

# Analisis Soal Ujian Mata Pelajaran Bahasa Arab Kelas 12 di Madrasah Aliyyah

**Iza Zainal Ambiya**

Institut Agama Islam Negeri (IAIN) Syekh Nurjati Cirebon

Email: [izaambiya99@gmail.com](mailto:izaambiya99@gmail.com)

**Sopwan Mulyawan**

Institut Agama Islam Negeri (IAIN) Syekh Nurjati Cirebon

Email: [sopmulyawan@gmail.com](mailto:sopmulyawan@gmail.com)

**Hasan Saefuloh**

Institut Agama Islam Negeri (IAIN) Syekh Nurjati Cirebon

Email: [hazansf@yahoo.com](mailto:hazansf@yahoo.com)

Diterima: 1 Mei 2022

Publish: 25 Juni 2022

## Abstrak

*Analisis butir soal merupakan alat untuk mengetahui apakah soal tersebut sudah cukup dengan kategori soal yang baik atau belum. Tujuan analisis butir soal untuk membantu meningkatkan kualitas tes melalui revisi atau membuang soal yang tidak efektif, serta untuk mengetahui informasi diagnostik tingkat pemahaman siswa pada materi yang telah diajarkan. Dalam hal ini masih banyak nilai siswa di bawah standar KKM. Tidak adanya analisis terhadap soal sehingga tidak akan tahu apakah soal tersebut dikatakan sebagai soal yang baik atau tidak. Sedangkan guru hanya langsung mengoreksi benar dan salah saja. kemudian tidak diujikan terlebih dahulu kerana berbagai hal. Oleh sebab itulah peneliti ingin melakukan analisis terhadap soal ujian madrasah yang dibuat oleh guru sendiri. Dalam penelitian ini peneliti menggunakan metode deskriptif kuantitatif. Peneliti menemukan bahwa hubungan antara soal dengan kemampuan berbahasa mencakup 2 kemampuan berbahasa yaitu: kemampuan “membaca” 52, 5 % dan “menulis” 47, 5%, masuk dalam kategori cukup dengan persentase 50%. Soal yang sesuai dengan Kompetensi Dasar (KD) mencakup 3 KD yaitu KD “3.1” 57, 5%. KD “3.2” 20%, dan KD “3.3” 22,5%, masuk dalam ketegori baik dengan persentase 75%. Kualitas soal ujian dari aspek validitas dan reliabilitas. Yang termasuk soal yang valid ialah 28 soal. Sedangkan yang termasuk soal tidak valid ialah 12 soal. Soal yang dibuat oleh guru memiliki reliabilitas yang tinggi dengan angka  $0,711 > r$  tabel  $0,254$ . Butir soal ujian ditinjau dari tingkat penalaran yang lebih tinggi. Butir tes menunjukkan tingkat domain kognitif, C1 “mengingat” mendapat 37, 5%, C2 “memahami” 20%, C3 “menerapkan” 22, 5%, C4 “menganalisis” 20%. Adapun soal yang menunjukkan tingkat HOTS ialah 20%, sedangkan soal yang menunjukkan tingkat LOTS adalah 80%.*

**Kata kunci:** analisis butir soal, soal ujian, bahasa arab

## **Abstract**

*Item analysis is a tool to find out whether the question is sufficient with a good question category or not. In addition, the purpose of item analysis is also to help improve the test through revision or removing ineffective questions, as well as to find out diagnostic information to students whether or not they have understood the material that has been taught. In this case, there are still a lot of student scores that are still below the KKM standard. There is no analysis of the problem so that one will not know whether the question is said to be a good question or not. Meanwhile, the teacher only corrects right and wrong. then not tested first for various reasons. That's why the researcher wanted to analyze the madrasa exam questions made by the teacher himself. In this research, the researcher uses quantitative descriptive method, which is a research that describes the data analysis in numbers. The researcher found that the relationship between questions and language skills included 2 language skills, namely: the ability to "read" 52.5% and "write" 47.5%, included in the sufficient category with a percentage of 50%. The questions that are in accordance with the Basic Competence (KD) include 3 KD, namely KD "3.1" 57.5%. KD "3.2" 20%, and KD "3.3" 22.5%, are in the good category with a percentage of 75%. The quality of the exam questions from the aspect of validity and reliability. Which includes valid questions are 28 questions. Meanwhile, 12 questions that are not valid. The questions made by the teacher have high reliability with the number  $0.711 > r$  table  $0.254$ . The test items are viewed from a higher level of reasoning. The test items showed the level of cognitive domain, C1 "remembered" got 37.5%, C2 "understood" 20%, C3 "applied" 22.5%, C4 "analyzed" 20%. The questions that show the HOTS level are 20%, while the questions that show the LOTS level are 80%.*

**Keywords:** *item analysis, exam questions, arabic language*

## **Pendahuluan**

Tes merupakan salah satu dari berbagai metode penilaian, dan merupakan sarana utama yang mengukur tingkat pencapaian siswa, dan mengidentifikasi sejauh mana kurikulum telah mencapai tujuan. dengan demikian, tes dapat bekerja untuk memperbaiki dan mengembangkan proses pendidikan dan memungkinkannya mencapai tujuan yang diinginkan sehingga dapat membuat suatu keputusan (Madkūr, 2010: 545).

Dewasa ini berbagai studi melaporkan bahwa untuk menghadapi tantangan dan perkembangan abad modern ini diperlukan bukan sekedar pengetahuan konseptual semata, melainkan keterampilan mengaplikasikan pengetahuan dan berbagai keterampilan berpikir.

Terkait dengan hal tersebut, *Partnership for 21st Century Skills* merumuskan beberapa keterampilan yang selanjutnya disebut sebagai kecakapan abad 21 atau sering disebut *21st Century Skills*. Beberapa keterampilan yang termuat dalam kecakapan abad 21 tersebut diantaranya yaitu kreativitas, keterampilan berpikir kritis dan pemecahan masalah. Keterampilan-keterampilan keterampilan berpikir tingkat tinggi mulai diperhatikan dengan diterapkannya Kurikulum 2013. Dengan demikian keterampilan berpikir tingkat tinggi atau *higher order thinking skills (HOTS)* menjadi tujuan utama dalam proses pembelajaran.

Hasil studi internasional *Programme for International Student Assessment (PISA)* tahun 2015 dan 2018 menunjukkan prestasi literasi membaca (*reading literacy*), literasi matematika (*mathematical literacy*), dan literasi sains (*scientific literacy*) yang dicapai siswa Indonesia sangat rendah. Pada umumnya kemampuan siswa Indonesia sangat rendah dalam (1) memahami informasi yang kompleks; (2) teori, analisis dan pemecahan masalah; (3) pemakaian alat, prosedur dan pemecahan masalah; dan (4) melakukan investigasi.

Berdasarkan fakta-fakta di atas, maka penilaian yang dikembangkan oleh guru madrasah diharapkan dapat mendorong peningkatan kemampuan berpikir tingkat tinggi, meningkatkan kreativitas, dan membangun kemandirian peserta didik untuk menyelesaikan masalah. Oleh karena itu, Direktorat Jendral Pendidikan Islam menyusun Petunjuk Teknis Penulisan Soal *Higher Oredor Thinking Skills (HOTS)* (Direktur Jendral Pendidikan Islam, 2021: 5-6).

Dalam muatan kurikulum 2013, sudah mulai mengimplementasikan soal berbasis *HOTS*. Soal-soal *HOTS* merupakan instrumen yang digunakan untuk mengukur keterampilan berpikir tingkat tinggi, yaitu keterampilan berpikir yang tidak sekadar mengingat (*remembering*), memahami (*understanding*), atau menerapkan (*applying*).

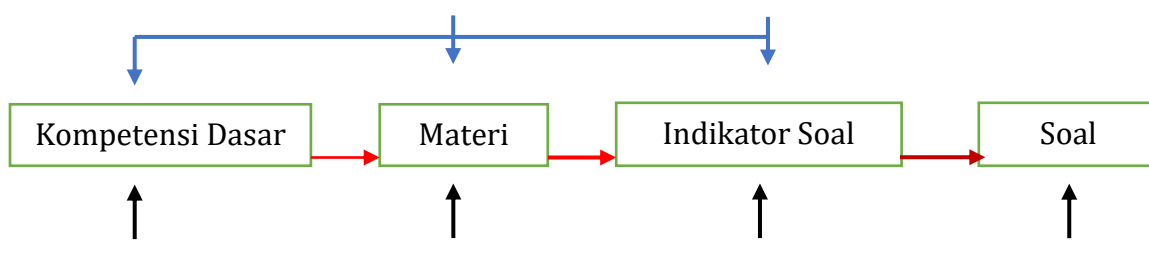
Untuk keberhasilan proses pendidikan, tujuan umum dan tujuan khusus harus ditentukan, dan skala yang sesuai dengan masing-masing tujuan ini harus dipilih, dengan mempertimbangkan kondisi tes yang baik. (Muhammad Al-Jagub, 2002: 26).

Keberhasilan penilaian hasil belajar sangatlah berkaitan dengan instrumen tes yang digunakan. Hal ini menjadikan kualitas intrumen tes sebagai suatu hal yang penting dalam proses penilaian hasil belajar agar benar-benar mengukur hal-hal yang dicantumkan dalam tujuan kurikuler. Untuk mencapai hal tersebut maka instrumen tes harus benar-benar

memenuhi syarat minimum sebagai tes yang baik yakni mengenai validitas, reliabilitas, daya pembeda dan tingkat kesukaran (Tilaar, 2019: 57-68).

Dalam penilaian soal tes prestasi belajar, misalnya ulangan harian, tes formatif, sumatif, dan ujian sekolah, penulis soal perlu memiliki pengetahuan tentang proses penjabaran kompetensi dasar menjadi indikator soal. Indikator soal dibuat untuk melihat ketercapaian kompetensi dasar yang dituntut dalam kurikulum (Tim Pusat Penilaian Pendidikan, 2019: 12).

Berikut ini adalah diagram yang menggambarkan proses penjabaran kompetensi dasar menjadi indikator:



- Menunjukkan adanya kesesuaian soal dengan indikator soal, materi dan kompetensi dasar
- Menunjukkan kesesuaian indikator soal dengan materi yang hendak diukur yang mengacu pada kompetensi dasar.
- Menunjukkan alur penulisan soal.

Menurut Umar kesesuaian soal dari aspek materi perlu diperhatikan. Salah satu kaidah dalam penulisan soal disebutkan bahwa soal haruslah sesuai dengan indikator (Jahja Umar, 1996: 36).

Dalam menyusun soal bahasa, harus mencakup dua hal yang sangat penting yaitu kompetensi bahasa dan unsur bahasa. Diantara kompetensi bahasa yaitu *istima* (mendengar), *kalam* (berbicara), *qiraah* (membaca) dan *kitabah* (menulis). Sedangkan unsur bahasa diantaranya *ashwat* (bunyi), *mufradat* (kosakata) dan *qawaid* (gramatikal). Seperti yang diungkapkan Ibrahim dan Yanti Kompetensi berbahasa terdiri atas kompetensi menyimak dan kompetensi membaca yang bersifat aktif reseptif serta berbicara dan menulis yang bersifat aktif produktif, sedangkan unsur kebahasaan terkait dengan pengetahuan tentang

sistem bahasa yaitu kosakata, struktur tata bahasa, dan ejaan (Nini Ibrahim dan Prima Gusti Yanti, 2017: 95).

Setelah guru membuat soal yang sesuai dengan KD perlu dilakukan kegiatan menganalisis soal. Untuk mengetahui sejauh mana soal yang dibuat apakah sudah bisa dikatakan sebagai soal yang baik atau tidak.

Kegiatan menganalisis butir soal merupakan suatu kegiatan yang harus dilakukan guru untuk meningkatkan mutu soal yang telah ditulis. Kegiatan ini merupakan proses pengumpulan, peringkasan, dan penggunaan informasi dari jawaban siswa untuk membuat keputusan tentang setiap penilaian. Analisis butir soal ini berisi uji validitas, uji reliabilitas, tingkat kesukaran, daya pembeda dan fungsi pengecoh atau *distractor* soal. Tujuan penelaahan adalah untuk mengkaji dan menelaah setiap butir soal agar diperoleh soal yang bermutu sebelum soal digunakan. Di samping itu, tujuan analisis butir soal juga untuk membantu meningkatkan tes melalui revisi atau membuang soal yang tidak efektif, serta untuk mengetahui informasi diagnostik pada siswa apakah mereka sudah atau belum memahami materi yang telah diajarkan (Dirjen Menejamen Pendidikan Dasar dan Menengah, 2008: 166).

Berdasarkan hasil wawancara dengan guru mata pelajaran Bahasa Arab yakni Bapak Deni Hilman Abdillah M. Ag, kelas XII MAN 3 Majalengka, masih banyak sekali siswa yang kurang mampu dalam mengerjakan soal-soal mata pelajaran Bahasa Arab sehingga nilai siswa masih di bawah KKM. Tidak adanya pengujian terhadap soal khususnya dalam uji validitas dan reliabilitas, guru hanya langsung mengoreksi benar dan salah saja. kemudian tidak diujikan terlebih dahulu kerana berbagai hal, seperti waktu, tenaga, biaya dan kemampuan guru dalam menganalisisnya sendiri.

Berdasarkan permasalahan di atas, maka perlu diadakan penelitian analisis butir soal. Soal yang digunakan dalam penelitian ini, yaitu soal Ujian Madrasah (UM) yang dibuat oleh guru sendiri. Peneliti mengambil kelas XII karena jenjang studi bahasa Arab yang sudah ditempuh selama kurang lebih tiga tahun. Dengan tujuan dapat mengetahui kemampuan siswa pada mata pelajaran bahasa Arab sehingga dapat memberikan sumbangan pada proses perkembangan perbaikan evaluasi atau penilaian di Madrasah Aliyah Negeri 3 Majalengka. Namun pada penelitian ini, peneliti fokus pada analisis soal Ujian Madrasah (UM) pada aspek hubungan soal dengan kompetensi berbahasa, kesesuaian antara soal dengan kompetensi

dasar, kualitas soal dari segi validitas dan reliabilitas, dan soal dalam persepektif *HOTS* (*Higher Oreder Thinking Skill*).

### **Metode Penelitian**

Penelitian ini menggunakan metode deskriptif kuantitatif, yaitu Dari pemaparan tersebut, maka penelitian deskriptif kuantitatif adalah penelitian yang pendeskripsian analisis datanya dinyatakan dalam angka-angka. Menurut Ibnu dalam Ainin penelitian deskriptif adalah suatu penelitian yang analisis datanya hanya sampai pada deskripsi variabel satu demi satu (Ainin, 2010: 12).

Populasi yang digunakan dalam penelitian ini adalah soal-soal ujian akhir bahasa Arab kelas XII di MAN 3 Majalengka Tahun Ajaran 2020/2021. Sedangkan sampel dalam penelitian ini adalah siswa kelas 12 yaitu kelas MIPA 1 dan kelas MIPA 2 di MAN 3 Majalengka tahun ajaran 2020/2021 dengan jumlah siswa kelas 12 MIPA 1 terdiri dari 30 siswa dan kelas 12 dan MIPA 2 terdiri dari 27 siswa, sehingga jumlah sampel dalam penelitian ini adalah 57 siswa.

Teknik pengumpulan data yang digunakan dalam penelitian ini adalah teknik wawancara dan dokumentasi. Wawancara adalah kuesioner lisan, apa adanya (percakapan terarah antara peneliti dan orang atau orang lain dengan tujuan untuk sampai pada fakta atau situasi tertentu, yang peneliti coba identifikasi untuk mencapai tujuan penelitian). Wawancara dengan guru bahasa Arab kelas 12 di MAN 3 Majalengka, untuk mengetahui informasi tentang status dan atau kondisi siswa kelas 12 di MAN 3 Majalengka tahun ajaran 2020/202.

Metode dokumentasi adalah pelaksanaan pengumpulan data melalui informasi atau data yang akan dianalisis dalam bentuk dokumen. Teknik dokumentasi digunakan untuk mengumpulkan data dari sumber non-manusia. Sumber-sumber tersebut terdiri dari dokumen dan rekaman (Syamsudin dan Vismaia S. Damaianti, 2015: 108). Adapun hasil dari metode dokumentasi ini diantaranya: lembar jawaban siswa soal ujian akhir bahasa Arab kelas XII di Madrasah Aliyah Negeri 3 Majalengka tahun ajaran 2020/2021. Kunci jawaban soal ujian akhir bahasa arab ke XII di Madrasah Aliyah Negeri 3 Majalengka tahun ajaran 2020/2021. Kisi-kisi soal ujian akhir Bahasa Arab di Madrasah Aliyah Negeri 3 Majalengka tahun ajaran 2020/2021.

Analisis data dilakukan dengan menggunakan lembar analisis dokumen yang didalamnya memuat tabel dengan format kolom nomor soal, dan kemunculan aspek tertentu di dalam soal ujian madrasah. Data yang diperoleh kemudian diolah dengan menggunakan rumus persentase (Aa Juhanda, 2016: 61–66).

$$\% \text{ kemunculan} = \frac{\Sigma \text{ kemunculan aspek tertentu di dalam soal}}{\Sigma \text{ seluruh soal}} \times 100\%$$

Nilai persentase kemunculan kemudian dikelompokkan berdasarkan kategori sebagai berikut: persentase 81-100% dikategorikan dengan predikat: baik sekali; 61-80% (baik); 41-60% (cukup); 21-40% (kurang); dan persentase < 21% dikategorikan sebagai kurang sekali.

Untuk menganalisis validitas dan menganalisis reliabilitas. Metode yang digunakan dalam menganalisis data ini adalah dengan menganalisis butir soal dengan bantuan software Spss Statistic 26. Spss Statistic 26 merupakan software pengolah data yang cukup terkenal di dunia pendidikan. Sangat mudah bagi guru untuk memahami penggunaannya di sekolah. Semua data ditulis dalam format program yang diberikan (Statistik Spss 26) (Dirjen Menejamen Pendidikan Dasar dan Menengah, 2008: 17).

## **Pembahasan**

### **Analisis**

Analisis adalah penggunaan alat bantu bimbingan dalam usaha mencapai pokok atau sub tujuan yang telah ditetapkan oleh sistem pemecahan masalah. Soal adalah alat yang mewakili blok bangunan tes, baik tertulis maupun lisan. Selama pertanyaan itu ditemukan, pasti ada yang bertanya (guru atau siswa), dan harus ada yang menjawab (siswa atau guru) (Majdi 'Aziz Ibrāhīm, 2009 : 641).

Analisis butir soal adalah pengkajian pertanyaan-pertanyaan dalam suatu perangkat tes agar diperoleh perangkat penilaian yang memiliki kualitas yang memadai (Ida Faridah, 2019: 154). Analisis butir soal merupakan analisis hubungan antara skor butir soal dengan skor keseluruhan, di mana jawaban siswa membandingkan suatu butir soal dengan jawaban keseluruhan.

Berdasarkan beberapa definisi dari analisis butir soal di atas, analisis butir soal merupakan alat untuk mengetahui apakah soal tersebut sudah cukup dengan kategori soal yang baik atau belum. Butir soal yang baik adalah butir yang memenuhi persyaratan sebagai kriteria sifat-sifat butir yang baik. Analisis butir soal dilakukan terhadap sejumlah besar butir soal dalam pemilihan hasil belajar. Analisis butir soal akan menggugurkan sebagian butir soal

yang dianalisis karena sifatnya yang tidak memenuhi syarat sebagai butir yang baik sehingga tidak memiliki kemampuan untuk mengukur hasil belajar yang baik.

### **Tes/Ujian**

Tes berasal dari bahasa latin *testum* yang berarti alat untuk mengukur bumi. Dalam bahasa Prancis Kuno, kata *test* berarti ukuran perbedaan antara emas, perak, dan logam lainnya (M. Chabib Thoha, 1991: 43).

*Test* dalam bahasa dari kata *Ikhtabar – yakhtabiru – ikhtibaran* dalam arti tes. Ujian dan atau test: jumlah pertanyaan yang dirumuskan oleh guru untuk mengukur penguasaan pelajar dari kompetensi bahasa yang diajarkan dalam program bahasa mereka (.Usāmah Zakī al-Sayyid ‘Ali, 2016: 21).

Tes merupakan sarana evaluasi akademik, artinya merupakan kegiatan yang mengukur kemampuan siswa dalam pendidikan. Tes adalah sejumlah pertanyaan yang harus dijawab oleh peserta tes secara lisan atau tertulis. Tes menjadi proses mengukur dan mengevaluasi proses yang direpresentasikan dalam semua pekerjaan yang dilakukan guru untuk menilai tingkat pencapaian, pemahaman, dan pemahaman siswa tentang topik yang dipelajarinya (Abī Labīd, 2009: 401).

Tes memainkan peran penting dalam pendidikan. Oleh karena itu, tes yang baik sangat penting untuk pengajaran dan pembelajaran yang baik. Agar tes dapat dikatakan baik dan bebas dari kesenjangan, maka harus dicirikan oleh beberapa karakteristik yang diidentifikasi oleh ahli tes, yang terpenting adalah (‘Abdu al-Rahman Ibrāhīm Alfauzān, 2011: 93).

### **Validitas**

Validitas atau *kesahihan* menunjukkan bahwa tes itu benar-benar mengukur fungsi yang hendak diukurnya, tanpa mengukur fungsi lain yang berdampingan atau sebagai penggantinya (Saifu al-islām Sa’dun Umar, 2009: 102). Validitas adalah tes yang mengukur yang seharusnya diukur. Jika tes tersebut bertujuan untuk mengukur pemahaman, maka tidak tepat memuat pertanyaan yang mengukur ejaan atau tata bahasa.

### **Reliabilitas**



Tes yang baik ditandai dengan reliabilitas, dan uji reliabilitas adalah pengujian yang memberikan hasil yang mendekati atau hasil yang sama jika diterapkan lebih dari satu kali dalam kondisi yang sama, dua atau tiga minggu kemudian (‘Abīdāt wa akharūn, 2015: 160). Reliabilitas tes merupakan prasyarat dari alat skala yang efektif dalam mengukur fenomena yang sedang diukur. Reliabilitas tes dimaksudkan untuk memberikan tes yang identik atau hasil yang dekat dalam pengukuran aspek perilaku jika ukuran itu digunakan lebih dari sekali atau jika digunakan dengan cara lain, tes tersebut terletak pada kemampuannya untuk mendeteksi perbedaan kinerja antar individu.

### **Mata Pelajaran**

Mata pelajaran adalah sekumpulan fakta, konsep, generalisasi, hukum dan teori yang berkaitan dengan bidang atau subjek tertentu (fisika, kimia, teknik, sejarah, sastra, dll.) (Sūniyā Hānim ‘Ali Qazāmil, 2013: 64-65).

Mata pelajaran Bahasa Arab merupakan suatu mata pelajaran yang diarahkan untuk mendorong, membimbing, dan membina kemampuan serta menumbuhkan sikap positif terhadap bahasa Arab baik reseptif maupun produktif.

Adapun jenis-jenis tes yang dapat diterapkan dalam bidang linguistik, tes bahasa terutama bertujuan untuk mengukur berbagai kemampuan berbahasa dalam semua kegiatan membaca, menulis, dan bahasa ekspresif (.Walīd Ahmad Jabir, 2002: 401).

### **Tes keterampilan mendengarkan**

Tes keterampilan mendengarkan adalah tes yang menuntut siswa untuk fokus pada tiga aspek: membedakan suara, memahami unsur-unsur tertentu, dan keakraban umum dengan apa yang didengar siswa. Tujuan dari tes ini adalah agar siswa memahami semua yang mereka dengar tentang bahasa Arab (‘Ābdu al-khāliq Muhammad, 1089: 107).

### **Tes keterampilan berbicara**

Tes keterampilan berbicara adalah kemampuan mengucapkan bunyi artikulasi atau kata-kata untuk mengekspresikan pikiran berupa ide, pendapat, keinginan, atau perasaan kepada mitra bicara. Tujuan dari tes ini adalah agar siswa mencapai tingkat kefasihan yang wajar, yaitu kemampuan untuk mengungkapkan dengan cara yang dapat dimengerti (Ansā Muhammad Ahmad, 2003: 167).

### **Tes keterampilan membaca**

Kegiatan membaca merupakan usaha memahami informasi yang disampaikan melalui lambang tulisan. Agar dapat menggali informasi tertulis, diperlukan pengetahuan tentang struktur dan kosakata bahasa yang bersangkutan selain sistem ejaan. Kegiatan membaca adalah memahami konteks ekstralinguistik melalui sarana penglihatan (Acep Hermawan, 2021: 78). Biasanya, guru meminta siswa untuk membaca dalam hati sebuah teks tertulis dan kemudian menjawab pertanyaan-pertanyaan berikutnya.

### **Tes keterampilan menulis**

Tes menulis biasanya merupakan tes produktivitas, bukan pengenalan, karena menulis itu sendiri adalah keterampilan produktif, bukan keterampilan reseptif. Tes tulis bisa objektif atau subjektif (esai). Tes menulis, seperti semua tes lainnya, memiliki tingkatan: yang terendah adalah tes menulis otomatis, yaitu tulisan tangan, dan yang tertinggi adalah tulisan penelitian. Dan di antara keduanya, yang terendah dan tertinggi, ada banyak tingkatan seperti penulisan surat, penulisan kata, penulisan kalimat sederhana, penulisan kalimat majemuk, penulisan paragraf, penulisan esai, dan penulisan penelitian (‘Ali Al-Khuli, 2000: 133).

### **Soal Keterampilan Berpikir Tingkat Tinggi (HOTS)**

Di Indonesia, muatan kurikulum diarahkan pada pengembangan, khususnya keterampilan berpikir berbeda dan keterampilan berpikir tingkat tinggi yang muncul dengan penerapan kurikulum 2013. Dengan demikian, kemampuan berpikir tinggi menjadi tujuan utama dalam proses pembelajaran.

Fisher dalam Purnomo Berbicara mengenai tujuan pembelajaran, dalam dunia pendidikan hal tersebut biasanya mengacu kepada taksonomi tujuan pembelajaran. Salah satu taksonomi yang paling terkenal yaitu taksonomi Bloom yang dikemukakan oleh Benjamin S. Bloom pada tahun 1956 (Bloom, 1956). Jika dikaitkan dengan proses kognitif dalam taksonomi Bloom tersebut, istilah HOTS sering dikontraskan dengan istilah *LOTS (Lower Order Thinking Skills)*. Proses kognitif analisis (*analysis*), sintesis (*synthesis*), dan evaluasi (*evaluation*) dikategorikan sebagai *HOTS*, sedangkan pengetahuan (*knowledge*), pemahaman (*comprehension*), dan aplikasi (*application*) dikategorikan sebagai *LOTS* (Pajar Purnomo, 2019: 32).

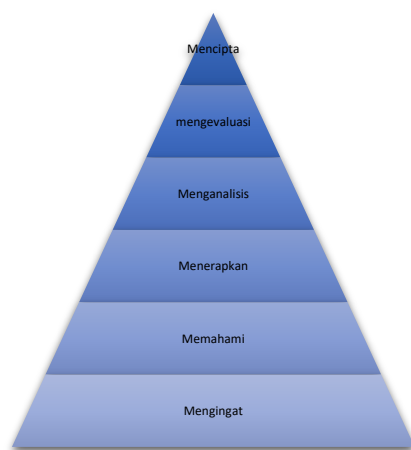
Menurut Jumiati tingkat berpikir yang lebih tinggi adalah kemampuan yang tidak hanya membutuhkan kemampuan mengingat tetapi juga membutuhkan kemampuan lain yang lebih

tinggi seperti kemampuan berpikir kreatif dan kritis. King menyebutkan dalam Sri Rahayu bahwa kemampuan berpikir tingkat tinggi meliputi berpikir kritis, logis, reflektif, metakognitif, dan kreatif (Sri Rahayu, dkk. 2020: 128).

Soal Tingkat Berpikir Tinggi (*Hots*) adalah alat ukur yang digunakan untuk mengukur efisiensi berpikir tingkat tinggi, efisiensi berpikir tingkat tinggi tidak hanya terbatas pada mengingat, memahami atau menerapkan.

Berdasarkan kutipan Shofiya dalam Hikmatu (2019: 5) mengatakan bahwa taksonomi berasal dari bahasa Yunani *taxis* yang berarti pengaturan dan *nomos* yang berarti ilmu pengetahuan. Taksonomi merupakan suatu tipe sistem klasifikasi yang berdasarkan data penelitian ilmiah mengenai hal-hal yang digolongkan-golongkan dalam sistematika itu.

Bidang ini mencakup bentuk-bentuk aktivitas intelektual manusia, terutama proses mental menghafal, memahami, dan menganalisis. Tujuan pendidikan yang berupaya mengembangkan proses mental ini termasuk dalam bidang ini. Dia (Bloom) membagi domain kognitif menjadi enam sub-level berbeda yang disusun secara hierarkis dari yang sederhana hingga yang paling kompleks, dan setiap level berisi level sebelumnya. Dengan demikian, penguasaan operasi sederhana sangat penting untuk menguasai operasi yang lebih kompleks (Tu'aimah, 2004: 37). Berikut taksonomi bloom setelah direvisi:



### Hasil Penelitian

#### Hubungan antara Soal dengan Keterampilan Berbahasa

Keterampilan Berbahasa	Nomor Soal	Jumlah	Peresentase
Mendengar	-	-	-
Berbicara	-	-	-
Membaca	1, 2, 3, 4, 5, 7, 8, 9, 10, 11, 13, 14, 16, 27, 28, 29, 34, 35, 39	21	25, 5 %

Menulis	6, 12, 15, 18, 19, 20, 22, 23, 24, 25, 26, 30, 31, 32, 33, 36, 37, 38, 40	19	47,5 %
---------	---	----	--------

Dari data di atas peneliti menemukan bahwa soal yang berkaitan dengan keterampilan berbahasa mencakup 2 keterampilan berbahasa yaitu: kemampuan membaca dan menulis. Sedangkan kemampuan berbahasa mendengar dan berbicara tidak tercakup dalam soal yang dibuat guru. Berdasarkan kategori persentase yang dikemukakan Arikunto, hubungan antara soal dengan *kemahiran* berbahasa ini masuk dalam kategori cukup dengan persentase 50%.

### Kesesuaian antara Soal dengan Kompetensi Dasar

No	Kompetensi Dasar	Nomor Soal	Jumlah	Persentase
1	3. 1. Memahami fungsi sosial, struktur teks dan unsur kebahasaan (bunyi, kata, dan makna) dari teks yang berkaitan dengan tema : التحيات والتعارف، الأسرة والبيت، المدرسة والبيئة المحيطة بها، الهوية. التسوق، الصحة، الحج والعمرة، تكنولوجيا الإعلام والاتصال، الأديان في إندونيسيا. الرياض، الشباب، الشعري العربي، الحضارة الإسلامية، الدراسة في الجامعة Yang melibatkan tindak tutur mengenalkan diri dan orang lain, meminta maaf, berterimakasih, dan berpamitan dengan memperhatikan تقسيم الكلمة – الأرقام ١ – ١٠٠، المفرد والمثنى والجمع، أدوات الاستفهام. عدد ألف ومليون ومليار وبلليون، حروف الجر وحروف العطف تصريف الفعل الماضي اللغوي، تصريف الفعل المضارع اللغوي، الجملة الاسمية والجملة الفعلية. النعت – الإضافة، اسم التفضيل، الأسماء الخمسة والأفعال الخمسة، المضارع المرفوع والمضارع المنصوب والمضارع المجزوم.	1, 2, 3, 4, 5, 7, 8, 9, 10, 11, 13, 14, 16, 17, 21, 22, 27, 28, 29, 33, 34, 35, 39,	23	57,5 %

2	3. 2. Menganalisis gagasan dari teks bahasa Arab yang berkaitan dengan tema:	12, 19, 20, 25, 31, 32, 36, 40	8	20 %
	التحيات والتعارف (التحيات، البيانات الشخصية، التعريف بالنفس وبالغير). الحياة اليومية (الأنشطة اليومية – الساعة – العبادة اليومية). الهواية (أنواع الهواية، الترويح في الإسلام). التسوق (في السوق التقليدي في السوربرمانكيت)، الصحة (المستشفى والصيدلية، عيادة المريض)، الحج والعمرة مكة والمدينة. الرياض (أنواع الرياضة، مسابقات رياضية)، الشباب (مرحلة الشباب، كن متفاضلا)، الحضارة الإسلامية (الحضارة الإسلامية قديما وحديثا علماء الإسلام)			
	Dengan memperhatikan bentuk, makna, dan fungsi dari susunan gramatikal			
	تقسيم الكلمة – الأرقام ١ – ١٠٠، أقسام الفعل – المذكور والمؤنث. عدد ألف ومليون ومليار وبلليون والحروف الجر وحروف العطف، التصريف اللغوي للفعل الماضي. النعت، الإضافة، الفعل المبني المعلوم – المبني للمجهول، الاسماء الخمسة والأفعال الخمسة			
3	3. 3. Menerapkan kaidah tentang bentuk, makna, dan fungsi dari susunan gramatikal:	6, 15, 18, 23, 24, 26, 30, 37, 38	9	22, 5 %
	الاسم النكرة والاسم المعرفة			
	Dalam menyusun teks bahasa Arab yang berkaitan dengan tema:			
	السفر والسياحة (وسائل المواصلات، المناطق السياحية)			
4	3. 4. Mengevaluasi teks bahasa Arab yang berkaitan dengan tema:	-	-	-
	المدرسة (مرافق المدرسة، الأدوات المدرسية، الأنشطة في المدرسة)، تكنولوجيا الإعلام والاتصال (الحاسوب، ولجوال، الأنلاين)، الأديان في إندونيسيا			

(الأديان والمعابد التسامح). الشباب  
 (مرحلة الشباب، كن متفاناً)، الشعر  
 العربي (الشعر، الشاعر البوصري)،  
 الدراسة في الجامعة (الدراسة في جامعات  
 إندونيسيا، الدراسة في خارج البلاد).

Dengan memperhatikan bentuk  
 makna, makna, dan fungsi dari  
 susunan gramatikal

المفرد والمثنى والجمع، تصريف الفعل  
 المضارع اللغوي الجملة الاسمية والجملة  
 الفعلية الفعل المبني للمعلم – الفعل  
 المبني للمجهول، اسم التفضيل، المضارع  
 المرفوع والمضارع المنصوب  
 والمضارع المجزوم.

Dari data di atas peneliti menemukan bahwa soal yang sesuai dengan Kompetensi Dasar (KD) mencakup 3 KD yaitu KD 3.1 dengan memperoleh 57, 5%, KD 3.2 dengan memperoleh 20%, dan KD 3.3 dengan memperoleh 22, 5%. Akan tetapi peneliti tidak menemukan dalam soal yang dibuat guru yang mencakup KD 3.4. Berdasarkan kategori persentase yang dikemukakan Arikunto, kesesuaian antar soal dengan kompetensi dasar masuk dalam kategori baik dengan persentase 75%.

### Kualitas Soal dari Aspek Validitas dan Reliabilitas

#### Validitas Soal

No	Validitas Item	Nomor Butir Soal	Jumlah
1	Valid	1, 2, 5, 6, 7, 9, 10, 12, 14, 16, 18, 20, 21, 23, 24, 25, 26, 27, 28, 30, 31, 32, 33, 34, 37, 38, 39, 40	28
2	Tidak Valid	3, 4, 8, 11, 13, 15, 17, 19, 22, 29, 35, 36	12

Dari data di atas menyatakan bahwa dari soal ujian madrasah yang dibuat oleh guru, yang termasuk soal yang valid ialah 28 soal. Sedangkan soal ujian yang dibuat guru yang termasuk soal tidak valid ialah 12 soal.

No	Kriteria	Nomor Butir Soal	Jumlah
1	Tinggi Sekali		
2	Tinggi	1, 14, 16	3
3	Sedang	5, 6, 7, 9, 12, 18, 27, 31, 32, 33, 37, 38, 39, 40	14
4	Rendah	2, 10, 13, 20, 21, 22, 23, 24, 25, 26, 28, 30, 34, 35, 36	15
5	Rendah Sekali	3, 4, 8, 11, 15, 17, 19, 29	8

Dari data di atas, yang termasuk soal valid dengan kategori validitas tinggi terdapat 3 soal, yang termasuk validitas sedang terdapat 14 soal, yang

termasuk validitas rendah terdapat 15 soal, dan yang termasuk validitas rendah sekali terdapat 8 soal.

### Reliabilitas

	Nilai Alpa Cronbach	R tabel	Keterangan
<b>Reliabilitas</b>	0,711	0,254	Item soal memiliki reliabilitas tinggi

Dari data di atas menunjukkan bahwa soal ujian madrasah yang dibuat oleh guru memiliki reliabilitas yang tinggi dengan angka  $0,711 > r \text{ tabel } 0,254$ .

### Soal dalam Persepektif *HOTS*

No	Tingkat Kognitif	Nomor Soal	Jumlah	Peresentase
1	Mengingat	1, 2, 3, 4, 5, 6, 8, 9, 10, 16, 17, 27, 29, 34, 35, 39	15	37,5 %
2	Memahami	7, 11, 13, 14, 21, 22, 28, 33	8	20 %
3	Menerapkan	15, 18, 23, 24, 26, 30, 37, 38	9	22,5 %
4	Menganalisis	12, 19, 20, 25, 31, 32, 36, 40	8	20 %
5	Mengevaluasi	-	-	-
6	Mencipta	-	-	-

Dari data di atas menyatakan bahwa butir soal ujian dalam persepektif *HOTS*. Butir tes menunjukkan tingkat domain kognitif, C 1 “mengingat” mendapat 37,5%. Butir tes menunjukkan tingkat domain kognitif C2 “memahami” 20%. Butir tes menunjukkan tingkat domain kognitif C3 “menerapkan” 22,5%. Butir tes menunjukkan tingkat domain kognitif C4 “menganalisis” 20%. Peneliti tidak menemukan butir tes yang menunjukkan tingkat domain kognitif C5 “mengevaluasi” dan C6 “mencipta”.

No	Soal	Tingkat Kognitif	Persentase
		Mencipta	
1	Hots	Mengevaluasi	
		Menganalisis	20 %
		Menerapkan	22,5 %
2	Lots	Memahami	20 %
		Mengingat	37,5 %

Dari data di atas menunjukkan bahwa soal ujian yang dibuat oleh guru yang termasuk dalam kategori soal *HOTS* ialah 20%. Sedangkan soal-soal yang termasuk dalam kategori soal *LOTS* ialah 80%.

### Kesimpulan

Hubungan antara soal ujian dengan kompetensi berbahasa, peneliti menemukan bahwa soal yang berkaitan dengan kemampuan berbahasa mencakup 2 kemampuan berbahasa yaitu: kemampuan membaca dan menulis. Sedangkan kemampuan berbahasa mendengar dan berbicara tidak tercakup dalam soal yang dibuat guru. Berdasarkan kategori persentase yang dikemukakan Arikunto, hubungan antara soal dengan *kemahiran* berbahasa ini masuk dalam kategori cukup dengan persentase 50%. Kesesuaian antara soal ujian dengan kompetensi dasar, peneliti menemukan bahwa soal yang sesuai dengan Kompetensi Dasar (KD) mencakup 3 KD yaitu KD 3.1 dengan memperoleh 57, 5%, KD 3.2 dengan memperoleh 20%, dan KD 3.3 dengan memperoleh 22, 5%. Akan tetapi peneliti tidak menemukan dalam soal yang dibuat guru yang mencakup KD 3.4. Berdasarkan kategori persentase yang dikemukakan Arikunto, kesesuaian antar soal dengan kompetensi dasar masuk dalam kategori baik dengan persentase 75%.

Kualitas soal ujian dari aspek validitas dan reliabilitas, yaitu: (a) Validitas dari pembahasan di atas menunjukkan bahwa soal ujian madrasah yang dibuat oleh guru, yang termasuk soal yang valid ialah 28 soal. Sedangkan soal ujian yang dibuat guru yang termasuk soal tidak valid ialah 12 soal, dan (b) Reliabilitas, dari pembahasan di atas menunjukkan bahwa soal yang dibuat oleh guru memiliki reliabilitas yang tinggi dengan angka  $0,711 > r_{tabel} 0,254$

Butir soal ujian madrasah dalam persepektif *HOTS*, Butir soal ujian ditinjau dari tingkat penalaran yang lebih tinggi. Butir tes menunjukkan tingkat domain kognitif, C 1 “mengingat” mendapat 37, 5%. Butir tes menunjukkan tingkat domain kognitif C2 “memahami” 20%. Butir tes menunjukkan tingkat domain kognitif C3 “menerapkan” 22, 5%. Butir tes menunjukkan tingkat domain kognitif C4 “menganalisis” 20%. Peneliti tidak menemukan butir tes yang menunjukkan tingkat domain kognitif C5 “mengevaluasi” dan C6 “mencipta”. Adapun soal-soal yang menunjukkan tingkat *HOTS* ialah 20%, sedangkan soal yang menunjukkan tingkat *LOTS* adalah 80%.

### **Daftar Pustaka**

- ‘Abās, Muhammad Khalīl waakharūna. 2014. *Madkhal ila manāhiju al-bahsi fī al-tarbiyyah wa ‘ilmu al-naḥsi*. ‘Amān: Dār al-Muyassarah lil-Nasyri wal-Tauzi’i
- ‘Abīdāt, waakharūna. 2015. *Al-bahsu al-‘ilmiy mafhūmuhu waadawātuahu waasālībuhu*. ‘Amān: Dār al-fikri
- Ainin. 2010. *Metodologi Penelitian Bahasa Arab*. Malang: Hilal Pustaka
- Al-Diyār, Abu Mas’ud. 2012. *Al-Qiyās wal-tasykhīṣu lizawiy su’ūbāt al-ta’alum*. Kuwait: Maktabah Kuwait al-watiniyyah asnā al-nasyri.



- Al-fauzān, ‘Abdu al-Rahman Ibrāhīm. 2011. *Idāāt limu’alimiy al-lugah al-‘arabiyyah ligairi al-naṭiqīna bihā*. Al-Riyād: Al-Mamlakatu al-‘arabiyyah al-su’udiyah.
- ‘Ali, Usāmah Zakī al-Sayyid. 2016. *Al-Ikhtibārāt al-lugawiyah muqārabah manhajiyah tatbiqiyah*. Al-Riyād: Dār wujūh lil-nasyri wal-tawzi’i
- Al-Jāgub, Muhammad. 2002. *Al-manhaj al-taqwim fi mihnah al-ta’lim*. ‘Aman: Dar wail lil-nasyri wal-tawzi’i.
- Al-Khuli, Muhammad ‘Ali. 2000. *Al-ikhtibarat al-lugawiyah*. Ardan: Dar al-falah lil-nasyri wal-tawzi’
- Al-Syarbīnī, Zakariyā waakharūna. 2013. *Manāhiju al-bahsi fi al-‘ulūmi al-tarbawiyah wal-nafsiyyah wal-ijtimā’iyyah*. Al-Riyād: Maktabah al-Syarqawi.
- Anetha. L. F, Tilaar dan H. 2019. *Analisis Butir Soal Semester Ganjil Mata Pelajaran Matematika*. JP3I: Jurnal Pengukuran Psikologi dan Pendidikan Indonesia. 8 (1). 57-68.
- Anis, Ibrāhīm waākharūna. 1979. *Al-Mu’jam al-wasit*. Turki: Dar Da‘wah
- Dawidri, Raja Wahīd. *Al-bahsu al-‘ilmi asasiyatuhu al-naẓariyatuhu wa mumarisatuhu al-‘ilmiyyah*. Damaskus: Dar al-fikri
- Depdiknas. 2008. *Panduan Analisis Butir Soal*. Jakarta: Dirjen Menejamen Pendidikan Dasar dan Menengah..
- Faridah, Ida. 2019. *EVALUASI Pembelajaran Berdasarkan Kurikulum Nasional*. Bandung: PT Remaja Rosda Karya.
- Hermawan, Acep. 2014. *Metodologi Pembelajaran Bahasa Arab*. Bandung: PT Remaja Rosda Karya.
- . 2021. *Penilaian Pembelajaran Bahasa Arab Prinsip dan Operasional*. Bandu: PT Remaja Rosda Karya.
- Ibrāhīm, Majdī ‘Aziz. 2009. *Mu’jam Muṣṭalahāt wamafāhīmi al-ta’lim wa-ta’alum*. Qahirah: ‘Alimul al-kutub.
- Ibrahim, Nini dan Prima Gusti Yanti. 2017 . *Bahan Ajar Evaluasi Pembelajaran Bahasa dan Sastra Indonesia untuk Perguruan Tinggi dan Umum*. Jakarta: Universitas Muhammadiyah Prof. Dr. Hamka.
- Jabir, Walid Ahmad. 2002. *Tadrisu al-lugah al-‘arabiyyah mafāhīmu naẓariyyatuhu wataṭbiqāti ‘amaliyyah*. ‘Aman: Dar al-fikri.
- Juhanda, Aa. 2016. *Analisis Soal Jenjang Kognitif Taksonomi Bloom Revisi pada Buku Sekolah Elektronik (BSE) Biologi SMA*. Vol. 21. No. 1. 61-66.
- Jumiati. 2016. *IPA dan Pembelajaran Beripikir Tingkat Tinggi (telaah buku siswa MI/SD kelas VI Tema. Karya Afriki, dkk)*. Vol. 2. No. 1. 1-10.
- Keputusan Direktur Jendral Pendidikan Islam. 2021. *Tentang Petunjuk Teknis Penyusunan Soal Higher Order Thinking Skills (HOTS) Di Madrasah*. Jakarta: Direktur Jendral Pendidikan Islam.
- Labīd, Abī. 2009. *Ṭuruqu al-tadrīs waasālību al-imtihān*. Pakistan: Al-jāmi’ah al-fārūqiyah
- Madkūr, ‘Ali waakhrarūna. 2010. *Al-marji’ fi manāhiji al-lugah al-‘arabiyyah lil-nāṭiqīna bilugātin ukhra*. Al-qāhirah: Dār al-fikri al-‘araby
- Modul Penyusunan Soal. 2019. *Modul Penyusunan Soal Keterampilan Berpikir Tingkat Tinggi (Higher Order Thinking Skills) Bahasa dan Sastra Arab*. Jakarta: Direktorat Pembinaan Sekolah Menengah Atas Direktorat Jenderal Pendidikan Dasar dan Menengah Kementrian Pendidikan dan Kebudayaan.
- Muhammad, ‘Ābdu al-khāliq. 1089. *Ikhtibār al-lugah*. Al-riyād: Jāmi’atu al-muluk sa’ūd.

- Nurgiyantoro, Burhan. 2011. *Penilaian Pembelajaran Bahasa*. Yogyakarta: BPFE-Yogyakarta.
- Purnomo, Pajar. 2019. *Penilaian Pembelajaran HOTS (Higher Order Thinking Skill)*. Cilacap: Candradimuka Pers.
- Purwanto. 2009. *Evaluasi Hasil Belajar*. Yogyakarta: Pustaka Pelajar.
- Qāsim, Ansā Muhammad Ahmad. 2003. *Al-furūqu al-fardiyyah wal-taqwīm*. ‘Amān: Dār al-fikri
- Qazāmil, Sūniyā Hānim ‘Ali. 2013. *Al-mu’jam al-‘asriy fī al-tarbiyyah*. Al-qāhirah: ‘Ālimul al-kutub
- Rahayu, Sri. dkk. 2020. *Pengembangan Soal High Order Thinking Skill untuk Meningkatkan Kemampuan Berpikir Tingkat Tinggi Matematika Siswa Sekolah Dasar*. Pedadidaktika: Jurnal Ilmiah Pendidikan Guru Sekolah Dasar. Vol. 7. No. 2. Hal. 127-137.
- Ruwaida, Hikmatu. “Proses Kognitif dalam Taksonomi Bloom Revisi: Analisis Kemampuan Mencipta (C6) pada Pembelajaran Fikih di MI Miftahul Anwar Desa Banua Lawas.” *AL-Madrasah: Jurnal Ilmiah Pendidikan Madrasah Ibtidaiyah* 4, no. 1 (2019): 51–76.
- Sudjana, Nana. 2019. *Penilaian Hasil Proses Belajar Mengajar*. Bandung: PT Remaja Rosda Karya.
- Sugiyono. 2016. *Metodologi Penelitian Pendiidkan Kuantitatif, Kualitatif, dan R&D*. Yogyakarta: Alfabeta.
- Syamsudin dan Vismaia S. Damaianti. 2015. *Metode Penelitian Pendidikan Bahasa*. Bandung: PT Remaja Rosda Karya.
- Thoha, M. Chabib . 1991. *Teknik Evaluasi Pendidikan*. Jakarta: Rajawali Pers.
- Tim Pusat Penilaian Pendidikan,. 2019. *Panduan Penilaian Tes Tertulis*. Jakarta: Pusat Penilaian Pendidikan.
- Tu’aimah, Rusydī Ahmad. 2000. *Namāzujū al-ikhtibarātu al-mawduiyyah fī al-lughah lil-marhalati al-sānawiyyah*. Al-Qāhirah: Dār Al-fikri al-‘Arabiy
- , 2004. *Tahlīl al-muhtawā fī al-‘ulūmu al-insāniyyah mafhūmuhu ususuhu asikhdāmuhu*. Al-Qāhirah: Dār Al-fikri al-‘Arabiy.
- Umar, Jahja. 1996. *Bahasa Penataran Pengujian Pendidikan*. Jakarta: Balitbang Dikbud.
- Umar, Mahmūd Ahmad waakharūna. 2010. *Al-Qiyās al-nafsi wal-tarbawiyi*. Aman: Dār al-muyassarah lil-nasyri wal-tawzi’i
- Umar, Saifu al-islām Sa’dun. 2009. *Al-muwajiz fī manhaj al-bahsi al-ilmīy fī al-tarbiyyah wal-‘ulum al-insāniyyah*. Dimasyqa: Dār al-fikri