pendahuluan**

Kedudukan Evaluasi dalam Pembelajaran

Arikunto & Jabar (2014: 1-2)

membedakan antara penilaian

(*assessment*), pengukuran (*measurement*) dan evaluasi (*evaluation*).

Penilaian yang merupakan kata benda dari 'nilai' berkaitan dengan

hal yang bersifat kualitatif. Sedangkan pengukuran berikaitan dengan hal-hal yang bersifat kuantitatif. Adapun evaluasi adalah kegiatan pengumpulan data dan informasi tentang bekerja dan berjalannya sesuatu yang mana data dan informasi tersebut selennjutnya digunakan untuk menentukan kebijakan. (Ihsan & Ihsan, 2008).

Penilaian merupakan kegiatan melekat dari pengajaran. Kegiatan penilaian terjadi baik pada awal, proses, maupun pada akhir pembelajaran. Pada awal pembelajaran, penilaian dilakukan untuk tujuan diagnosa yakni untuk mengetahui tingkat kemampuan dan modal pertama yang dimiliki siswa atau juga berfungsi sebagai penentuan kebijakan dalam hal penempatan (*placement*) siswa pada kelompok belajar tertentu. Pada saat pembelajaran berlangsung, penilaian berfungsi untuk mengukur ketercapaian tujuan pembelajaran yang telah ditetapkan dan hasilnya digunakan sebagai *feedback* atas kegiatan pembelajaran yang dilakukan (*formative*). Penilaian pada akhir belajar dilakukan untuk

mengukur ketercapaian keseluruhan tujuan kurikulum yang telah ditetapkan pada jenjang pendidikan tertentu (*summative*) dan hasilnya digunakan sebagai laporan kepada siswa tentang hasil belajarnya, kepada guru, orang tua siswa, masyarakat dan pemerintah sebagai wujud akuntabilitas penyelenggaraan pendidikan. (Nur-iyah, 2014: 79).

Secara prinsipil evaluasi merupakan suatu kegiatan penilaian yang bertujuan untuk mengukur tingkat efektifitas kegiatan dalam mencapai tujuan yang diharapkan. Secara fungsional kegiatan penilaian merupakan kegiatan mencari informasi yang akan dijadikan landasan menentukan kebijakan selanjutnya. (Chairawati, 2014: 16).

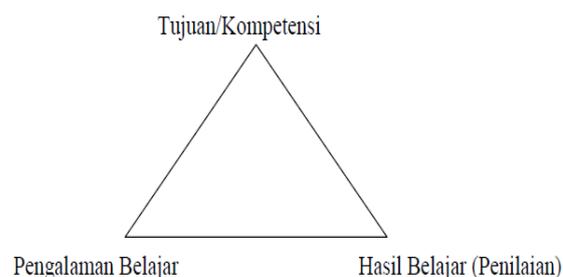
Berdasarkan fungsinya, evaluasi dapat diklasifikasikan menjadi dua bagian, yakni evaluasi proses dan evaluasi hasil pembelajaran. Evaluasi Proses merupakan kegiatan pengukuran yang dilaksanakan secara sistematis untuk memperoleh informasi tentang efektifitas aktifitas belajar

mengajar. Sedangkan evaluasi hasil belajar menunjuk pada aktifitas penilaian terhadap tingkat kualitas hasil belajar yang dicapai peserta didik. (Chairawati, 2014:16). Di sinilah maka benar apa yang dikatakan bahwa evaluasi adalah sebuah proses pengukuran dan penilaian. (Ihsan & Ihsan, 2008)

Siskandar (2016: 117) dalam penelitiannya menyimpulkan bahwa efektifitas monitoring dan evaluasi merupakan salah satu faktor yang menentukan kesuksesan implementasi Kurikulum 2013 di samping Kelengkapan Infrastruktur, Kompetensi guru, Prinsip kepemimpinan, Fasilitas guru, Lingkungan dan budaya sekolah, and Efektifitas monitoring dan Evaluasi Implementasi Kurikulum. Dan kesalahan dalam penerapan atau pelaksanaan penilaian (*Assessment*) merupakan salah satu faktor dari kegagalan implementasi kurikulum 2013 di samping kesalahan persepsi tentang Kurikulum 2013, dan kurang diterapkannya Rencana Pelaksanaan Pembelajaran, ruang kelas pembelajaran, penggunaan media IT,

fasilitas mengajar dan manajemen madrasah.

Hubungan antara penilaian dan pembelajaran secara ringkas digambarkan Nuriyah (2014: 75) sebagai berikut ini:



Gambar 1: Hubungan Penilaian dan Pembelajaran

Secara umum, alat penilaian (*instrument*) dapat dikategorikan ke dalam dua bentuk yakni: 1) Tes; dan 2) Bukan Tes (*non-test*). Alat pengukuran yang termasuk ke dalam kategori non tes adalah: a) Kuisisioner; b) Wawancara; c) Daftar Cocok (*check list*); d) Pengamatan atau Observasi ; e) Penugasan; f) Portofolio; g) Jurnal; h) Inventori; i) Penilaian diri (*Self Assesment*); j) Penilaian oleh teman (*peer assessment*). Sedangkan tes adalah sejumlah pertanyaan yang harus dijawab, atau pertanyaan-pertanyaan yang harus dipilih atau

ditanggapi, tugas-tugas yang harus dikerjakan oleh orang yang diuji pada waktu tertentu. Tes merupakan sejumlah pertanyaan yang memiliki jawaban benar atau salah, pertanyaan yang membutuhkan jawaban atau diberi tanggapan untuk mengukur tingkat kemampuan seseorang dalam aspek tertentu. (Wening, 2012 : 4).

Adapun prinsip-prinsip penilaian hasil belajar peserta didik adalah sebagai berikut: 1) Mendidik; 2) Terbuka / Transparan; 3) Menyeluruh; 4) Terpadu dengan pembelajaran; 5) Obyektif; 6) Sistematis; 7) Berkesinambungan; 8) Adil (*fair*); 9) Menggunakan acuan kriteria. (Wening, 2012: 3).

Dalam penilaian kompetensi pengetahuan, pendidik dapat menggunakan tes tulis, tes lisan dan penugasan (*resitasi*). Tes tulis dapat berbentuk soal-soal yang bersifat obyektif seperti: pilihan ganda, isian, jawaban singkat, benar-salah dan menjodohkan. Atau bersifat subyektif seperti soal uraian. Instrument tes lisan dapat berupa daftar pertanyaan. Instrument penugasan berupa pekerjaan rumah

atau proyek yang dikerjakan secara individu atau kelompok sesuai dengan karakteristik tugas.

Sebuah tes dapat dikatakan layak sebagai alat pengukur jika memiliki beberapa persyaratan sebagai berikut: (1) *Valid* (Ketepatan); (2) *Reliabel* (Handal); (3) Obyektif; (4) Praktikabilitas; (5) Ekonomis. Valid berarti sebuah tes mampu mengukur apa yang seharusnya diukur. Sekumpulan soal ujian adalah valid untuk mengukur prestasi dan tidak valid jika digunakan untuk mengukur tingkat partisipasi. Reliabel berarti sebuah instrumen harus dapat diandalkan sebagai alat pengukur. Jika sebuah instrumen diujikan secara berkali-kali dan hasilnya adalah sama atau mendekati sama, maka instrumen tersebut dapat dikatakan reliabel atau dapat diandalkan. Obyektif artinya bahwa sebuah instrumen penilaian haruslah bebas dari pandangan diri penilai. Praktikabilitas berarti sebuah instrumen haruslah mudah dioperasikan (dijalankan). Ekonomis berarti bahwa sebuah instrumen haruslah mudah dijang-

kau dalam hal pembiayaannya. (Wening, 2012: 4-6).

### Tingkat Kesulitan (Difficulty Level)

#### Test

Pemberian tes adalah sebuah pendekatan yang diambil oleh para ahli yang menganut model pengukuran (*measurement*) klasik dalam evaluasi. Model ini berpandangan bahwa pengukuran adalah suatu kegiatan yang ilmiah dan dapat diterapkan dalam berbagai bidang termasuk pendidikan. Obyek dari pengukuran ini adalah hasil belajar peserta didik. Sedang tujuannya adalah untuk mengungkapkan perbedaan individual. (Ibrahim & Ali, 2007: 106-108).

Kemudian, mengingat bahwa tujuan dari pengukuran adalah untuk mengungkapkan perbedaan individual peserta didik dilihat dari hasil belajarnya, maka dalam penyusunan soal tes harus memperhatikan masalah daya pembeda dan tingkat kesukaran masing-masing soal dan daya pembeda dari sebuah test. Daya pembeda adalah tingkat kemampuan soal untuk mengung-

kapkan perbedaan kemampuan individual peserta didik. Setiap soal yang dapat dijawab dengan benar oleh hampir semua siswa adalah soal dengan daya pembeda yang rendah. Demikian pula soal yang tidak dapat dijawab dengan benar oleh hampir semua siswa juga memiliki daya pembeda yang rendah pula.

Tingkat kesukaran (*difficulty level*) suatu soal adalah proporsi atau presentase subyek yang menjawab butir tes tertentu dengan benar. Sedang angka yang menunjukkan sukar atau tidaknya butir soal dalam tes disebut dengan indeks (dilambangkan dengan  $p$ ). Indeks ini terentang antara 0 sampai 1. Diharapkan dalam sebuah soal tes disusun dengan tingkat kesukaran yang merata (proporsional) antara soal yang tingkat mudah, sedang dan sulit. (Ibrahim & Ali, 2007: 107-108); (Rasyid & Mansur, 2008: 239); (Rasyid & Mansur, 2008: 245). Artikel ini fokus membahas tentang tingkat kesulitan test.

Dalam membicarakan tentang tingkat kesukaran atau tingkat

kesulitan butir soal, maka seorang guru pembuat soal haruslah memahami dan memperhatikan butir soal dan peserta tes. (Rasyid & Mansur, 2008: 239). Yang dimaksud dengan butir soal adalah materi pelajaran yang dibuat soal. Sedangkan yang dimaksud dengan peserta test adalah kemampuan menjawab (*intake*) peserta didik terhadap materi. Bisa saja sebuah butir soal adalah mudah bagi sebagian peserta didik namun dianggap sulit bagi sebagian yang lain.

Asumsi yang digunakan dalam menyusun sebuah soal tes yang dengan kualitas yang baik adalah adanya keseimbangan dalam tingkat kesulitannya disamping dalam masalah validitas dan reliabilitasnya. Keseimbangan dalam tingkat kesulitan yang dimaksud adalah bahwa dalam sekian jumlah soal harus ada butir-butir soal yang bersifat sulit, sedang dan mudah secara proporsional. (Rasyid & Mansur, 2008: 240).

Persoalan yang lain dalam menentukan tingkat kesulitan (demikian pula daya pembeda) soal

adalah ukuran untuk menentukan apakah soal tersebut memiliki daya pembeda rendah, sedang atau tinggi. Apakah soal tersebut mudah, sedang atau sulit. Dalam hal ini, maka dapat dilaksanakan uji otoritas, yakni otorits gurulah yang menentukan ukuran rendah-tingginya daya pembeda atau mudah sulitnya sebuah soal. Hal ini dengan mempertimbangkan: 1) Kemampuan subyek tes (*testee*) dalam menjawab butir soal; 2) sifat materi yang diujikan atau dinyatakan; 3) isi bahan yang dinyatakan sesuai dengan bidang keilmuannya, baik luasnya maupun kedalamannya; 4) bentuk tes obyektif apakah pilihan ganda, benar-salah atau mencocokkan. Tipe benar salah adalah lebih mudah dari pada pilihan ganda. (Rasyid & Mansur, 2008: 240).

Namun demikian penentuan otoritatif sebagaimana dijelaskan di atas tersebut belumlah bersifat secara mutlak. Sebab masih memerlukan pembuktian secara empirik. Karena bisa jadi sebuah soal dianggap sulit berdasarkan keputusan *judgement* seorang guru

(pemangku otoritas) namun secara empirik adalah sebaliknya. Untuk itu diperlukan beberapa teknik komputasi dalam menentukannya.

Untuk mengukur tingkat kesulitan soal (selanjutnya disimbolkan dengan:  $p$ ) digunakan beberapa cara, yaitu: 1) proporsi menjawab benar; 2) skala kesukaran linera; 3) indeks davis; 4) skala bivariat. (Rasyid & Mansur, 2008: 241).

### Metode Penelitian

Analisis dalam penelitian ini menggunakan 2 (dua) metode yaitu: (1) Analisis statistik deskriptif atau yang disebut juga dengan statistik sosial (Ritzer, 2001) pada data yang berupa prosentase; (2) metode analisis tingkat kesulitan (*difficulty level*) tes dengan teknik analisis proporsi menjawab benar sebagaimana dijelaskan dalam Rasyid & Mansur (2008: 241).

### Pembahasan

Penghitungan tingkat kesulitan test dalam artikel ini dilakukan pada 20 (dua puluh) soal dalam bentuk pilihan ganda (*multiple choice*) pada

latihan Bab 1 dan 2 pada Buku Guru Kurikulum 2013 mata pelajaran Sejarah Kebudayaan Islam (SKI) Kelas 9 (Sembilan) yang diterbitkan oleh Kementerian Agama Jakarta Revisi tahun 2016. Dua bab yang dimaksud adalah pembahasan tentang: (1) Jejak Peradaban Dinasti Abbasiyah; dan (2) Kecemerlangan Ilmuwan Muslim Dinasti Abbasiyah. Soal latihan dari kedua bab tersebut terdiri dari 10 (sepuluh) butir soal yang berasal dari Bab 1, dan 10 (sepuluh) soal dari Bab 2. Butir soal disusun secara berurutan yakni 10 (sepuluh) soal secara berurutan adalah salah yang diambil dari Bab 1, dilanjutkan 10 (sepuluh) soal yang diambil dari Bab 2. Subyek penelitian ini adalah siswa kelas 9 (Sembilan) MTs Nahdlatul Ulama Kepuharjo Karangploso Malang sebanyak 36 siswa yang diambil secara populasi tanpa menggunakan sample.

Terkait dengan penelitian ini, peneliti menggunakan cara mencari proporsi menjawab benar dengan rumus:

$$p_i = \frac{\sum x_i}{Sm_i N}$$

Dimana:

$p_i$  = tingkat kesukaran butir  $i$  atau proporsi menjawab benar butir  $i$

$\sum x_i$  = banyaknya peserta tes yang menjawab benar butir  $i$

$S_{mi}$  = skora maksimum

$N$  = jumlah peserta tes

Peneliti menghitung masing-masing tingkat kesulitan ke 20 (dua puluh) soal tersebut dengan menggunakan bantuan aplikasi Microsoft Excell 2010 yang dijalankan di bawah sistem operasi Winodows 10. Penghitungan dimulai dengan melakukan analisis butir soal dan memaparkan proporsi jawaban benar siswa terhadap setiap butir soal. Data ini dipaparkan dalam bentuk prosentase.

Penelitian dilanjutkan dengan menentukan kriteria tingkat kesulitan setiap soal tes. Yang mana kriteria ini didasarkan pada apa yang dijelaskan oleh Rsyid & Mansur (2008) yakni: (a)  $p \leq 0.30$  = butir soal sulit; (b)  $0.30 < p \leq 0.70$  = butir soal sedang; dan (c)  $p > 0.70$  = butir soal sulit.

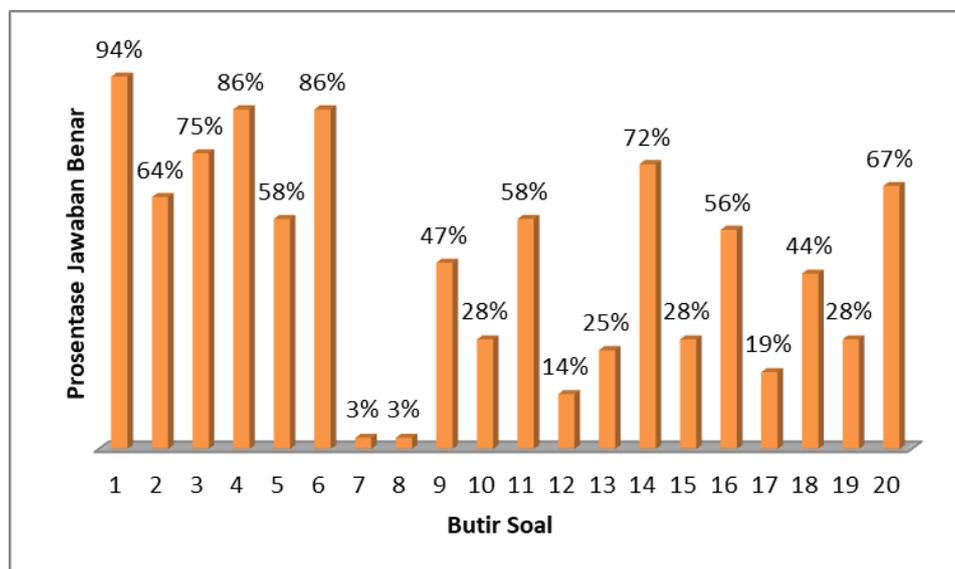
Berdasarkan hasil penghitungan prosentase jawaban benar siswa terhadap setiap butir soal, dapat diapaparkan data sebagai berikut: (1) sebesar 94% (34 siswa) dari

keseluruhan siswa menjawab benar pada butir soal nomor 1; (2) sebesar 64% (23 siswa) dari keseluruhan siswa menjawab benar pada butir soal nomor 2; (3) sebesar 75% (27 siswa) dari keseluruhan siswa menjawab benar pada butir soal nomor 3; (4) sebesar 86% (31 siswa) dari keseluruhan siswa mampu menjawab benar pada butir soal nomor 4; (5) sebesar 58% (21 siswa) dari keseluruhan siswa mampu menjawab benar pada butir soal nomor 5; (6) sebesar 86% (31 siswa) dari keseluruhan siswa mampu menjawab benar pada butir soal nomor 6; (7) sebesar 3% (1 siswa) dari keseluruhan siswa mampu menjawab benar pada butir soal nomor 7; (8) sebesar 3% (1 siswa) dari keseluruhan siswa mampu menjawab benar pada butir soal nomor 8; (9) sebesar 47% (17 siswa) dari keseluruhan siswa mampu menjawab benar pada butir soal nomor 9; (10) sebesar 28% (10 siswa) dari keseluruhan siswa mampu menjawab benar pada butir soal nomor 10; (11) sebesar 58% (21 siswa) dari keseluruhan siswa mampu menjawab benar pada butir soal nomor 11; (12) sebesar 14% (5 siswa) dari keseluruhan

siswa mampu menjawab benar pada butir soal nomor 12; (13) sebesar 25% (9 siswa) dari keseluruhan siswa mampu menjawab benar pada butir soal nomor 13; (14) sebesar 72% (26 siswa) dari keseluruhan siswa mampu menjawab benar pada butir soal nomor 14; (15) sebesar 28% (10 siswa) dari keseluruhan siswa mampu menjawab benar pada butir soal nomor 15; (16) sebesar 56% (20 siswa) dari keseluruhan siswa mampu menjawab benar pada butir soal nomor 16; (17) sebesar 19% (7 siswa) dari keseluruhan siswa mampu

menjawab benar pada butir soal nomor 17; (18) sebesar 44% (16 siswa) dari keseluruhan siswa mampu menjawab benar pada butir soal nomor 18; (19) sebesar 28% (10 siswa) dari keseluruhan siswa mampu menjawab benar pada butir soal nomor 19; (20) sebesar 67% (24 siswa) dari keseluruhan siswa mampu menjawab benar pada butir soal nomor 20.

Data-data tersebut jika disajikan dalam bentuk diagram batang maka akan terlihat sebagai berikut:

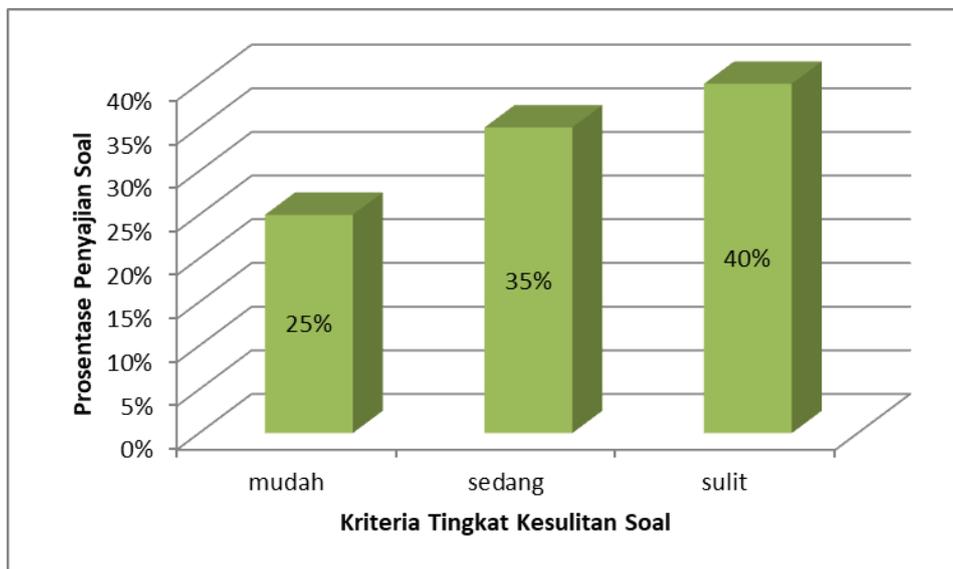


Gambar 2. Diagram batang prosentase jawaban benar siswa

Berdasarkan hasil analisis, dapat dipaparkan data bahwa: (a) Sebanyak 25% atau sejumlah 5 soal termasuk dalam kriteria mudah; (b)

Sebanyak 35% atau sejumlah 7 soal termasuk dalam kriteria sedang; dan (c) sebanyak 40% atau sejumlah 8 soal termasuk dalam kriteria sulit.

Jika disajikan ke dalam bentuk diagram batang maka data tersebut akan terlihat sebagai berikut ini:



Gambar 3. Prosentase Penyajian Soal

Butir-butir soal yang termasuk ke dalam kategori mudah adalah butir soal nomor: 1, 3, 4, 6, dan 14. Butir soal yang termasuk ke dalam kategori sedang adalah: 2, 5, 9, 11, 16, 18 dan 20. Sedangkan butir soal yang termasuk ke dalam kategori sulit adalah butir soal pada nomor: 7, 8, 10, 12, 13, 15, 17 dan 19.

Dengan demikian, berdasarkan hasil analisis data di atas dengan menggunakan kacamata teori yang dinyatakan oleh Ibrahim & Ali (2007) dapat disimpulkan bahwa soal-soal tersebut dapat dikatakan ideal sebagai soal latihan.

### Simpulan

Berdasarkan pembahasan di atas dapat disimpulkan bahwa soal mata pelajaran Sejarah Kebudayaan Islam (SKI) untuk kelas 9 MTs pada pembahasan: (1) Jejak Peradaban Dinasti Bani Abbasiyah; dan (2) Kecemerlangan Ilmuwan Muslim Dinasti Bani Abbasiyah, pada Buku Guru yang diterbitkan oleh Kementerian Agama tahun 2015 disajikan dengan proporsi: (1) 25% soal dalam kategori mudah; (2) 35% soal dalam kategori sedang; dan (3) 40% soal dalam kategori sulit (sukar). Dengan demikian, bahwa soal-soal

tersebut dapat dikatakan ideal sebagai soal latihan.

---

## DAFTAR PUSTAKA

- Arikunto, Suharsimi & Jabar, Cepi Safrudin. (2014). *Evaluasi Program Pendidikan, Pedoman Teoritis bagi Mahasiswa dan Praktisi Pendidikan*. Jakarta: Bumi Aksara.
- Chairawati, Fajri. (2014). Evaluasi Pembelajaran pada Kelas Internasional Fakultas Dakwah IAIN Ar-Raniry. *Jurnal Al-Bayan*. 20 (29).
- Dokumen Kurikulum 2013. (2012). Kementerian Pendidikan dan Kebudayaan
- Ibrahim, R. & Ali, M. (2007). Teori Evaluasi Pendidikan. Dalam: Ali, M., Ibrahim, R., Sukmadinata, N.D., Sudjana, D., & Rasjidin, W. (ed). *Ilmu dan Aplikasi Pendidikan*. Bandung: PT. Imperial Bhakti Utama.
- Ihsan, Hamdani & Ihsan, Fuad. (2008). *Filsafat Pendidikan Islam*. Bandung: CV. Aneka
- Kartowagiran, Badrun. (2014). *Penilaian Berbasis Kurikulum 2013 (Makalah disampaikan pada: Pelatihan Penilaian Otentik Bagi Guru SMP di Wonosari yang Mengimplementasikan Kurikulum 2013, pada 29 Agustus 2014)*. Yogyakarta: Universitas Negeri Yogyakarta
- Lampiran Peraturan Menteri Pendidikan dan Kebudayaan Nomor 22 Tahun 2016 Tentang Standar Proses Pendidikan dasar dan menengah
- Nuriyah, Nunung. (2014). Evaluasi Pembelajaran: Sebuah Kajian Teori. *Jurnal Edueksos*. 03 (1).
- Rasyid, Harun & Mansur. (2008). *Penilaian Hasil Belajar*. Bandung: CV. Wacana Prima.
- Ritzer, George (ed). (2001). *Handbook Teori Sosial*. Jakarta: Nusamedia
- Siskandar. (2016). Evaluasi Implementasi Kurikulum 2013 di Madrasah Aliyah. Dalam: *Cendekia*. 10 (2).
- Wening, Sri. (2012). *Materi Evaluasi Pembelajaran*. Yogyakarta: Fakultas Teknik Universitas Negeri Yogyakarta.