

BAB I

PENDAHULUAN

A. Latar Belakang

Di dalam dunia pendidikan, evaluasi memegang peran yang sangat penting, karena secara umum evaluasi mempunyai tujuan yaitu : (1) Untuk memperoleh data pembuktian, yang akan menjadi petunjuk sampai dimana tingkat kemampuan dan tingkat keberhasilan peserta didik dalam mencapai tujuan-tujuan kurikuler, setelah mereka menempuh proses pembelajaran dalam jangka waktu yang telah ditentukan, (2) Untuk mengukur dan menilai sampai manakah efektifitas mengajar dan metode-metode mengajar yang telah diterapkan atau dilaksanakan oleh pendidik, serta kegiatan belajar yang dilaksanakan oleh peserta didik (Anas Sudjiono, 2005: 16).

Selain itu, evaluasi dalam bidang pendidikan juga mempunyai tujuan khusus yaitu : (1) Untuk merangsang kegiatan peserta didik dalam menempuh program pendidikan. Tanpa adanya evaluasi maka tidak mungkin timbul kegairahan atau rangsangan pada peserta didik untuk memperbaiki dan meningkatkan prestasinya masing-masing, (2) Untuk mencari dan menemukan faktor-faktor penyebab keberhasilan dan ketidakberhasilan peserta didik dalam mengikuti program pendidikan, sehingga dapat dicari dan ditemukan jalan keluar atau cara perbaikannya (Anas Sudjiono, 2005: 17).

Proses belajar mengajar prestasi belajar siswa sebagai tolak ukur keberhasilan pembelajaran dapat diukur melalui penilaian. Alat yang sering digunakan untuk penilaian tersebut lazim disebut dengan soal. Alat yang baik dapat mengukur apa yang seharusnya diukur, dengan kata lain soal yang baik adalah soal yang dapat mengukur apa yang seharusnya diukur, sesuai dengan keadaan yang sebenarnya. Selain itu, soal yang baik juga dapat untuk membedakan antara siswa yang pandai dengan siswa yang belum pandai, memiliki taraf kesukaran yang berdistribusi normal serta mengungkapkan konsep-konsep yang telah diajarkan kepada siswa secara profesional sesuai dengan tujuan pembelajaran yang hendak dicapai.

Sebaliknya soal yang tidak baik adalah soal yang tidak dapat mengukur dengan tepat apa yang seharusnya diukur sesuai dengan keadaan yang sebenarnya. Selain itu, soal yang tidak baik sulit untuk membedakan siswa yang Pandai dengan siswa yang belum pandai, memiliki taraf kesukaran yang tidak berdistribusi normal, dalam arti soal tersebut terlalu sulit atau terlalu mudah untuk dikerjakan siswa, serta tidak mengungkapkan konsep yang telah diajarkan kepada siswa secara profesional sesuai dengan tujuan pembelajaran yang hendak dicapai.

Alat untuk mengukur keberhasilan siswa berupa soal, harus memiliki karakteristik-karakteristik tertentu untuk menjamin hasilnya. Adapun sebuah tes dapat dikatakan baik sebagai alat pengukur harus

memenuhi persyaratan tes, yaitu memiliki : validitas, Reliabilitas, dan objektifitas, praktibilitas, dan ekonomis (Suharsini Arikunto, 1996,56).

Keberhasilan belajar yang dapat dicapai siswa dapat diukur dengan mengerjakan soal-soal ulangan harian, ulangan umum, maupun ujian akhir. Selama ini soal untuk ulangan umum dan soal ujian akhir atau ulangan akhir semester hanya memiliki validitas secara logis, yang dilakukan dengan pembuatan kisi-kisi soal yang digunakan untuk memberikan informasi tentang keterukuran tujuan pembelajaran. Padahal soal-soal ulangan umum dan soal ujian akhir sangat terkait dengan validitas, daya beda, taraf kesukaran butir soal, serta pengungkapan proporsi konsep-konsep pelajaran tertentu sesuai dengan tujuan pembelajaran.

Pada sistem pendidikan, ulangan akhir sekolah (UAS) dijadikan salah satu bentuk pengukuran dan penilaian hasil belajar pada akhir pendidikan dasar dan menengah, hasil pengukuran tersebut selain untuk kepentingan siswa dalam sertifikasi, juga berguna sebagai bahan masukan dalam pengambilan keputusan berbagai kebijakan, serta sebagai fungsi pemantau mutu pendidikan disekolah.

Hasil pengukuran juga diharapkan dapat memberikan gambaran atau informasi yang akurat tentang tingkat penguasaan siswa terhadap suatu materi pelajaran serta daya serap siswa agar dapat diperoleh informasi yang akurat. Ulangan akhir sekolah sebagai tes prestasi

belajar dituntut memenuhi segala persyaratan bagi sebuah alat ukur yang baik.

Informasi hasil pengukuran sangat berguna dalam pengambilan keputusan serta dapat digunakan sebagai salah satu dasar untuk mengadakan penilaian dalam dalam rangka menentukan kebijakan. Mutu informasi yang didapatkan dari hasil pengetesan merupakan hal yang sangat penting. Mutu informasi tersebut ditentukan oleh mutu tes, dan mutu tes ditentukan mutu setiap butirnya yang dikemas dalam sebuah perangkat tes.

Begitu pentingnya ulangan akhir semester sebagai pengukuran penilaian hasil belajar dalam pendidikan, maka masalah yang berkaitan dengan perangkat tes UAS perlu mendapatkan perhatian yang serius.

Soal Ulangan Akhir Semester kelas 6 (enam) di SD Negeri se-Kecamatan Karangmojo khususnya soal Pendidikan Agama Islam belum pernah dianalisis, sehingga soal tersebut tidak dapat memberi gambaran apakah soal tersebut terlalu mudah atau terlalu sulit yang mengakibatkan soal tersebut tidak dapat untuk mengukur siswa yang pandai dengan siswa yang belum pandai. Selain itu juga kadang tidak terlalu memperhatikan validitas, daya beda, taraf kesukaran butir soal, serta proporsi konsep-konsep pelajaran sesuai dengan tujuan pembelajaran.

Berdasarkan kenyataan di atas maka sangat diperlukan dilakukannya penelitian terhadap butir-butir soal Pendidikan Agama

Islam pada kelas 6 (enam) di SD Negeri se-Kecamatan Karangmojo dikarenakan belum ada penelitian yang dilakukan oleh guru atau pihak lain berkenaan dengan soal UAS mata pelajaran Pendidikan Agama Islam tersebut.

Disamping itu, ada beberapa alasan mengapa soal UAS mata pelajaran Pendidikan agama Islam SD kelas 6 harus dianalisis butir soalnya, antara lain yaitu : (1) untuk mengetahui kekuatan dan kelemahan butir soal UAS, sehingga dapat menemukan butir soal yang baik, yang harus direvisi dan butir soal yang harus dibuang; (2) Untuk menyediakan spesifikasi butir soal UAS secara lengkap; (3) Untuk mengetahui masalah yang terkandung dalam butir soal seperti : kemenduan soal, kesalahan meletakkan jawaban, soal yang terlalu mudah atau terlalu sukar, atau soal yang tidak dapat membedakan peserta UAS yang mempersiapkan diri dengan baik dan yang tidak mempersiapkan diri saat menghadapi UAS.

Pengujian butir soal tersebut dapat dilakukan dengan dua cara yaitu kualitatif dan kuantitatif, analisis kualitatif bisa dilakukan dengan menelaah dari segi materi, konstruksi butir soal, dan aspek kognitif tiap butir soal, sedangkan analisis kuantitatif meliputi : Reliabilitas, tingkat kesukaran, daya beda, dan distribusi jawaban, dapat dilakukan dengan menggunakan program ITEMAN.

B. Rumusan Masalah

Dari latar belakang, tujuan dan kegunaan penelitian di atas, dapat dirumuskan permasalahannya sebagai berikut :

1. Bagaimanakah validitas soal mata pelajaran Pendidikan Agama Islam kelas 6 di SD Negeri se-Kecamatan Karangmojo ?
2. Bagaimanakah reliabilitas soal tersebut ?
3. Bagaimanakah daya beda soal tersebut ?
4. Bagaimanakah taraf kesukaran soal tersebut ?
5. Bagaimanakah pengecoh (*distraktor*) butir soal tersebut ?
6. Bagaimanakah kualitas soal UAS tersebut ?

C. Tujuan dan Kegunaan Penelitian

1. Tujuan penelitian ini masih sangat sederhana, sesuai dengan permasalahannya maka penelitian ini bertujuan untuk mengetahui soal UAS pendidikan Agama Islam kelas enam di SD Negeri se-Kecamatan Karangmojo Gunungkidul. Adapun tujuannya adalah sebagai berikut :
 - a. Untuk mengetahui apakah butir soal mata pelajaran Pendidikan Agama Islam di SD Negeri se-Kecamatan Karangmojo Mempunyai butir soal yang valid dan handal atau tidak.
 - b. Untuk mengetahui apakah butir soal mata pelajaran Pendidikan Agama Islam di SD Negeri se-Kecamatan Karangmojo Reliabilitasnya tinggi atau rendah.

- c. Untuk mengetahui apakah butir soal mata pelajaran Pendidikan Agama Islam di SD Negeri se-Kecamatan Karangmojo mempunyai daya beda yang tinggi atau rendah.
 - d. Untuk mengetahui apakah butir soal mata pelajaran Pendidikan Agama Islam di SD Negeri se-Kecamatan Karangmojo mempunyai taraf kesukarannya berdistribusi normal atau tidak.
 - e. Untuk mengetahui apakah butir soal mata pelajaran Pendidikan Agama Islam di SD Negeri se-Kecamatan Karangmojo mempunyai pengecoh yang berfungsi atau tidak.
 - f. Untuk mengetahui kualitas soal UAS secara keseluruhan sudah baik atau belum.
2. Kegunaan penelitian

Secara umum hasil penelitian ini diharapkan bisa memberi masukan terhadap kualitas butir soal Ulangan Akhir Semester I (UAS) di SD Negeri se-Kecamatan Karangmojo sebagai salah satu alat untuk mengevaluasi hasil belajar siswa. Secara operasional yang diharapkan dari hasil penelitian ini adalah :

- a. Bagi Guru, yaitu :
 - a) Dapat memperkenalkan cara menganalisis butir soal dengan menggunakan komputer khususnya program ITEMAN.
 - b) Dapat melakukan analisis hasil tes untuk kebaikan pengajaran berikutnya.

- c) Dapat membuat soal yang berkualitas, yaitu soal yang dapat mengukur apa yang seharusnya diukur sesuai dengan keadaan sebenarnya.
- b. Bagi Sekolah, hasil penelitian ini diharapkan dapat memberikan masukan mengenai kualitas butir soal tes UAS.
- c. Bagi Dinas Pendidikan, Pemuda, dan Olahraga Kabupaten Gunungkidul, dapat memberikan informasi yang sangat berharga dalam pembuatan kebijakan baru, khususnya pembuatan soal-soal mata pelajaran Pendidikan Agama Islam untuk UAS.
- d. Memberikan kontribusi yang positif pada ilmu pendidikan khususnya analisis butir soal.
- e. Dapat dijadikan kajian bagi peneliti-peneliti lain lebih lanjut dalam bidang analisis butir soal.

D. Tinjauan Pustaka

Nurdin (2006:13) dalam penelitian Kualitas butir Soal Tes Seleksi Penerimaan Mahasiswa Baru Program D2 PGSD di Universitas Khairun Tahun 2005, menyimpulkan bahwa dari keseluruhan soal tes yang berjumlah 60 soal, jumlah butir soal yang diterima 36 butir (60%), 1 butir soal direvisi (1,7%) dan 23 soal yang ditolak (38,3%)

Sudjani (1997:87) dalam penelitian tentang karakteristik internal perangkat soal Fisika pada EBTANAS SMU di Propinsi Jawa Barat

menemukan bahwa 50% butir soal baik, 15% butir soal perlu direvisi dan 35% soal jelek atau ditolak.

Fachrur Rozikin (2010:196-220) dalam penelitiannya tentang analisis butir soal evaluasi pembelajaran pendidikan agama Islam (alQur'an Hadist) pada kelas XII IPA dan IPS di MAN I Bojonegoro, dalam penelitian tersebut dari 50 soal yang dianalisis ditemukan 52% soal dalam kategori mudah, 16% dalam kategori sedang dan 34% dalam kategori sukar. Dan dari segi kualitas butir soal dari 50 soal yang dianalisis ditemukan 14 soal ditolak/jelek, 24 soal baik dan 12 soal direvisi.

Lilis Tri Ariyana (2011:44) dalam penelitiannya tentang analisis butir soal Ulangan Akhir Semester Gasal mata pelajaran Ilmu Pendidikan Alam (IPA) kelas IX di SMP se-Kabupaten Grobogan tahun pelajaran 2010/2011, menyimpulkan bahwa soal IPA SMP Kelas IX di Kabupaten Grobogan (1) memiliki validitas logis, karena sudah sesuai dengan soal standar, tetapi perlu perbaikan aspek konstruksi pada beberapa soal.(2) reliabel dengan kategori tinggi. (3) memiliki tingkat kesukaran sedang.(4) Grobogan memiliki daya beda baik.(5) memiliki efektifitas pengecoh berfungsi.(6) memiliki kualitas sesuai standar.

Dedi Wahyudi (2011: 96-100) dalam penelitiannya Analisis kualitas butir soal PAI dalam pencapaian kompetensi siswa di SMA Negeri 2 Kebumen menemukan bahwa Soal ulangan akhir semester genap mata pelajaran pendidikan agama islam kelas X SMA Negeri 2

kebumen tahun pelajaran 2010-2011 mempunyai validitas jelek karena dari 50 soal hanya ada 21 soal yang valid, tingkat kesukarannya terlalu mudah dan mempunyai daya beda yang kurang baik. Akan tetapi soal tersebut memiliki Reliabilitas yang tinggi dan dilihat dari pencapaian kompetensinya telah mampu mengukur semua kompetensi yang harus dicapai siswa pada semester genap. Dan dilihat dari kualitas soalnya soal tersebut masuk dalam kategori jelek karena hanya ada 5 butir soal yang baik dan bisa dimasukkan ke bank soal.

Purwana (2005:97-100) dalam penelitiannya Analisis Butir Soal Ujian Akhir Sekolah Pada Mata Pelajaran ISMUBA di SMA Muhammadiyah 3 Yogyakarta Tahun Pelajaran 2003/2004 dan 2004/2005 menyimpulkan bahwa soal ISMUBA tersebut tidak memenuhi validitas logis, terdapat kesalahan pada kontruksi butir soal, distribusi aspek kognitif belum ideal, distribusi tingkat kesukaran butir soal belum memenuhi distribusi kesukaran yang ideal, rata-rata butir soal dalam kategotri baik, distribusi pengecoh dalam kategori baik, kualitas soal belum memenuhi kualitas soal yang baik.

Dari hasil penelitian yang dilakukan oleh Nurdin dan Sudjani, dan Fachrur rozikin, Lilis Tri Ariyana dan Hedi wahyudi dapat di informasikan bahwasanya semua yang mereka teliti adalah mengenai analisis butir soal. Dan dari kelima peneliti di atas diperoleh hasil atau kesimpulan bahwa masih banyak soal yang belum memenuhi kategori

soal yang baik, atau bisa dikatakan masih banyak sekali soal yang perlu direvisi maupun soal yang tertolak atau soal yang jelek.

Sebagaimana yang kita ketahui bahwasanya kualitas butir soal yang kurang baik ataupun jelek akan memberikan pengaruh pada hasil ujian peserta, sehingga perlu sekali dilakukan penelitian sejenis secara terus menerus dan berkesinambungan agar kualitas butir soal dari waktu ke waktu akan semakin baik.

Soal-soal ulangan atau ujian baik SD, SMP, dan SMA harus selalu dilakukan penelitian terhadap semua mata pelajaran termasuk soal pendidikan agama Islam di SD. Karena soal-soal ulangan maupun soal-soal ujian sekolah maupun nasional sangat penting, sebagaimana yang telah dipaparkan di atas bahwa soal-soal tersebut menjadi sebuah alat ukur siswa mana yang sudah menguasai materi dan yang belum menguasai materi pelajaran.

E. Landasan Teori

1. Pengukuran dan Penilaian

Pengukuran kuantitatif pada hakikatnya adalah membandingkan sesuatu dengan atau atas dasar ukuran tertentu. Misalnya, mengukur suhu badan dengan ukuran berupa Thermometer: hasilnya: 36°Celsius, 38°Celsius, 39°Celciu:, dan seterusnya (Anas Sudjiono, 2005: 4)

Pengukuran yang bersifat kuantitatif itu, dapat dibedakan menjadi tiga macam, yaitu : (1) Pengukuran yang dilakukan bukan untuk menguji sesuatu: misalnya; pengukuran yang dilakukan oleh penjahit pakaian mengenai panjang lengan, panjang kaki, lebar bahu, ukuran pinggang dan sebagainya. (2) Pengukuran yang dilakukan untuk menguji sesuatu; misalnya: pengukuran untuk menguji daya tahan per baja terhadap tekanan berat, pengukuran untuk menguji daya tahan nyala lampu pijar, dan sebagainya. (3) Pengukuran untuk menilai, yang dilakukan dengan jalan menguji sesuatu; misalnya: mengukur kemajuan belajar peserta didik dalam rangka mengisi nilai rapor yang dilakukan dengan menguji mereka dalam bentuk tes hasil belajar. Pengukuran jenis ketiga inilah yang biasa dikenal dalam dunia pendidikan (Anas Sudjiono, 2005:4)

Penilaian adalah usaha mengunpulkan berbagai informasi secara kesinambunngan dan menyeluruh, tentang proses dan hasil belajar mengajar yang telah dicapai oleh siswa melalaui proses belajar mengajar yang dilakukan (Depdikbud 1991: 3).

Penilaian merupakan komponen penting dalam sistem pendidikan karena mencerminkan perkembangan dan kemajuan hasil pendidikan dari satu waktu ke waktu lain. Selain itu penilaian merupakan intepretasi dari hasil pengukuran dengan jalan membandingkan dengan suatu patokan atau kriteria. Penilaian dalam pendidikan dilakukan terhadap konsep dan proses yang

dilakukan. Penilaian proses merupakan penilaian terhadap kegiatan dan kemajuan siswa pada saat berlangsungnya kegiatan pembelajaran di kelas yang dilakukan selama pembelajaran. Kegiatan ini memberikan umpan balik bagi guru untuk menentukan apakah pembelajaran dapat dilanjutkan atau justru diulang kembali, atau bagi siswa perlu bantuan guru atau bantuan tambahan pelajaran (Depdikbud 1994: 9).

Asas pelaksanaan dalam sebuah penilaian terdiri dari tiga pokok yaitu objektif, menyeluruh dan berkesinambungan (Depdikbud 1991:3). Dikatakan objektif apabila penilaian yang digunakan dapat menggambarkan yang sesungguhnya. Penilaian dikatakan menyeluruh apabila mencakup proses maupun hasil belajar serta menggambarkan perubahan tingkah laku yang mencakup ranah kognitif, afektif dan psikomotorik. Dan suatu penilaian dikatakan berkesinambungan apabila pelaksanaan penilaian dilakukan secara terus menerus, terencana dan bertahap.

Penyelenggaraan evaluasi atau penilaian hasil belajar siswa merupakan salah satu mendapatkan umpan balik bagi guru tentang sejauh mana tujuan intruksional pengajaran telah dicapai, sehingga guru dengan demikian mengetahui apakah guru masih harus memperbaiki langkah yang harus ditempuh dalam penyelenggaraan keegiatan pengajaran dikelas (B. Suryobroto, 2005: 143)

a. Jenis Penilaian

Dalam buku tatalaksana kurikulum, terdapat beberapa macam evaluasi di sekolah yaitu test formatif, test sub-sumatif, test sumatif, dan evaluasi belajar tahap akhir / EBTA (B. Suryobroto, 2005: 143).

- 1) Test Formatif, ialah evaluasi atau usaha penilaian hasil belajar yang berupa test (soal-soal pertanyaan) yang diberikan kepada siswa setelah satu pokok bahasan selesai dipelajari.
- 2) Test Sub-Sumatif, ialah test yang diberikan kepada siswa dengan bahan atau materi meliputi beberapa pokok bahasan yang sejenis. Test ini sering disebut pula sebagai test unit untuk mengungkap hasil belajar siswa terhadap satu unit bahan pelajaran. Biasanya apabila guru merencanakan akan mengadakan test sub-sumatif maka test formatifnya tidak diselenggarakan.
- 3) Test Sumatif, ialah evaluasi atau usaha penilaian hasil belajar yang berupa test (soal-soal pertanyaan) yang dilakukan setelah kegiatan belajar mengajar berlangsung dalam satuan waktu.
- 4) Evaluasi Belajar Tahap akhir (EBTA), Evaluasi belajar ini merupakan usaha penilaian yang terakhir dilakukan untuk mengungkap hasil belajar siswa secara keseluruhan selama

siswa belajar di suatu sekolah. EBTA oleh masyarakat umum dikenal sebagai ujian akhir.

b. Penyusunan Butir Soal

Jika diinginkan hasil penilaian yang baik, maka harus diperhatikan unsur-unsur penting yang ada dalam penilaian, diantaranya adalah penyusunan, alat penilaian, prosedur penilaian, dan suasana penilaian. Alat penilaian yang baik harus memenuhi syarat-syarat, yaitu : valid, reliabel, mempunyai tingkat kesukaran yang memadai, dan mempunyai daya beda yang baik, serta memiliki pengecoh yang berfungsi.

Bentuk soal yang digunakan dalam penialain berdasarkan pada tujuan yang akan dicapai. Penyusunan soal merupakan penyusunan butir soal karena butir soal adalah yang membangun suatu soal. Biasanya digunakan bentuk soal objektif (pilihan ganda) dan uraian dalam berbagai penilaian.

Ada beberapa prinsip dasar yang harus diperhatikan dalam penyusunan soal tes hasil belajar agar tes tersebut dapat mengukur tujuan intruksional khusus untuk mata pelajaran yang telah diajarkan selama proses belajar mengajar yaitu, (Anas Sudjiono, 2005: 97-99) :

- 1) Tes hasil belajar harus dapat mengukur secara jelas hasil belajar yang telah ditetapkan sesuai dengan tujuan intruksional.

- 2) Butir-butir tes hasil belajar harus merupakan sample yang representatif dari populasi bahan pelajaran yang telah diajarkan, sehingga dapat mewakili dianggap mewakili seluruh performance yang telah diperoleh selama peserta didik mengikuti suatu unit pelajaran.
- 3) Bentuk soal yang dikeluarkan dalam tes hasil belajar harus bervariasi , sehingga cocok untuk mengukur hasil belajar yang sesuai dengan tujuan tes itu sendiri.
- 4) Tes hasil belajar harus didesain sesuai dengan kegunaan untuk memperoleh hasil yang diinginkan.
- 5) Tes hasil belajar harus memiliki Reliabilitas yang yang dapat diandalkan.
- 6) Tes hasil belajar disamping harus dapat mengukur keberhasilan belajar siswa, juga harus bisa dijadikan alat untuk memperoleh informasi yang berguna untuk memperbaiki cara belajar siswa dan cara mengajar guru.

Dilihat dari bentuk soalnya, tes hasil belajar ada dua macam yaitu test subjectif dan test objektif (Anas Sudjiono, 2005: 99)

a) Butir Soal Subjektif (Uraian)

Tes uraian adalah suatu tes yang menghendaki siswa untuk menjawab pertanyaan soal-soal tes yang menghendaki jawaban panjang dengan cara menguraikan atau menerangkan atau menceritakan hal-hal lainnya. Tes uraian ini mempunyai

ciri khas dimulai dengan kata perintah seperti : uraikan, terangkan, mengapa, beri alasan, bandingkan, apa bedanya, apa yang kamu ketahui tentang dsb. (B. Suryosubroto: 145)

Ada beberapa petunjuk operasional dalam penyusunan tes uraian yang perlu diperhatikan, yaitu (Anas Sudjiono: 104-105):

- 1) Butir-butir soal dapat mencakup ide-ide pokok dari materi pelajaran yang sudah diajarkan.
- 2) Untuk menghindari kecurangan maka susunan kalimat pertanyaan hendaknya dibuat berlainan dengan dengan kalimat yang ada dalam buku pelajaran atau sumber lainnya.
- 3) Butir soal dirumuskan dan disusun secara tegas bagaimana dan seperti apa jawaban benar yang dikehendaki tester.
- 4) Pertanyaan-pertanyaan butir soal dibuat bervariasi.
- 5) Kalimatnya disusun secara ringkas, padat dan jelas.
- 6) Dibuat petunjuk yang jelas cara mengerjakan soal.

b) Butir Soal Objektif

Tes objektif juga bisa dikatakan adalah salah satu jenis penilaian hasil belajar yang terdiri dari butir-butir soal (items) yang dapat dijawab oleh peserta tes dengan jalan memilih

salah satu atau lebih diantara beberapa kemungkinan jawaban yang telah dipasangkan oleh masing-masing item, atau dengan mengisikan jawaban berupa simbol atau kata-kata tertentu pada tempat yang telah disediakan untuk masing-masing soal (item) yang bersangkutan untuk diisi (Anas Sudjiono, 2005: 106).

Sebagai salah satu jenis tes hasil belajar, tes objektif dibedakan menjadi lima golongan, yaitu (Anas Sudjiono, 2005: 107-118).

1) Tes Objektif Berbentuk Benar-Salah (*True-False Test*)

Adalah suatu tes di mana butir-butir soal yang diajukan dalam tes berupa pernyataan yang mempunyai dua kemungkinan jawaban yaitu benar atau salah, dan dalam tes tersebut sudah dibubuhi tanda (simbol) tertentu. Untuk menjawab soal tersebut cukup mencoret atau melingkari simbol huruf B (benar) dan S (salah) atas pernyataan yang diberikan dalam soal.

2) Tes Objektif Berbentuk Menjodohkan (*Matching Test*)

Adalah merupakan salah satu tes yang didalamnya disediakan dua kelompok bahan yaitu satu kelompok seri pertanyaan dan satu kelompok seri jawaban, dalam tes ini

peserta tes harus mencari pasang-pasangan yang sesuai antara kelompok pertanyaan dan kelompok jawaban yang dibeikan dalam tes tersebut.

3) Tes Objektif Berbentuk Isian (*Fill in Test*)

Adalah salah satu bentuk tes yang biasanya berbentuk cerita atau karangan. Kata-kata penting yang ada dalam cerita atau karangan tersebut dikosongkan oleh pembuat soal dan tugas peserta tes adalah mengisi kata-kata yang telah dikosongkan dalam soal tersebut.

4) Tes Objektif Berbentuk Melengkapi (*Completion Test*)

Tes Completion Test ini hampir mirip dengan Fill in, akan tetapi terdapat perbedaan yaitu pada Fill in bahan yang diteskan merupakan satu kesatuan cerita, sedangkan completion tidak harus demikian dengan kata lain butir soal dalam test ini dibuat berbeda atau berlainan antara yang satu dengan yang lain.

5) Tes Objektif Berbentuk Pilihan Ganda (*Multiple Choice Item Test*)

Adalah salah satu tes objektif yang didalamnya terdiri atas pertanyaan atau pernyataan yang belum selesai, untuk menyelesaikannya peserta tes harus memilih satu dari beberapa option jawaban yang telah disediakan pada setiap butir soal yang bersangkutan.

Oleh sementara pendidik, soal pilihan ganda dianggap paling bermanfaat dan paling luwes di antara semua jenis tes karena dapat digunakan untuk menguji sebagian terbesar mata pelajaran. Bentuk soal dari soal pilihan ganda adalah suatu stem. Stem dapat berupa pertanyaan atau pernyataan. Stem diikuti alternatif-alternatif atau pilihan-pilihan jawaban.

Ada beberapa petunjuk operasional dalam penyusunan tes objektif yang perlu diperhatikan, yaitu (Anas Sudjiono: 136-138):

- 1) Dalam pembuatan tes yang bermutu tinggi seorang guru atau dosen harus selalu berlatih dalam merancang dan menyusun test objektif secara berkesinambungan agar soal yang dibuat semakin baik.
- 2) Soal tes yang sudah dipergunakan, hendaknya dilakukan analisis item untuk mengidentifikasi mana soal yang sudah baik, kurang baik dan tidak baik.
- 3) Untuk mencegah kerjasama yang tidak sehat dan spekulasi tebakan maka dipersiapkan sanksi berupa pengurangan skor untuk setiap jawaban salah dengan harapan peserta tes akan mengerjakan dengan jujur dan hati-hati.
- 4) Butir soal yang akan diteskan sebaiknya dibuatkan kisi-kisi materi pelajaran yang telah diajarkan.

- 5) Dalam penyusunan kalimat menggunakan bahasa atau istilah-istilahnya hendaknya cukup sederhana, ringkas, jelas dan mudah dipahami oleh peserta tes.
- 6) Soal yang diteskan disusun harus satu makna, jangan sampai terjadi makna ganda yang akan mengakibatkan kerancuan dalam memberi jawaban pada soal.
- 7) Cara memenggal dan memutus kalimat, pembubuhan tanda baca, penulisan tanda Aljabar hendaknya ditulis secara benar agar tidak mengganggu konsentrasi peserta tes.
- 8) Diberi pedoman atau petunjuk pengerjaan soal tes tersebut agar peserta tes dapat bekerja sesuai dengan petunjuk yang telah ditentukan.

c. Taksonomi Tujuan Pendidikan

Taksonomi adalah ilmu tentang klasifikasi secara umum, dan juga klasifikasi spesifik mengenai suatu hal dengan aturan aturan tertentu, contohnya taksonomi tujuan pengajaran. Taksonomi tujuan pengajaran membantu guru untuk menetapkan arah dan tujuan yang akan dicapai dalam kegiatan pembelajaran. Taksonomi tujuan pengajaran disusun menurut prinsip struktur tertentu yang makin lama makin sukar. Hal ini dapat dipahami karena psikologi, mengingat adalah proses berpikir yang lebih sederhana daripada memahami.

Bloom, seperti dikutip oleh Anas Sudjiono dalam bukunya Pengantar Evaluasi Pendidikan, membagi tujuan pendidikan atau pengajaran menjadi tiga aspek yaitu :

- 1) Ranah Proses Berpikir (*Cognitife Domain*), adalah ranah yang mencakup kegiatan mental (otak). Meliputi Pengetahuan (*Knowledge*), Pemahaman (*Comprehension*), Penerapan (*Applikation*), Analisis (*Analysis*), Sintesis (*Synthesis*), Penilaian (*Evaluation*) (Anas Sudjiono, 2005: 49-40).
- 2) Ranah Nilai atau Sikap (*Afektif Domain*), adalah ranah yang berkaitan dengan sikap dan nilai. Meliputi penerimaan atau memperhatikan (*Receiving* atau *Attending*), menanggapi (*Responding*), menilai atau menghargai (*Valuing*), mengatur atau mengorganisasikan (*Organization*), karakterakterisasi dengan nilai kompleks atau kompleks nilai (*Characterization by a Value Complex*) (Anas Sudjiono, 2005: 54).
- 3) Ranah Ketrampilan (*Psykomotor Domain*), adalah ranah yang berkaitan dengan ketrampilan (*skill*), atau kemampuan orang bertindak setelah memperoleh pengalaman belajar tertentu. Hasil belajar Psikomotorik sebenarnya kelanjutan dari

hasil belajar kognitif (memahami sesuatu) dan hasil belajar Afektif (yang baru tampak dalam bentuk kecenderungan-kecenderungan untuk berperilaku. Hasil belajar Kognitif dan afektif akan menjadi hasil belajar Psikomotor apabila peserta didik telah menunjukkan perilaku atau perbuatan tertentu sesuai dengan makna yang terkandung dalam ranah Kognitif dan Afektifnya (Anas Sudjiono, 2005: 49-40).

Dalam konteks evaluasi hasil belajar, maka ketiga ranah atau domain tersebut yang harus dijadikan sasaran dalam setiap kegiatan evaluasi hasil belajar yaitu : (1) apakah peserta didik sudah dapat memahami semua bahan atau materi pelajaran yang telah diberikan kepada mereka? (2) Apakah semua peserta didik sudah bisa menghayatinya? (3) Apakah semua materi pelajaran yang telah diberikan itu sudah dapat diamalkan secara kongkrit dalam praktek atau dalam kehidupan sehari-hari? (Anas Sudjiono, 2005: 49)

d. Ranah Kognitif

Dalam penelitian ini digunakan klasifikasi dari Bloom, karena klasifikasi tersebut sangat rinci, dan banyak dipakai sebagai acuan dalam dunia pendidikan di Indonesia.

Ranah kognitif adalah ranah yang mencakup kegiatan mental (otak). Menurut Bloom, segala upaya yang menyangkut aktifitas otak adalah termasuk dalam ranah kognitif. Dalam ranah kognitif itu terdapat enam jenjang proses berpikir, mulai dari jenjang yang terendah sampai dengan jenjang yang paling tinggi, yaitu (Anas Sudjiono, 2005: 50) :

Pengetahuan (*Knowledge*) adalah kemampuan seseorang untuk mengingat kembali (*Recall*) atau mengenali kembali tentang nama, istilah, ide, gejala, rumus-rumus, dan sebagainya, tanpa mengharap kemampuan untuk menggunakannya. Pengetahuan atau ingatan ini adalah merupakan proses berpikir yang paling rendah (Anas Sudjiono, 2005: 50).

Pemahaman (*Comprehension*) adalah kemampuan seseorang untuk mengerti atau memahami sesuatu setelah sesuatu itu ia ketahui dan ingat. Dengan kata lain, memahami adalah mengetahui tentang sesuatu dan dapat

melihat dari berbagai segi. Pemahaman merupakan jenjang kemampuan berpikir yang setingkat lebih tinggi dari ingatan atau hafalan (Anas Sudjiono, 2005: 50)

Aplikasi (*Application*) adalah kesanggupan seseorang untuk menerapkan atau menggunakan ide-ide umum, tata cara atau metode-metode, prinsip-prinsip, rumus-rumus, teori-teori dan sebagainya, dalam situasi yang baru dan kongkret. Aplikasi atau penerapan ini adalah merupakan proses berpikir setingkat lebih tinggi ketimbang pemahaman (Anas Sudjiono, 2005: 51).

Analisis (*Analysis*) adalah kemampuan seseorang untuk merinci atau menguraikan suatu bahan atau keadaan menurut bagian-bagian yang lebih kecil dan mampu memahami hubungan diantara bagian-bagian atau faktor-faktor yang satu dengan faktor-faktor lainnya. Jenjang analisis adalah setingkat lebih tinggi ketimbang jenjang aplikasi (Anas Sudjiono, 2005: 51).

Sintesis (*Synthesis*) adalah kemampuan berfikir yang merupakan kebalikan dari proses berpikir analisis. Sintesis merupakan suatu proses yang memadukan bagian-bagian atau unsur-unsur secara logis, sehingga menjelma menjadi suatu pola yang berstruktur atau berbentuk pola baru. Jenjang sintesis

kedudukannya setingkat lebih tinggi ketimbang jenjang analisis (Anas Sudjiono, 2005: 51).

Evaluasi (*Evaluation*) adalah merupakan jenjang berfikir paling tinggi dalam ranah kognitif menurut Taksonomi Bloom. Soal termasuk dalam ranah evaluasi jika dalam pertanyaan tersebut terkandung maksud agar siswa mampu membuat pendugaan atau penilaian terhadap materi, pekerjaan, fenomena, atau tingkah laku berdasarkan suatu kriteria tertentu (Anas Sudjiono, 2005: 51)

2. Analisis Butir Soal

Analisis adalah suatu kemampuan untuk merinci atau menguraikan suatu bahan atau keadaan menurut bagian-bagian yang lebih kecil atau faktor-faktor penyebabnya dan mampu memahami hubungan diantara bagian atau faktor yang satu dengan bagian atau faktor yang lainnya (Depdikbud, 1991: 16). Analisis butir soal bertujuan untuk mengetahui validitas isi yang berupa kebenaran konsep, tingkat kesukaran, daya beda, fungsi distraktor, Reliabilitas serta kualitas soalnya yang dilakukan dengan cara memperbaiki, menyeleksi, mengganti dan merevisi.

Penelaahan soal secara kuantitatif maksudnya adalah penelaahan butir soal didasarkan pada data empirik dari butir soal yang bersangkutan. Data empirik ini diperoleh dari soal yang telah

diujikan (Depdiknas, 2008: 11). Analisis butir soal secara klasik adalah proses penelaahan butir soal melalui informasi dari jawaban peserta didik guna meningkatkan mutu butir soal yang bersangkutan dengan menggunakan teori tes klasik (Depdiknas, 2008: 11).

Analisis butir soal teori tes klasik dilakukan dengan bantuan program komputer Item and Test Analysis (ITEMAN). Analisis ini akan menghasilkan butir soal dan skala statistik perangkat tes. Statistik butir soal meliputi ; tingkat kesukaran butir, daya pembeda dan efektifitas distraktor, sedangkan skala statistik perangkat tes antara lain; rerata, median, keandalan, kemencengan dan kesalahan baku pengukuran.

a. Validitas

Validitas dapat didefinisikan sebagai tingkat ketepatan yang dimiliki alat penilaian untuk mengukur sesuatu terhadap kelompok tertentu. Suatu pengukur dapat dikatakan alat ukur tersebut mengukur dapat mengukur apa yang hendak diukur secara tepat.

Menurut Suharsimi Arikunto (1996: 65-67) secara garis besar validitas tes terdiri atas (1) Validitas Logis/Rasional (instrumen yang mempunyai persyaratan valid berdasarkan hasil penalaran) yang terdiri dari validitas isi dan validitas konstruk. (2) Validitas Empiris (instrumen yang mempunyai syarat valid

apabila sudah diuji dari pengalamannya) yang terdiri dari validitas “ada sekarang” dan validitas ramalan.

1) Validitas Isi (*Content Validity*)

Validitas isi artinya kejituan daripada suatu tes ditinjau dari isi tes tersebut. Validitas isi dari suatu tes hasil belajar adalah validitas yang diperoleh setelah dilakukan penganalisisan, penelusuran atau pengujian terhadap isi yang terkandung dalam tes hasil belajar tersebut.

Validitas isi adalah validitas yang dilihat dari segi isi itu sendiri sebagai alat pengukur hasil belajar yaitu: sejauh mana tes hasil belajar sebagai alat pengukur hasil belajar peserta didik, isinya telah dapat mewakili secara representatif terhadap keseluruhan materi atau bahan pelajaran yang seharusnya ditekankan (diujikan) (Anas Sudjiono, 2005: 164)

2) Validitas Konstruksi (*Construct validity*)

Validitas konstruksi artinya kejituan daripada suatu tes ditinjau dari susunan tes tersebut. Sebuah tes dikatakan memiliki validitas konstruksi apabila butir-butir soal yang membangun tes tersebut mengukur setiap aspek berpikir seperti yang disebutkan dalam tujuan instruksional khusus. Dengan kata lain jika butir-butir soal mengukur aspek berpikir tersebut sudah sesuai dengan aspek berpikir yang

menjadi tujuan instruksional (Suharsimi Arikunto, 1996: 67-68).

Validitas konstruksi dari suatu tes hasil belajar dapat dilakukan penganalisisanya dengan jalan melakukan pencocokan antara aspek-aspek berpikir yang terkandung dalam tes hasil belajar tersebut, dengan aspek-aspek berpikir yang dikehendaki untuk diungkap oleh tujuan instruksional khusus. Dengan demikian, maka tes hasil belajar tersebut dapat dinyatakan sebagai tes hasil belajar yang valid dari segi susunannya atau telah memiliki validitas konstruksi (Anas Sudjiono, 2005: 167).

3) Validitas “ada sekarang” (*Concurrent validity*)

Validitas “ada sekarang” artinya kejituan daripada suatu tes dilihat dari korelasinya terhadap kecakapan yang telah dimilikinya saat ini secara riil. Tes sebagai alat pengukur dapat dikatakan telah memiliki validitas bandingan apabila tes tersebut dalam kurun waktu yang sama dengan secara tepat telah mampu menunjukkan adanya hubungan yang searah, antara tes pertama dengan tes berikutnya (Anas Sudjiono, 2005: 176-177)

4) Validitas Prediksi/Ramalan (*Prediktif Validity*)

Validitas ramalan artinya ketepatan daripada suatu alat pengukur ditinjau dari kemampuan tes tersebut untuk

meramalkan prestasi yang dicapainya kemudian. Sebuah tes dikatakan memiliki validitas prediksi atau validitas ramalan apabila mempunyai kemampuan untuk meramalkan apa yang akan terjadi pada masa yang akan datang (Suharsimi Arikunto, 1996: 69).

b. Reliabilitas

Reliabilitas adalah kestabilan skor yang diperoleh orang yang sama ketika diuji ulang dengan tes yang sama pada situasi yang berbeda atau dari suatu pengukuran ke pengukuran yang lainnya. Jadi reliabilitas dapat dikatakan sebagai tingkat konsistensi atau kemantapan hasil dari hasil dua pengukuran terhadap hal yang sama. Hasil pengukuran itu diharapkan akan sama apabila pengukuran itu diulangi. Dengan perangkat tes yang reliabel, apabila tes itu kita berikan dua kali pada orang yang sama, tetapi dalam selang waktu yang berbeda, sepanjang tidak ada perubahan kemampuan, maka skor yang diperoleh akan konstan (Hayat dkk., 1997: 22).

Reliabilitas dan validitas merupakan suatu hal yang penting dalam suatu tes. Reliabilitas mendukung validitas. Suatu tes, mungkin saja reliabel tetapi belum tentu valid. Sebaliknya, tes yang valid sudah pasti reliabel. Reliabilitas memiliki dua konsistensi. Konsistensi pertama, adalah konsistensi internal yakni tingkat sejauhmana soal itu homogen baik dari tingkat

segi kesukaran maupun bentuk soalnya. Konsistensi yang kedua, yaitu konsistensi eksternal yakni tingkat sejauh mana skor yang dihasilkan tetap sama sepanjang kemampuan orang yang diukur belum berubah.

Apabila hasil skor tes pertama sama dengan hasil skor tes kedua, maka tes dikatakan memiliki reliabilitas yang tinggi atau terdapat korelasi yang tinggi antara hasil tes pertama dengan hasil tes kedua. Kalau antara hasil tes pertama dan kedua tidak terdapat hubungan atau hubungan rendah, maka tes itu dikatakan tidak reliabel (Hayat dkk., 1997: 22).

Reliabilitas dalam arti konsistensi atau homogenitas tes, merupakan koefisien korelasi yang menunjukkan seberapa jauh suatu perangkat tes homogen, untuk mengukur suatu mata pelajaran atau bidang studi yang sama. Reliabilitas yang paling baik, bila dicapai angka koefisien 1,00. Dalam pengukuran angka koefisien biasanya kurang dari 1,00, yang disebabkan oleh sifat soal, situasi pada saat pengukuran, keadaan subjek, dan sebagainya (Izzak Latunussa, 1988: 37)

Untuk menentukan tinggi rendahnya reliabilitas soal, digunakan acuan sebagai berikut (Izzak Latunussa, 1988: 37) :

Tabel 1

Kriteria Tingkat Reliabilitas Butir Soal

No	Rentang Nilai	Keputusan
1	$rh < 0,200$	Tidak Reliabel
2	$0,200 < rh < 0,399$	Rendah
3	$0,400 < rh < 0,699$	Sedang
4	$0,700 < rh < 0,899$	Tinggi
5	$rh > 0,900$	Sangat Tinggi

c. Daya Beda

Daya beda suatu soal berfungsi untuk menentukan dapat tidaknya suatu soal membedakan kelompok dalam aspek yang diukur sesuai dengan perbedaan yang ada pada kelompok itu. Tujuan dari pengujian daya beda adalah untuk melihat kemampuan butir soal dalam membedakan antara peserta didik yang berkemampuan tinggi dengan peserta didik yang berkemampuan rendah (Hayat dkk., 1997: 19).

Ada beberapa cara yang digunakan untuk menentukan daya pembeda, antara lain dengan menggunakan (Hayat dkk., 1997: 19) : (1) indeks diskriminasi, (2) indeks korelasi, (3) indeks keselarasan.

Daya beda biasanya disimbolkan dengan D (huruf kapital), langkah-langkah untuk menentukan daya beda adalah :

- 1) Menyusun lembar jawaban dari skor tertinggi ke skor terendah sesuai dengan N (jumlah siswa).
- 2) Mengalikan N dengan 27%, hasil pembulatan diperoleh adalah n.
- 3) Menghitung n kelompok atas (lembar jawaban dengan skor tertinggi di hitung dari atas) dan n kelompok bawah (lembar jawaban dengan skor terendah dihitung dari bawah).
- 4) Menentukan proporsi butir soal yang dijawab dengan benar untuk masing-masing kelompok. Kelompok atas (P_H) dan kelompok bawah (P_L) dengan cara membagi jumlah jawaban yang benar dengan n.

Daya beda butir soal (D), merupakan selisih proporsi butir soal yang dijawab dengan benar antara kelompok atas (P_H) dengan kelompok bawah (P_L).

$$D = (P_H - P_L)$$

Untuk menentukan keputusan soal diterima, direvisi atau ditolak maka menggunakan kriteria parametrik soal yang digambarkan dalam tabel berikut (Pakpahan, 1990: 76) :

Tabel 2

Kriteria Daya Beda Butir Soal

Parameter	Koefisien	Keputusan
Daya Beda	$> 0,30$	Diterima
	$0,10 - 0,29$	Direvisi
	$< 0,10$	Ditolak

d. Tingkat Kesukaran Butir Soal

Tingkat kesukaran pada analisis butir soal secara klasikal, tingkat kesukaran (p) dapat diperoleh dengan beberapa cara antara lain : (1) skala kesukaran linier; (2) Skala Bivariat; (3) Indeks Davis; dan (4) Proporsi menjawab benar. Cara yang paling mudah dan paling umum digunakan adalah skala rata-rata atau proporsi menjawab benar atau *Proportion correct* (p), yaitu jumlah peserta tes yang menjawab benar pada soal yang dianalisis dibandingkan dengan peserta tes seluruhnya. persamaan yang digunakan untuk menentukan tingkat kesukaran (p) ini adalah, (Hayat dkk., 1997: 17)

$$p = \frac{\sum B}{N}$$

p = proporsi menjawab benar pada butir soal tertentu

$\sum B$ = banyaknya peserta tes menjawab benar

N = jumlah peserta tes yang menjawab.

Besarnya tingkat kesukaran (p) berkisar antara 0 sampai dengan 1. Tingkat kesukaran dikategorikan menjadi tiga bagian seperti tampak pada tabel 1 berikut ini :

Tabel 3

Kriteria Tingkat Kesukaran Butir Soal

<i>Proportion Correct</i> (p)	Kategori Soal
$p > 0,70$	Mudah
$0,30 < p < 0,70$	Sedang
$p < 0,30$	Sukar

e. Efektivitas Pengecoh

Setiap tes pilihan ganda memiliki satu pertanyaan serta beberapa pilihan jawaban. Diantara pilihan jawaban yang ada, hanya satu yang benar. Selain jawaban yang benar tersebut, adalah jawaban yang salah. Jawaban yang salah itulah yang dikenal dengan *distractor* (pengecoh). Dengan demikian, efektivitas distraktor adalah seberapa baik pilihan yang salah tersebut dapat mengecoh peserta tes yang memang tidak mengetahui kunci jawaban yang tersedia. Semakin banyak peserta tes yang memilih distraktor tersebut, maka distaktor itu dapat menjalankan fungsinya dengan baik (Djunaidi Lababa, 2008:29)

Tujuan utama dari pemasangan distraktor pada setiap butir item itu adalah agar dari sekian banyak peserta tes yang mengikuti tes hasil belajar ada yang tertarik atau terangsang untuk memilihnya, sebab mereka menyangka bahwa distraktor yang mereka pilih merupakan jawaban yang betul (Anas Sudjiono, 2005:410).

Untuk menentukan berfungsi tidaknya pengecoh, diadakan analisis butir soal. Cara yang mudah untuk menentukan berfungsi tidaknya pengecoh, dapat dilihat pada print out computer hasil analisis program ITEMAN dengan melihat tanda “*minus*” dan “*plus*” pada kolom *Prop. Endorsing* dan *Point Biser*. Jawaban yang baik adalah jika kunci jawaban positif dan distraktor negatif.

f. Kualitas Soal

Dalam penentuan kualitas butir soal Perangkat UAS mata pelajaran Pendidikan Agama Islam, dilakukan dengan cara melihat dari karakteristik butir dan jumlah butir yang jelek diketahui, maka dilanjutkan dengan penentuan kualitas perangkat tes UAS. Kriteria penentuan kualitas tes ini, Penulis kembangkan sendiri dengan mempertimbangkan presentasi butir soal yang jelek.

3. Ulangan Akhir Semester

Penilaian pendidikan dasar diselenggarakan untuk memperoleh ketereangan tentang proses belajar-mengajar dan upaya pencapaian tujuan pendidikan dasar dalam rangka pembinaan dan pengembangannya, serta menentukan akreditasi satuan pendidikan dasar yang bersangkutan. Penilaian pendidikan dasar tersebut mencakup (1) Kegiatan dan kemajuan belajar siswa (2) Pelaksanaan kurikulum (3) Guru dan tenaga kependidikan lain (4) Satuan pendidikan sebagai satu keseluruhan (Husain, Abdul Rajak, 1995: 36)

Ulangan Akhir Semester (UAS) dalam sistem pendidikan nasional, dijadikan salah satu bentuk pengukuran dan penilaian hasil belajar pada akhir semester jenjang pendidikan dasar dan menengah. Semua paket soal UAS Sekolah Dasar (SD) yang di dalamnya termasuk soal pendidikan agama Islam dibuat oleh Dinas Pendidikan, Pemuda dan Olah Raga.

Dalam kategori evaluasi hasil belajar disekolah UAS masuk dalam kategori tes sumatif. Tes sumatif adalah evaluasi atau usaha penilaian hasil belajar yang berupa tes (soal-soal pertanyaan) yang dilaksanakan setelah kegiatan belajar-mengajar berlangsung dalam satuan waktu tertentu misalnya setelah satu catur wulan (di SD) atau semester (di sekolah menengah) (B. Suryosubroto, 2005: 144).

Depdikbud (1991: 45) menyatakan bahwa penilaian sumatif adalah penilaian yang dilakukan pada akhir semester.

Yang menjadi tujuan utama dari tes sumatif atau UAS adalah untuk menentukan nilai yang melambangkan keberhasilan peserta didik setelah mereka menempuh proses pembelajaran dalam waktu tertentu, sehingga dapat ditentukan : (1) Kedudukan masing-masing peserta didik di tengah-tengah kelompoknya, (2) Dapat atau tidaknya peserta didik mengikuti program pengajaran berikutnya), (3) Kemajuan peserta didik, untuk menginformasikan kepada pihak orang tua, petugas bimbingan dan konseling, lembaga-lembaga pendidikan lainnya, atau pasaran kerja yang tertuang dalam bentuk Rapor dan Surat Tanda Tamat Belajar (Anas Sudjiono, 2005: 72-73).

Mutu butir soal tipe pilihan ganda (*Multiple Choice*) untuk UAS sangat bergantung pada kemampuan orang yang mengkontruksi butir soal. Butir soal yang dibuat secara serampangan atau dibuat oleh orang yang tidak terlatih, akan berakibat yang negatif bagi proses pendidikan secara keseluruhan, karena mengarah pada intepretasi yang salah terhadap hasil belajar peserta ulangan. Jadi, pelatihan dan pengetahuan tentang prinsip penyusunan butir soal pilihan ganda, akan sangat menentukan pengukuran hasil belajar siswa.

F. Metode Penelitian

1. Jenis Penelitian

Jenis penelitian ini adalah penelitian lapangan (*Field Research*) dan dilakukan dengan pendekatan kuantitatif karena fokus penelitiannya adalah menganalisis butir soal dengan menggunakan program komputer ITEMAN sebagai alat untuk mengukurnya. Dan penelitian ini bersifat evaluatif, yaitu penelitian yang memaparkan hasil penilaian terhadap objek, yang disesuaikan dengan kriteria-kriteria yang sudah baku.

2. Definisi Operasional

Untuk menghindari kesalahpahaman pengertian, yang terdapat pada penelitian ini perlu diberi batasan sebagai berikut :

- a. Validitas Soal, adalah validitas isi yang terdiri dari validitas logis yang dilihat dari kesesuaian antara materi soal dengan kisi-kisi soal dan validitas muka yang dilihat dari kesesuaian antara materi soal dengan silabus program pengajaran.
- b. Reliabilitas soal adalah tingkat ketepatan soal untuk mengukur, sehingga soal dapat dipercaya.
- c. Daya beda butir soal adalah selisih proporsi butir soal yang dijawab benar antara kelompok siswa (atas) yang pandai dengan kelompok siswa (bawah) yang belum pandai.
- d. Taraf kesukaran butir soal adalah proporsi total kelompok siswa yang menjawab butir soal dengan benar.

- e. Distraktor adalah pengecoh yang mendampingi kunci jawaban yang terdapat pada 4 option jawaban.
- f. Kualitas butir soal adalah kualitas soal yang dilihat dari segi banyaknya perbandingan antara soal yang baik dengan soal yang jelek.

3. Populasi dan Sample

a. Populasi

Populasi adalah wilayah generalisasi yang terdiri atas : objek/subjek yang mempunyai kualitas dan karakteristik tertentu yang diterapkan oleh peneliti untuk dipelajari dan ditarik kesimpulannya (Sogiyono, 2010: 117).

Karena penelitian ini adalah tentang analisis butir soal Pendidikan Agama Islam Kelas 6 (enam), maka yang menjadi subjek penelitian adalah siswa kelas 6 (Enam) di SD Negeri se-Kecamatan Karangmojo, Gunungkidul yang beragama Islam dengan kata lain yang menjadi populasi adalah semua siswa kelas 6 (enam) di SD Negeri se-Kecamatan Karangmojo Gunungkidul sebagai berikut :

Tabel 4

Populasi Siswa Kelas 6 SD Negeri Beragama Islam
se-Kecamatan Karangmojo Gunungkidul

No	Nama SD Negeri	Siswa Beragam Islam		
		L	P	Jumlah
1	Sokoliman I	11	7	18
2	Banyu Bening I	12	21	33
3	Grogol IV	19	17	36
4	Banyu Bening III	6	10	16
5	Grogol I	9	9	18
6	Gelaran I	11	14	25
7	Gelaran II	20	10	30
8	Gelaran III	5	4	9
9	Wiladeg	17	11	28
10	Bendungan I	6	9	15
11	Bendungan II	8	9	17
12	Bendungan III	6	4	10
13	Slametan	8	8	16
14	Kelor	2	3	5
15	Karangduwet III	4	6	10
16	Karangmojo II	7	8	15
17	Karangmojo IV	4	9	13

18	Karangmojo I	6	4	10
19	Karangduwet I	6	4	10
20	Ngagel	9	4	13
21	Tlogowareng	5	5	10
22	Karangmojo III	12	13	25
23	Gedangan I	8	13	21
24	Pangkah	11	8	19
25	Karanganom II	3	15	18
26	Karanganom I	3	7	10
27	Ngawis	3	6	9
28	Candi I	6	9	15
29	Candi II	6	14	20
30	Candi Baru I	11	5	16
31	Candi Baru II	10	10	20
32	Pengkol	7	19	26
33	Karangmojo V	8	4	12
34	Bejiharjo II	7	6	13
35	Jetis	8	12	20
36	Gedangan	8	8	16
37	Karangwetan	3	2	5
Jumlah Siswa Islam		295	327	622

Sedangkan untuk objek penelitiannya adalah lembar jawab soal Pendidikan Agama Islam yang sudah dikerjakan oleh semua siswa kelas 6 (enam) yang beragama Islam di SD Negeri Se-Kecamatan Karangmojo Gunungkidul tersebut.

b. Sample

Sample adalah bagian dari jumlah dan karakteristik yang dimiliki oleh populasi tersebut. Bila populasi besar, dan peneliti tidak mungkin mempelajari semua yang ada pada populasi, misalnya karena keterbatasan dana, tenaga dan waktu, maka peneliti dapat menggunakan sample yang diambil dari populasi itu. Apa yang dipelajari dari sample itu, kesimpulannya akan dapat diberlakukan untuk populasi. Untuk itu sample yang diambil dari populasi harus betul-betul representatif (mewakili) (Sugiyono, 2010: 118). Menurut catatan administrasi siswa yang berjumlah lebih dari 100 maka peneliti hanya mengambil sebagian dari populasi yang ada. Karena penelitian ini adalah penelitian sample, maka peneliti mengambil sebagian dari semua siswa di SD Negeri se-Kecamatan Karangmojo yang beragama Islam.

Seperti apa yang disampaikan oleh Suharsimi Arikunto “untuk sekedar ancer-ancer, maka apabila subjeknya kurang dari 100, lebih baik diambil semua sehingga

penelitiannya adalah penelitian populasi. Tetapi apabila jumlah subjeknya besar, dapat diambil 10-15% atau 20-25% atau lebih". Dengan berdasarkan hal di atas, maka dalam penelitian ini peneliti mengambil sample 25 % dari 622 siswa. Maka dari 622 siswa tersebut yang diteliti hanya 124 siswa. Adapun soal mata pelajaran Pendidikan Agama Islam semester ganjil di SD Negeri se-Kecamatan Karangmojo tahun pelajaran 2011/2012 bisa dilihat pada lampiran 1.

4. Teknik Pengumpulan data

Data diperoleh dengan metode dokumentasi, data tersebut adalah lembar jawaban Ulangan Akhir Semester (UAS) butir soal Pendidikan Agama Islam di SD Negeri Se-Kecamatan Karangmojo Gunungkidul tahun pelajaran 2011/2012.

5. Analisis Data

a. Validitas Butir Soal

Cara yang dilakukan dalam penentuan validitas soal adalah dengan menggunakan validitas Rasional/Logis yang terdiri dari validitas isi dan validitas konstruksi.

b. Penentuan Reliabilitas Butir Soal

Reliabilitas butir soal diperoleh dengan menggunakan program analisis butir soal ITEMAN. Hasil dari penghitungan tersebut dikonsultasikan pada tabel sebagai berikut :

Tabel 5

Kriteria Reliabilitas Butir Soal

No	Rentang Nilai	Keputusan
1	$rh < 0,200$	Tidak Reliabel
2	$0,200 < rh < 0,399$	Rendah
3	$0,400 < rh < 0,699$	Sedang
4	$0,700 < rh < 0,899$	Tinggi
5	$rh > 0,900$	Sangat Tinggi

c. Penentuan Daya Beda Butir Soal

Cara yang digunakan untuk menentukan daya beda butir soal adalah menggunakan program komputer analisis butir soal ITEMAN. Hasil dari penghitungan tersebut dikonsultasikan pada tabel 5 sebagai berikut :

Tabel 6

Kriteria Daya Beda Butir Soal

Parameter	Koefisien	Keputusan
Daya Beda	$> 0,10$	Diterima
	$0,10 - 0,29$	Direvisi
	$< 0,10$	Ditolak

Butir soal yang baik adalah yang dapat membedakan mereka yang pandai dan kurang pandai. Indeks daya beda yang baik adalah 0,30 ke atas. Indeks daya beda yang angkanya

“*minus*” berarti tidak baik, sebab mereka yang kurang pandai justru lebih baik dari pada yang pandai. Dalam print out komputer, hasil analisis program ITEMAN, daya beda dapat dilihat pada kolom *Point Biser*.

d. Penentuan Taraf Kesukaran Butir Soal

Cara yang digunakan dalam menentukan taraf kesukaran butir soal adalah dengan menggunakan program komputer analisis butir soal ITEMAN. Hasil dari perhitungan tersebut dikonsultasikan pada tabel 4 sebagai berikut :

Tabel 7

Kriteria Tingkat Kesukaran Butir Soal

Parameter	Koefisien	Keputusan
Tingkat Kesukaran	0,30 – 0,70	Diterima
	0,10 – 0,29 atau 0,71 – 0,90	Direvisi
	< 0,10 atau > 0,90	Ditolak

Indeks kesulitan yang baik adalah yang “*sedang*”, artinya tidak terlalu sulit atau terlalu mudah. Seberapa besar kategori baik itu, bergantung pada tujuan diadakannya tes. Untuk kepentingan analisis butir soal pada UAS ini, Penulis tetapkan 0,30 – 0,70. Dalam print out computer hasil analisis

program ITEMAN, indeks kesulitan dapat dilihat pada kolom *Prop. Correct*.

e. Penentuan Efektifitas Pengecoh Butir Soal

Cara yang mudah untuk menentukan berfungsi tidaknya pengecoh dapat dilihat pada print out komputer hasil analisis program ITEMAN, dengan melihat tanda “*minus*” dan “*plus*” pada kolom *Prop. Endorsing* dan *Point Biser*. Jawaban yang baik, adalah jika kunci jawaban positif dan distraktor negatif.

f. Penentuan Kualitas Perangkat Tes UAS

Setelah karakteristik butir dan jumlah butir yang jelek diketahui, maka dilanjutkan dengan penentuan kualitas perangkat tes UAS. Kriteria penentuan kualitas tes ini, Penulis kembangkan sendiri dengan mempertimbangkan presentasi butir soal yang jelek. Kualitas perangkat tes ditemukan dengan kriteria sebagaimana dalam tabel berikut :

Tabel 8

Kriteria Kualitas Butir Soal

No	Jumlah Butir Yang Ditolak	Kualitas
1	0% - 20%	Sangat Baik
2	21% - 40%	Baik
3	41% - 60%	Cukup
4	61% - 80%	Kurang Baik
5	Diatas 80%	Kurang Sekali

G. Sistematika Pembahasan

Sistematika penulisan Skripsi ini akan penulis sampaikan dalam 5 bab yang terdiri dari :

Bab I. Pendahuluan, diuraikan tentang adanya persoalan bahwa selama ini soal-soal ulangan Akhir Semester (UAS) di SD se-Kecamatan Karangmojo belum pernah diadakan analisis terhadap soal-soal ulangan tersebut untuk mengetahui kualitas butir soal UAS tersebut baik dilakukan oleh pihak guru maupun pihak lainnya.

Pada bagian tujuan dan kegunaan penelitian akan diungkapkan tujuan dari penelitian serta mengungkapkan kegunaan penelitian, karena dengan mengetahui tujuan dan kegunaan penelitian akan diketahui bahwa penelitian ini akan mengungkapkan dan menemukan sesuatu yang sama sekali baru.

Dari latar belakang masalah, tujuan dan kegunaan penelitian selanjutnya dirumuskan rumusan masalah satu persatu, dalam bentuk pertanyaan, dengan harapan pertanyaan tersebut dapat dipecahkan setelah dilakukan penelitian.

Pada bagian tinjauan pustaka, memuat tentang hasil hasil penelitian yang perolah peneliti terdahulu, dilihat dari hubungan-hubungan, perbedaan maupun kesamaan dengan penelitian yang akan dilakukan, hal ini penting, karena akan diketahui keunikan atau kekhasan dari penelitian yang dilakukan oleh para peneliti.

Pada bagian landasan teori memuat beberapa pernyataan para pakar pendidikan khususnya yang berhubungan dengan persoalan-persoalan evaluasi pendidikan dan analisis butir soal. Pernyataan tersebut diharapkan dapat dijadikan landasan teori dalam penelitian.

Pada bagian Metode Penelitian. Pada bab ini disajikan metode yang tepat untuk melakukan penelitian skripsi, hal ini penting, karena dengan menggunakan metode Penelitian yang tepat maka tujuan dari penelitian akan dicapai.

Bab II. Adalah gambaran umum tentang SD Negeri di kecamatan Karangmojo.

Bab III. Hasil Penelitian dan Pembahasan. Pada bab ini akan diuraikan secara detail proses penelitian sejak awal pengumpulan data hingga akhir pengolahan data. Setelah data diolah dan dianalisis, akan disajikan sebagai hasil penelitian yang selanjutnya akan dibahas hasil penelitian dengan teori yang terdapat pada tinjauan pustaka. Dari sinilah, akan diperoleh kesimpulan dari penelitian.

Bab IV. Kesimpulan dan Saran. Pada bab ini, akan diuraikan kesimpulan akhir dari penelitian serta saran-saran untuk penelitian selanjutnya.