

BAB II

LANDASAN TEORI

A. Deskripsi Teori

1. Tes Hasil Belajar

a. Pengertian

Tes merupakan alat ukur untuk proses pengumpulan data di mana dalam memberikan respon atas pertanyaan dalam instrumen, peserta didorong untuk menunjukkan kemampuan maksimalnya. Peserta diharuskan mengeluarkan kemampuan semaksimal mungkin agar data yang diperoleh dari hasil jawaban peserta didik benar-benar menunjukkan kemampuannya.⁷

Tes hasil belajar juga merupakan tes penguasaan, karena tes ini berfungsi mengukur penguasaan peserta didik terhadap materi yang diajarkan oleh guru atau dipelajari oleh peserta didik. Tes diujikan setelah peserta didik memperoleh sejumlah materi sebelumnya dan pengujian dilakukan untuk mengetahui penguasaan peserta didik atas materi tersebut.⁸ Karenanya, tes hasil belajar yang baik harus Mampu mengukur kemampuan peserta didik dalam memahami materi-materi yang diajarkan. Terkait dengan

⁷ Purwanto, *Evaluasi Hasil Belajar*, (Yogyakarta: Pustaka Pelajar, 2009), hlm. 64.

⁸ Purwanto, *Evaluasi Hasil Belajar*, hlm. 66.

evaluasi tes hasil belajar tersebut akan mengukur nilai dan efektifitas dari bagian tertentu dalam pendidikan.⁹

Dalam dunia pendidikan, tes hasil belajar adalah kegiatan yang sering dilakukan. Tes hasil belajar dilakukan untuk mengukur sejauh mana kemampuan peserta didik dalam memahami materi-materi pembelajaran.

Tes hasil belajar merupakan sumber data bagi guru untuk mengetahui berapakah nilai peserta didik. Tes hasil belajar juga dapat dijadikan sebagai evaluasi bagi guru maupun pihak sekolah. Dengan tes tersebut peserta didik dapat mengetahui dimana posisinya jika dibandingkan dengan teman-temannya.

Ujian madrasah merupakan salah satu bentuk tes hasil belajar. Ujian madrasah dilaksanakan ketika peserta didik telah menyelesaikan materi-materi pembelajaran di sekolah/madrasah. Ujian ini dilaksanakan sebelum peserta didik meninggalkan sekolah.

b. Macam

Tes hasil belajar dapat dikelompokkan menjadi beberapa macam. Menurut peranan fungsionalnya dalam pembelajaran, tes hasil belajar dibagi menjadi empat macam, yaitu:¹⁰

⁹ A. V. Kelly, *The Curriculum Theory and Practice*, (London, Sage Production, 2006), hlm. 137

¹⁰ Purwanto, *Evaluasi Hasil Belajar*, hlm. 67-69.

1) Tes Formatif

Tes formatif digunakan untuk mengetahui sejauh mana kemampuan peserta didik setelah mengikuti proses belajar mengajar. Tes formatif diujikan setelah peserta didik menyelesaikan materi-materi tertentu. Tes formatif dalam praktik pembelajaran dikenal sebagai ulangan harian.

2) Tes Sumatif

Tes sumatif merupakan sebagai tes yang digunakan untuk mengetahui penguasaan siswa atas semua jumlah materi yang disampaikan dalam satuan kurun waktu tertentu seperti caturwulan atau semester. Dalam praktik pengajaran tes sumatif dikenal sebagai ujian akhir semester atau caturwulan tergantung satuan waktu yang digunakan untuk menyelesaikan materi.

3) Tes Diagnostik

Evaluasi hasil belajar mempunyai fungsi diagnostik. Tes hasil belajar yang digunakan sebagai dasar untuk melakukan evaluasi diagnostik adalah tes diagnostik. Dalam evaluasi diagnostik, tes hasil belajar digunakan untuk mengidentifikasi siswa-siswa yang mengalami masalah dan menelusuri jenis masalah yang dihadapi.

4) Tes Penempatan

Tes penempatan (placement test) adalah tes hasil belajar yang dilakukan untuk menempatkan peserta didik dalam kelompok yang sesuai dengan kemampuan ataupun bakat minatnya. Pengelompokan dilakukan agar pemberian layanan pembelajaran dapat dilakukan sesuai kemampuan maupun bakat minat peserta didik. Dalam praktik pembelajaran penempatan merupakan hal yang banyak dilakukan, misalnya tes penempatan peserta didik ke dalam kelompok IPA, IPS, atau Bahasa.

Ujian madrasah tergolong dalam tes sumatif. Dikatakan tes sumatif karena ujian madrasah ini dimaksudkan untuk mengetahui penguasaan peserta didik terhadap materi yang telah disampaikan dalam kurun waktu tertentu. Dalam jenjang Madrasah Aliyah, ujian madrasah dilaksanakan ketika peserta didik sudah kelas XII. Ujian madrasah dilaksanakan sebelum peserta didik melaksanakan Ujian Nasional sebagai syarat kelulusan

c. Bentuk

Berdasarkan bentuk pertanyaannya, tes hasil belajar dapat berbentuk tes objektif dan esai. Tes esai adalah suatu bentuk tes yang terdiri dari pertanyaan atau yang menghendaki jawaban yang berupa uraian-uraian. Tes

dirancang agar peserta didik mampu menjawab pertanyaan dengan susunan kalimat disusun sendiri oleh peserta didik.

Tes objektif adalah tes keseluruhan informasi yang diperlukan untuk menjawab tes telah tersedia. Butir soal pada tes objektif mengandung jawaban yang harus dipilih oleh siswa. Kemungkinan jawaban telah dipasok oleh pengkonstruksi tes dan peserta hanya memilih jawaban dari kemungkinan jawaban yang telah disediakan.

Tes objektif mempunyai beberapa keunggulan. Pertama, penilaiannya yang sangat objektif. Sebuah pertanyaan hanya mempunyai dua kemungkinan, benar atau salah. Kunci jawaban memberikan informasi apakah jawaban siswa benar atau salah. Namun tes objektif juga mempunyai beberapa kelemahan. Pertama, tes objektif diragukan kemampuannya untuk mengukur hasil belajar yang kompleks. Kedua, peluang peserta didik melakukan tebakan sangat tinggi.¹¹

Bentuk pertanyaan yang digunakan dalam ujian madrasah menggunakan soal pilihan ganda. Dalam soal pilihan ganda tersebut terdapat soal dan pilihan jawaban. Dari pilihan jawaban yang disediakan, terdapat satu pilihan jawaban yang benar dan yang lainnya merupakan pilihan jawaban salah. Pilihan jawaban salah tersebut dikenal sebagai pengecoh.

¹¹ Purwanto, Evaluasi Hasil Belajar, hlm. 70-71.

d. Komponen

Tes hasil belajar mempunyai beberapa komponen. Pada tes hasil belajar bentuk esai, komponen dapat berupa perangkat soal, petunjuk pengerjaan, dan soal. Lebih dari itu, tes objektif mempunyai sejumlah komponen selain yang ada dalam tes esai, yaitu pilihan, kunci jawaban, dan pengecoh. Masing-masing komponen dibahas berikut:¹²

- 1) Perangkat soal, perangkat soal adalah keseluruhan butir pertanyaan atau pertanyaan berikut segala kelengkapannya.
- 2) Petunjuk pengerjaan, petunjuk pengerjaan mendeskripsikan detail petunjuk yang harus dilakukan dalam mengerjakan soal, misalnya: memberikan tanda silang, melingkari, memberikan jawaban singkat, dan sebagainya.
- 3) Butir soal, soal merupakan pertanyaan atau pernyataan yang menimbulkan situasi masalah yang harus dipecahkan oleh siswa. Penguasaan siswa diketahui dari kemampuannya membuat pemecahan masalah. Satuan untuk soal adalah butir sehingga tiap item pertanyaan atau pernyataan dikenal sebagai butir soal.
- 4) Pilihan, soal objektif adalah soal yang segala kemungkinan jawaban telah disediakan dan tugas peserta tes adalah memilih satu pilihan yang merupakan

¹² Purwanto, Evaluasi Hasil Belajar, hlm. 74-75.

jawaban atas pertanyaan. Sejumlah alternatif yang ditawarkan dinamakan pilihan (options).

- 5) Kunci jawaban, kunci jawaban adalah pilihan yang merupakan jawaban atas pertanyaan yang diajukan dalam soal.
- 6) Pengecoh, pengecoh adalah pilihan yang bukan merupakan kunci jawaban. Misalnya: pada soal objektif jenis benar-salah, bila kunci jawabannya adalah salah maka benar merupakan pengecoh. Pada soal objektif pilihan ganda dengan empat pilihan a, b, c, d dan kunci jawabannya adalah c maka a, b, d merupakan pengecoh.

2. Kualitas Tes

Tes yang digunakan untuk mengukur kemampuan peserta didik haruslah tes yang berkualitas baik. Tes yang berkualitas baik tentunya mampu menjadi tolok ukur yang baik untuk mengukur kemampuan peserta didik. Untuk mengetahui bagaimanakah kualitas tes, maka dilakukan analisis kualitas tes. Dengan analisis kualitas tes kita dapat mengetahui bagaimana kondisi soal yang digunakan untuk tes.

Setidaknya terdapat empat karakteristik yang harus dimiliki oleh tes hasil belajar, sehingga tes tersebut dapat dinyatakan sebagai tes yang baik. Keempat karakteristik tersebut adalah sebagai berikut:¹³

¹³ Anas Sudijono, Pengantar Evaluasi Pendidikan, hlm. 93.

a. Validitas

Valid diartikan dengan istilah tepat, benar, shahih, absah. Jadi kata validitas sering diartikan dengan ketepatan, kebenaran, atau kesahihan, atau keabsahan.

Sebuah tes dikatakan valid apabila tes tersebut dengan secara tepat, secara benar, secara shahih, atau secara absah dapat mengukur apa yang seharusnya diukur.

b. Reliabilitas

Tes hasil belajar yang baik memiliki reliabilitas atau bersifat reliabel. Reliabilitas diartikan sebagai keajegan atau kemantapan. Sebuah tes dinyatakan reliabel apabila hasil-hasil pengukuran yang dilakukan dengan menggunakan tes tersebut secara berulang kali terhadap subyek yang sama senantiasa menunjukkan hasil yang tetap sama atau sifatnya ajeg dan stabil.

Keandalan mengacu pada konsistensi atau presisi atau ketergantungan dari pengukuran penilaian. Bagaimana hasil yang konsisten yang satu dengan yang lainnya.

c. Obyektif

Tes hasil belajar dikatakan obyektif bila tes tersebut disusun dan dilaksanakan dengan apa adanya. Dalam penyusunan, tes disusun berdasarkan materi dan bahan pelajaran yang diajarkan. Dalam pemberian skor juga apa adanya dan tidak ada subyektivitas dalam kegiatan tersebut.

d. Praktis

Tes hasil belajar dikatakan praktis jika tes tersebut dilaksanakan secara sederhana dan tidak membutuhkan peralatan maupun persyaratan-persyaratan yang sulit pengadaannya. Pengerjaan soal juga tidak membutuhkan waktu yang lama serta pedoman skoring yang tidak mempersulit pengoreksi.

Dari keempat karakteristik tadi, peneliti menggunakan karakteristik pertama dan kedua yaitu valid dan reliabel untuk penelitian terkait kualitas tes soal ujian madrasah yang telah dilaksanakan di MAN Wonosobo. Dengan diadakannya penelitian tersebut diharapkan kualitas soal pada tes tersebut dapat diketahui.

3. Kualitas Butir Tes

Analisis butir tes hasil belajar dilakukan untuk menguji bagaimanakah keadaan butir-butir soal yang digunakan untuk suatu tes. Pada analisis butir soal, kita menguji keadaan soal-soal tersebut tiap butirnya.

Analisis terhadap butir-butir tes yang telah dijawab oleh peserta didik mempunyai dua tujuan penting. Pertama, jawaban soal itu merupakan informasi diagnostik untuk meneliti pelajaran dari kelas itu dan kegagalan-kegagalan belajarnya, serta selanjutnya untuk membimbing ke arah cara belajar yang lebih baik. Kedua, jawaban-jawaban terhadap soal yang terpisah dan

perbaikan soal-soal yang didasarkan atas jawaban-jawaban itu merupakan basis bagi Penyiapan tes-tes yang lebih baik untuk tahun berikutnya.¹⁴

Ada dua teori yang dapat digunakan untuk melakukan analisis butir soal. Kedua teori tersebut berupa teori klasik dan modern. Teori klasik adalah teori mengenai analisis butir tes di mana analisis dilakukan dengan memperhitungkan kedudukan butir dalam suatu kelas atau kelompok. Sedangkan teori modern yang lebih dikenal dengan teori respon butir dewasa ini dikembangkan untuk mengatasi keterbatasan pada teori klasik. Akan tetapi dalam pelaksanaannya, teori klasik masih sering digunakan karena lebih mudah dalam menggunakannya.

Dalam teori klasik, ada sejumlah karakteristik butir yang diuji yaitu tingkat kesukaran, daya beda, dan fungsi pengecoh. Setiap butir akan diperiksa mutunya dalam tiga karakteristik tersebut. Butir yang baik adalah butir yang mempunyai tingkat kesukaran sedang, daya beda tinggi, dan pengecoh yang berfungsi efektif.¹⁵

a. Tingkat Kesukaran

Tingkat kesukaran didefinisikan sebagai proporsi peserta tes yang menjawab benar. Soal yang baik haruslah tidak terlalu sukar dan terlalu mudah. Soal yang baik

¹⁴ Mulyadi, *Evaluasi Pendidikan*, (Malang: UIN Maliki Press, 2010), hlm. 109.

¹⁵ Purwanto, *Evaluasi Hasil Belajar*, hlm. 97.

setidaknya memiliki tingkat kesukaran yang sedang. Sehingga soal tersebut dapat dikerjakan oleh peserta didik yang pandai dan yang kurang pandai. Untuk menentukan tingkat kesukaran butir dapat dilakukan dengan cara mencari jumlah peserta didik yang menjawab benar dibagi dengan jumlah peserta didik yang mengikuti tes tersebut.

$$TK = \frac{\sum B}{\sum P}$$

Keterangan:

TK = Tingkat Kesukaran

$\sum B$ = Jumlah peserta didik yang menjawab benar

$\sum P$ = Jumlah peserta tes

Nilai tingkat kesukaran berkisar antara 0 sampai 1. Semakin tinggi nilai tingkat kesukaran berarti semakin mudah soal tersebut, semakin rendah nilai tingkat kesukaran berarti semakin sulit soal tersebut. Nilai tingkat kesukaran 0 terjadi apabila seluruh peserta tes tidak ada yang menjawab benar soal tersebut dan nilai tingkat kesukaran 1 terjadi ketika seluruh peserta tes menjawab benar soal tersebut.

Butir soal yang baik memiliki tingkat kesukaran sedang, artinya soal itu tidak terlalu sulit dan tidak terlalu mudah. Untuk menentukan kriteria sedang tergantung jumlah kategori yang diinginkan. Misalnya kategori sukar, sedang, dan mudah maka kriteria sedang adalah antara 0,33

sampai 0,66. Berikut pembagian kategori TK ke dalam tiga kelompok: ¹⁶

Rentang TK	Kategori
0,00 – 0,32	Sukar
0,33 – 0,66	Cukup
0,67 – 1,00	Mudah

b. Daya Beda

Daya beda adalah kemampuan butir soal tes membedakan peserta didik yang mempunyai kemampuan tinggi dan rendah. Butir yang baik mampu membedakan peserta didik yang mempunyai kemampuan tinggi dan kemampuan rendah. Untuk menentukan daya beda, terlebih dahulu kita akan menentukan peserta didik kelompok atas dan kelompok bawah. Untuk menentukannya, kita melihat skor yang didapatkan oleh peserta didik. Peserta didik kelompok atas adalah peserta didik dengan skor tinggi, sedangkan peserta didik kelompok bawah adalah peserta didik dengan skor rendah.

Untuk menentukan nilai daya beda digunakan rumus sebagai berikut: ¹⁷

$$DB = \frac{\sum T_B}{\sum T} - \frac{\sum R_B}{\sum R}$$

¹⁶ Purwanto, Evaluasi Hasil Belajar, hlm. 99.

¹⁷ Purwanto, Evaluasi Hasil Belajar, hlm. 102.

Keterangan,

$\sum T_B$ = jumlah peserta didik yang menjawab benar pada kelompok peserta didik yang mempunyai kemampuan tinggi

$\sum T$ = jumlah kelompok peserta didik yang mempunyai kemampuan tinggi

$\sum R_B$ = jumlah peserta yang menjawab benar pada kelompok peserta didik yang mempunyai kemampuan rendah

$\sum R$ = jumlah kelompok peserta didik yang mempunyai kemampuan rendah.

Nilai daya beda yang baik adalah yang bernilai positif, semakin positif nilai daya beda tersebut, maka butir soal tersebut semakin dapat membedakan antara peserta didik yang pandai dengan yang kurang pandai.

c. Fungsi pengecoh

Butir soal dalam soal pilihan ganda terdiri dari soal dan pilihan jawaban. Pilihan jawaban terdiri dari jawaban-jawaban yang benar dan yang salah. Pilihan jawaban pada tingkat SMP dan SMA biasanya terdiri lima pilihan yaitu a, b, c, d, dan e. Dari kelima pilihan jawaban tersebut terdapat satu jawaban yang benar, sementara empat pilihan jawaban yang lain merupakan jawaban yang salah. Empat pilihan jawaban yang salah tersebut dikenal dengan istilah pengecoh.

Butir soal yang baik adalah butir yang pengecohnya efektif, efektif dalam hal ini berarti pengecoh yang ada telah dipilih oleh peserta tes. Pengecoh yang baik adalah pengecoh yang mirip dengan jawaban, sehingga pengecoh tersebut mampu mengelabui peserta tes untuk memilih pengecoh tersebut. Pengecoh berfungsi dengan baik apabila dari seluruh peserta tes terdapat 5% yang memilih pengecoh tersebut.

d. Validitas Butir

Sebelum tes hasil belajar digunakan untuk mengumpulkan data, terlebih dahulu harus diperiksa bahwa tes hasil belajar telah valid. Hal itu diperlukan untuk menjamin adanya kesesuaian antara tes hasil belajar dengan hasil belajar yang diukur. Pengumpulan data menggunakan tes hasil belajar yang tidak valid menghasilkan data hasil belajar yang tidak valid.

Validitas butir adalah ketepatan suatu butir dapat mengukur secara tepat sesuatu yang ingin diukur. Pengujian terhadap validitas butir ini dilakukan untuk menentukan validitas tiap-tiap butir. Untuk menghitung validitas butir, kita menggunakan korelasi poin biserial.

4. Validitas Tes

Validitas berhubungan dengan kemampuan untuk mengukur secara tepat sesuatu yang ingin diukur.¹⁸ Validitas tes

¹⁸ Purwanto, Evaluasi Hasil Belajar, hlm. 114.

berarti ketepatan soal tes dalam mengukur kemampuan peserta didik. Pengertian lain disebutkan bahwa, “Validity is the extent to which an assessment measure does the job for which it is intended” yang berarti Validitas adalah sejauh mana suatu ukuran penilaian melakukan pekerjaan yang dimaksudkan.¹⁹

Soal yang akan digunakan dalam tes harus sudah diuji validitasnya terlebih dahulu. Hal tersebut dimaksudkan agar soal tersebut benar-benar mampu mengukur kemampuan peserta didik yang akan diukur. Uji validitas soal merupakan hal yang penting, karena soal yang belum diuji validitasnya belum jelas apakah soal tersebut benar-benar mengukur kemampuan peserta didik sesuai materi yang diajarkan atau belum.

Tes hasil belajar yang valid akan memberikan data yang valid, namun tes yang tidak valid juga akan memberikan data yang tidak valid. Oleh karena itu, tes yang digunakan dalam ujian haruslah tes yang valid agar memberikan data yang valid dan tepat.

Validitas isi tidak ditunjukkan oleh suatu indeks, akan tetapi oleh kesamaan butir-butir tes di dalam mewakili tingkah laku yang akan diukur. Suatu tes hasil belajar mempunyai validitas isi yang tinggi apabila butir-butirnya mewakili

¹⁹ Charity Akuadi Okonkwo, “ Sustainable Assessment and Evaluation Strategies”, http://tojde.anadolu.edu.tr/tojde40/articles/article_6.htm, diakses 27 November 2013

kemampuan peserta didik yang akan diukur sesuai dengan apa yang dipelajarinya.²⁰

Dalam penelitian ini, validitas yang akan diuji yaitu validitas isi. Esensi dari validitas ini adalah sejauh manakah suatu alat ukur itu mencerminkan atau menjadi sampel dari bahan atau materi yang akan dijadikan pijakan interpretasi terhadap hasil pengukuran dengan instrumen tersebut. Sejauh mana tingkat validitas isi suatu instrumen bergantung pada sejauh mana aspek materi terwakili dalam alat ukur.²¹

Suatu tes dikatakan mempunyai validitas isi apabila dapat mengukur tujuan khusus tertentu yang sejajar dengan materi atau isi pelajaran. Untuk menguji validitas isi instrumen tes dapat dilakukan dengan membandingkan antara isi instrumen dengan materi pelajaran yang telah diajarkan.²² Validitas isi ditentukan berdasarkan kesesuaian terhadap Standar Kompetensi dan Kompetensi dasar yang digunakan.

5. Reliabilitas Tes

Reliabilitas berhubungan dengan masalah kepercayaan. Suatu tes dapat dikatakan mempunyai taraf kepercayaan yang

²⁰ Dewa Ketut Sukardi dan Desak P. E. Nila Kusmawati, *Analisis Tes Psikologis Tes Psikologis Teori dan Praktik*, (Jakarta, 2009: Rineka Cipta), hlm. 264.

²¹ Shodiq Abdullah, *Evaluasi Pembelajaran*, (Semarang, 2012: Pustaka Rizqi Putra), hlm. 78.

²² Eko Putro Widoyoko, *Evaluasi Program Pembelajaran*, (Yogyakarta, 2009: Pustaka Pelajar), hlm. 129.

tinggi jika tes tersebut dapat memberikan hasil yang tetap. Maka pengertian reliabilitas tes berhubungan dengan ketetapan hasil tes.²³

Tes hasil belajar dikatakan memiliki reliabilitas jika soal tersebut bersifat konsisten ataupun relatif memberikan hasil yang tetap meskipun dikerjakan berkali-kali. Karena itu, soal tes yang baik harus reliabel agar ketika diujikan lagi akan mendapatkan hasil yang relatif sama.

Banyak metode yang dapat dipilih untuk menguji reliabilitas. Metode-metode itu secara garis besar dapat dikelompokkan menjadi dua kelompok berdasarkan perbedaan mendefinisikan reliabilitas.²⁴

- a. Reliabilitas adalah kestabilan hasil pengukuran apabila tes hasil belajar diujikan beberapa kali. Reliabilitas sebagai stabilitas eksternal ini memandang bahwa tes hasil belajar dikatakan reliabel bila diujikan beberapa kali akan memberikan hasil pengukuran yang relatif konsisten. Tergolong dalam kelompok ini adalah metode tes ulang dan metode paralel.
- b. Reliabilitas merupakan konsistensi internal hasil pengukuran butir-butir tes hasil belajar. Tes hasil belajar dikatakan reliabel apabila di antara butir tes hasil belajar memberikan

²³ Suharsimi Arikunto, *Dasar-Dasar Evaluasi Pendidikan*, (Jakarta, 2011: PT. Bumi Aksara), hlm. 86

²⁴ Purwanto, *Evaluasi Hasil Belajar*, hlm. 155.

hasil pengukuran yang konsisten. Metode dalam kelompok ini dapat dibagi menjadi dua berdasarkan jumlah butirnya. Apabila jumlah butir genap maka metode dapat menggunakan metode belah dua, Flanagan dan Rulon, sedangkan bila jumlah butir ganjil maka metode yang dapat digunakan adalah metode Kuder Richardson, Hoyt dan Alpha Cronbach.

Tes dikatakan reliabel apabila tes tersebut konsisten atau ajeg dalam hasil ukurnya sehingga dapat dipercaya. Instrumen yang reliabel tidak bersifat tendensius yang mengarahkan responden untuk memilih jawaban-jawaban tertentu.²⁵

Dalam penelitian ini, untuk menentukan reliabilitas tes ujian madrasah, peneliti menggunakan metode pengujian reliabilitas Kuder Richardson 20.

B. Kajian Pustaka

Dalam penelitian terkait dengan kualitas tes ujian madrasah, peneliti menggunakan beberapa skripsi di bawah ini sebagai rujukan:

1. Skripsi berjudul Analisis dan Relevansi Soal-Soal Ujian Nasional Matematika Sekolah Menengah Atas Program Studi Ilmu Pengetahuan Alam Kabupaten Batang Tahun Pelajaran

²⁵ Trianto, Pengantar Penelitian Pendidikan bagi Pengembangan Profesi Pendidikan dan Tenaga Kependidikan, (Jakarta, 2011: Kencana Prenada Media Group), hlm. 271.

2008/2009 dan 2009/2010 dengan Kurikulum Tingkat Satuan Pendidikan Matematika SMA/MA. Skripsi yang ditulis oleh Ilman Nafi'a dari Fakultas MIPA IKIP PGRI Semarang tersebut membahas tentang sejauh mana adanya keterkaitan antara soal Ujian Nasional Matematika program IPA dengan kurikulum yang digunakan. Dalam penelitian tersebut ditemukan bahwa tidak semua butir soal relevan dengan kurikulum SMA/MA dipandang dari validitas isi dan konstruksinya.²⁶

Pada skripsi di atas, pembahasan difokuskan kepada analisis soal ujian nasional terkait dengan kurikulum yang digunakan. Skripsi tersebut menjadi rujukan peneliti dalam daftar pustaka, khususnya tentang validitas isi. Peneliti kemudian menggunakannya untuk analisis soal ujian madrasah.

2. Skripsi berjudul Analisis Soal Ujian Akhir Semester Bidang Ilmu Pengetahuan Alam (IPA) Kelas VIII Semester Gasal Sekolah Menengah Pertama Negeri 1 Sukorejo Kabupaten Kendal Tahun Pelajaran 2006/2007. Skripsi tersebut ditulis oleh Mujiyanto, mahasiswa Jurusan Kurikulum dan Teknologi Pendidikan, Fakultas Ilmu Pendidikan Universitas Negeri Semarang.²⁷

²⁶ Ilman Nafi'a, "Analisis dan Relevansi Soal-Soal Ujian Nasional Matematika Sekolah Menengah Atas Program Studi Ilmu Pengetahuan Alam Kabupaten Batang Tahun Pelajaran 2008/2009 dan 2009/2010 dengan Kurikulum Tingkat Satuan Pendidikan", Skripsi (Semarang: FMIPA IKIP PGRI Semarang, 2011), abstraksi

²⁷ Mujiyanto, "Analisis Soal Ujian Akhir Semester Bidang Ilmu Pengetahuan Alam (IPA) Kelas VIII Semester Gasal Sekolah Menengah

Hasil penelitian pada skripsi tersebut ditemukan bahwa soal-soal yang digunakan untuk Ujian Akhir Semester kualitasnya baik untuk soal essay, namun belum baik untuk soal pilihan ganda. Pada soal pilihan ganda, kualitas belum baik setelah ditinjau dari segi validitas, reliabilitas, serta analisis butir soalnya. Terkait dengan penelitian tersebut, peneliti menjadikannya sebagai rujukan dalam analisis butir soal, validitas, serta reliabilitasnya dalam ujian madrasah.

3. Skripsi berjudul Analisis Butir Soal Ulangan Akhir Semester Gasal IPA kelas IX SMP di Kabupaten Grobogan. Skripsi tersebut ditulis oleh Lilis Tri Ariyana mahasiswa Jurusan Biologi FMIPA Universitas Negeri Semarang.

Dalam penelitian tersebut ditemukan bahwa soal Ujian Akhir Semester Gasal IPA kelas IX SMP di Kabupaten Grobogan masih didominasi oleh tes yang belum divalidasi. Penelitian tersebut bertujuan untuk mengukur validitas, tingkat kesukaran, daya beda, efektifitas pengecoh, dan reliabilitas soal dengan menggunakan program ITEMAN.²⁸ Terkait dengan skripsi tersebut, peneliti menjadikan hasil penelitian tersebut sebagai sumber rujukan untuk meneliti ujian madrasah.

Pertama Negeri 1 Sukorejo Kabupaten Kendal Tahun Pelajaran 2006/2007”, skripsi, (Semarang: FIP Universitas Negeri Semarang, 2007), abstraksi

²⁸ Ariyana, Lilis Tri. “Analisis Butir Soal Ulangan Akhir Semester Gasal IPA kelas IX SMP di Kabupaten Grobogan” Skripsi, Semarang: FMIPA Universitas Negeri Semarang, 2011), abstraksi

C. Kerangka Berpikir

Naskah soal yang digunakan untuk ujian madrasah dibuat oleh MGMP Matematika di wilayah tersebut. Dalam kenyataannya ternyata soal yang digunakan belum sepenuhnya diujikan terlebih dahulu. Hal tersebut terbukti karena ada soal yang masih cacat, baik tidak ada jawabannya ataupun jawaban ganda. Dalam penelitian ini, peneliti ingin melakukan analisis terkait dengan kualitas soal ujian madrasah tersebut berdasarkan implementasi di MAN Wonosobo pada tahun pelajaran 2012/2013.

Untuk mengetahui kualitas tes, analisis yang digunakan oleh peneliti meliputi analisis secara kuantitatif dan kualitatif. Secara kuantitatif dilakukan dengan analisis butir tes meliputi tingkat kesukaran, daya beda, fungsi pengecoh, serta validitas butir. Selain dengan analisis butir tes, juga dilakukan uji reliabilitas tes. Sedangkan secara kualitatif dilakukan uji validitas isi tes.

Setelah dilakukan penelitian dan diketahui kualitas soal, diharapkan dapat dilakukan tindak lanjut terhadap soal-soal tersebut. Tindak lanjut yang dimaksud di sini adalah memperbaiki maupun mengganti soal-soal yang kualitasnya kurang baik. Untuk soal yang sudah berkualitas baik juga dapat dipertahankan dan dapat digunakan untuk ujian madrasah tahun-tahun setelahnya.

