

SKRIPSI

**ANALISIS BUTIR SOAL UJIAN AKHIR SEMESTER GASAL
PENDIDIKAN AGAMA ISLAM DI SMA NEGERI 1 KOTA
MUNGKID TAHUN AJARAN 2022/2023**

Diajukan kepada Fakultas Agama Islam Universitas Muhammadiyah Magelang
untuk Memenuhi Salah Satu Syarat guna Memperoleh Gelar Sarjana Pendidikan



Oleh:

Abdul Muiz

NIM: 19.0401.0073

**PROGRAM STUDI PENDIDIKAN AGAMA ISLAM
FAKULTAS AGAMA ISLAM
UNIVERSITAS MUHAMMADIYAH MAGELANG
2023**

SKRIPSI

**ANALISIS BUTIR SOAL UJIAN AKHIR SEMESTER GASAL
PENDIDIKAN AGAMA ISLAM DI SMA NEGERI 1 KOTA
MUNGKID TAHUN AJARAN 2022/2023**

Diajukan kepada Fakultas Agama Islam Universitas Muhammadiyah Magelang
untuk Memenuhi Salah Satu Syarat guna Memperoleh Gelar Sarjana Pendidikan



Oleh:

Abdul Muiz

NIM: 19.0401.0073

**PROGRAM STUDI PENDIDIKAN AGAMA ISLAM
FAKULTAS AGAMA ISLAM
UNIVERSITAS MUHAMMADIYAH MAGELANG
2023**

BAB I

PENDAHULUAN

A. Latar Belakang Masalah

Pendidikan nasional berfungsi mengembangkan kemampuan watak dan peradaban bangsa yang bermartabat dalam rangka mencerdaskan kehidupan bangsa dan mengembangkan potensi peserta didik agar menjadi manusia yang beriman dan bertakwa kepada Tuhan Yang Maha Esa, sehat, berilmu, cakap, kreatif, dan bertakwa. menjadi warga negara yang demokratis dan bertanggung jawab, sesuai dengan fungsi dan tujuan Sisdiknas no 20 Tahun 2003, pasal 3.¹

Perkembangan pendidikan di Indonesia tidak terlepas dari evaluasi pendidikan yang dilakukan oleh sekolah formal di seluruh Indonesia untuk mengetahui ketercapaian murid dalam memahami pembelajaran sesuai dengan tujuan pembelajaran, selain itu juga evaluasi pendidikan yang dilakukan untuk mengetahui ketercapaian kompetensi yang telah dirancang sedemikian rupa agar siswa dapat memahami materi dan mencapai tujuan pembelajaran tersebut.

Oleh karena itu guru bertugas melaksanakan suatu kegiatan yaitu penilaian atau evaluasi prestasi belajar siswa dan siswa yang terlibat aktif dalam memperoleh perubahan perlu adanya evaluasi pada siswa setelah mengikuti proses belajar mengajar. Guru dituntut untuk dapat mengevaluasi

¹ Undang-undang R.I. Tentang Sistem Pendidikan Nasional No. 20 Tahun 2003, (Cet. V, Jakarta, Sinar Grafika, 2013), h. 3

prestasi belajar siswa karena evaluasi merupakan salah satu komponen penting dari kegiatan belajar mengajar. Selain itu, guru harus mampu menyusun bahan pelajaran dan menyajikan materi untuk mengkondisikan kegiatan belajar siswa. evaluasi pembelajaran bertujuan untuk membuat proses belajar mengajar menjadi lebih baik. Pelaksanaan evaluasi pembelajaran menghasilkan informasi yang digunakan untuk meningkatkan proses belajar mengajar.²

Begitu pula dalam implementasi pembelajaran perlu adanya evaluasi di setiap tahapan agar tercapai tujuan pembelajaran sesuai dengan perencanaan, pelaksanaan, tujuan, dan muaranya pada evaluasi yang menjadi tolak ukur sebuah keberhasilan dalam proses belajar mengajar, namun dewasa ini timbul berbagai problematika berkaitan dgn proses belajar mengajar yg membuat tergesernya tatanan proses belajar mengajar yg mana salah satunya wabah *covid-19* yg berlangsung kurang lebih 2 tahun yang membuat institusi sekolah berfikir keras untuk meningkatkan proses belajar mengajarnya, dan setelah kenormalan baru ini sekolah perlu mengembalikan dan meningkatkan proses belajar mengajar di sekolah seperti biasanya yaitu tatap muka secara langsung.

Ditinjau dari Al-Qur'an bahwa setiap apa yang dilakukan manusia terdapat pahala ataupun balasan, oleh karena itu dalam Islam manusia harus melakukan evaluasi kepada diri sendiri untuk meningkatkan keimanan, Dalam Al-Qur'an surat Al- Ankabut ayat 2-3 dijelaskan bahwa;

² Akhmad Riadi, "Problematika Sistem Evaluasi Pembelajaran" 15, no. 27 (2017): 1–12.

أَحْسِبَ النَّاسُ أَنْ يُتْرَكُوا أَنْ يَقُولُوا ءَامَنَّا وَهُمْ لَا يُفْتَنُونَ. وَلَقَدْ فَتَنَّا الَّذِينَ مِنْ قَبْلِهِمْ³

فَلْيَعْلَمَنَّ اللَّهُ الَّذِينَ صَدَقُوا وَلْيَعْلَمَنَّ الْكٰذِبِينَ

Artinya: " Apakah manusia itu mengira bahwa mereka dibiarkan (saja) mengatakan: "Kami telah beriman", sedang mereka tidak diuji lagi?, Dan sesungguhnya kami telah menguji orang-orang yang sebelum mereka, maka sesungguhnya Allah mengetahui orang-orang yang benar dan sesungguhnya Dia mengetahui orang-orang yang dusta".³

Penjelasan dari dalil tersebut bahwa pentingnya evaluasi bagi seorang pendidik dengan menanamkan nilai-nilai kesadaran kepada peserta didik agar dapat mengevaluasi dirinya sendiri sesuai dengan dalil ayat tersebut, bahwa keberhasilan suatu pendidikan akan terdongkrak ketika peserta didik didorong untuk melakukan evaluasi.⁴

Dalam hal ini seorang guru harus bisa meningkatkan proses belajar dan mengevaluasi ketercapaian menurut Undang-Undang Nomor 14 Tahun 2005 tentang Guru dan Dosen, kompetensi adalah seperangkat pengetahuan, keterampilan, dan perilaku yang harus dimiliki, dihayati, dan dikendalikan oleh guru dan dosen untuk melaksanakan tanggung jawab profesionalnya. Selain itu kompetensi guru sebagaimana dimaksud dalam Pasal 8 meliputi kompetensi pedagogik, kompetensi kepribadian, kompetensi sosial, dan kompetensi

³ Q.S Al- Ankabut ayat: 2-3

⁴ Nurul Hasanah, Lailatur Badriyah, and Wena Diarmard Selia, "Evaluasi Pembelajaran Ditinjau Dari Al-Qur'an Surah Al-Ankabut Ayat 2-3," *JPT (Jurnal Pendidikan Tematik)* 1, no. 2 (2020): 18–19.

profesional yang diperoleh melalui pendidikan profesi,” sesuai ayat 1 Pasal 10.⁵

Kompetensi yang dimiliki guru sebagaimana yang dijelaskan diatas bahwa salah satu kompetensi yang harus dikuasai yaitu kompetensi pedagogik yang mana menjadi berkaitan dengan kapasitas untuk memahami siswa, merencanakan dan melaksanakan pengajaran, menilai hasil belajar, dan membantu siswa menyadari potensi penuh mereka. Oleh karena itu guru harus memahami bahwa tidak hanya merancang dan membuat alat evaluasi untuk mengetahui ketercapaian siswa saja, akan tetapi mengevaluasi apakah alat evaluasi tersebut dapat memenuhi ketercapaian sebagai alat evaluasi yang baik. menurut Sudijono, evaluasi pendidikan, adalah suatu proses atau kegiatan untuk menentukan nilai pendidikan agar dapat diketahui mutu dan hasilnya.⁶

Secara garis besar, maka alat evaluasi yang digunakan dapat digolongkan menjadi dua macam, yaitu tes dan nontes.⁷Oleh karena itu perlunya ada analisis butir soal sebagai bagian dari proses evaluasi untuk mengetahui pencapaian peserta didik dan juga kualitas soal itu sendiri yang berkaitan dengan validitas, reliabilitas, dan tingkat kesukaran butir soal dan dua beda. sebagaimana dinyatakan oleh Arikunto bahwa, jika tes tersebut valid, reliabel, objektif, praktis, dan hemat biaya, maka tes tersebut memenuhi

⁵ Meita Fitriawanati, “Peran Analisis Butir Soal Guna Meningkatkan Kualitas Butir Soal, Kompetensi Guru Dan Hasil Belajar Peserta Didik,” *JPT : Jurnal Pendidikan Tematik*, 2017.

⁶ Santi Wahndianti and Sumarsih, “Analisis Butir Soal Ujian Akhir Semester Gasal Mata Pelajaran Hitung Dagang” 7, no. 1 (2018): 1–14.

⁷ Farida and Anna Musyarofah, “Validitas Dan Reliabilitas Dalam Analisis Butir Soal,” *Al-Mu’arrib: Jurnal Pendidikan Bahasa Arab* 1, no. 1 (2021): 34–44.

syarat sebagai alat ukur yang baik.⁸ Jika tes yang diberikan dengan kualitas yang buruk, maka hasilnya akan buruk dan Jika tes tersebut mampu mengukur apa yang seharusnya diukur, maka dikatakan valid, dan dikatakan reliabel jika dapat dipercaya dan konsisten.

Berdasarkan uraian diatas bahwa evaluasi sangat erat kaitanya dengan peningkatan pembelajaran yang ada di sekolah, yang mana dalam evaluasi tersebut akan memberikan dampak yang baik, baik untuk guru maupun peserta didik, perkembangan hasil belajar siswa dan tingkat kemampuan siswa sama-sama dapat dilihat oleh guru. pada akhirnya, pengajar akan mengetahui tentang kelangsungan pengalaman yang berkembang. Selain itu, guru dapat memutuskan apa yang harus dilakukan terhadap siswa yang hasil belajarnya bersifat positif atau negatif.

Dengan demikian peran guru sebagai evaluator menunjukkan dua hal: kemampuan untuk mengevaluasi keberhasilan mengajarnya dan kemampuan siswa untuk menguasai kompetensi berbasis kurikulum, yang berdampak signifikan terhadap kompetensinya.⁹ Dalam hal ini tidak semua sekolah menerapkan analisis butir soal, di SMA Negeri 1 Kota Mungkid penerapan analisis soal ujian akhir semester dilaksanakan dengan uji validitas oleh pihak ahli sekolah, sedangkan untuk perumusan dan penyusunan soal dibuat oleh guru mata pelajaran itu sendiri, dan bukan dari hasil musyawarah beberapa guru, khususnya guru pendidikan agama Islam.

⁸ Fitriawanati, "Peran Analisis Butir Soal Guna Meningkatkan Kualitas Butir Soal, Kompetensi Guru Dan Hasil Belajar Peserta Didik."

⁹ Ibid.

Oleh karena itu perlu adanya uji kualitas soal, untuk menentukan apakah soal tersebut layak atau belum dan sudah sesuai dengan kaidah dan tujuan pembelajaran pada kurikulum 2013, uji validitas soal hanya dikumpulkan dan belum ada pengujian mendalam dalam sampai dengan analisis butir soal berkaitan dengan uji validitas, reliabilitas, tingkat kesukaran butir dan daya beda soal ujian akhir semester di SMA Negeri 1 Kota Mungkid, kenyataan di lapangan menunjukkan bahwa selama ini kegiatan analisis butir-butir soal jarang dilakukan. Itulah sebabnya materi, validitas, reliabilitas, dan analisis butir soal yang terdiri dari tingkat kesukaran, daya pembeda tidak diketahui secara pasti.

Peneliti dalam hal ini memahami bahwa penting adanya uji kualitas soal untuk mengetahui seberapa jauh ketercapaian materi yang disampaikan oleh guru dan sebagai tolak ukur ketercapaian siswa tidak hanya dengan hasil penilaian saja, akan tetapi guru perlu mengetahui dan menjadi bahan evaluasi tentang seberapa valid, reliabel daya beda dan tingkat kesukaran soal yang telah diujikan untuk menjadi bahan perancangan soal di kemudian hari.

Dalam hal ini peneliti tertarik untuk meneliti uji validitas, reliabilitas, daya beda dan tingkat kesukaran butir soal sesuai dengan fenomena di atas, oleh karena itu peneliti tertarik untuk meneliti tentang **“Analisis Butir Soal Ujian Akhir Semester Gasal Pendidikan Agama Islam Pada Kelas XII di SMA Negeri 1 Kota Mungkid Tahun Ajaran 2022/2023”**.

B. Batasan Masalah

Sesuai dengan latar belakang, bahwa batasan masalah yang ada yaitu berkaitan dengan uji validitas, releabilitas, tingkat kesukaran butir soal dan daya beda, sesuai dengan fenomena diatas, oleh karena itu peneliti tertarik untuk meneliti tentang“**Analisis Butir Soal Ujian Akhir Semester Gasal Pendidikan Agama Islam Pada Kelas XII di SMA Negeri 1 Kota Mungkid Tahun Ajaran 2022/2023**”.

C. Rumusan Masalah

Sesuai dengan batasan masalah diatas bahwa penulis merumuskan sub masalah sebagai berikut:

1. Bagaimana validitas butir soal Ujian Akhir Semester Gasal Pendidikan Agama Islam, pada kelas XII di SMA Negeri 1 Kota Mungkid Tahun Ajaran 2022/2023.
2. Bagaimana releabilitas butir soal Ujian Akhir Semester Gasal Pendidikan Agama Islam, pada kelas XII di SMA Negeri 1 Kota Mungkid Tahun Ajaran 2022/2023.
3. Bagaimana tingkat kesukaran dan daya beda butir soal Ujian Akhir Semester Gasal Pendidikan Agama Islam, pada kelas XII di SMA Negeri 1 Kota Mungkid Tahun Ajaran 2022/2023.

D. Tujuan Penelitian

Berdasarkan rumusan masalah yang ada, penelitian ini bertujuan untuk mengetahui :

1. Untuk mengetahui kualitas butir soal Ujian Akhir Semester Gasal Pendidikan Agama Islam, pada kelas XII di SMA Negeri 1 Kota Mungkid Tahun Ajaran 2022/2023 dari segi validitasnya.
2. Untuk mengetahui kualitas butir soal Ujian Akhir Semester Gasal Pendidikan Agama Islam, pada kelas XII di SMA Negeri 1 Kota Mungkid Tahun Ajaran 2022/2023 dari segi reliabilitasnya.
3. Untuk mengetahui tingkat kesukaran dan daya beda butir soal Ujian Akhir Semester Gasal Pendidikan Agama Islam, pada kelas XII di SMA Negeri 1 Kota Mungkid Tahun Ajaran 2022/2023.

E. Kegunaan Penelitian

1. Kegunaan teoritis
 - a. Dapat dijadikan acuan dalam penelitian yang akan mendatang, dan menambah khazanah keilmuan untuk meningkatkan pendidikan di indonesia sekaligus menjadi bahan masukan untuk meningkatkan kualitas soal evaluasi di mata pelajaran Pendidikan Agama Islam dengan merumuskan soal evaluasi dengan baik.
 - b. Dapat menjadi referensi penelitian berkaitan dengan analisis buti soal Ujian Akhir Semester Gasal Pendidikan Agama Islam, pada kelas XII di SMA Negeri 1 Kota Mungkid Tahun Ajaran 2022/2023.
2. Kegunaan praktis

- a. Dapat digunakan untuk bahan pertimbangan dan peningkatan akademik SMA Negeri 1 Kota Mungkid pada jangka waktu yang pendek maupun jangka waktu yang panjang.
- b. Hasil penelitian dapat dijadikan sebagai bahan evaluasi dan menambah wawasan guru dan karyawan dalam peningkatan soal evaluasi berkaitan dengan analisis butir soal Ujian Akhir Semester Gasal Pendidikan Agama Islam, pada kelas XII di SMA Negeri 1 Kota Mungkid Tahun Ajaran 2022/2023. Juga menjadi motivasi bagi guru untuk meningkatkan soal evaluasi guna meningkatkan kualitas evaluasi khususnya pada mata pelajaran Pendidikan Agama Islam.
- c. Hasil penelitian ini dapat menjadi bahan penelitian dan juga untuk menambah wawasan bagi pihak-pihak yang membutuhkan berkaitan dengan analisis butir soal Ujian Akhir Semester Gasal Pendidikan Agama Islam, pada kelas XII di SMA Negeri 1 Kota Mungkid Tahun Ajaran 2022/2023.

BAB II

KAJIAN PUSTAKA

A. Kajian Teori

1. Evaluasi

a. Pengertian

Secara harfiah evaluasi berasal dari kata bahasa Inggris “*evaluation*” yang berarti “penilaian” atau “*assessment*”. Istilah pengukuran (*measurement*) dan penilaian (*assessment*) sering digunakan untuk mengartikan evaluasi itu sendiri selain evaluasi.¹⁰ Karena makna etimologinya yang mirip, ketiga istilah ini terkadang digunakan secara bergantian. Namun, menurut terminologi, ketiganya memiliki arti yang berbeda, dalam bahasa Arab *al-Taqdir*, dalam bahasa Indonesia berarti: evaluasi. Frasa dasarnya adalah *value*, dalam bahasa Arab: *al-Qimah*; dalam bahasa Indonesia artinya : nilai. Dengan demikian secara harfiah penilaian instruksional (*instructional assessment*) = *al-Taqdir al-Tarbawiy*) dapat diartikan sebagai evaluasi dalam mata pelajaran atau evaluasi mata pelajaran yang mengacu pada kegiatan pembelajaran.¹¹

Pendidikan nasional berfungsi mengembangkan kemampuan watak dan peradaban bangsa yang bermartabat dalam rangka

¹⁰ John M. Echols and Hassan Shadilly, *An English-Indonesian Dictionary* (Cet. XXV; Jakarta: PT. Gramedia, 2003), h. 220

¹¹ Gito Supriadi, *Pengantar Dan Teknik Evaluasi Pembelajaran*, 1st ed. (Malang: Intimedia, 2011).

mencerdaskan kehidupan bangsa dan mengembangkan potensi peserta didik agar menjadi manusia yang beriman dan bertakwa kepada Tuhan Yang Maha Esa, sehat, berilmu, cakap, kreatif, dan bertakwa. menjadi warga negara yang demokratis dan bertanggung jawab, sesuai dengan fungsi dan tujuan Sisdiknas no 20 Tahun 2003, pasal 3.¹²

Menurut Mehren dan Lehmann evaluasi adalah prosedur merencanakan, memperoleh, dan menyediakan informasi yang diperlukan untuk membuat alternatif keputusan.¹³ Dapat kita pahami bahwa evaluasi menjadi alat untuk mengumpulkan informasi yang telah direncanakan untuk memperoleh data yang menjadi bahan dalam menentukan keputusan.

Ditinjau dari Al-Qur'an bahwa setiap apa yang dilakukan manusia terdapat pahala ataupun balasan, oleh karena itu dalam Islam manusia harus melakukan evaluasi kepada diri sendiri untuk meningkatkan keimanan, Dalam Al-Qur'an surat Al- Ankabut ayat 2-3 dijelaskan bahwa;

أَحْسِبَ النَّاسُ أَنْ يُتْرَكُوا أَنْ يَقُولُوا ءَامَنَّا وَهُمْ لَا يُفْتَنُونَ. وَلَقَدْ فَتَنَّا الَّذِينَ مِنْ قَبْلِهِمْ ط

فَلْيَعْلَمَنَّ اللَّهُ الَّذِينَ صَدَقُوا وَلْيَعْلَمَنَّ الْكٰذِبِينَ

Artinya: " Apakah manusia itu mengira bahwa mereka dibiarkan (saja) mengatakan: "Kami telah beriman", sedang mereka tidak diuji

¹² Undang-undang R.I. Tentang Sistem Pendidikan Nasional No. 20 Tahun 2003, (Cet. V, Jakarta, Sinar Grafika, 2013), h. 3

¹³ Rusydi Ananda and Tien Rafida, *Pengantar Evaluasi Program Pendidikan*, ed. Candra Wijaya, *Perdana Publishing*, 1st ed., vol. 53 (Medan: Perdana Publishing, 2017).

*lagi?, Dan sesungguhnya kami telah menguji orang-orang yang sebelum mereka, maka sesungguhnya Allah mengetahui orang-orang yang benar dan sesungguhnya Dia mengetahui orang-orang yang dusta”’.*¹⁴

Penjelasan dari dalil tersebut bahwa pentingnya evaluasi bagi seorang pendidik dengan menanamkan nilai-nilai kesadaran kepada peserta didik agar dapat mengevaluasi dirinya sendiri, sesuai dengan ayat sebelumnya. keberhasilan suatu pendidikan akan terdongkrak ketika peserta didik didorong untuk melakukan evaluasi .¹⁵

Alkin mendefinisikan bahwa evaluasi sebagai proses metodelis pengumpulan, analisis, dan pelaporan data yang dapat digunakan untuk membuat keputusan tentang program atau proyek yang sedang dievaluasi. Selanjutnya Stark dan Thomas mengemukakan bahwa evaluasi adalah proses pemilihan, pengumpulan, analisis, dan penyajian informasi yang dapat dijadikan landasan untuk pengambilan keputusan dan penyusunan program¹⁶

b. Ruang Lingkup Penilaian Pendidikan

Abas Sudjono menjelaskan bahwa ruang lingkup penilaian pendidikan terlihat dari ruang lingkup program pembelajaran, kegiatan/proses pembelajaran dan hasil pembelajaran.¹⁷

¹⁴ Q.S Al- Ankabut ayat: 2-3

¹⁵ Hasanah, Badriyah, and Selia, “Evaluasi Pembelajaran Ditinjau Dari Al-Qur’an Surah Al-Ankabut Ayat 2-3.”

¹⁶ Ananda and Rafida, *Pengantar Evaluasi Program Pendidikan*, vol. 53, p. .

¹⁷ Tanwir, “Dasar-Dasar Dan Ruang Lingkup Evaluasi Pendidikan,” *Jurnal Pendidikan Islam* 13, no. 1 (2015): 56–57.

1) Evaluasi program studi

Dalam mengevaluasi suatu program pembelajaran, ada tiga hal pokok yang harus digunakan untuk evaluasi. evaluasi isi program pendidikan; evaluasi strategi belajar mengajar.¹⁸

2) Penilaian kegiatan/proses pembelajaran

Proses penyampaian pendidikan dapat melibatkan beberapa hal, seperti: Kesesuaian antara proses belajar mengajar yang berlangsung di GBPP, motivasi guru dalam melaksanakan program pengajaran, kemauan siswa untuk berpartisipasi dalam proses pembelajaran, minat atau perhatian siswa setelah mengikuti pelajaran, aktivitas siswa atau partisipasi siswa selama proses pembelajaran. Di mana pembelajaran berlangsung, peran membimbing, menasihati siswa tentang apa yang mereka butuhkan, komunikasi dua arah antara guru dan siswa selama proses pembelajaran, mendorong dan memotivasi siswa, penerapan teori yang diperoleh di kelas Dengan memberikan tugas kepada siswa yang berkaitan dengan upaya tersebut untuk menghilangkan efek negatif yang dihasilkan dari kegiatan sekolah..¹⁹

3) Penilaian hasil belajar

Penilaian hasil belajar siswa meliputi: Peringkat, dalam kaitannya dengan kemampuan siswa terhadap tujuan khusus yang

¹⁸ Supriadi, *Pengantar Dan Teknik Evaluasi Pembelajaran*.

¹⁹ Ibid.

ingin dicapai dalam satuan kurikulum terbatas, menilai tingkat penyalahgunaan tujuan umum pelajaran oleh siswa;

Dari ketiga bidang penilaian pendidikan tersebut diuraikan, penilaian pendidikan bukan sekedar kumpulan teknik yang diperlukan guru untuk mengukur hasil belajar siswa, melainkan suatu proses berkesinambungan yang membentuk keseluruhan proses pendidikan. jelas.

.Dalam ruang lingkup Pendidikan Agama Islam ada beberapa pembahasan berkaitan dengan tujuan dan lingkup capaian yang harus dicapai oleh peserta didik disamping ruang lingkup evaluasi program pendidikan berkaitan dengan ketiga lingkup evaluasi pendidikan.

2. Pendidikan Agama Islam

a. Pengertian

Pendidikan Agama Islam merupakan salah satu usaha edukatif dan materi mendidik Pelajaran Agama yang ditujukan untuk penataan karakter siswa sesuai Pelajaran Agama Islam. Pesantren juga merupakan suatu pekerjaan sebagai pembimbing dan pengasuhan bagi siswa atau siswi agar nantinya setelah menyelesaikan sekolah mereka menjadi mandiri dan siap untuk mengambil kesimpulan tentang gaya

hidup (*lifestyle*) ataupun menjadi pedoman serta tujuan hidup mereka.²⁰

Sedangkan Pendidikan Islam adalah peningkatan jiwa manusia dan rencana tingkah laku dan perasaan dalam pandangan agama Islam yang sepenuhnya dimaksudkan untuk memahami tujuan Islam dalam eksistensi manusia dan masyarakat, khususnya dalam segala bidang kehidupan.

Pendidikan Islam menurut Al-Ghazali adalah mencari dan membentuk kepribadian seseorang itu baik dunia dan akhiratnya kelak. Ini akan terjadi ketika ilmu pengetahuan yang didapat di sekolah dicari dan dilatih untuk keuntungannya untuk membuatnya lebih dekat dengan Tuhannya dan mencapai kebahagiaan di akhirat.²¹ Pendidikan Islam pada jalur pendidikan informal, nonformal dan formal memiliki peran strategis untuk mencapai tujuan pendidikan Islam yang telah dirumuskan.²²

Seperti yang diungkapkan Munzir Hitami bahwa alasan pendidikan agama Islam harus mencakup tiga hal, khususnya: tujuan utamanya adalah teleologi, secara spesifik kembali kepada Tuhan, tujuan selanjutnya adalah optimis, untuk menghadapi segala sesuatu yang tela

²⁰ Rena Citra, "Kompetensi Guru Pendidikan Agama Islam Dalam Meningkatkan Motivasi Belajar Siswa (Studi Deskriptif Di Smpn 1 Singingi Hilir)," *Jurnal AL-HIKMAH* 51, no. 1 (2018): 173–183.

²¹ Khoirotul Ni'amah, "Paradigma Pendidikan Islam Perspektif Al-Ghazali," *Heutagogia: Journal of Islamic Education* 1, no. 1 (2021): 59–71.

²² Ahmad Darlis, "Hakikat Pendidikan Islam: Telaah Antara Hubungan Pendidikan Informal, Non Formal Dan Formal" XXIV, no. 1 (1992): 174.

ditetapkan di dunia, dan tujuan ketiga adalah Amanat untuk menjadi hamba Allah SWT.²³

b. Aspek pendidikan Agama Islam

Sesuai dengan keputusan kementerian agama indonesia nomor 13 tahun 2014 tentang pendidikan keagamaan Islam diantaranya:

1) Al- Qur'an dan Hadits

Mengedepankan pada aspek membaca, menulis, dan menerjemahkan dengan baik dan benar

2) Tauhid

Mengedepankan pada aspek pemahaman terhadap keyakinan beragama, serta mengamalkan nilai-nilai ketauhidan atau nilai-nilai yang terkandung di asm'ul husna maupun sifat-sifat yang dimiliki oleh Allah SWT.

3) Fikih

Mengedepankan pada pemahaman ibadah dan muamalah dengan baik dan benar

4) Tarikh

Mengedepankan pada aspek pemahaman dan mengambil ibrah atau pelajaran dari suatu kisah, peristiwa-peristiwa yang bersejarah dalam islam, serta meneladani ulama-ulama serta tokoh-tokoh muslim yang membawa perubahan.

5) Akhlaq

²³ Ade Imelda Frimayanti, "Implementasi Pendidikan Nilai Dalam Pendidikan Agama Islam," *Al-Tadzkiyyah: Jurnal Pendidikan Islam* 8, no. 11 (2017): Hal. 240.

Mengedepankan pada aspek keteladanan pada sikap terpuji dan menjauhi segala hal perilaku yang tercela.²⁴

Dalam menyelesaikan penilaian Pendidikan agama Islam, Al-Qur'an memberikan beberapa petunjuk sebagai berikut:

- a) Standar koherensi (kemajuan), pelajaran dari pendidikan agama Islam, sangat memperhatikan tentang pedoman dalam kesesuaiannya dan mematuhi standar tersebut untuk memutuskan pilihan dari seseorang (Q.S. 46:13 - 14).
- b) Pedoman luas, aturan yang memandang semua sudut pandang, terdiri dari karakter, hafalan, mencari tahu keaslian, keahlian, disposisi membantu, kewajiban (Q.S. 99: 7-8).
- c) Kaidah objektivitas, dalam menilai dengan melihat realitas yang sebenarnya, tidak boleh terpengaruh oleh hal-hal yang bersifat irrasional maupun emosional (Q.S 5:8).²⁵

Dalam pembahasan yang sudah dipaparkan diatas bahwa perlunya ada evaluasi terhadap proses belajar siswa pada mata pelajaran Pendidikan Agama Islam, oleh karena itu dalam proses evaluasi pendidikan terdapat beberapa teknik evaluasi. Secara garis besar alat evaluasi yang digunakan dapat digolongkan menjadi dua macam, yaitu tes dan nontes.²⁶

Metode pengujian dapat berupa tertulis atau tidak tertulis. Metode nontes digunakan untuk menilai sikap, keterampilan, perilaku, dan

²⁴ Kemenag RI, *Peraturan Menteri Agama Nomor 13 Tentang Pendidikan Keagamaan Islam*, 2014.

²⁵ Frimayanti, "Implementasi Pendidikan Nilai Dalam Pendidikan Agama Islam."

²⁶ Farida and Musyarofah, "Validitas Dan Reliabilitas Dalam Analisis Butir Soal."

sebagainya. Tes dalam pendidikan adalah suatu alat penilaian atau asesmen yang sistematis, valid, reliabel, dan objektif untuk menentukan kompetensi, keterampilan, dan tingkat pengetahuan siswa tentang materi, dalam berupa tugas atau pertanyaan yang diajukan oleh siswa itu sebuah cara atau sekelompok siswa.

Berdasarkan objektivitas evaluasi, tes dibedakan menjadi tes objektif dan tes subjektif. Tes objektif dapat berupa tes kebenaran, tes menjodohkan, tes pelengkap, tes isian, dan tes pilihan ganda. Tes subyektif, di sisi lain, termasuk tes esai dan tes esai. Tes pilihan ganda (*multiple-choice*) adalah bentuk teks dengan jawaban yang benar atau paling tepat. dalam tes ini, jawabannya mutlak.

Oleh karena itu perlunya ada analisis butir soal sebagai bagian dari proses evaluasi untuk mengetahui pencapaian peserta didik dan juga kualitas soal itu sendiri yang berkaitan dengan validitas, reliabilitas, dan tingkat kesukaran butir soal. sebagaimana dinyatakan oleh Arikunto bahwa, jika tes tersebut valid, reliabel, objektif, praktis, dan hemat biaya, maka tes tersebut memenuhi syarat sebagai alat ukur yang baik.²⁷ Jika tes yang diberikan dengan kualitas yang buruk, maka hasilnya akan buruk dan Jika tes tersebut mampu mengukur apa yang seharusnya diukur, maka dikatakan valid, dan dikatakan reliabel jika dapat dipercaya dan konsisten.

Alat yang baik (dalam bentuk tes atau non tes) harus valid dan reliabel. Efektif berarti membutuhkan efektivitas internal dan eksternal.

²⁷ Fitrianawati, "Peran Analisis Butir Soal Guna Meningkatkan Kualitas Butir Soal, Kompetensi Guru Dan Hasil Belajar Peserta Didik."

Suatu ukuran memiliki validitas internal atau wajar jika kriteria yang terkandung dalam ukuran tersebut secara wajar (secara teoritis) mencerminkan apa yang sedang diukur. Menurut Sutrisno Hadi, validitas internal sarana berupa tes harus memenuhi validitas konstruk dan validitas isi. Untuk instrumen non tes yang digunakan untuk mengukur sikap cukup memenuhi validitas konstruk.²⁸

c. Analisis butir soal

a. Pengertian

Menurut Arikunto, “analisis soal adalah suatu prosedur yang sistematis, yang akan memberikan informasi-informasi yang sangat khusus terhadap butir tes yang kita susun”. Selain itu menurut Nana Sudjana, Analisis butir soal atau analisis item adalah pengkajian pertanyaan-pertanyaan tes agar diperoleh perangkat pertanyaan yang memiliki kualitas yang memadai. Analisis soal juga dilakukan untuk mengetahui berfungsi tidaknya sebuah soal. Analisis pada umumnya dilakukan melalui dua cara, yaitu analisis kualitatif (*qualitatif control*) dan analisis kuantitatif (*quantitatif control*).²⁹

b. Fungsi analisis butir soal

Dalam *Journal of Multiple Choice Item Analysis from Linguistic Aspects*, Arikunto menyatakan bahwa fungsi dari item analysis adalah: Ini membantu Anda mengidentifikasi item yang rusak, mendapatkan informasi yang dapat Anda gunakan untuk lebih mempersempit

²⁸ Sugiono, *Metode Penelitian Kuantitatif Kualitatif Dan R&D*, 19th ed. (Bandung: ALFABETA, 2013).

²⁹ Farida and Musyarofah, “Validitas Dan Reliabilitas Dalam Analisis Butir Soal.”

pertanyaan yang Anda minati, dan mendapatkan wawasan tentang situasi di mana item tersebut ditempatkan.³⁰

Analisis butir oleh Thorndik dan Hagnr dikutip oleh Purwanto Dalam jurnal *Multiple Choice Item Analysis from Language Aspects* tertulis: jawaban atas pertanyaan memberikan informasi diagnostik untuk memeriksa pelajaran kelas dan kesalahan belajar mereka, yang mengarah ke pembelajaran yang lebih baik sebuah. Dengan menjawab setiap pertanyaan dan meningkatkan pertanyaan berdasarkan jawaban tersebut, Anda akan meletakkan dasar untuk persiapan yang lebih baik untuk ujian tahun depan.³¹

c. Karakteristik Butir Soal

1. Validitas

Validitas adalah sejauh mana suatu tes mengukur seberapa baik (tes atau nontes) valid dan reliabel. Efektif berarti membutuhkan efektivitas internal dan eksternal.

- a) Validitas internal dapat dilakukan dengan dua cara yaitu analisis faktor dan analisis butir. Analisis faktor. Analisis faktor dilakukan dengan mengkorelasikan skor faktor dengan skor total. Metode korelasi yang digunakan adalah metode korelasi *product-moment*. Korelasi positif yang signifikan

³⁰ Nurjanah and Noni Marlianingsih, "Analisi Butir Soal Pilihan Ganda Dari Aspek Kebahasaan," *Faktor Jurnal Ilmu Kependidikan* II, no. 1 (2015): 69–78.

³¹ Ibid.

berarti item untuk faktor tersebut dianggap valid dan analisis butir dengan mengkorelasikan skor.³²

b) Validitas eksternal

Validitas eksternal dilakukan dengan cara mengkorelasikan skor hasil uji coba instrumen yang dibuat guru dengan instrumen yang sudah baku. Menguji signifikansi angka korelasi dengan mencari angka korelasi antara skor-skor tes buatan dengan skor- skor tes toufel. Teknik korelasi yang digunakan adalah teknik korelasi *Product Moment*, jika angka korelasi yang diperoleh ternyata signifikan, berarti tes yang dibuat guru dapat dianggap valid.³³

1) Pengujian Validitas Tes Berbentuk Objektif Test

Tes berbentuk objektif seperti pilihan ganda, benar-salah, dan menjodohkan merupakan tes dengan skor dikotomi peringkat 0 dan 1. Pengecekan validitas dilakukan untuk korelasi positif dan sesuai dengan uji korelasi *product-moment*. Dalam metode tersebut, perbedaan yang signifikan antara skor item dan skor total berarti item tersebut dianggap valid.³⁴

Besarnya korelasi Arikunto memberikan kriteria penafsiran sebagai berikut: Antara 0,800 sampai dengan

³² Asrul, Rusydi Ananda, and Rosinta, *Evaluasi Pembelajaran, Ciptapustaka Media*, 2nd ed. (Medan: Citapustaka Media, 2015).

³³ Ibid.

³⁴ Ibid.

1,00 = sangat tinggi, antara 0,600 sampai dengan 0,800 = tinggi, antara 0,400 sampai dengan 0,600 = cukup, antara 0,200 sampai dengan 0,400 = rendah, antara 0,00 sampai dengan 0,200 = sangat rendah.³⁵

2. Reliabilitas

Reliabilitas berasal dari kata *reliability* berarti sejauh mana hasil suatu pengukuran dapat dipercaya. Suatu hasil pengukuran dapat dipercaya apabila dalam beberapa kali pelaksanaan pengukuran terhadap kelompok subyek yang sama, diperoleh hasil pengukuran yang relatif sama, selama aspek yang diukur dalam diri subyek memang belum berubah.³⁶

Menurut Allen dan Yen dalam jurnal *Reliabilitas Instrumen Pendidikan* bahwa koefisien reliabilitas besarnya ditentukan oleh satu dikurangi dengan perbandingan varians kesalahan pengukuran dan varians skor tampak. Atas dasar demikian semakin kecil varians kesalahan pengukuran, maka akan semakin tinggi koefisien reliabilitas, sehingga varians tampak atas dasar demikian semakin kecil varians kesalahan pengukuran, maka akan semakin tinggi koefisien reliabilitas, sehingga varians tampak atas dasar demikian semakin kecil

³⁵ Suharsimi Arikunto, *Dasar-Dasar Evaluasi Pendidikan*, Revisi. (Jakarta: Bumi Aksara, 2010).

³⁶ Nina Helpiana, "Analisis Butir Soal Ujian Akhir Semester Genap Tahun Ajaran 2018/2019 Mata Pelajaran PAI Kelas XI SMAN 2 Gunung Meriah Kabupaten Aceh Singkil" (2020).

varians kesalahan pengukuran, maka akan semakin tinggi koefisien reliabilitas, sehingga varians tampak dapat digunakan sebagai atribut koefisien reliabilitas yang sesungguhnya.³⁷

Dalam menentukan reliabilitas, menggunakan rumus alpha dengan rentang skor 0 - 10, 0 - 100 atau berbentuk skala 1 - 3, 1 - 5 atau 1 - 10.

Sebelum dicari angka reliabilitasnya, perlu terlebih dahulu dicari varians butir dan varians skor total dengan rumus:

$$\sigma^2 = \frac{\sum X^2 - \frac{(X^2)}{N} N}{N}$$

$$r_{11} = \left[\frac{k}{k-1} \right] \left[1 - \frac{\sum S_b^2}{S_t^2} \right]$$

Keterangan:

r_{11} = Reliabilitas instrumen

k = Jumlah butir soal

$\sum s_b^2$ = Jumlah varians butir

s_t^2 = Varians total

Interpretasi terhadap hasil perhitungan koefisien reliabilitas tes (r_{11}) pada umumnya digunakan patokan sebagai berikut. Jika angka 0.800-1.000 tingkat realibilitas sangat tinggi, Jika angka 0.600-0.800 tingkat realibilitasnya tinggi, Jika angka 0.400-0.600 tingkat realibilitasnya cukup, Jika angka 0.200-0.400 tingkat realibilitasnya rendah, Jika angka 0.000-0.200

³⁷ Ibid.

tingkat realibilitasnya sangat rendah.³⁸ Hasil perhitungan menggunakan rumus tersebut diinterpretasikan dengan tingkat koefisien korelasi sebagai berikut:

Penelitian reliabilitas ini akan rumus meliputi KR-20, *Alpha Cronbach*, *Spearman Brown* dengan dua software yaitu Microsoft Office Excel 2010 dan software IBM SPSS 20. Untuk interpretasi reliabilitas instrumen adalah sebagai berikut: 0,80 – 1,00 = tinggi 0,60 – 0,80 = cukup 0,40 – 0,60 = agak rendah 0,20 – 0,40 = rendah 0,00 – 0,20 = sangat rendah, Pada penelitian ini instrumen dikatakan reliabel jika tingkat koefisien korelasinya 0,600 atau lebih.

3. Tingkat Kesukaran

Soal yang baik adalah soal yang tidak terlalu mudah atau tidak terlalu sukar.³⁹ Soal yang terlalu mudah akan membuat siswa menyepelkan soal tersebut dan tidak merangsang siswa untuk menyelesaikan problematika soal tersebut, di samping itu soal yang terlalu sulit akan membuat siswa tidak nyaman dan putus asa untuk memecahkan problematika dalam soal tersebut karena jauh dari jangkauan pemahamannya, sehingga soal yang tidak terlalu mudah dan tidak terlalu sukar akan

³⁸ Ardilah Muluki, "Analisis Kulaitas Butir TES Semester Ganjil Mata Pelajaran Ilmu Pengetahuan Alam Kelas IV MI Radhiatul Adawiyah" (Universitas Muhammadiyah Makassar, 2020).

³⁹ Asrul, Ananda, and Rosinta, *Evaluasi Pembelajaran*.

menjadikan evaluasi dan pengukuran tingkat ketercapaian siswa, serta tujuan pembelajaran dapat tercapai dengan baik.

Indeks Kesukaran Angka yang menunjukkan tingkat kesulitan dan kemudahan suatu soal. Indeks kesukarannya antara 0,00 dan disebut indeks kesukaran. tingkat kesulitan ke 1.0. Indeks kesukaran ini menunjukkan tingkat kesukaran soal. Antara 0,00 dan 1,0. Indeks kesukaran ini menunjukkan tingkat kesukaran soal. Tampilkan masalah dengan indeks kesulitan 0,0

Soal dengan indeks kesukaran 0,0 menunjukkan bahwa tugas tersebut terlalu sulit. Menunjukkan bahwa tugasnya terlalu sulit. Jika tidak, indeks 1,0 menunjukkan bahwa tugasnya terlalu mudah. Untuk penilaian indeks kesukaran ini diberi lambang P (P kapital), dan untuk penilaian indeks kesukaran ini diberi lambang P (P singkatan dari big). Jadi masalahnya adalah pada $P = 0,20$. Sebaliknya, soal $P = 0,30$ lebih sulit daripada soal $P = 0,80$.⁴⁰

⁴⁰ Ibid.

Adapun rumus mencari P adalah

$$P = \frac{B}{JS}$$

Keterangan:

P = Indeks kesukaran.

B = Banyaknya siswa yang menjawab soal itu dengan betul

JS = Jumlah seluruh siswa peserta tes

Hasil perhitungan menggunakan rumus (P) dengan dua software yaitu Microsoft Office Excel 2010 dan software IBM SPSS 20. Soal dengan P 0,00 sampai 0,30 merupakan soal sukar, soal dengan P 0,30 sampai 0,70 merupakan soal sedang, dan soal dengan P 0,70 sampai 1,00 merupakan soal mudah. Meskipun demikian, mereka berpendapat bahwa soal yang dianggap baik, atau cukup sulit, adalah soal dengan indeks peringkat 0,30 hingga 0,70..⁴¹

Tabel 1. Interpretasi Indeks kesukaran item

Besarnya P	Interpretasi
P 0,00 sampai 0,30	Sukar
P 0,30 sampai 0,70	Sedang
P 0,70 sampai 1,00	Mudah

4. Daya Beda

Kekhasan suatu soal adalah kemampuannya membedakan siswa yang pandai (berkemampuan tinggi) dan siswa yang

⁴¹ Ibid.

kurang pandai (berkemampuan rendah). Angka yang menunjukkan besarnya perbedaan karakteristik disebut indeks diskriminasi, disingkat D. Penelitian ini menggunakan rumus D(Daya Beda) dengan dua software yaitu Microsoft Office Excel 2010 dan software IBM SPSS 20. Mirip dengan indeks empati, indeks identifikasi (identitas) berkisar antara 0,00 hingga 1,00. Tanda negatif indeks kewaspadaan digunakan ketika ada sesuatu yang "terbalik" dan menunjukkan kualitas pengujian. Jadi anak cerdas itu bodoh dan anak bodoh itu cerdas..⁴²

Untuk kelompok kecil, keseluruhan kelompok pengujian dibagi menjadi dua bagian yang sama besar, 50% untuk kelompok atas dan 50% untuk kelompok bawah. Untuk grup besar, dua kutub, atau 27% teratas, biasanya diambil sebagai grup teratas, mengingat biaya dan waktu analisis. (JA) dan 27% terbawah adalah kelompok terbawah (JB).

⁴² Ibid.

Rumus untuk menentukan indeks Daya beda adalah :

$$D = \frac{B_A}{J_A} - \frac{B_B}{J_B} = PA - PB$$

Keterangan:

J = Jumlah peserta tes

JA = Banyaknya peserta kelompok atas

JB = Banyaknya peserta kelompok bawah

BA = Banyaknya peserta kelompok atas yang menjawab soal itu dengan benar.

BB = Banyaknya peserta kelompok bawah yang menjawab soal itu

Tabel 2. Interpretasi Indeks Daya Beda Item

Besarnya Angka Indeks Diskriminasi Item (J)	Interpretasi
Kurang dari 0,20	Lemah
0,20 – 0,40	Cukup
0,41 – 0,70	Baik
0,71 – 1,00	Sangat Baik

B. Penelitian Terdahulu

Studi yang dilakukan oleh Siti Fathimah Al Fathiyah bahwa butir soal pilihan ganda Ulangan Umum Semester Ganjil TP. 2014/2015 bidang studi Bahasa Arab siswa kelas XI-IPS MA Roudlotul Ulum Pagak Malang, dapat disimpulkan bahwa, *Pertama*, tingkat validitas soal masih dikategorikan dalam soal yang kurang valid,. *Kedua*, tingkat reliabilitas soal ini masuk dalam kategori Sedang. untuk tingkat kesukarannya antara butir soal nomor 1-15 ada 8 butir soal yang Mudah dan 3 butir soal termasuk soal yang Sedang dan 4

butir soal yang termasuk dalam soal yang Sukar. *Ketiga*, tingkatan daya pembeda soal ini masih banyak yang tidak bisa membedakan antara siswa dari kelompok atas dengan siswa kelompok bawah. *Keempat*, kualitas dari pengecoh, masih banyak pengecoh yang perlu direvisi.⁴³

Terkait penelitian terdahulu yang dilakukan oleh Siti Fathimah Al Fathiyah memiliki persamaan yaitu sama- sama menggunakan metode pendekatan kuantitatif analisis butir soal mulai dari validitas, reliabilitas, tingkat kesukaran, daya beda , sedangkan perbedaan dalam penelitian ini yaitu analisis yang diteliti yaitu berkaitan dengan soal Ujian Akhir Semester Gasal Pendidikan Agama Islam, pada kelas XII di SMA Negeri 1 Kota Mungkid Tahun Ajaran 2022/2023.

Studi yang dilakukan oleh Isnani Mua'arifah bahwa, soal Ulangan Akhir Semester mata pelajaran Biologi pada Kelas X dan XI MAN Kota Palangka Raya tahun ajaran 2016/2017 memiliki kualitas yang cukup baik, Validitas soal pada Kelas X sebanyak 52% atau 26 butir soal valid sedangkan soal yang tidak valid sebanyak 48% atau 24 butir soal dan XI 65% sebanyak 58% atau 29 butir soal valid sedangkan soal yang tidak valid sebanyak 42% atau 21 butir soal, Reliabilitas soal pada kelas X dan XI digolongkan tinggi atau reliabel yaitu pada kelas X sebesar 0.708 dan kelas XI sebesar 0.793., Fungsi distraktor soal pada kelas X sebanyak 5 soal kategori sangat baik, 15 soal kategori baik, 14 soal kategori cukup, 8 soal kategori kurang baik dan 8 soal kategori tidak baik. Untuk kelas XI IPA sebanyak 2 soal kategori sangat

⁴³ Siti Fathimah Al Fathiyah, "Analisis Butir Soal Pelajaran Bahasa Arab Di MA Roudlotul Ulum Pagak Malang," *Tarbiyatuna* 4, no. 1 (2019): 78–100.

baik, 8 soal kategori baik, 19 soal kategori cukup, 12 soal kategori kurang baik dan 9 soal kategori tidak baik, Tingkat kesukaran soal pada Kelas X sebanyak 6 soal kategori sukar, 25 soal kategori sedang dan 19 soal kategori mudah dan pada kelas XI sebanyak 7 soal kategori sukar, 14 soal kategori sedang dan 29 soal kategori mudah, Daya pembeda soal pada Kelas X dengan daya pembeda kategori baik sekali sebanyak 0%, kategori baik 8%, kategori sedang 24%, dan kategori jelek 68%. Kemudian daya pembeda pada kelas XI IPA dengan daya pembeda kategori baik sekali sebanyak 0%, kategori baik 8%, kategori sedang 32% dan kategori jelek 60%.⁴⁴

Terkait penelitian terdahulu yang dilakukan oleh Isnani Mua'arifah memiliki persamaan yaitu sama- sama menggunakan metode pendekatan kuantitatif analisis butir soal mulai dari validitas, reliabilitas, tingkat kesukaran, daya beda , sedangkan perbedaan dalam penelitian ini yaitu analisis yang diteliti yaitu berkaitan dengan soal Ujian Akhir Semester Gasal Pendidikan Agama Islam, pada kelas XII di SMA Negeri 1 Kota Mungkid Tahun Ajaran 2022/2023.

Dalam penelitian yang dilakukan oleh Nelly Agustin analisis soal dari segi daya pembeda dan tingkat kesukaran dari soal Ulangan Akhir Semester Ganjil Mata Pendidikan Agama Islam pada kelas VIII di SMP Islam Terpadu Bina Insani Metro Tahun Ajaran 2018/2019, Secara keseluruhan kualitas soal berdasarkan tingkat kesukaran adalah kurang baik, kualitas soal berdasarkan

⁴⁴ Isnani Mua'arifah, "Analisis Kualitas Butir Soal Pilihan Ganda Mata Pelajaran Biologi Kelas X Dan XI Pada Ulangan Akhir Semester Tahun Ajaran 2016/2017 Di MAN Kota Palangka Raya" (2017): 1-14.

daya pembeda cukup baik, secara keseluruhan soal berkualitas cukup baik karena 21 dari 40 soal berkualitas baik.⁴⁵

Terkait penelitian terdahulu yang dilakukan oleh Nelly Agustin memiliki persamaan yaitu sama- sama menggunakan metode pendekatan kuantitatif sedangkan perbedaan dalam penelitian ini yaitu variabel yang diteliti yaitu berkaitan dengan soal Ujian Akhir Semester Gasal Pendidikan Agama Islam, pada kelas XII di SMA Negeri 1 Kota Mungkid Tahun Ajaran 2022/2023

Dalam penelitian yang dilakukan oleh Fauziah Umami analisis soal dari segi dimensi telaah materi, konstruksi, dapat disimpulkan bahwa kualitas butir soal Soal Tes Pilihan Ganda Ujian Nasional Tahun Pelajaran 2018-2019 Mata Pelajaran Bahasa Indonesia Jenjang Sltp (Sekolah Lanjutan Tingkat Pertama) dinyatakan memiliki kualitas soal yang “sangat baik” pada dimensi telaah materi (98%) serta “baik” pada dimensi telaah konstruksi (79,33%) dan bahasa (80%). Adapun kualitas butir soal pada tahun 2019 memiliki kualitas yang “sangat baik” terhadap dimensi telaah materi (97,33%) dan bahasa (85%) serta “baik” terhadap dimensi telaah konstruksi (80%).⁴⁶

Terkait penelitian terdahulu yang dilakukan oleh Fauziah Umami memiliki perbedaan dalam penelitian ini yaitu metode yang digunakan kuantitatif dengan variabel yang diteliti yaitu berkaitan dengan validitas,

⁴⁵ Nelly Agustin, “Analisis Soal Ujian Akhir Semester Ganjil Mata Pelajaran Pendidikan Agama Islam Ditinjau Dari Tingkat Kesukaran Dan Daya Pembeda Pada Kelas VIII Di SMP Islam Terpadu Bina Insani Metro” (2019).

⁴⁶ Fauziah Umami, “Analisis Butir Soal Tes Pilihan Ganda Ujian Nasional Tahun Pelajaran 2018-2019 Mata Pelajaran Bahasa Indonesia Jenjang Sltp (Sekolah Lanjutan Tingkat Pertama)” (2019).

reliabilitas, tingkat kesukaran, dan daya beda soal Ujian Akhir Semester Gasal Pendidikan Agama Islam, pada kelas XII di SMA Negeri 1 Kota Mungkid Tahun Ajaran 2022/2023.

C. Kerangka Berpikir

Kerangka berfikir adalah alur dalam penelitian berkaitan dengan hubungan, pengaruh antar variabel yang diuji, Dalam pembahasan yang sudah dipaparkan diatas bahwa perlunya ada evaluasi terhadap proses belajar siswa pada mata pelajaran Pendidikan Agama Islam, oleh karena itu dalam proses evaluasi pendidikan terdapat beberapa teknik evaluasi, Secara garis besar, maka alat evaluasi yang digunakan dapat digolongkan menjadi dua macam, yaitu tes dan nontes.⁴⁷

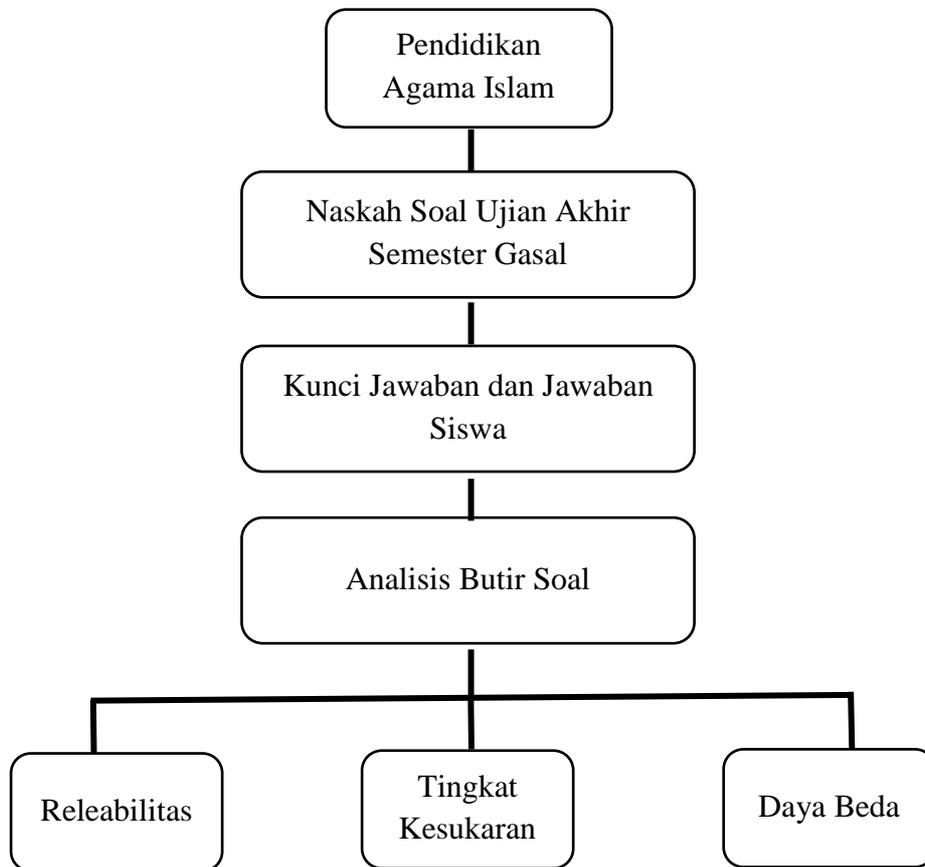
Oleh karena itu perlunya ada analisis butir soal sebagai bagian dari proses evaluasi untuk mengetahui pencapaian peserta didik dan juga kualitas soal, sebagaimana dinyatakan oleh Arikunto bahwa, jika tes tersebut valid, reliabel, objektif, praktis, dan hemat biaya, maka tes tersebut memenuhi syarat sebagai alat ukur yang baik.⁴⁸

Dalam penelitian ini fokus kajian ataupun penelitian yaitu berkaitan dengan analisis uji validitas, reliabilitas, tingkat kesukaran butir soal dan daya beda Soal Ujian Akhir Semester Gasal Pendidikan Agama Islam, pada kelas XII di SMA Negeri 1 Kota Mungkid Tahun Ajaran 2022/2023.

⁴⁷ Farida and Musyarofah, "Validitas Dan Reliabilitas Dalam Analisis Butir Soal."

⁴⁸ Fitrianawati, "Peran Analisis Butir Soal Guna Meningkatkan Kualitas Butir Soal, Kompetensi Guru Dan Hasil Belajar Peserta Didik."

Upaya peneliti agar penelitian ini mudah dipahami, maka peneliti menuangkan kerangka berfikir, kerangka berfikir pada penelitian ini adalah:



Gambar 1. Skema Kerangka Berpikir

BAB III

METODE PENELITIAN

A. Jenis dan Pendekatan Penelitian

Sesuai dengan masalah yang ingin dikaji maka penelitian ini merupakan jenis penelitian analisis isi yang mana penelitian ini dilakukan bukan kepada orang, tetapi lebih kepada instrumen evaluasi.⁴⁹ Penelitian ini menggunakan metode pendekatan deskriptif kuantitatif metode penelitian yang berdasarkan pada filsafat positivisme, penggunaan data menggunakan instrumen penelitian analisis data yang bersifat kuantitatif/statistik.⁵⁰ Dengan mengkaji butir soal Ujian Akhir Semester Gasal Pendidikan Agama Islam pada kelas XII di SMA Negeri 1 Kota Mungkid Tahun Ajaran 2022/2023, berkaitan dengan validitas soal, reliabilitas soal, tingkat kesukaran dan daya pembeda, format penelaahan soal bentuk pilihan ganda dan uraian menggunakan bantuan aplikasi komputer yaitu IBM SPSS Statistics versi 20, Microsoft Office Excel dan ZYA CBT.

B. Definisi Operasional

Pada penelitian ini yang menggunakan pendekatan kuantitatif dengan Variabel dalam penelitian ini adalah analisis validitas butir soal, tingkat kesukaran dan daya pembeda. Analisis butir soal merupakan suatu serangkaian kegiatan pengumpulan, peringkasan, dan penggunaan informasi dari jawaban siswa untuk membuat keputusan tentang setiap penilaian yang dilakukan untuk

⁴⁹ Priyono, *Metode Penelitian Kuantitatif*, ed. Teddy Chandra, Revisi 200. (Sidoarjo: ZIFATAMA PUBLISHING, 2008).

⁵⁰ Sugiono, *Metode Penelitian Kuantitatif Kualitatif Dan R&D*.

mengetahui mutu soal yang telah ditulis. Dalam analisis butir soal mencakup beberapa aspek yaitu sebagai berikut:

1. Validitas

Validitas adalah sejauh mana suatu tes mengukur seberapa baik (tes atau nontes) valid dan reliabel. Efektif berarti membutuhkan efektivitas internal dan eksternal.

Validitas internal dapat dilakukan dengan dua cara yaitu analisis faktor dan analisis butir.

- a. Analisis faktor.

Analisis faktor dilakukan dengan mengkorelasikan skor faktor dengan skor total. Metode korelasi yang digunakan adalah metode korelasi product-moment. Korelasi positif yang signifikan berarti item untuk faktor tersebut dianggap valid.

- b. Analisis butir dengan mengkorelasikan skor.⁵¹

Validitas eksternal dilakukan dengan cara mengkorelasikan skor hasil uji coba instrumen yang dibuat guru dengan instrumen yang sudah baku. Menguji signifikansi angka korelasi dengan mencari angka korelasi antara skor-skor tes buatan dengan skor-skor tes toulou. Teknik korelasi yang digunakan adalah teknik korelasi *Product Moment*, jika angka korelasi yang diperoleh ternyata signifikan, berarti tes yang dibuat guru dapat dianggap valid.⁵²

⁵¹ Asrul, Ananda, and Rosinta, *Evaluasi Pembelajaran*.

⁵² Ibid.

1) Pengujian Validitas Tes Berbentuk Objektif Test

Tes berbentuk objektif seperti pilihan ganda, benar-salah, dan menjodohkan merupakan tes dengan skor dikotomi peringkat 0 dan 1. Pengecekan validitas dilakukan untuk korelasi positif dan sesuai dengan uji korelasi *product-moment*. Dalam metode tersebut, perbedaan yang signifikan antara skor item dan skor total berarti item tersebut dianggap valid.⁵³ Besarnya korelasi Arikunto memberikan kriteria penafsiran sebagai berikut: Antara 0,800 sampai dengan 1,00 = sangat tinggi, antara 0,600 sampai dengan 0,800 = tinggi, antara 0,400 sampai dengan 0,600 = cukup, antara 0,200 sampai dengan 0,400 = rendah, antara 0,00 sampai dengan 0,200 = sangat rendah.⁵⁴

2. Reliabilitas

Reliabilitas berasal dari kata *reliability* berarti sejauh mana hasil suatu pengukuran dapat dipercaya. Suatu hasil pengukuran dapat dipercaya apabila dalam beberapa kali pelaksanaan pengukuran terhadap kelompok subyek yang sama, diperoleh hasil pengukuran yang relatif sama, selama aspek yang diukur dalam diri subyek memang belum berubah.⁵⁵

Menurut Allen dan Yen dalam jurnal *Reliabilitas Instrumen Pendidikan* bahwa koefisien reliabilitas besarnya ditentukan oleh

⁵³ Ibid.

⁵⁴ Arikunto, *Dasar-Dasar Evaluasi Pendidikan*.

⁵⁵ Helpiana, "Analisis Butir Soal Ujian Akhir Semester Genap Tahun Ajaran 2018/2019 Mata Pelajaran PAI Kelas XI SMAN 2 Gunung Meriah Kabupaten Aceh Singkil."

satu dikurangi dengan perbandingan varians kesalahan pengukuran dan varians skor tampak. Atas dasar demikian semakin kecil varians kesalahan pengukuran, maka akan semakin tinggi koefisien reliabilitas, sehingga varians tampak atas dasar demikian semakin kecil varians kesalahan pengukuran, maka akan semakin tinggi koefisien reliabilitas, sehingga varians tampak atas dasar demikian semakin kecil varians kesalahan pengukuran, maka akan semakin tinggi koefisien reliabilitas, sehingga varians tampak dapat digunakan sebagai atribut koefisien reliabilitas yang sesungguhnya.⁵⁶

Dalam menentukan reliabilitas, menggunakan rumus alpha dengan rentang skor 0 - 10, 0 - 100 atau berbentuk skala 1 - 3, 1 - 5 atau 1 - 10. sebelum dicari angka reliabilitasnya, perlu terlebih dahulu dicari varians butir dan varians skor total . Interpretasi terhadap hasil perhitungan koefisien reliabilitas tes (r_{11}) pada umumnya digunakan patokan sebagai berikut. Jika angka 0.800-1.000 tingkat realibilitas sangat tinggi, Jika angka 0.600-0.800 tingkat realibilitasnya tinggi, Jika angka 0.400-0.600 tingkat realibilitasnya cukup, Jika angka 0.200-0.400 tingkat realibilitasnya rendah, Jika angka 0.000-0.200 tingkat realibilitasnya sangat rendah.⁵⁷ Hasil perhitungan menggunakan

⁵⁶ Ibid.

⁵⁷ Muluki, "Analisis Kulaitas Butir TES Semester Ganjil Mata Pelajaran Ilmu Pengetahuan Alam Kelas IV MI Radhiatul Adawiyah."

rumus tersebut diinterpretasikan dengan tingkat koefisien korelasi sebagai berikut:

Penelitian reliabilitas ini akan menggunakan rumus $K - R$ untuk soal bentuk pilihan ganda, dan rumus Alpha untuk soal bentuk uraian. Untuk interpretasi reliabilitas instrumen adalah sebagai berikut: $0,80 - 1,00 =$ tinggi $0,60 - 0,80 =$ cukup $0,40 - 0,60 =$ agak rendah $0,20 - 0,40 =$ rendah $0,00 - 0,20 =$ sangat rendah, Pada penelitian ini instrumen dikatakan reliabel jika tingkat koefisien korelasinya $0,600$ atau lebih.

3. Tingkat Kesukaran

Soal yang baik adalah soal yang tidak terlalu mudah atau tidak terlalu sukar.⁵⁸ soal yang terlalu mudah akan membuat siswa menyepelkan soal tersebut dan tidak merangsang siswa untuk menyelesaikan problematika soal tersebut, di samping itu soal yang terlalu sulit akan membuat siswa tidak nyaman dan putus asa untuk memecahkan problematika dalam soal tersebut karena jauh dari jangkauan pemahamannya, sehingga soal yang tidak terlalu mudah dan tidak terlalu sukar akan menjadikan evaluasi dan pengukuran tingkat ketercapaian siswa, serta tujuan pembelajaran dapat tercapai dengan baik.

Indeks Kesukaran Angka yang menunjukkan tingkat kesulitan dan kemudahan suatu soal. Indeks kesukarannya antara $0,00$ dan disebut

⁵⁸ Asrul, Ananda, and Rosinta, *Evaluasi Pembelajaran*.

indeks kesukaran. tingkat kesulitan ke 1.0. Indeks kesukaran ini menunjukkan tingkat kesukaran soal. Antara 0,00 dan 1,0. Indeks kesukaran ini menunjukkan tingkat kesukaran soal. Tampilkan masalah dengan indeks kesulitan 0,0

Soal dengan indeks kesukaran 0,0 menunjukkan bahwa tugas tersebut terlalu sulit. Menunjukkan bahwa tugasnya terlalu sulit. Jika tidak, indeks 1,0 menunjukkan bahwa tugasnya terlalu mudah. Untuk penilaian indeks kesukaran ini diberi lambang P (P kapital), dan untuk penilaian indeks kesukaran ini diberi lambang P (P singkatan dari big). Jadi masalahnya adalah pada $P = 0,20$. Sebaliknya, soal $P = 0,30$ lebih sulit daripada soal $P = 0,80$.⁵⁹

Soal dengan P 0,00 sampai 0,30 merupakan soal sukar, soal dengan P 0,30 sampai 0,70 merupakan soal sedang, dan soal dengan P 0,70 sampai 1,00 merupakan soal mudah. Meskipun demikian, mereka berpendapat bahwa soal yang dianggap baik, atau cukup sulit, adalah soal dengan indeks peringkat 0,30 hingga 0,70..⁶⁰

4. Daya Beda

Kekhasan suatu soal adalah kemampuannya membedakan siswa yang pandai (berkemampuan tinggi) dan siswa yang kurang pandai (berkemampuan rendah). Angka yang menunjukkan besarnya perbedaan karakteristik disebut indeks diskriminasi, disingkat D. Mirip dengan indeks empati, indeks identifikasi (identitas) berkisar

⁵⁹ Ibid.

⁶⁰ Ibid.

antara 0,00 hingga 1,00. tanda. Tanda negatif indeks kewaspadaan digunakan ketika ada sesuatu yang "terbalik" dan menunjukkan kualitas penguji. Jadi anak cerdas itu bodoh dan anak bodoh itu cerdas..⁶¹

Untuk kelompok kecil, keseluruhan kelompok penguji dibagi menjadi dua bagian yang sama besar, 50% untuk kelompok atas dan 50% untuk kelompok bawah. Untuk grup besar, dua kutub, atau 27% teratas, biasanya diambil sebagai grup teratas, mengingat biaya dan waktu analisis. (JA) dan 27% terbawah adalah kelompok terbawah (JB).

C. Populasi dan Sampel

Populasi adalah jumlah keseluruhan obyek/subjek yang akan di analisis dan disinyalir dapat menunjang keberlangsungan proses penelitian karena memiliki karakteristik dan kualitas serta sesuai dengan tujuan penelitian yang dilaksanakan oleh peneliti yang akan dicari analisis serta hasil kesimpulan dari penelitian yang dilaksanakan terhadap suatu generasi atau obyek/subyek secara menyeluruh atau sebagian.

Sebagaimana dinyatakan oleh Nawawi populasi adalah keseluruhan subjek yang terdiri dari orang, benda, makhluk, tumbuhan, efek samping atau kejadian yang terjadi sebagai sumber. populasi juga merupakan subjek ujian secara keseluruhan ⁶²Populasi dicirikan sebagai "semua individu dari suatu

⁶¹ Ibid.

⁶² Tukiran Taniredja and Hidayati Mustafidah, *Penelitian Kuantitatif (Sebuah Pengantar)* (Bandung: ALFABETA, 2011).

pertemuan" peristiwa atau artikel yang telah ditandai dengan jelas. Atau sebaliknya perkumpulan yang lebih besar yang menjadi tujuan spekulasi⁶³

Sedangkan sampel dapat diuraikan sebagai bagian atau delegasi dari populasi yang diteliti, Arikunto menyatakan bahwa uji eksplorasi diambil secara tidak lengkap dari keseluruhan artikel yang berada di bawah konsentrasi yang dipandang sebagai gambaran seluruh penduduk dan diambil dengan menggunakan metode-metode tertentu. Tes juga berarti bagian dari populasi, atau kumpulan kecil yang diperhatikan.⁶⁴

Sehingga dari populasi dan sampel yang telah didapat oleh peneliti ketika populasi yang ada telah melebihi kuota sampel yang sesuai dengan karakteristik dan juga batasan-batasan peneliti yang telah di buat ataupun keterbatasan tenaga, waktu, dana dsb. Maka peneliti mengambil sampel tidak dari keseluruhan populasi yang ada.

Data responden menggunakan sampling jenuh berjumlah 205 siswa, meliputi siswa SMA Negeri 1 Kota Mungkid kelas XII MIPA (1, 2, 3) dan IPS(2, 3, 4).Daftar jumlah siswa yang merupakan sampel dalam penelitian ini sebagai berikut:

⁶³ Asmaul Husna and Budi Suryana, *Metodologi Penelitian Dan Statistik*, vol. 1, 2017.

⁶⁴ Taniredja and Mustafidah, *Penelitian Kuantitatif (Sebuah Pengantar)*.

Tabel 3. Daftar siswa populasi penelitian

No	Kelas	Jumlah siswa
1.	Kelas XII MIPA 1	36
2.	Kelas XII MIPA 2	33
3.	Kelas XII MIPA 3	36
4.	Kelas XII IPS 2	30
5.	Kelas XII IPS 3	35
6.	Kelas XII IPS 4	35
7.	Jumlah	205

D. Jenis dan Sumber Data

Penelitian yang dilakukan tidak terlepas dari data objek sebagai penunjang informasi yang lebih spesifik untuk menunjang penelitian dalam menyusun setiap data dan fakta yang didapatkan agar informasi yang didapatkan dapat diolah dan dijadikan hasil dalam suatu keperluan penelitian.

Terdapat beberapa hal yang dapat mempengaruhi sebuah penelitian yaitu kualitas instrumen data, berkaitan dengan cara menghitung validitas, reliabilitas instrumen dan juga instrumen pengumpulan data. Dalam penelitian ini jenis pendekatan yang dilakukan yaitu jenis pendekatan kuantitatif dengan sumber data yang didapatkan dari dua sumber informasi, yaitu informasi esensial dan informasi opsional. Informasi esensial adalah informasi yang dikumpulkan langsung oleh analis dan biasanya dikumpulkan dengan menggunakan teknik studi, persepsi, atau dokumentasi. Sedangkan informasi opsional adalah informasi yang tidak dikumpulkan oleh peneliti secara

langsung tetapi diambil dari catatan cetak atau elektronik yang berbeda.⁶⁵

Dalam penelitian ini sumber informasi esensial yaitu responden sedangkan sumber informasi opsional yaitu instrumen evaluasi pada mata pelajaran Pendidikan Agama Islam di SMA Negeri 1 Kota Mungkid..

E. Teknik Pengumpulan Data

Pengumpulan data adalah cara peneliti dalam melaksanakan penelitian berupa pengumpulan data. Pengumpulan data yang ditempuh peneliti merupakan langkah pertama dalam mencari dan mendapatkan data karena tujuan awal penelitian adalah mendapatkan data dan menguji data tersebut. Setiap penelitian yang dilaksanakan perlu adanya pemilihan metode yang tepat karena berkaitan dengan hasil penelitian yang akan dihasilkan berupa uji hipotesis, menjawab problematika, ataupun menguji hubungan atau pengaruh pada variabel-variabel yang diuji hubungan, pengaruh.⁶⁶

Penelitian ini menggunakan metode dokumentasi dengan telaah data dokumentasi yang didapat langsung dari data primer guru mata pelajaran dan panitia pelaksanaan ujian dengan menelusuri arsip-arsip data, teknik dokumentasi ini digunakan untuk mendapatkan dokumen dan arsip berupa kisi-kisi soal, soal, kunci jawaban, hasil skor jawaban siswa pada soal Ujian Akhir Semester Gasal Mata Pelajaran Pendidikan Agama Islam pada Kelas XII untuk menjadi bahan analisis data untuk mengetahui isi dan kualitas soal berkaitan dengan uji validitas, reliabilitas, tingkat kesukaran butir soal dan daya beda yang di analisis sesuai dengan kriteria yang sudah ditentukan.

⁶⁵ Asmaul Husna and Suryana, *Metodologi Penelitian Dan Statistik*, vol. 1, p. .

⁶⁶ Syahrudin and Salim, *Metodologi Penelitian Kuantitatif*, ed. Rusdi Ananda (Bandung: Citapustaka Media, 2017).

F. Teknik Analisis Data

1. Pengujian Validitas Tes Berbentuk Objektif Test

Tes berbentuk objektif seperti pilihan ganda, benar-salah, dan menjodohkan merupakan tes dengan skor dikotomi peringkat 0 dan 1. Pengecekan validitas dilakukan untuk korelasi positif dan sesuai dengan uji korelasi *product-moment* IBM SPP versi 20. Dalam metode tersebut, perbedaan yang signifikan antara skor item dan skor total berarti item tersebut dianggap valid.⁶⁷ Besarnya korelasi Arikunto memberikan kriteria penafsiran sebagai berikut: Antara 0,800 sampai dengan 1,00 = sangat tinggi, antara 0,600 sampai dengan 0,800 = tinggi, antara 0,400 sampai dengan 0,600 = cukup, antara 0,200 sampai dengan 0,400 = rendah, antara 0,00 sampai dengan 0,200 = sangat rendah.⁶⁸

2. Uji Reliabilitas Instrument

Reliabilitas menyiratkan ketergantungan instrumen estimasi dengan asumsi instrumen estimasi digunakan untuk menyelesaikan estimasi rehashed.⁶⁹

Dalam reliabilitas terdapat data yang reliabel yang mana digunakan untuk menguji responden yang sama di waktu yang berbeda dengan beberapa uji responden dan dihasilkan hasil yang sama. Uji reliabilitas instrumen dengan teknik belah 2.

Dengan melibatkan instrumen yang valid dan reliabel dalam pengumpulan informasi, diyakini pemeriksaan ini akan substansial dan

⁶⁷ Asrul, Ananda, and Rosinta, *Evaluasi Pembajalaran*.

⁶⁸ Arikunto, *Dasar-Dasar Evaluasi Pendidikan*.

⁶⁹ Asmaul Husna and Suryana, *Metodologi Penelitian Dan Statistik*, vol. 1, p. .

dapat diandalkan. Ketergantungan mengacu pada tingkat dari sesuatu instrument. Dengan cara ini, instrument substansial secara positif merupakan prasyarat untuk mendapatkan hasil yang valid dan reliabel.

Sedangkan reliabilitas instrumen dalam penelitian ini menggunakan rumus koefisien alpha sebagai berikut:

$$\sigma^2 = \frac{\sum X^2 - \frac{(X^2)}{N}}{N}$$

$$r_{11} = \left[\frac{k}{k-1} \right] \left[1 - \frac{\sum S_b^2}{S_t^2} \right]$$

Keterangan:

r_{11} = Reliabilitas instrumen

k = Jumlah butir soal

$\sum S_b^2$ = Jumlah varians butir

S_t^2 = Varians total

Interpretasi terhadap hasil perhitungan koefisien reliabilitas tes (r_{11}) pada umumnya digunakan patokan sebagai berikut. Jika angka 0.800-1.000 tingkat realibilitas sangat tinggi, Jika angka 0.600-0.800 tingkat realibilitasnya tinggi, Jika angka 0.400-0.600 tingkat realibilitasnya cukup, Jika angka 0.200-0.400 tingkat realibilitasnya rendah, Jika angka 0.000-0.200 tingkat realibilitasnya sangat rendah.⁷⁰

Hasil perhitungan menggunakan rumus meliputi KR-20, *Alpha Cronbach*, *Spearman Brown* dengan dua software yaitu Microsoft Office

⁷⁰ Muluki, "Analisis Kulaitas Butir TES Semester Ganjil Mata Pelajaran Ilmu Pengetahuan Alam Kelas IV MI Radhiatul Adawiyah."

Excel 2010 dan software IBM SPSS 20, diinterpretasikan dengan tingkat koefisien korelasi sebagai berikut:

Interpretasi terhadap hasil perhitungan koefisien reliabilitas tes (r_{11}) pada umumnya digunakan patokan sebagai berikut. Jika angka 0.800-1.000 tingkat realibilitas sangat tinggi, Jika angka 0.600-0.800 tingkat realibilitasnya tinggi, Jika angka 0.400-0.600 tingkat realibilitasnya cukup, Jika angka 0.200-0.400 tingkat realibilitasnya rendah, Jika angka 0.000-0.200 tingkat realibilitasnya sangat rendah.⁷¹ Hasil perhitungan menggunakan rumus tersebut diinterpretasikan dengan tingkat koefisien korelasi sebagai berikut:

Penelitian reliabilitas ini akan menggunakan rumus K – R. untuk soal bentuk pilihan ganda, dan rumus Alpha untuk soal bentuk uraian. Untuk interpretasi reliabilitas instrumen adalah sebagai berikut: 0,80 – 1,00 = tinggi 0,60 – 0,80 = cukup 0,40 – 0,60 = agak rendah 0,20 – 0,40 = rendah 0,00 – 0,20 = sangat rendah, Pada penelitian ini instrumen dikatakan reliabel jika tingkat koefisien korelasinya 0,600 atau lebih.

Pada penelitian ini instrumen dikatakan reliabel jika tingkat koefisien korelasinya 0,600 atau lebih. Kedua uji tersebut dilakukan dengan bantuan program *Microsoft Office Excel* 2016 dan SPSS v20.⁷²

⁷¹ Ibid.

⁷² Aini Nur Ma'rifah, "Hubungan Kecerdasan Intelektual dan Spiritual Dengan Hasil Belajar Aqidah Ahlak Peserta Didik Ma Al-Mubarak Kecamatan Bandar Mataram Kabupaten Lampung Tengah," 2018.

3. Tingkat Kesukaran

Soal yang baik adalah soal yang tidak terlalu mudah atau tidak terlalu sukar.⁷³ soal yang terlalu mudah akan membuat siswa menyepelekan soal tersebut dan tidak merangsang siswa untuk menyelesaikan problematika soal tersebut, di samping itu soal yang terlalu sulit akan membuat siswa tidak nyaman dan putus asa untuk memecahkan problematika dalam soal tersebut karena jauh dari jangkauan pemahamannya, sehingga soal yang tidak terlalu mudah dan tidak terlalu sukar akan menjadikan evaluasi dan pengukuran tingkat ketercapaian siswa, serta tujuan pembelajaran dapat tercapai dengan baik.

Indeks Kesukaran Angka yang menunjukkan tingkat kesulitan dan kemudahan suatu soal. Indeks kesukarannya antara 0,00 dan disebut indeks kesukaran. tingkat kesulitan ke 1.0. Indeks kesukaran ini menunjukkan tingkat kesukaran soal. Antara 0,00 dan 1,0. Indeks kesukaran ini menunjukkan tingkat kesukaran soal. Tampilkan masalah dengan indeks kesulitan 0,0

Soal dengan indeks kesukaran 0,0 menunjukkan bahwa tugas tersebut terlalu sulit. Menunjukkan bahwa tugasnya terlalu sulit. Jika tidak, indeks 1,0 menunjukkan bahwa tugasnya terlalu mudah. Untuk penilaian indeks kesukaran ini diberi lambang P (P kapital), dan untuk penilaian indeks kesukaran ini diberi lambang P (P singkatan dari big). Jadi masalahnya

⁷³ Asrul, Ananda, and Rosinta, *Evaluasi Pembelajaran*.

adalah pada $P = 0,20$. Sebaliknya, soal $P = 0,30$ lebih sulit daripada soal $P = 0,80$.⁷⁴

Adapun rumus mencari P adalah

$$P = \frac{B}{JS}$$

Keterangan:

P = Indeks kesukaran.

B = Banyaknya siswa yang menjawab soal itu dengan betul

JS = Jumlah seluruh siswa peserta tes

Hasil perhitungan menggunakan rumus (P) dengan dua software yaitu Microsoft Office Excel 2010 dan software IBM SPSS 20. Soal dengan P 0,00 sampai 0,30 merupakan soal sukar, soal dengan P 0,30 sampai 0,70 merupakan soal sedang, dan soal dengan P 0,70 sampai 1,00 merupakan soal mudah. Meskipun demikian, mereka berpendapat bahwa soal yang dianggap baik, atau cukup sulit, adalah soal dengan indeks peringkat 0,30 hingga 0,70.⁷⁵

4. Daya Beda

Kekhasan suatu soal adalah kemampuannya membedakan siswa yang pandai (berkemampuan tinggi) dan siswa yang kurang pandai (berkemampuan rendah). Angka yang menunjukkan besarnya perbedaan karakteristik disebut indeks diskriminasi, disingkat D . Mirip dengan indeks empati, indeks identifikasi (identitas) berkisar antara 0,00 hingga

⁷⁴ Ibid.

⁷⁵ Ibid.

1,00. tanda. Tanda negatif indeks kewaspadaan digunakan ketika ada sesuatu yang "terbalik" dan menunjukkan kualitas penguji. Jadi anak cerdas itu bodoh dan anak bodoh itu cerdas..⁷⁶

Untuk kelompok kecil, keseluruhan kelompok penguji dibagi menjadi dua bagian yang sama besar, 50% untuk kelompok atas dan 50% untuk kelompok bawah. Untuk grup besar, dua kutub, atau 27% teratas, biasanya diambil sebagai grup teratas, mengingat biaya dan waktu analisis. (JA) dan 27% terbawah adalah kelompok terbawah (JB).

Rumus untuk menentukan indeks Daya beda adalah :

$$D = \frac{B_A}{J_A} - \frac{B_B}{J_B} = PA - PB$$

Keterangan:

J = Jumlah peserta tes

JA = Banyaknya peserta kelompok atas JB : Banyaknya peserta kelompok bawah

BA = Banyaknya peserta kelompok atas yang menjawab soal itu dengan benar.

BB = Banyaknya peserta kelompok bawah yang menjawab soal itu

Tabel 5. Interpretasi Indeks Daya Beda Item

Besarnya Angka Indeks Diskriminasi Item (J)	Interpretasi
Kurang dari 0,20	Lemah
0,20 – 0,40	Cukup
0,41 – 0,70	Baik
0,71 – 1,00	Sangat Baik

⁷⁶ Ibid.

BAB V

PENUTUP

A. Simpulan

Hasil penelitian dan pembahasan analisis butir soal ujian akhir semester gasal pendidikan Agama Islam pada kelas XII di SMA Negeri 1 Kota Mungkid tahun ajaran 2022/2023, yang terdiri dari validitas soal, reliabilitas soal, tingkat kesukaran soal dan daya beda soal dapat disimpulkan sebagai berikut:

1. Dari hasil analisis validitas soal menggunakan rumus *Product Moment* software IBM SPSS versi 20, bahwa analisis menghasilkan 38 soal valid atau sekitar 84 % butir soal valid dengan kriteria validitas sangat rendah berjumlah 6, validitas rendah 21 soal, validitas cukup 11 soal, dan 7 butir soal dinyatakan tidak valid atau sekitar 16%.

Sehingga Validitas soal ujian akhir semester gasal Pendidikan Agama Islam kelas XII di SMA Negeri 1 Kota Mungkid dinyatakan Valid dengan rumus *Product Moment* software IBM SPSS versi 20.

2. Dari hasil analisis tingkat koefisien reliabilitas dengan rumus KR-20 menunjukkan nilai $r(11)$ yaitu 0,740, sedangkan pada rumus *Alpha Cronbach*, indek reliabilitasnya adalah 0,722, dan pada rumus *Spearman Brown*, indek reliabilitasnya adalah 0,759.

Sehingga Reliabilitas soal ujian akhir semester gasal Pendidikan Agama Islam kelas XII di SMA Negeri 1 Kota Mungkid dinyatakan Reliabel sangat baik sesuai dengan ketiga rumus tersebut.

3. Tingkat Kesukaran dan Daya Beda

a. Tingkat Kesukaran

Hasil analisis menggunakan rumus (P) menunjukkan bahwa dari 38 soal terdapat 29 butir soal (72%) dalam kategori mudah, 7 butir soal (18 %) dalam kategori sedang, dan 3 butir soal (10%).

Hasil analisis dengan menggunakan software IBM SPSS versi 20 menunjukkan seluruh rumus menyatakan terdapat 29 butir soal (72%) dalam kategori mudah. 3 butir soal (10%) dalam kategori sukar, 7 butir soal (18%) dalam kategori sedang.

Sehingga Tingkat Kesukaran soal ujian akhir semester gasal Pendidikan Agama Islam kelas XII di SMA Negeri 1 Kota Mungkid dapat disimpulkan bahwa, butir-butir soal Pendidikan Agama Islam yang berbentuk *multiple choice* memiliki kualitas soal yang kurang baik karena belum memiliki keseimbangan.

b. Daya Beda

Hasil analisis menggunakan rumus D(Daya Beda) menunjukkan bahwa terdapat soal dalam kategori lemah yaitu 5 butir soal atau 13%, soal dalam kategori baik yaitu 9 butir soal atau 24%, soal dalam kategori cukup yaitu 24 butir soal atau 63%.

Berdasarkan hasil analisis dengan menggunakan software IBM SPSS versi 20 menunjukkan bahwa terdapat soal dalam kategori lemah yaitu 13 butir soal atau 24%, soal dalam kategori cukup yaitu

22 butir soal atau 56%, dan soal dalam kategori baik yaitu 3 butir soal atau 8%.

Sehingga Daya Beda soal ujian akhir semester gasal Pendidikan Agama Islam kelas XII di SMA Negeri 1 Kota Mungkid dapat disimpulkan bahwa, butir-butir soal Pendidikan Agama Islam yang berbentuk *multiple choice* memiliki kualitas soal yang cukup.

B. Saran

Berdasarkan hasil, pembahasan, dan simpulan penelitian ini maka saran yang hendak diajukan peneliti sebagai berikut:

1. Bagi peneliti yaitu menginformasikan kepada guru pembuat soal terkait hasil penelitian yang telah dilakukan. Jika kelak menjadi seorang guru dapat melakukan analisis terlebih dahulu sebelum mengujikan butir soal kepada peserta didik.
 - a. Dapat dijadikan acuan dalam penelitian yang akan mendatang, dan menambah khazanah keilmuan untuk meningkatkan pendidikan di Indonesia sekaligus menjadi bahan masukan untuk meningkatkan kualitas soal evaluasi di mata pelajaran Pendidikan Agama Islam dengan merumuskan soal evaluasi dengan baik
 - b. Perlu adanya penelitian lanjutan dan juga sudut pandang yang berbeda berkaitan dengan instrument soal evaluasi ujian akhir semester pada mata pelajaran Pendidikan Agama Islam.

2. Bagi Guru

- a. Soal yang berkualitas baik pilihan ganda, dapat disimpan di bank soal untuk digunakan kembali.
- b. Soal yang berkualitas kurang baik sebaiknya dilakukan revisi agar dapat digunakan kembali.
- c. Soal yang berkualitas tidak baik seharusnya dibuang karena membutuhkan revisi yang signifikan.
- d. Guru perlu memperhatikan indikator kualitas soal, agar soal yang dibuat lebih berkualitas. Guru juga perlu melakukan uji coba terhadap soal sebelum digunakan pada ujian sesungguhnya
- e. Hasil penelitian dapat dijadikan sebagai bahan evaluasi dan menambah wawasan guru dan karyawan dalam peningkatan soal evaluasi

3. Bagi Sekolah

Sekolah perlu meningkatkan keterampilan guru dalam melakukan evaluasi pembelajaran yaitu dengan memberikan pengetahuan dan kemampuan kepada guru mengenai analisis butir soal, sehingga guru dapat membuat instrumen tes yang berkualitas. Soal yang berkualitas tidak baik seharusnya dibuang karena membutuhkan revisi yang signifikan dan dapat digunakan untuk bahan pertimbangan dan peningkatan akademik SMA Negeri 1 Kota Mungkid pada jangka waktu yang pendek maupun jangka waktu yang panjang.

DAFTAR PUSTAKA

- Agustin, Nelly. "Analisis Soal Ujian Akhir Semester Ganjil Mata Pelajaran Pendidikan Agama Islam Ditinjau Dari Tingkat Kesukaran Dan Daya Pembeda Pada Kelas VIII Di SMP Islam Terpadu Bina Insani Metro" (2019).
- Ahmad Darlis. "Hakikat Pendidikan Islam: Telaah Antara Hubungan Pendidikan Informal, Non Formal Dan Formal" XXIV, no. 1 (1992): 174.
- Ananda, Rusydi, and Tien Rafida. *Pengantar Evaluasi Program Pendidikan*. Edited by Candra Wijaya. *Perdana Publishing*. 1st ed. Vol. 53. Medan: Perdana Publishing, 2017.
- Arikunto, Suharsimi. *Dasar-Dasar Evaluasi Pendidikan*. Revisi. Jakarta: Bumi Aksara, 2010.
- Asmaul Husna, and Budi Suryana. *Metodologi Penelitian Dan Statistik*. Vol. 1, 2017.
- Asrul, Rusydi Ananda, and Rosinta. *Evaluasi Pembelajaran*. *Ciptapustaka Media*. 2nd ed. Medan: Ciptapustaka Media, 2015.
- Farida, and Anna Musyarofah. "Validitas Dan Reliabilitas Dalam Analisis Butir Soal." *Al-Mu'arrib: Jurnal Pendidikan Bahasa Arab* 1, no. 1 (2021): 34–44.
- Al Fathiyah, Siti Fathimah. "Analisis Butir Soal Pelajaran Bahasa Arab Di MA Roudlotul Ulum Pagak Malang." *Tarbiyatuna* 4, no. 1 (2019): 78–100.
- Fitrianawati, Meita. "Peran Analisis Butir Soal Guna Meningkatkan Kualitas Butir Soal, Kompetensi Guru Dan Hasil Belajar Peserta Didik." *JPT: Jurnal Pendidikan Tematik*, 2017.
- Frimayanti, Ade Imelda. "Implementasi Pendidikan Nilai Dalam Pendidikan Agama Islam." *Al-Tadzkiyyah: Jurnal Pendidikan Islam* 8, no. 11 (2017): Hal. 240.
- Hasanah, Nurul, Lailatur Badriyah, and Wena Diarmard Selia. "Evaluasi Pembelajaran Ditinjau Dari Al-Qur'an Surah Al-Ankabut Ayat 2-3." *JPT (Jurnal Pendidikan Tematik)* 1, no. 2 (2020): 18–19.
- Helpiana, Nina. "Analisis Butir Soal Ujian Akhir Semester Genap Tahun Ajaran 2018/2019 Mata Pelajaran PAI Kelas XI SMAN 2 Gunung Meriah Kabupaten Aceh Singkil" (2020).

- Kemenag RI. *Peraturan Menteri Agama Nomor 13 Tentang Pendidikan Keagamaan Islam*, 2014.
- Ma'rifah, Aini Nur. "Hubungan Kecerdasan Intelektual dan Spiritual Dengan Hasil Belajar Aqidah Ahlak Peserta Didik Ma Al-Mubarak Kecamatan Bandar Mataram Kabupaten Lampung Tengah," 2018.
- Mua'arifah, Isnani. "Analisis Kualitas Butir Soal Pilihan Ganda Mata Pelajaran Biologi Kelas X Dan XI Pada Ulangan Akhir Semester Tahun Ajaran 2016/2017 Di MAN Kota Palangka Raya" (2017): 1–14.
- Muluki, Ardilah. "Analisis Kualitas Butir TES Semester Ganjil Mata Pelajaran Ilmu Pengetahuan Alam Kelas IV MI Radhiatul Adawiyah." Universitas Muhammadiyah Makassar, 2020.
- Ni'amah, Khoirotul. "Paradigma Pendidikan Islam Perspektif Al-Ghazali." *Heutagogia: Journal of Islamic Education* 1, no. 1 (2021): 59–71.
- Nurjanah, and Noni Marlianingsih. "Analisis Butir Soal Pilihan Ganda Dari Aspek Kebahasaan." *Faktor Jurnal Ilmu Kependidikan* II, no. 1 (2015): 69–78.
- Priyono. *Metode Penelitian Kuantitatif*. Edited by Teddy Chandra. Revisi 200. Sidoarjo: ZIFATAMA PUBLISHING, 2008.
- Rena Citra. "Kompetensi Guru Pendidikan Agama Islam Dalam Meningkatkan Motivasi Belajar Siswa (Studi Deskriptif Di Smpn 1 Singingi Hilir)." *Jurnal AL-HIKMAH* 51, no. 1 (2018): 173–183.
- Riadi, Akhmad. "Problematika Sistem Evaluasi Pembelajaran" 15, no. 27 (2017): 1–12.
- Sugiono. *Metode Penelitian Kuantitatif Kualitatif Dan R&D*. 19th ed. Bandung: ALFABETA, 2013.
- Supriadi, Gito. *Pengantar Dan Teknik Evaluasi Pembelajaran*. 1st ed. Malang: Intimedia, 2011.
- Susanto, Hery, Achi Rinaldi, and Novalia. "Analisis Validitas Reabilitas Tingkat Kesukaran Dan Daya Beda Pada Butir Soal Ujian Akhir Semester Ganjil Mata Pelajaran Matematika." *Al-Jabar: Jurnal Pendidikan Matematika* 6, no. 2 (2015): 203–217.
- Syahrudin, and Salim. *Metodologi Penelitian Kuantitatif*. Edited by Rusdi Ananda. Bandung: Citapustaka Media, 2017.

- Taniredja, Tukiran, and Hidayati Mustafidah. *Penelitian Kuantitatif (Sebuah Pengantar)*. Bandung: ALFABETA, 2011.
- Tanwir. “Dasar-Dasar Dan Ruang Lingkup Evaluasi Pendidikan.” *Jurnal Pendidikan Islam* 13, no. 1 (2015): 56–57.
- Umami, Fauziah. “Analisis Butir Soal Tes Pilihan Ganda Ujian Nasional Tahun Pelajaran 2018-2019 Mata Pelajaran Bahasa Indonesia Jenjang Sltip (Sekolah Lanjutan Tingkat Pertama)” (2019).
- Wahndianti, Santi, and Sumarsih. “Analisis Butir Soal Ujian Akhir Semester Gasal Mata Pelajaran Hitung Dagang” 7, no. 1 (2018): 1–14.

