

BAB I

PENDAHULUAN

A. Latar Belakang Masalah

Pendidikan sebagai suatu program yang melibatkan sejumlah komponen yang bekerja sama dalam sebuah proses untuk mencapai tujuan yang telah diprogramkan. Untuk mengetahui apakah penyelenggaraan program dapat mencapai tujuannya secara efektif dan efisien, maka perlu dilakukan evaluasi (Purwanto, 2011:1). Kegiatan evaluasi mempunyai peranan penting dalam pendidikan dan pembelajaran, karena dengan evaluasi dapat diketahui seberapa jauh peserta didik telah menguasai bahan atau materi yang telah dipelajari sehingga dapat berfungsi sebagai umpan balik (*feedback*) bagi perbaikan belajar mengajar di kelas.

Berdasarkan Peraturan Pemerintah (PP) No. 19 tahun 2005 tentang Standar Nasional Pendidikan, evaluasi hasil belajar dilakukan dengan penilaian hasil belajar yang bertujuan untuk memantau proses, kemajuan dan perbaikan hasil belajar serta untuk menilai pencapaian kompetensi peserta didik, sebagai bahan penyusunan laporan kemajuan hasil belajar dan untuk memperbaiki proses pembelajaran. Penilaian hasil belajar dilakukan dalam bentuk ulangan harian, ulangan tengah semester dan ulangan kenaikan kelas. Salah satu alat yang dapat digunakan sebagai sarana untuk melakukan penilaian hasil belajar adalah tes (Arifin, 2009:46).

Tes sebagai alat evaluasi dalam pendidikan, mempunyai peranan penting dalam mengukur prestasi hasil belajar peserta didik dalam kurun waktu tertentu. Suatu tes dikatakan berkualitas apabila memenuhi validitas dan reliabilitas. Untuk memperoleh tes hasil belajar yang valid dan reliabel maka seorang pendidik harus melakukan analisis butir soal yang bertujuan untuk mengkaji dan menelaah setiap butir soal agar diperoleh soal yang berkualitas. Soal yang berkualitas jika digunakan dalam penilaian dan evaluasi hasil belajar akan memberikan informasi yang tepat sesuai dengan tujuannya. Analisis kualitas butir soal mencakup analisis secara kualitatif (validitas isi dan konstruk) dan analisis kuantitatif (validitas empiris dan reliabilitas soal).

Soal ulangan akhir semester (UAS) gasal mata pelajaran fiqih di MTs. Negeri Nglipar, dibuat oleh musyawarah guru mata pelajaran (MGMP) Fiqih Gunungkidul. Selanjutnya soal direvisi oleh kelompok kerja kepala Madrasah Tsanawiyah (K3MTs.) Kabupaten Gunungkidul, yang kemudian diujikan secara serentak di Madrasah Tsanawiyah se-Gunungkidul. Soal UAS gasal fiqih tersebut, tidak diketahui apakah sudah termasuk butir-butir soal yang telah memenuhi syarat sebagai alat ukur yang baik atau belum, karena belum pernah diujicobakan dan dianalisis secara kualitatif maupun kuantitatif, sehingga belum diketahui tingkat validitas logis, validitas empiris dan reliabilitasnya (Wawancara dengan Bapak Suparno S.Pd.I. M.Si. selaku guru mata pelajaran fiqih, sekaligus wakil ketua MGMP fiqih Gunungkidul, tanggal 17 Desember 2013).

Kriteria ketuntasan minimal (KKM) untuk mata pelajaran fiqih di MTs. Negeri Nglipar sebesar 75. Untuk mencapai nilai KKM tersebut, maka guru yang bersangkutan telah memberikan kisi-kisi soal fiqih untuk setiap kelas. Tetapi berdasarkan hasil UAS gasal fiqih yang telah dilaksanakan pada hari selasa, tanggal 3 Desember 2013, menunjukkan bahwa nilai prestasi siswa masih rendah. Hal ini ditandai dengan ditemukannya banyak siswa yang mendapat nilai di bawah KKM, yaitu 75. Bahkan ada beberapa siswa yang mengikuti remidi karena nilainya masih jauh di bawah KKM.

Berdasarkan uraian tersebut, maka perlu diadakan penelitian dengan judul “Analisis Kualitas Butir Soal Ulangan Akhir Semester Gasal pada Mata Pelajaran Fiqih di MTs. Negeri Nglipar Gunungkidul”. Dari penelitian ini dapat diketahui mengenai kualitas butir soalnya, meliputi validitas logis, validitas empiris (tingkat kesukaran, daya pembeda dan efektifitas pengecoh) serta reliabilitasnya. Hasil penelitian ini juga dapat dijadikan sebagai referensi dalam memilih soal-soal, dapat membantu melihat terukur tidaknya kompetensi yang diharapkan tercapai melalui soal tersebut dan sebagai bahan pertimbangan MGMP Fiqih Gunungkidul dalam pembuatan soal yang akan datang sehingga dapat menyempurnakan atau memperbaiki kualitas soal yang akan diujikan. Sementara soal yang sudah dianalisis dan hasilnya berkualitas dalam arti memiliki validitas dan reliabilitas yang tinggi, serta mampu mengukur kompetensi yang diharapkan, dapat dijadikan sebagai bank soal.

B. Rumusan Masalah

Berdasarkan latar belakang masalah yang telah dipaparkan, maka permasalahan yang dikaji adalah:

1. Bagaimanakah tingkat kesukaran butir soal ulangan akhir semester gasal mata pelajaran fiqih kelas VII, VIII dan IX di MTs. Negeri Nglipar?
2. Bagaimanakah daya pembeda butir soal ulangan akhir semester gasal mata pelajaran fiqih kelas VII, VIII dan IX di MTs. Negeri Nglipar?
3. Apakah pengecoh butir soal ulangan akhir semester gasal mata pelajaran fiqih kelas VII, VIII dan IX di MTs. Negeri Nglipar telah efektif?
4. Apakah butir soal ulangan akhir semester gasal mata pelajaran fiqih kelas VII, VIII dan IX di MTs. Negeri Nglipar telah reliabel?
5. Apakah butir soal ulangan akhir semester gasal mata pelajaran fiqih kelas VII, VIII dan IX di MTs. Negeri Nglipar memiliki validitas logis?

C. Tujuan dan Kegunaan Penelitian

Tujuan dari penelitian ini adalah:

1. Untuk mengukur tingkat kesukaran butir soal ulangan akhir semester gasal mata pelajaran fiqih kelas VII, VIII dan IX di MTs. Negeri Nglipar.
2. Untuk mengukur daya pembeda butir soal ulangan akhir semester gasal mata pelajaran fiqih kelas VII, VIII dan IX di MTs. Negeri Nglipar.
3. Untuk mengukur efektifitas pengecoh butir soal ulangan akhir semester gasal mata pelajaran fiqih kelas VII, VIII dan IX di MTs. Negeri Nglipar.

4. Untuk mengukur reliabilitas butir soal ulangan akhir semester gasal mata pelajaran fiqih kelas VII, VIII dan IX di MTs. Negeri Nglipar.
5. Untuk mengukur validitas logis butir soal ulangan akhir semester gasal mata pelajaran fiqih kelas VII, VIII dan IX di MTs. Negeri Nglipar.

Adapun kegunaan dari penelitian ini sebagai berikut:

1. Kegunaan Secara Teoritis

Menambah pengetahuan dan wawasan mengenai analisis butir soal melalui pendekatan kualitatif dan kuantitatif dalam menentukan kualitas butir soal UAS gasal mata pelajaran fiqih di MTs. Negeri Nglipar.

2. Kegunaan Secara Praktis

- a. Bagi penulis, memberikan kontribusi pengetahuan dan menambah wacana keilmuan khususnya dalam menganalisis kualitas suatu tes.
- b. Bagi guru, sebagai bahan informasi dan pertimbangan dalam rangka menganalisis kualitas butir soal mata pelajaran fiqih baik analisis secara kualitatif dan kuantitatif.
- c. Bagi madrasah, sebagai acuan bagi lembaga madrasah untuk memperhatikan kualitas butir soal yang dijadikan sebagai evaluasi hasil belajar mata pelajaran fiqih.
- d. Bagi instansi terkait, hasil penelitian ini diharapkan menjadi bahan masukan bagi MGMP Fiqih dan K3MTs. Gunungkidul, untuk melakukan analisis kualitas butir soal di masa mendatang sehingga dapat dilakukan program yang tepat sasaran, misalnya memasukkan soal yang berkualitas pada bank soal.

D. Tinjauan Pustaka

Pada dasarnya urgensi tinjauan pustaka penelitian dijadikan sebagai bahan auto kritik terhadap penelitian yang ada mengenai kelebihan maupun kekurangannya, sekaligus sebagai bahan perbandingan terhadap kajian yang terdahulu. Untuk menghindari terjadinya pengulangan hasil temuan yang membahas permasalahan yang sama, baik dalam bentuk penelitian, buku dan dalam bentuk tulisan lainnya, maka penulis akan memaparkan beberapa bentuk tulisan yang sudah ada. Penulis berpendapat bahwa beberapa bentuk tulisan yang penulis temukan, masing-masing menunjukkan perbedaan dari segi pembahasan meliputi, tempat penelitian, subyek dan objek penelitian serta hasil penelitian. Beberapa penelitian yang sudah teruji keshahihannya meliputi:

Penelitian Awan Rahmadi NIM 20060720159, *Analisis Butir Soal Mata Ujian Al-Quran Hadits Pada Ujian Akhir Sekolah Tahun 2008 Dengan Kasus Testee Siswa SMK Muhammadiyah I Playen Gunungkidul Yogyakarta*. Mahasiswa Fakultas Agama Islam, Jurusan Pendidikan Agama Islam, Universitas Muhammadiyah Yogyakarta. Hasil penelitiannya menunjukkan bahwa sebesar 40% dari jumlah soal dalam tingkat kesukaran sedang, 29% dalam kategori sukar dan 31% dalam kategori baik. Ditinjau dari daya pembeda, sebanyak 24 atau 53% dalam kategori baik dan sebanyak 21 atau 47% termasuk kategori jelek. Sedangkan sebanyak 10 pengecoh telah efektif dan sebanyak 35 tidak efektif.

Penelitian Rofi Fajariyah NIM 20070720165, *Analisis Butir Soal UAS Mata Pelajaran Pendidikan Agama Islam SMP Negeri 2 Wonosari Tahun Ajaran 2008/2009*. Mahasiswa Fakultas Agama Islam, Jurusan Pendidikan Agama Islam, Universitas Muhammadiyah Yogyakarta. Hasil penelitiannya menunjukkan bahwa soal UAS PAI yang telah diujikan mempunyai tingkat kesukaran dalam kategori sedang sebanyak 19 atau 42% dan sebanyak 26 atau 58% dalam kategori mudah. Daya pembeda soal yang termasuk kategori cukup sebanyak 10 atau 22% dan sebanyak 35 atau 78% mempunyai daya pembeda yang jelek. Sedangkan pengecoh yang berfungsi sebanyak 88 atau 65% dan pengecoh yang tidak berfungsi sebanyak 47 atau 35%.

Penelitian Sarjana NIM 20090720241, *Analisis Butir Soal Ulangan Akhir Semester Gasal Kelas VI Mata Pelajaran Pendidikan Agama Islam di SD Negeri se-Kecamatan Karang Mojo Tahun Pelajaran 2011/2012*. Mahasiswa Fakultas Agama Islam, Jurusan Pendidikan Agama Islam, Universitas Muhammadiyah Yogyakarta. Hasil penelitiannya menunjukkan bahwa soal UAS gasal PAI yang telah diujikan, belum valid, mempunyai reliabilitas sedang, daya pembeda rendah, 60% distraktor tidak berfungsi dan banyaknya soal yang gugur sebanyak 52%.

Perbedaan penelitian ini dengan penelitian-penelitian lain, yaitu terletak pada subyek penelitian, objek penelitian dan tempat penelitian. Penelitian yang disusun ini, tempat penelitiannya di MTs. Negeri Nglipar, dengan subyek penelitiannya adalah peserta didik kelas VII, VIII dan IX, yang telah mengikuti UAS gasal mata pelajaran fiqih, sedangkan objek penelitian ini

meliputi soal, kunci jawaban dan lembar jawaban dari peserta didik pada UAS gasal mata pelajaran fiqih tersebut. Adapun metode analisis data yang digunakan dalam penelitian ini menggunakan *item and test analysis* (ITEMAN) untuk analisis kuantitatif sedangkan untuk analisis kualitatif menggunakan lembar penelaahan butir soal pilihan ganda.

E. Landasan Teori

1. Evaluasi Pendidikan

Menurut Purwanto (2011:1), evaluasi adalah pengambilan keputusan berdasarkan hasil pengukuran dan standar kriteria. Kegiatan evaluasi setelah dilakukan pengukuran dan keputusan evaluasi dilakukan berdasarkan hasil pengukuran. Evaluasi diharapkan dapat menjadi umpan balik untuk program yang telah dijalankan (*feedback*) dan memberikan informasi yang diperlukan untuk menjalankan program dimasa yang akan datang (*feedforward*).

Kaitannya dalam dunia pendidikan, kegiatan evaluasi sering digunakan karena selama satu periode pendidikan berlangsung, orang perlu mengetahui hasil atau prestasi yang telah dicapai baik oleh pihak pendidik maupun oleh peserta didik. Hal ini dapat dirasakan dalam semua bentuk dan jenis pendidikan, baik pendidikan formal, nonformal maupun informal. Di sekolah, guru sering mengadakan evaluasi, mulai dari ulangan harian, ulangan tengah semester (UTS), ulangan akhir semester (UAS) dan sebagainya.

2. Evaluasi Hasil Belajar Fiqih

Mata pelajaran fiqih di Madrasah Tsanawiyah merupakan salah satu mata pelajaran PAI yang mempelajari tentang fiqih ibadah, terutama menyangkut pengenalan dan pemahaman tentang cara-cara pelaksanaan rukun Islam dan pembiasaannya dalam kehidupan sehari-hari, serta fiqih muamalah yang menyangkut pengenalan dan pemahaman sederhana mengenai ketentuan tentang makanan dan minuman yang halal dan haram, khitan, qurban, serta tata cara pelaksanaan jual beli dan pinjam meminjam. Secara substansial mata pelajaran fiqih memiliki kontribusi dalam memberikan motivasi kepada peserta didik untuk mempraktikkan dan menerapkan hukum Islam dalam kehidupan sehari-hari sebagai perwujudan keserasian, keselarasan dan keseimbangan hubungan manusia dengan Allah SWT, dengan diri manusia itu sendiri, sesama manusia, makhluk lainnya ataupun lingkungannya.

Tujuan mata pelajaran fiqih di Madrasah Tsanawiyah adalah untuk membekali peserta didik agar dapat:

- a. Mengetahui dan memahami cara-cara pelaksanaan hukum Islam baik yang menyangkut aspek ibadah maupun muamalah untuk dijadikan pedoman hidup dalam kehidupan pribadi dan sosial.
- b. Melaksanakan dan mengamalkan ketentuan hukum Islam dengan benar dan baik, sebagai perwujudan dari ketaatan dalam menjalankan ajaran agama Islam baik dalam hubungan manusia dengan Allah SWT, dengan diri manusia itu sendiri, sesama

manusia, dan makhluk lainnya maupun hubungan dengan lingkungannya.

Adapun ruang lingkup mata pelajaran fiqih di Madrasah Tsanawiyah meliputi:

- a. Fiqih ibadah, yang menyangkut: pengenalan dan pemahaman tentang cara pelaksanaan rukun Islam yang benar dan baik, seperti: tata cara thaharah, shalat, puasa, zakat dan ibadah haji.
- b. Fiqih muamalah, yang menyangkut: pengenalan dan pemahaman mengenai ketentuan tentang makanan dan minuman yang halal dan haram, khitan, qurban, serta tata cara pelaksanaan jual beli dan pinjam meminjam.

Jadi dalam evaluasi mata pelajaran fiqih merupakan kegiatan atau proses penentuan nilai mata pelajaran fiqih, sehingga dapat diketahui mutu atau hasil-hasilnya. Setelah diketahui hasilnya maka tahap selanjutnya ialah mengambil langkah perbaikan pada program pembelajaran yang akan datang.

3. Pengertian Ulangan Akhir Semester

Peraturan Pemerintah (PP) No. 19 tahun 2005, tentang standar nasional pendidikan, bahwa yang disebut ulangan adalah proses yang dilakukan untuk mengukur kompetensi peserta didik secara berkelanjutan dalam proses pembelajaran, untuk memantau kemajuan, melakukan perbaikan pembelajaran dan penentuan keberhasilan belajar peserta didik. UAS adalah kegiatan oleh pendidik untuk mengukur pencapaian

kompetensi peserta didik diakhir semester. Cakupan ulangan semester meliputi seluruh indikator yang mempresentasikan KD pada semester tersebut (Arifin, 2009:45).

Bentuk soal yang dipakai dalam ulangan semester dapat berupa pilihan ganda, campuran pilihan ganda dan uraian atau semuanya bentuk uraian. Materi yang diujikan berdasarkan kisi-kisi soal. Tingkat berfikir yang terlibat mulai dari ingatan (C1) sampai dengan evaluasi (C6).

Ulangan akhir semester (UAS) merupakan salah satu bentuk alat evaluasi yang akan digunakan dalam mengambil keputusan pada program pembelajaran selanjutnya, sehingga alat evaluasi yang digunakan dalam UAS harus memperhatikan validitas logis, reliabilitas, tingkat kesukaran item, daya pembeda dan fungsi pengecoh. Soal UAS fiqih menggunakan bentuk tes obyektif berupa *Multi Choice Item* dengan 4 alternatif jawaban sejumlah 40 butir soal dan tes subyektif/uraian sejumlah 5 butir soal. Dalam penelitian ini, yang akan dijadikan obyek analisis oleh peneliti adalah soal pilihan ganda.

Menurut Arifin (2009:138), soal tes bentuk pilihan ganda dapat digunakan untuk mengukur hasil belajar yang lebih kompleks dan berkenaan dengan aspek ingatan, pengertian, aplikasi, analisis, sintesis dan evaluasi. Soal tes bentuk pilihan ganda terdiri atas pembawa pokok persoalan dan pilihan jawaban. Pembawa pokok persoalan dapat dikemukakan dalam bentuk pertanyaan maupun pernyataan (*statement*) yang belum sempurna yang sering disebut *stem* sedangkan pilihan

jawaban itu bisa berbentuk perkataan, bilangan atau kalimat yang disebut *option*. Pilihan jawaban terdiri atas jawaban yang benar yang selanjutnya disebut kunci jawaban dan kemungkinan jawaban yang salah yang dinamakan pengecoh (*distractor* atau *decoy* atau *fails*).

Untuk menghasilkan soal pilihan ganda yang benar-benar berkualitas, menurut Arifin (2009:143) ada beberapa hal yang harus diperhatikan dalam penyusunannya, yaitu sebagai berikut:

- a. Harus mengacu pada kompetensi dasar dan indikator soal.
- b. Berilah petunjuk mengerjakan dengan jelas.
- c. Jangan memasukkan materi soal yang tidak relevan dengan apa yang sudah dipelajari oleh peserta didik.
- d. Pernyataan pada soal seharusnya merumuskan persoalan yang jelas dan berarti.
- e. Pernyataan dan pilihan hendaknya merupakan kesatuan kalimat yang tidak terputus.
- f. Alternatif jawaban harus berfungsi, homogen dan logis.
- g. Panjang pilihan jawaban hendaknya lebih pendek dari pada itemnya.
- h. Usahakan agar pernyataan dan pilihan tidak mudah diasosiasikan.
- i. Alternatif jawaban yang betul hendaknya jangan sistematis.
- j. Harus diyakini benar bahwa hanya ada satu kunci jawaban yang benar.

4. Analisis Kualitas Butir Soal

Peningkatan mutu pendidikan tidak terlepas dari penerapan penilaian yang secara tepat mengukur hasil akhir dari suatu pembelajaran. Artinya untuk menilai hasil akhir dalam pembelajaran diperlukan alat penilaian yang berkualitas. Salah satu alat penilaian yang sering digunakan adalah tes. Untuk mengetahui kualitas tes maka perlu dilakukan analisis soal, sebelum soal tersebut diberikan kepada peserta tes.

Analisis kualitas tes merupakan suatu tahap yang harus ditempuh untuk mengetahui derajat kualitas suatu tes baik tes secara keseluruhan maupun butir soal yang menjadi bagian dari tes tersebut. Dalam penilaian hasil belajar, tes diharapkan dapat menggambarkan sampel perilaku dan menghasilkan nilai yang obyektif serta akurat. Jika tes yang digunakan guru kurang baik, maka hasil yang diperolehpun tentunya kurang baik. Hal ini dapat merugikan peserta didik itu sendiri. Artinya hasil yang diperoleh peserta didik menjadi tidak obyektif dan tidak adil. Oleh sebab itu, tes yang digunakan guru harus memiliki kualitas yang baik dilihat dari berbagai segi. Tes hendaknya disusun sesuai dengan prinsip dan prosedur penyusunan tes. Setelah digunakan perlu diketahui apakah tes tersebut berkualitas baik atau tidak. Untuk mengetahui apakah suatu tes yang digunakan termasuk baik atau tidak maka perlu dilakukan analisis kualitas tes (Arifin, 2009:128).

Untuk mengetahui kualitas suatu tes yang meliputi validitas dan reliabilitas, dapat dianalisis dengan pendekatan kualitatif dan kuantitatif.

Pendekatan kualitatif (*logical validity*) dilakukan dengan menelaah butir soal meliputi aspek materi, konstruksi dan bahasa. Analisis kualitatif ini dilakukan sebelum tes itu digunakan/diujicobakan, seperti menggunakan kartu telaah. Sedangkan pendekatan kuantitatif (*empirical validity*) merupakan metode penelaahan data empiris meliputi tingkat kesukaran, daya pembeda, efektifitas pengecoh dan reliabilitasnya. Analisis dengan pendekatan kuantitatif ini dilakukan berdasarkan data hasil ujian atau hasil ujicoba yang diperoleh dari peserta tes (Arifin, 2009:132).

5. Uji Terhadap Butir Soal

Analisis kualitas tes berkaitan dengan pertanyaan apakah tes sebagai suatu alat ukur benar-benar mengukur apa yang hendak dan seharusnya diukur serta sampai mana tes tersebut dapat diandalkan dan berguna. Kedua pertanyaan ini sebenarnya menunjuk pada dua hal pokok, yaitu validitas dan reliabilitas. Kedua hal tersebut merupakan ciri atau karakteristik alat ukur yang baik. Berdasarkan hal tersebut, maka uji terhadap butir soal dalam penelitian ini meliputi beberapa hal, yaitu:

a. Validitas

Validitas berasal dari kata "*valid*" yang berarti tepat, benar, shahih dan absah. Jadi validitas diartikan sebagai suatu ketepatan, kebenaran, keshahihan dan keabsahan. Suatu tes dikatakan valid apabila benar-benar mengukur apa yang hendak dan seharusnya diukur, serta sejauh mana alat ukur tersebut dapat diandalkan dan berguna (Arifin, 1991:109).

Validitas sebuah tes dapat diketahui dari hasil pemikiran dan dari pengalaman. Hal yang pertama diperoleh adalah validitas logis (*logical validity*) dan hal yang kedua diperoleh validitas empiris (*empirical validity*) (Arikunto, 2013:80). Secara garis besar ada dua macam validitas, yaitu validitas logis dan validitas empiris.

1) Validitas Logis

Validitas logis terbagi menjadi dua macam, yaitu:

a) Validitas Isi (*content validity*)

Menurut Arikunto (2013:81), validitas isi bagi sebuah instrumen menunjuk suatu kondisi sebuah instrumen yang disusun berdasarkan isi materi pelajaran yang akan dievaluasi. Validitas isi menunjukkan sejauh mana item-item dalam tes mencakup keseluruhan kawasan isi yang hendak diukur oleh tes tersebut. Selain harus komprehensif, tetapi isinya juga harus relevan dan tidak keluar dari batasan tujuan pengukuran.

Sedangkan menurut Purwanto (2011:120), validitas isi adalah pengujian validitas dilakukan atas isinya untuk memastikan apakah butir tes mengukur secara tepat keadaan yang ingin diukur. Validitas isi disebut juga validitas kurikuler. Oleh karena itu, validitas ini erat kaitannya dengan materi yang akan diukur dalam tes. Tentu saja materi yang dimaksud adalah materi yang terdapat dalam kurikulum. Validitas isi mencerminkan sejauh mana butir-butir dalam tes mencerminkan

materi yang disajikan dalam kurikulum. Sebuah tes dikatakan memiliki validitas isi jika butir-butir tes bersifat representatif terhadap isi materi dalam kurikulum tersebut.

Pengujian validitas isi tidak melalui prosedur pengujian secara statistik, melainkan melalui analisis secara rasional. Pengetahuan terhadap kurikulum menjadi dasar berpijak yang penting untuk dapat melakukan analisis validitas isi. Cara yang praktis untuk melakukan analisis validitas isi adalah dengan melihat apakah butir-butir tes telah disusun sesuai dengan *blue-print* (kisi-kisi) yang sudah dirancang sebelumnya. *Blue print* menjadi acuan dalam menuangkan domain atau ranah dan indikator yang akan diukur dalam tes.

b) Validitas Kontruksi (*construct validity*)

Validitas kontruksi sebuah instrumen menunjuk suatu kondisi sebuah instrumen yang disusun, berdasarkan kontruksi aspek-aspek kejiwaan yang seharusnya dievaluasi. Hasil belajar dikontruksi oleh sejumlah ranah, baik ranah kognitif, afektif dan psikomotorik.

Cara melakukan pengujian validitasnya, yaitu dengan melihat kesesuaian konstruksi butir yang ditulis dengan hasil belajar, indikator dan butir-butir instrumen yang direncanakan di dalam kisi-kisinya (Purwanto, 2011:128).

2) Validitas Empiris

Istilah “validitas empiris” memuat kata “empiris” yang berarti pengalaman. Suatu instrumen dikatakan memiliki validitas empiris apabila sudah diuji dari pengalaman. Validitas empiris terdiri atas:

a) Tingkat Kesukaran

Menurut Arikunto (2013:225), soal yang baik adalah soal yang tidak terlalu mudah atau tidak terlalu sukar. Soal yang terlalu mudah tidak merangsang peserta didik untuk mempertinggi usaha memecahkannya. Sebaliknya soal yang terlalu sukar dapat menyebabkan peserta didik menjadi putus asa dan tidak mempunyai semangat untuk mencoba lagi karena di luar jangkauannya.

Bilangan yang menunjukkan sukar dan mudahnya suatu soal disebut indeks kesukaran (*difficulty index*). Di dalam istilah evaluasi, indeks kesukaran ini diberi simbol P (p besar), singkatan dari kata “proporsi”. Besarnya tingkat kesukaran antara 0,00 sampai dengan 1,0. Menurut Purwanto (2011:101), nilai indeks tingkat kesukaran diklasifikasikan sebagai berikut:

Soal dengan P 0,00 sampai 0,19 adalah soal sangat sukar.

Soal dengan P 0,20 sampai 0,39 adalah soal sukar.

Soal dengan P 0,40 sampai 0,59 adalah soal sedang.

Soal dengan P 0,60 sampai 0,79 adalah soal mudah.

Soal dengan P 0,80 sampai 1,00 adalah soal sangat mudah.

b) Daya Pembeda

Daya pembeda soal adalah kemampuan suatu soal untuk membedakan antara peserta didik yang pandai (berkemampuan tinggi) dengan peserta didik yang bodoh (berkemampuan rendah). Angka yang menunjukkan besarnya daya pembeda disebut indeks diskriminasi, disingkat D (d besar). Indeks diskriminasi (daya pembeda) ini berkisar antara 0,00 sampai 1,00. Hanya bedanya, indeks kesukaran tidak mengenal tanda negatif (-), tetapi pada indeks diskriminasi ada tanda negatif. Tanda negatif pada indeks diskriminasi digunakan jika suatu soal “terbalik” menunjukkan kualitas *testee*. Yaitu anak pandai disebut bodoh dan anak bodoh disebut pandai (Arikunto, 2013:226).

Adapun klasifikasi nilai daya pembeda soal sebagai berikut:

D : 0,00 – 0,20 : jelek (*poor*)

D : 0,21 – 0,40 : cukup (*satisfactory*)

D : 0,41 – 0,70 : baik (*good*)

D : 0,71 – 1,00 : baik sekali (*excellent*)

D : negatif, adalah daya pembeda sangat jelek dan soal dibuang saja.

c) Analisis Pengecoh (*distractor*)

Pada soal pilihan ganda terdapat alternatif jawaban (*option*) yang disebut sebagai pengecoh (*distractor*). Pengecoh yang tidak dipilih sama sekali oleh *testee* berarti bahwa pengecoh itu jelek, terlalu menyolok menyesatkan. Sebaliknya sebuah pengecoh dapat dikatakan berfungsi dengan baik, apabila pengecoh tersebut mempunyai daya tarik yang besar

bagi pengikut-pengikut tes yang kurang memahami konsep atau kurang menguasai bahan. Suatu *distractor* dapat dikatakan berfungsi baik jika paling sedikit dipilih oleh 5% dari pengikut tes (Arikunto, 2013:233).

b. Reliabilitas

Reliability berasal dari kata “*rely*” yang artinya percaya dan *reliable* yang artinya dapat dipercaya. Sebagai alat ukur, tes hasil belajar (THB) harus memenuhi persyaratan reliabilitas. Keterpercayaan berhubungan dengan ketepatan dan konsistensi, THB dikatakan dapat dipercaya apabila memberikan hasil pengukuran hasil belajar yang relatif tetap dan konsisten (Purwanto, 2011:153). Sedangkan menurut Arifin (2009:258), reliabilitas adalah tingkat atau derajat konsistensi dari suatu instrumen. Reliabilitas tes berkenaan dengan pertanyaan, apakah suatu tes teliti dan dapat dipercaya sesuai dengan kriteria yang telah ditetapkan. Suatu tes dapat dikatakan reliabel jika selalu memberikan hasil yang sama apabila diteskan pada kelompok yang sama pada waktu atau kesempatan yang berbeda.

Menurut Gronlund yang diacu Arifin (2009:258), mengemukakan ada dua faktor yang dapat mempengaruhi tinggi-rendahnya reliabilitas, yaitu panjang tes dan tingkat kesukaran. Panjang tes berarti banyaknya soal tes, ada kecenderungan semakin panjang suatu tes akan lebih tinggi tingkat reliabilitas tes karena semakin banyak soal maka akan semakin banyak sampel yang diukur dan proporsi jawaban yang benar semakin banyak sehingga faktor tebakan akan semakin rendah. Sedangkan tingkat kesukaran soal yang baik dan ideal adalah soal yang tidak terlalu mudah dan tidak terlalu sukar. Tinggi

rendahnya reliabilitas, secara empirik ditunjukkan oleh suatu angka yang disebut koefisien reliabilitas. Tinggi rendahnya reliabilitas dicerminkan oleh tinggi rendahnya korelasi antara dua distribusi skor dari dua alat ukur yang paralel yang dikenakan pada kelompok individu yang sama.

Adapun pengertian reliabilitas menurut Arikunto (2013:104), adalah ketetapan suatu tes apabila diteskan pada subyek yang sama. Nilai dari reliabilitas ini diberi simbol dengan huruf “r” (r kecil). Adapun klasifikasi nilai reliabilitas sebagai berikut:

$0,800 < r \leq 1,000$: sangat tinggi

$0,600 < r \leq 0,800$: tinggi

$0,400 < r \leq 0,600$: cukup

$0,200 < r \leq 0,400$: rendah

$0,000 < r \leq 0,200$: sangat rendah.

6. Tes Sebagai Hasil Belajar Kognitif

Hasil belajar kognitif adalah perubahan perilaku yang terjadi dalam kawasan kognisi. Proses belajar kognitif meliputi kegiatan dari penerimaan stimulus eksternal oleh sensori, penyimpanan dan pengolahan dalam otak menjadi informasi hingga pemanggilan kembali informasi ketika diperlukan untuk menyelesaikan masalah (Purwanto, 2011:50).

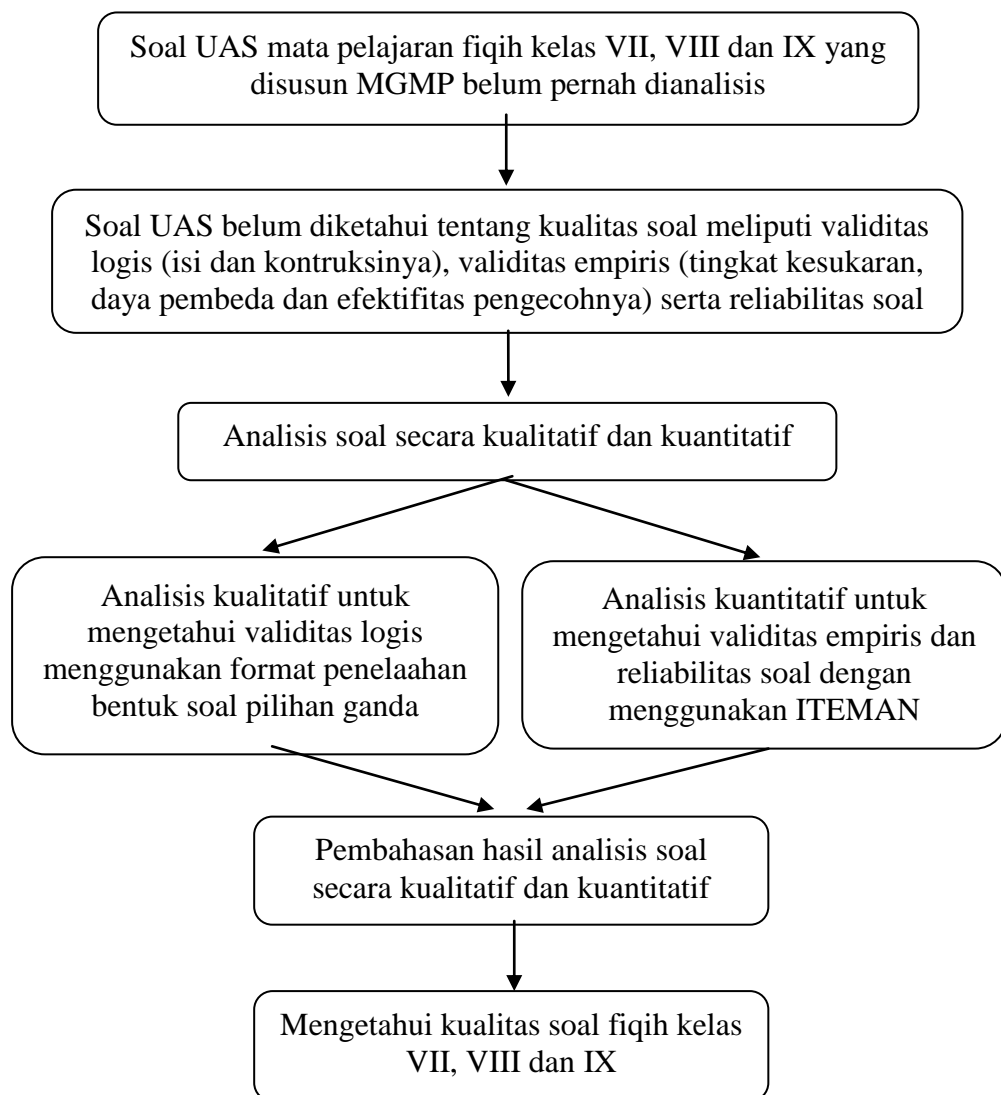
Hasil belajar kognitif merupakan kemampuan yang menimbulkan perubahan perilaku dalam domain kognitif meliputi beberapa tingkat atau jenjang. Banyak klasifikasi dibuat para ahli psikologi dan pendidikan, namun klasifikasi yang paling banyak digunakan adalah yang dibuat oleh Benjamin

S. Bloom (Arikunto, 2013:130). Bloom membagi dan menyusun secara hirarki tingkat hasil belajar kognitif mulai dari yang paling rendah dan sederhana yaitu hapalan sampai yang paling tinggi dan kompleks yaitu evaluasi. Keenam tingkatan kognitif tersebut sebagai berikut:

- a. Hapalan atau *recognition* (C1), kemampuan menghafal merupakan kemampuan kognitif yang paling rendah. Kemampuan ini merupakan kemampuan memanggil kembali fakta yang disimpan dalam otak digunakan untuk merespon suatu masalah.
- b. Pemahaman atau *comprehension* (C2), kemampuan pemahaman adalah kemampuan untuk melihat hubungan fakta dengan fakta. Menghafal fakta tidak lagi cukup karena pemahaman menuntut kemampuan akan fakta dan hubungannya.
- c. Penerapan atau *application* (C3), kemampuan penerapan adalah kemampuan kognitif untuk memahami aturan, hukum, rumus dan sebagainya dan digunakan untuk memecahkan suatu masalah.
- d. Analisis atau *analysis* (C3), kemampuan analisis adalah kemampuan memahami sesuatu dengan menguraikannya ke dalam unsur-unsur.
- e. Sintesis atau *synthesis* (C4), kemampuan sintesis adalah kemampuan memahami dengan mengorganisasikan bagian-bagian ke dalam kesatuan.
- f. Evaluasi atau *evaluation* (C6), kemampuan evaluasi adalah kemampuan membuat penilaian dan mengambil keputusan dari hasil penilaiannya.

F. Kerangka Berfikir

Dengan dilakukannya analisis ulangan akhir semester gasal mata pelajaran fiqih kelas VII, VIII dan IX di MTs. Negeri Nglipar yang disusun oleh MGMP Fiqih, diharapkan dapat diketahui validitas logis, validitas empiris dan reliabilitasnya. Memperhatikan uraian pada landasan teori, maka dapat disusun kerangka berfikir sebagai berikut:



G. Metode Penelitian

1. Jenis Penelitian

Jenis penelitian ini adalah penelitian lapangan (*field reseach*) yaitu penelitian yang mengungkapkan fakta di lapangan dengan lokasi di MTs. Negeri Nglipar. Penelitian ini dilakukan dalam situasi alamiah, akan tetapi didahului oleh semacam intervensi (campur tangan) dari pihak peneliti. Intervensi ini dimaksudkan agar fenomena yang dikehendaki oleh peneliti dapat segera tampak dan diamati (Azwar, 2010:21).

Ditinjau dari hadirnya peneliti, penelitian ini termasuk penelitian deskriptif. Penelitian deskriptif bertujuan menggambarkan secara sistematis, akurat fakta dan karakteristik mengenai populasi atau mengenai hal tertentu, yang hasilnya dipaparkan dalam bentuk laporan penelitian (Azwar, 2010:7). Adapun pendekatan dalam penelitian ini, berupa pendekatan kuantitatif, karena data-data penelitian berupa angka-angka (*numerical*) dan dianalisis dengan statistik menggunakan program *Item and test Analysis* (ITEMAN).

2. Penegasan Konsep

Untuk menghindari kesalahan persepsi dan perbedaan konsep, serta untuk menjelaskan variabel yang terdapat dalam penelitian ini, maka perlu adanya penegasan konsep/variabel dalam penelitian ini.

a. Analisis kualitas butir soal

Analisis soal adalah kegiatan menganalisis butir-butir soal sehingga dapat diketahui soal-soal mana yang perlu diperbaiki, diseleksi, diganti atau direvisi (Arifin, 1991:128). Analisis soal dalam penelitian ini merupakan prosedur yang sistematis untuk mengkaji kualitas pertanyaan dalam tes dari jawaban peserta didik, meliputi indikator-indikator sebagai berikut:

- 1) Validitas logis (isi & konstruksi).
- 2) Validitas empiris (tingkat kesukaran, daya pembeda dan efektifitas pengecoh).
- 3) Reliabilitas butir soal.

b. Ulangan akhir semester gasal mata pelajaran fiqih di MTs. Negeri Nglipar

Dalam penelitian ini yang dimaksud dengan ulangan akhir semester gasal mata pelajaran fiqih di MTs. Negeri Nglipar, merupakan ulangan akhir semester gasal mata pelajaran fiqih mulai dari kelas VII, VIII dan IX, yang dilaksanakan pada hari selasa, tanggal 3 Desember 2013.

3. Lokasi dan Waktu Penelitian

Penelitian ini dilaksanakan di MTs. Negeri Nglipar, yang berlokasi di Dusun Blembeman, Desa Natah, Kecamatan Nglipar, Gunungkidul. Waktu pelaksanaan penelitian ini setelah dilaksanakannya ulangan akhir semester gasal dan remidi, yaitu pada bulan Desember 2013.

4. Populasi dan Sampel

Populasi adalah keseluruhan subyek penelitian (Arikunto, 2010:173). Sedangkan menurut Azwar (2010:77), populasi didefinisikan sebagai kelompok subjek yang hendak dikenai generalisasi hasil penelitian. Sebagai suatu populasi, kelompok subjek ini harus memiliki ciri-ciri atau karakteristik-karakteristik bersama yang membedakannya dari kelompok subjek yang lain. Atas dasar di atas, maka populasi dalam penelitian ini adalah peserta didik di MTs. Negeri Nglipar, yaitu sebagai berikut:

Tabel 1.

Daftar jumlah peserta didik yang menjadi populasi penelitian

No.	Kelas	Jumlah Peserta didik
1.	VII A	17
2.	VII B	18
3.	VII C	17
4.	VIII A	22
5.	VIII B	22
6.	VIII C	23
7.	IX A	17
8.	IX B	18
9.	IX C	18
Jumlah		172

Untuk menentukan sampel yang representatif maka penulis menggunakan *purposive sample*. Menurut Arikunto (2013:183), *purposive sample* atau sampel bertujuan dilakukan dengan mengambil subjek bukan didasarkan atas strata, random atau daerah tetapi didasarkan atas adanya

tujuan tertentu. Oleh karena itu dalam penelitian ini yang menjadi *purposive sample* adalah kelas VII B, VIII C dan IX C.

5. Objek Penelitian

Adapun objek yang dianalisis dalam penelitian ini adalah soal UAS, kunci jawaban dan lembar jawab peserta didik kelas VII B, VIII C dan IX C mata pelajaran fiqih di MTs. Negeri Nglipar. Objek pada populasi selanjutnya diteliti, hasilnya dianalisis, disimpulkan dan kesimpulan itu berlaku untuk seluruh populasi (Arikunto, 2010:174).

6. Metode Pengumpulan Data

Pengumpulan data dalam penelitian ini berupa *interview* (wawancara) dan dokumentasi.

a. *Interview* (wawancara)

Jenis *interview* yang digunakan dalam penelitian ini adalah *inguided interview* (wawancara bebas). Wawancara bebas adalah suatu proses *interview* dimana pewawancara bebas menanyakan apa saja, tetapi juga mengingat akan data apa yang akan disampaikan (Arikunto, 2010:199).

Wawancara dalam penelitian ini dilaksanakan pada hari selasa, tanggal 17 Desember 2013, yaitu dengan Bapak Suparno S.Pd.I, M.SI. selaku guru mata pelajaran fiqih sekaligus sebagai wakil ketua MGMP fiqih Kab. Gunungkidul dan beberapa siswa kelas VII B, VIII C dan IX C, yang telah mengikuti UAS semester gasal di MTs. Negeri Nglipar.

b. Dokumentasi

Dokumentasi, yaitu mencari data mengenai hal-hal atau variabel yang berupa catatan, transkrip, buku, surat kabar, majalah, prasasti, notulen rapat, lengger, agenda dan sebagainya (Arikunto, 2010:274). Dokumentasi yang digunakan untuk memperoleh data dalam penelitian ini antara lain:

- 1) Lembar soal dan lembar kunci jawaban ulangan akhir semester gasal kelas VII, VIII dan IX mata pelajaran fiqih MTs. Negeri Nglipar.
- 2) Lembar jawaban peserta didik terhadap soal ulangan akhir semester gasal kelas VII, VIII dan IX mata pelajaran fiqih MTs. Negeri Nglipar.
- 3) Silabus, SK/KD dan kisi-kisi soal ulangan akhir semester gasal kelas VII, VIII dan IX mata pelajaran fiqih MTs. Negeri Nglipar.
- 4) Gambaran umum MTs. Negeri Nglipar tahun ajaran 2013/2014.

7. Analisis Data

a. Analisis Secara Kualitatif (*logical validity*)

Ghofur, dkk (2004:93), menjelaskan bahwa cara analisis dengan pendekatan kualitatif (*logical validity*), yaitu dengan cara mencermati butir-butir soal yang telah disusun dilihat dari kesesuaiannya dengan kisi-kisi soal dalam pemenuhan persyaratan baik aspek materi, konstruksi dan bahasa. Aspek yang harus diperhatikan diantaranya:

1) Aspek Materi

- a) Butir soal sesuai dengan indikator pembelajaran
- b) Indikator pembelajaran sesuai dengan kompetensi dasar
- c) Materi soal sesuai dengan jenjang kelas pada semester yang dicapai
- d) Hanya ada satu kunci jawaban yang benar
- e) Soal mengacu pada ranah kognitif

2) Aspek Kontruksi

- a) Pokok soal dirumuskan dengan singkat, jelas dan tegas
- b) Rumusan pokok soal dan pilihan jawaban merupakan pernyataan yang diperlukan saja
- c) Pokok soal tidak memberi petunjuk kunci jawaban
- d) Pokok soal dalam pernyataan negatif maka kata negatif tersebut di garis bawah atau ditulis tebal
- e) Pilihan jawaban homogen dan logis ditinjau dari materi
- f) Tulisan Arab dalam bentuk surat/ayat al-Quran dan al-Hadist terbaca, jelas dan berfungsi
- g) Panjang pilihan jawaban relatif sama
- h) Pilihan jawaban tidak menggunakan pernyataan "semua jawaban diatas benar/salah" dan sejenisnya
- i) Pilihan jawaban dalam bentuk angka/waktu diurutkan berdasarkan besar kecilnya atau kronologisnya
- j) Butir soal tidak bergantung pada jawaban soal sebelumnya

3) Bahasa/Budaya

- a) Menggunakan bahasa yang umum, bukan tabu atau berlaku di daerah setempat
- b) Menggunakan bahasa Indonesia yang baku sesuai dengan kaidah bahasa Indonesia
- c) Menggunakan bahasa yang komunikatif

Berdasarkan format penelaahan soal bentuk pilihan ganda di atas terdapat 18 kaidah yang telah ditentukan, dengan 5 kaidah pada aspek materi, 10 kaidah pada aspek konstruksi dan 3 kaidah pada aspek bahasa/budaya. Butir soal bentuk pilihan ganda dikatakan baik, apabila memenuhi keseluruhan kaidah-kaidah yang telah ditentukan dari ketiga aspek, yaitu aspek materi, konstruksi dan bahasa.

b. Analisis Kuantitatif

Menurut Purwanto (2011:97), analisis butir soal dapat dilakukan dengan salah satu dari dua cara tergantung teori yang digunakan. Teori analisis tes tersebut dapat berupa teori klasik atau modern. Teori tes klasik masih sering digunakan karena penggunaannya yang lebih mudah, disamping teori tes modern yang masih dalam proses pengembangan. Oleh karena itu analisis tes pada penelitian ini berupa teori tes klasik dengan menggunakan perangkat lunak (*software*), yaitu *item and test analysis* (ITEMAN).

Item dan Analisis Item (ITEMAN) merupakan perangkat lunak (*software*) yang dibuat melalui bahasa pemrograman komputer yang

diciptakan khusus untuk analisis statistik butir soal dan tes. Program ITEMAN lebih sederhana dan mudah digunakan bila dibandingkan dengan program lain (Abidin, 2008).

Program ITEMAN juga memberikan hasil skor untuk setiap peserta tes yang menunjukkan jumlah benar dari seluruh jawaban. Sebelum menggunakan program ITEMAN perlu diketahui bahwa terdapat 5 baris utama yang harus *dientrykan*. Data yang akan dianalisis diketik melalui *notepad* atau *Microsoft Office Word* dengan jenis font *Courier New*. File data yang akan *dientrykan* ke program ITEMAN terdiri atas 5 baris yaitu:

- 1) Baris pertama adalah baris pengontrol yang mendeskripsikan data
- 2) Baris kedua adalah daftar kunci jawaban setiap butir soal
- 3) Baris ketiga adalah daftar jumlah option untuk setiap butir soal
- 4) Baris keempat adalah daftar butir soal yang akan dianalisis (jika butir yang akan dianalisis diberi tanda Y, jika tidak dianalisis diberi tanda N)
- 5) Baris kelima dan seterusnya adalah data peserta didik dan jawaban pilihan peserta didik. Setiap pilihan jawaban peserta didik (untuk soal pilihan ganda) diketik dengan menggunakan huruf, misal ABCD atau 1234 untuk 4 pilihan jawaban, sedangkan untuk 5 pilihan jawaban yaitu ABCDE atau 12345.

Adapun langkah-langkah melakukan analisis soal dengan ITEMAN:

- 1) Membuat file data, file data ditulis dengan *notepad* atau *Microsoft Office Word* dengan jenis font *Courier New*).

Contoh:

025 o N 10

AABCDDCBBDDCCBCDDABACBCAB

44444444444444444444444444444444

YYYYYYYYYYYYYYYYYYYYYYYYYYYYYY

Agus AACBBDCBBAABBCBCDDCAABDBC

Andi ABACDDCCBBABCCDBACDACBBCC

Diah AADDCBBCADDDADCCBBCAACBDC

Elga BABBDDDACCBBADDAADCCBDDBD

Erna ABBDACDCBBCBCAADDCCBBCDAA

Keterangan:

Baris 1, meliputi jumlah soal, kode omit, kode tidak dijawab, jumlah karakter peserta didik.

Baris 2, berupa *option*/jawaban ditulis dengan angka atau huruf.

Baris 3, jumlah dari *option*/pilihan (a, b, c dan d).

Baris 4, soal yang akan dianalisis, bila tidak dianalisis ditulis N.

Baris 5, jawaban dari peserta didik bisa angka/huruf.

Langkah selanjutnya, simpan file notepad, misalnya FIQIH.

2) Menjalankan Program ITEMAN

- a) Double klik file program ITEMAN.
- b) Tulislah file data: contoh FIQIH.TXT, kemudian tekan enter.
- c) Ketik nama file hasil analisis, contoh HSLFIQIH.TXT, kemudian tekan enter.
- d) Ketik 'Y', kemudian tekan enter.
- e) Ketik file untuk total skor peserta didik, contoh SKORFIQIH.TXT, kemudian tekan enter.
- f) Analisis selesai.

3) Interpretasi Hasil Analisis

Hasil analisis dengan ITEMAN dapat dibedakan menjadi 2 bagian, yaitu statistik butir soal dan hasil analisis statistik tes/skala. Statistik butir soal adalah untuk tes yang terdiri dari butir-butir soal yang bersifat dikotomi misalnya pilihan ganda. Statistik berikut adalah output dari setiap butir soal yang dianalisis:

- a) *Seq.N* : adalah nomor urut butir soal dalam file data
- b) *Scala item* : nomor urut butir soal dalam tes
- c) *Prop.Correct* : proporsi peserta didik yang menjawab benar butir tes (indeks tingkat kesukaran soal secara klasikal).
- d) *Biser* : indeks daya pembeda soal dengan menggunakan koefisien korelasi biserial.

- e) *Point biserial* : juga indeks daya pembeda soal dan pilihan jawaban (alternatif) dengan menggunakan koefisien point biserial. Penafsirannya sama dengan statistik biserial.
- f) Statistik pilihan jawaban (alternatif) memberikan informasi yang sama dengan statistik butir soal. Perbedaannya adalah bahwa statistik pilihan jawaban dihitung secara terpisah.

Sedangkan hasil interpretasi statistik tesnya, sebagai berikut:

- a) *N of Items* : jumlah butir soal dalam tes yang dianalisis.
- b) *N of Examinees*: jumlah peserta tes
- c) *Mean* : skor atau rata-rata peserta tes
- d) *Variance* : varian dari distribusi skor peserta tes yang memberikan gambaran tentang sebaran skor peserta tes.
- e) *Std.Deviasi* : deviasi standar dari distribusi skor tes (akar dari varians)
- f) *Skew* : kemiringan distribusi skor peserta tes yang memberikan gambaran tentang bentuk distribusi skor peserta tes.
- g) *Kurtosis* : puncak distribusi skor yang menggambarkan kelandaian distribusi skor dibanding dengan distribusi normal.
- h) *Minimum* : skor terendah peserta tes
- i) *Maximum* : skor tertinggi peserta tes
- j) *Median* : skor tengah dimana 50% berada pada atau lebih rendah dari skor tersebut.
- k) *Alpha* : koefisien reliabilitas alpha untuk tes atau skala tersebut yang merupakan indeks homogenitas tes atau skala. Semakin besar reliabilitas

(alpha) berarti kualitas instrumen semakin kuat dalam menunjukkan tingkat keajegan atau konsistenan hasil jika instrumen diulang.

- l) *SEM* : kesalahan pengukuran standar untuk setiap tes atau skala. *SEM* merupakan estimasi dari deviasi standar kesalahan pengukuran dalam skor tes.
- m) *Mean P* : rerata tingkat kesukaran semua butir soal dalam tes secara klasikal dihitung dengan cara mencari rata-rata proporsi peserta tes yang menjawab benar untuk semua butir soal tes.
- n) *Mean item tot* : nilai rata-rata indeks daya pembeda dari semua soal dalam tes yang diperoleh dengan menghitung nilai rata-rata point biserial dari semua soal dalam tes
- o) *Mean biserial* : nilai rata-rata indek daya pembeda yang diperoleh dengan menghitung nilai rata-rata korelasi biserial dari semua butir soal.

H. Sistematika Pembahasan

Sistematika pembahasan dalam penelitian ini sebagai berikut:

Bagian awal meliputi: Halaman sampul, halaman judul, pernyataan keaslian, nota dinas, halaman pengesahan, halaman motto, halaman persembahan, kata pengantar, daftar isi, daftar tabel, daftar gambar, daftar lampiran dan abstrak.

Bagian utama pada penelitian ini ada empat bab, yaitu:

BAB I Berisi pendahuluan, meliputi latar belakang masalah, rumusan masalah, tujuan dan kegunaan penelitian, tinjauan pustaka, landasan teori, kerangka berfikir, metode penelitian dan sistematika pembahasan.

BAB II Gambaran umum MTs. Negeri Nglipar, meliputi letak geografis, sejarah berdirinya, visi, misi dan tujuan, data guru, karyawan dan peserta didik serta sarana dan prasarana.

BAB III Hasil penelitian dan pembahasan meliputi pemaparan data hasil penelitian dan interpretasi hasil analisis kualitas butir soal UAS fiqih MTs. Negeri Nglipar.

BAB IV Penutup, pada bagian ini berisi kesimpulan dan saran-saran.

Kemudian bagian akhir pada penelitian ini, meliputi daftar pustaka dan lampiran-lampiran.