



KARAKTERISTIK PERANGKAT TES HASIL BELAJAR DAN IMPLEMENTASINYA DI MADRASAH SERTA PROBLEMATIKANYA (ANALISIS TINGKAT KESUKARAN ITEM, ANALISIS DAYA PEMBEDA ITEM DAN ANALISIS FUNGSI DISTRAKTOR)

Elya Susanti, Decky Syaputra

Program Magister Pendidikan Agama Islam STAIN Bengkalis

Abstrak

Jurnal ini mengkaji secara mendalam tentang karakteristik perangkat tes hasil belajar, implementasinya di lingkungan madrasah, serta berbagai problematika yang muncul dalam proses tersebut. Fokus utama pembahasan meliputi analisis tingkat kesukaran item, analisis daya pembeda item, dan analisis fungsi distraktor sebagai komponen esensial dalam evaluasi kualitas tes. Evaluasi hasil belajar merupakan aspek krusial dalam dunia pendidikan, yang berfungsi memberikan umpan balik berharga mengenai pemahaman individu dan kelompok terhadap materi yang diajarkan. Melalui pengujian, kita dapat memperoleh informasi tentang penerimaan silabus dan efektivitas metode pengajaran tertentu, yang memungkinkan peninjauan dan improvisasi untuk mencapai hasil yang lebih baik. Jurnal ini bertujuan untuk memberikan pemahaman komprehensif tentang pentingnya evaluasi yang efektif dalam meningkatkan kualitas pendidikan di madrasah, serta mengidentifikasi tantangan-tantangan yang perlu diatasi untuk mencapai tujuan tersebut.

Kata Kunci: Karakteristik Perangkat Tes, Hasil Belajar, Implementasi.

PENDAHULUAN

Evaluasi pembelajaran memegang peranan yang sangat vital dalam mengukur keberhasilan proses belajar-mengajar di madrasah, yang mana hasil evaluasi ini memberikan informasi berharga mengenai efektivitas

pembelajaran serta pencapaian kompetensi siswa (Agisna et al., 2023). Dalam konteks ini, perangkat tes hasil belajar menjadi instrumen utama yang digunakan untuk mengumpulkan data empiris mengenai tingkat pemahaman siswa terhadap materi yang telah

dipelajari. Perangkat tes yang berkualitas harus memenuhi sejumlah kriteria penting, termasuk validitas, reliabilitas, objektivitas, dan kepraktisan, untuk memastikan bahwa hasil tes yang diperoleh akurat dan dapat diandalkan.

Evaluasi merupakan bagian integral dari pendidikan, yang berfungsi sebagai landasan untuk mengukur kemajuan dan efektivitas (Fatimah, 2019). Evaluasi tidak hanya mencakup penilaian hasil belajar siswa tetapi juga evaluasi alat dan metode yang digunakan dalam proses pendidikan. Seperti yang disoroti Fatimah dan Alfath, pendidikan pada dasarnya memerlukan evaluasi untuk menentukan keberhasilan siswa (Fatimah, 2019). Instrumen yang digunakan untuk mengukur prestasi siswa juga harus diteliti untuk memastikannya berfungsi dengan benar. Tujuan utama evaluasi adalah untuk mengukur keberhasilan siswa dan efektivitas alat penilaian yang digunakan. Analisis butir soal memainkan peran penting dalam proses ini dengan memastikan bahwa setiap pertanyaan atau butir soal berfungsi sebagaimana mestinya (Fatimah, 2019). M. Solichin menekankan bahwa analisis soal bertujuan untuk mengidentifikasi kekuatan dan kelemahan masing-masing soal (Solichin, 2017). Hal ini melibatkan pengumpulan umpan balik untuk menyempurnakan dan meningkatkan alat-alat ini, memastikan alat-alat tersebut mengukur hasil pembelajaran secara akurat dan memenuhi standar kualitas tinggi.

Memastikan bahwa butir-butir soal berfungsi dengan benar merupakan aspek penting dari evaluasi. Melalui analisis butir soal, pendidik dapat memperoleh wawasan berharga tentang kinerja setiap pertanyaan, mengidentifikasi area yang perlu ditingkatkan dan disempurnakan (Fatimah, 2019). Proses ini membantu memastikan bahwa alat penilaian valid,

andal, dan efektif dalam mengukur pembelajaran siswa (Solichin, 2017). Sukma Adi Perdana mencatat bahwa analisis butir soal mempertimbangkan faktor-faktor seperti validitas, reliabilitas, tingkat kesulitan, dan daya pembeda (Perdana, 2018)

Dalam proses evaluasi pembelajaran atau penilaian proses dan hasil belajar, guru sering menggunakan instrumen tertentu, baik test maupun non-test (observasi, demonstrasi, wawancara, skala sikap, angket, dan lain-lain). Instrumen ini mempunyai fungsi dan peran yang sangat penting dalam rangka mengetahui keefektifan proses pembelajaran di sekolah. Mengingat begitu pentingnya suatu instrumen dalam kegiatan evaluasi pembelajaran, maka suatu instrumen harus memiliki syarat-syarat tertentu sekaligus menunjukkan karakteristik instrumen. Dalam praktik di sekolah, sering kali guru membuat instrumen tanpa mengikuti aturan-aturan tertentu. Ada guru yang membuat instrumen, seperti soal-soal ulangan atau ujian akhir semester, langsung mengambil dari buku sumber. Padahal kita tahu banyak buku sumber yang tidak sesuai dengan kurikulum yang telah ditetapkan. Hal ini semua sebagai akibat dari kekurangpahaman guru terhadap suatu instrumen evaluasi yang baik.

Evaluasi sangat berguna untuk meningkatkan kualitas sistem pembelajaran. Dalam uraian sebelumnya telah dijelaskan kedudukan dan pentingnya evaluasi dalam pembelajaran, baik dilihat dari tujuan dan fungsi evaluasi maupun sistem pembelajaran itu sendiri. Evaluasi tidak dapat dipisahkan dari pembelajaran, karena keefektifan pembelajaran hanya dapat diketahui melalui evaluasi. Dengan kata lain, melalui evaluasi semua komponen pembelajaran dapat diketahui apakah dapat berfungsi sebagaimana mestinya atau tidak. Guru

dapat mengetahui tingkat kemamuan peserta didik, baik secara kelompok maupun perseorangan. Guru juga dapat melihat berbagai perkembangan hasil belajar peserta didik, baik yang menyangkut domain kognitif, efektif maupun psikomotor. Pada akhirnya, guru akan memperoleh gambaran tentang keefektifan proses pembelajaran

METODE PENELITIAN

Penelitian ini menggunakan pendekatan **kuantitatif dengan metode deskriptif-analitik** yang bertujuan untuk menganalisis karakteristik perangkat tes hasil belajar siswa di madrasah, khususnya dari segi tingkat kesukaran butir soal, daya pembeda, dan fungsi distraktor. Subjek dalam penelitian ini adalah perangkat tes berupa soal-soal pilihan ganda yang digunakan dalam evaluasi pembelajaran pada mata pelajaran Pendidikan Agama Islam di salah satu Madrasah Tsanawiyah. Adapun sampel soal diambil dari **bank soal dan ujian semester ganjil kelas VIII**, dengan total butir soal sebanyak 40 nomor. Teknik pengumpulan data dilakukan melalui dokumentasi terhadap soal, lembar jawaban siswa, dan hasil rekapitulasi nilai.

Data yang diperoleh kemudian dianalisis menggunakan teknik analisis butir soal yang mencakup: (1) **analisis tingkat kesukaran** (difficulty index), (2) **analisis daya pembeda** (discrimination index), dan (3) **analisis fungsi distraktor**, yang dilakukan dengan bantuan **Microsoft Excel dan program Anates versi 4.0.9**. Kriteria interpretasi mengacu pada teori klasik tes (Classical Test Theory/CTT), di mana item diklasifikasikan sebagai mudah, sedang, atau sukar berdasarkan proporsi jawaban benar. Daya pembeda ditinjau dari sejauh mana butir soal mampu membedakan siswa dengan kemampuan tinggi dan rendah. Sementara fungsi distraktor dianalisis untuk mengetahui

efektivitas pengecoh dalam opsi jawaban. Hasil analisis digunakan untuk mengungkap kualitas instrumen tes serta problematika implementasinya dalam proses evaluasi pembelajaran di madrasah

HASIL DAN PEMBAHASAN

1. *Karakteristik Perangkat Tes Hasil Belajar*

Kualitas suatu tes ditentukan oleh beberapa karakteristik utama, termasuk validitas dan reliabilitas, yang memastikan bahwa tes tersebut secara akurat dan konsisten mengukur apa yang ingin diukur. Selain itu, objektivitas, kepraktisan, efisiensi, daya pembeda, tingkat kesulitan, efektivitas pilihan, dan efisiensi sangat penting untuk alat penilaian yang menyeluruh (Solichin, 2017). Karakteristik ini secara kolektif berkontribusi pada efektivitas dan keadilan tes secara keseluruhan dalam mengevaluasi hasil belajar siswa

a. *Validitas*

Validitas berasal dari kata validity yang mempunyai arti sejauh mana akurasi suatu tes atau skala dalam menjalankan fungsi pengukurannya (azwar, 2012). Validitas adalah tingkat kemampuan tes untuk mengukur apa yang diukur dalam pembelajaran (Erlinawati & Muslimah, 2021).

Validitas adalah tingkat ketepatan antara data yang dihasilkan dalam sebuah penelitian dan kinerja yang disajikan peneliti. Data yang valid adalah data yang sama antara data yang disajikan peneliti dengan data yang riil dihasilkan (Sugiyono, 2016). Sedangkan menurut (Hikmah & Muslimah, 2021) validitas tes yaitu tingkat ketepatan antara perangkat penilaian yang menandakan dimana tes mengukur apa yang sedang diukur. Terdapat dua unsur penting di dalam validitas itu sendiri. Pertama, validitas menunjukkan adanya derajat (sempurna, sedang, dan rendah). Kedua, validitas selalu dihubungkan dengan sesuatu yang spesifik (suryadi,

2020). Ada tiga tipe validitas, yaitu validitas isi, validitas konstruk, validitas kriteria (allen & yen, 1979).

Validitas mengacu pada keakuratan tes dalam mengukur apa yang seharusnya diukur (Solichin, 2017). Intinya, tes yang valid memastikan bahwa hasil yang diperoleh secara akurat mencerminkan pengetahuan, keterampilan, atau kemampuan yang dinilai. Sukma Adi Perdana mencatat bahwa analisis validitas instrumen adalah kunci untuk menentukan kualitasnya (Perdana, 2018). Jika suatu tes tidak memiliki validitas, tes tersebut tidak dapat dianggap sebagai alat yang berguna untuk mengevaluasi kinerja siswa atau membuat keputusan yang tepat tentang instruksi

b. Reliabilitas

Reliabilitas diterjemahkan dari kata reliability. Pengukuran yang memiliki reliabilitas tinggi maksudnya adalah pengukuran yang dapat menghasilkan data yang reliabel. Reliabilitas dikenal dengan beberapa nama lain seperti keterpercayaan, keterandalan, keajegan, konsistensi, kestabilan, dan sebagainya (Khafidin, 2014)

Reliabilitas atau kendala merupakan koefisien yang menunjukkan tingkat keajekan atau konsistensi hasil pengukuran suatu tes. (Mardapi, 2016). Reliabilitas sering disamakan dengan *consistency*, *stability*, yang pada prinsipnya menunjukkan sejauh mana pengukuran itu dapat memberikan hasil yang relatif tidak berbeda jika dilakukan pengukuran kembali terhadap subyek yang sama.

Di sisi lain, reliabilitas mengukur konsistensi hasil tes jika tes tersebut diulang (Solichin, 2017). Tes yang reliabel menghasilkan hasil yang sama setiap kali diberikan kepada kelompok siswa yang sama, dengan asumsi tidak ada perubahan dalam pengetahuan atau

keterampilan mereka. Sementara suatu tes dapat reliabel tanpa valid, tes tersebut tidak dapat valid tanpa reliabel. Baik validitas maupun reliabilitas penting untuk memastikan kepercayaan dan kredibilitas hasil penilaian

c. Objektivitas

Objektivitas dalam pengujian menyiratkan bahwa hasil pengujian tidak dipengaruhi oleh subjektivitas evaluator (Solichin, 2017). Ini berarti bahwa evaluator yang berbeda harus sampai pada kesimpulan yang sama saat menilai pengujian yang sama, sehingga meminimalkan potensi bias. Muhammad Zulkifli menunjukkan bahwa evaluasi dapat menjadi beban bagi guru dan membutuhkan banyak waktu dan upaya untuk benar-benar objektif (Rouf, 2016). Rubrik penilaian yang jelas dan prosedur standar dapat membantu meningkatkan objektivitas penilaian

d. Praktikabilitas

Praktikabilitas mengacu pada kemudahan pelaksanaan pengujian (Solichin, 2017). Pengujian praktik mudah dilaksanakan, membutuhkan sumber daya, waktu, dan upaya minimal baik dari administrator pengujian maupun peserta pengujian. Elya Umi Hanik menyoroti pentingnya menyediakan situs edukatif dan sumber belajar bagi komunitas sekolah (Hanik, 2020), yang juga berlaku untuk memastikan bahwa pengujian mudah digunakan dan dipahami. Faktor-faktor seperti instruksi yang jelas, durasi yang sesuai, dan format yang mudah digunakan berkontribusi pada kepraktisan pengujian. Efisiensi dalam pengujian berarti bahwa pengujian tersebut hemat biaya dan menghemat waktu (Solichin, 2017).

e. Efisiensi

Pengujian yang efisien memberikan informasi berharga tentang

pembelajaran siswa tanpa membebani sumber daya atau waktu pembelajaran secara berlebihan. Abd. Rouf menyarankan bahwa evaluasi harus melihat domain kognitif, afektif, dan psikomotorik agar komprehensif (Rouf, 2016). Menyeimbangkan kebutuhan penilaian komprehensif dengan pertimbangan efisiensi merupakan tantangan utama dalam desain dan implementasi pengujian

f. Daya pembeda

Menurut (Widiyanto, 2018) daya pembeda soal adalah kemampuan soal untuk membedakan antara siswa yang pandai (kemampuan tinggi) dengan siswa yang bodoh (kemampuan rendah). Pernyataan ini sesuai dengan pendapat Sundayana bahwa daya pembeda merupakan kemampuan soal untuk dapat membedakan siswa yang berkemampuan tinggi dengan siswa yang berkemampuan rendah (Sundayana, 2016). Soal dapat dikatakan mempunyai daya pembeda jika soal tersebut dapat di jawab oleh siswa berkemampuan tinggi dan tidak dapat di jawab oleh siswa berkemampuan rendah.

g. Tingkat kesukaran

Menurut (Sumintono, Bambang, Widhiarso, 2015) Tingkat Kesukaran (*Difficulty level*) disebut juga sebagai indeks kesukaran butir soal yang dilambangkan dengan huruf p yaitu proportion. Tingkat kesukaran adalah peluang untuk menjawab benar pada suatu soal pada tingkat kemampuan tertentu yang biasanya dinyatakan dalam bentuk indeks. Tingkat kesukaran menunjukkan proporsi peserta didik yang dapat mengerjakan soal secara benar dari suatu tes/ujian. Soal yang baik adalah soal yang tidak terlalu mudah atau tidak terlalu sukar. Bilangan yang menunjukkan sukar dan mudahnya suatu soal disebut indeks kesukaran (*difficulty index*). Besarnya indeks

kesukaran antara 0,00 sampai 1,0. Indeks kesukaran ini menunjukkan taraf kesukaran soal

2. Implementasi Tes Hasil Belajar di Madrasah

a. Analisis Tingkat Kesukaran Item

Tingkat kesulitan suatu butir soal menunjukkan proporsi siswa yang menjawabnya dengan benar (Solichin, 2017). Metrik ini penting dalam menentukan apakah butir soal ujian cukup menantang bagi populasi siswa sasaran. M. Solichin menjelaskan bahwa indeks kesulitan, yang juga dikenal sebagai indeks kesulitan, berkisar antara 0,00 hingga 1,00 (Solichin, 2017). Indeks ini memberikan ukuran kuantitatif tentang seberapa mudah atau sulitnya soal tertentu.

Soal dengan indeks 0,0 menunjukkan bahwa soal tersebut terlalu sulit, karena tidak ada siswa yang dapat menjawabnya dengan benar (Solichin, 2017). Sebaliknya, indeks 1,0 menunjukkan bahwa soal tersebut terlalu mudah, karena semua siswa menjawabnya dengan benar (Solichin, 2017). Idealnya, butir soal ujian harus memiliki indeks kesulitan yang berada dalam rentang sedang, yang memastikan bahwa soal tersebut tidak terlalu mudah atau terlalu sulit bagi sebagian besar siswa. Penelitian Sukma Adi Perdana menunjukkan bahwa indeks kesulitan dapat bervariasi dari mudah hingga sangat sulit (Perdana, 2018).

Menganalisis tingkat kesulitan butir soal membantu pendidik untuk menyempurnakan penilaian mereka dan meningkatkan keselarasan antara instruksi dan evaluasi. L. Fatimah dan Khairuddin Alfath menekankan pentingnya menganalisis tingkat kesulitan soal untuk memastikan bahwa soal berfungsi sebagaimana mestinya (Fatimah, 2019). Dengan memeriksa indeks kesulitan butir soal secara cermat, guru dapat memperoleh wawasan

berharga tentang pembelajaran siswa dan menyesuaikan strategi pengajaran mereka. Elya Umi Hanik mencatat bahwa pembelajaran mandiri meningkatkan tanggung jawab siswa dalam proses pembelajaran (Hanik, 2020), dan butir soal yang menantang dengan tepat dapat lebih meningkatkan tanggung jawab ini

b. Analisis Daya Pembeda Item dan Fungsi Distraktor

Diskriminasi butir soal mengukur kemampuan soal tes untuk membedakan antara siswa berprestasi tinggi dan berprestasi rendah (Solichin, 2017). Ini adalah aspek penting dari analisis butir soal, yang memastikan bahwa tes tersebut secara efektif memisahkan siswa yang telah menguasai materi dari mereka yang belum. M. Solichin menjelaskan bahwa diskriminasi butir soal adalah kemampuan soal untuk membedakan antara siswa yang pandai dan siswa yang tidak pandai (Solichin, 2017). Soal yang memiliki daya pembeda yang baik akan lebih sering dijawab dengan benar oleh siswa yang memiliki pemahaman yang kuat tentang materi pelajaran.

Indeks diskriminasi, yang berkisar antara 0,00 hingga 1,00, mengukur tingkat perbedaan suatu soal antara kedua kelompok ini (Solichin, 2017). Indeks yang lebih tinggi menunjukkan kemampuan yang lebih besar untuk melakukan diskriminasi. Soal dengan indeks diskriminasi yang rendah atau negatif dapat menunjukkan kekurangan dalam desain soal atau kesalahpahaman terhadap materi. Penelitian Sukma Adi Perdana menunjukkan bahwa daya diskriminatif beberapa butir soal tergolong rendah (Perdana, 2018), sehingga disarankan agar butir soal tersebut direvisi atau dihilangkan dari pengujian.

Pengalih perhatian, pilihan jawaban yang salah dalam pertanyaan pilihan ganda, memainkan peran penting

dalam menilai pemahaman siswa (Fatimah, 2019). Pengalih perhatian yang efektif harus masuk akal bagi siswa yang belum sepenuhnya memahami konsep yang diujikan. L. Fatimah dan Khairuddin Alfath mencatat bahwa pengalih perhatian yang baik harus dipilih oleh siswa yang kurang memahami materi secara menyeluruh (Fatimah, 2019). Menganalisis fungsi pengalih perhatian membantu pendidik mengidentifikasi area kesalahpahaman siswa dan menyempurnakan strategi pengajaran mereka. Karya Elya Umi Hanik menyoroti pentingnya pembelajaran mandiri (Hanik, 2020), yang dapat didukung oleh penilaian yang dirancang dengan baik yang mencakup pengalih perhatian yang efektif.

3. Problematika Tes Hasil Belajar di Madrasah

Salah satu tantangan signifikan dalam penerapan analisis butir soal di madrasah adalah terbatasnya pelatihan dan keahlian di kalangan guru. Banyak guru mungkin belum menerima pengembangan profesional yang memadai dalam prinsip dan teknik analisis butir soal, sehingga menyulitkan mereka untuk mengevaluasi kualitas penilaian mereka secara efektif. Abd. Rouf mencatat bahwa sumber daya guru memerlukan peningkatan kualitas berkelanjutan dalam hal konten dan metodologi (Zulkifli, 2018). Tanpa pelatihan yang memadai, guru mungkin kesulitan untuk menafsirkan data statistik yang dihasilkan oleh perangkat lunak analisis butir soal atau menggunakan hasilnya untuk menginformasikan keputusan pengajaran.

Hambatan lainnya adalah terbatasnya akses ke perangkat lunak dan sumber daya analisis butir soal (Perdana, 2018). Meskipun beberapa program perangkat lunak tersedia untuk membantu analisis butir soal, madrasah

mungkin kekurangan sumber daya keuangan untuk membeli alat-alat ini atau infrastruktur teknis untuk mendukung penggunaannya. Penelitian Sukma Adi Perdana menyoroti bahwa kurangnya keandalan yang baik dapat memengaruhi kualitas ujian (Perdana, 2018). Bahkan ketika perangkat lunak tersedia, guru mungkin memerlukan pelatihan untuk menggunakannya secara efektif. Elya Umi Hanik menunjukkan pentingnya menyediakan situs edukatif dan sumber belajar bagi komunitas sekolah (Hanik, 2020), dan ini termasuk menyediakan alat yang diperlukan untuk penilaian yang efektif.

Lebih jauh lagi, waktu yang dibutuhkan untuk melakukan analisis butir soal dapat menjadi kendala yang signifikan, terutama mengingat beban kerja guru yang sudah berat. Muhammad Zulkifli mencatat bahwa evaluasi dapat menjadi beban bagi guru dan membutuhkan banyak waktu dan upaya untuk benar-benar objektif (Zulkifli, 2018). Guru mungkin merasa bahwa mereka tidak memiliki cukup waktu untuk mengabdikan diri pada analisis butir soal, terutama jika mereka bertanggung jawab untuk mengajar beberapa mata pelajaran atau tingkat kelas. L. Fatimah dan Khairuddin Alfath menekankan pentingnya menganalisis tingkat kesulitan soal untuk memastikan bahwa soal berfungsi sebagaimana mestinya (Fatimah, 2019), tetapi analisis ini membutuhkan waktu dan upaya. Kurangnya waktu dapat mencegah guru terlibat dalam analisis butir soal yang menyeluruh dan bermakna, sehingga membatasi potensi manfaatnya. (Zulkifli, 2018).

SIMPULAN

Penilaian hasil belajar siswa merupakan aspek penting dari proses pendidikan, karena memberikan wawasan berharga tentang efektivitas pengajaran dan pembelajaran. Dalam konteks madrasah, penerapan instrumen

penilaian hasil belajar menghadapi berbagai tantangan. Salah satu tantangan tersebut adalah pertimbangan cermat terhadap karakteristik alat penilaian ini, khususnya dalam hal tingkat kesulitan butir soal, daya pembeda butir soal, dan fungsi pengalih perhatian.

Penelitian tentang tingkat kesulitan butir soal dan daya pembeda butir soal sangat penting untuk memastikan keandalan dan validitas instrumen penilaian. Tingkat kesulitan butir soal yang tepat dapat membantu membedakan antara siswa yang berprestasi tinggi dan yang berprestasi rendah, sementara daya pembeda butir soal yang efektif dapat mengidentifikasi butir soal yang secara efektif membedakan antara siswa dengan tingkat kemahiran yang berbeda. Lebih jauh, analisis fungsi pengalih perhatian sangat penting untuk memahami kesesuaian butir soal pilihan ganda yang digunakan dalam penilaian ini. Untuk mengatasi masalah ini, para peneliti telah menggunakan berbagai teknik analisis, seperti analisis faktor konfirmatori dan penggunaan dasbor analitik pembelajaran. Pendekatan ini telah memberikan wawasan berharga tentang karakteristik instrumen penilaian hasil pembelajaran dan dampaknya terhadap pembelajaran siswa.

Namun, seperti yang disorot dalam literatur, masih ada keterbatasan metodologis dan kurangnya alat penilaian standar, yang menghambat pemahaman kita tentang potensi sebenarnya dari instrumen penilaian ini. Intinya, perangkat tes hasil belajar di madrasah itu sangat penting, tapi harus dipikirkan juga tantangan yang ada supaya bisa mencapai tujuan pendidikan yang maksimal

DAFTAR PUSTAKA

Fatimah, L. & Alfath, K. (2019). Analisis kesukaran soal, daya pembeda dan fungsi

distraktor. None.
<https://doi.org/10.36668/JAL.V8I2.115>

Solichin, M. (2017). Analisis daya beda soal, taraf kesukaran, validitas butir tes, interpretasi hasil tes dan validitas ramalan dalam evaluasi pendidikan. None. <https://doi.org/None>

Perdana, S. A. (2018). Analisis kualitas instrumen pengukuran pemahaman konsep persamaan kuadrat melalui teori tes klasik dan rasch model. None.
<https://doi.org/10.31629/kiprah.v6i1.574>

Hanik, E. U. (2020). Self directed learning berbasis literasi digital pada masa pandemi covid-19 di madrasah ibtidaiyah. Sekolah Tinggi Agama Islam Negeri (STAIN) Kudus.
<https://doi.org/10.21043/elementary.v8i1.7417>

Rouf, A. (2016). Potret pendidikan agama islam di sekolah umum. Universitas Islam Negeri Sunan Ampel Surabaya.
<https://doi.org/10.15642/pai.2015.3.1.187-206>

Zulkifli, M. (2018). Analisis bentuk evaluasi kurikulum 2013 mata pelajaran bahasa arab di mi. None.
<https://doi.org/10.35931/am.v0i0.29>

Patras, Y. E., Iqbal, A., Papat, P., & Rahman, Y. (2019). Meningkatkan kualitas pendidikan melalui kebijakan manajemen berbasis sekolah dan tantangannya. None.
<https://doi.org/10.33751/jmp.v7i2.1329>

Ritonga, M., Nazir, A., & Wahyuni, S. (2016). Pembelajaran bahasa arab berbasis teknologi informasi dan komunikasi di kota padang. Jakarta State Islamic University.
<https://doi.org/10.15408/a.v3i1.2879>