

PENGEMBANGAN INSTRUMEN PENILAIAN HASIL BELAJAR MATA PELAJARAN FIQH DI MADRASAH ALIYAH

Disusun Oleh :
DWI PRIYANTO, S.Ag., M.Pd



PENGEMBANGAN INSTRUMEN PENILAIAN HASIL BELAJAR MATA PELAJARAN FIQH DI MADRASAH ALIYAH

Copyright © Rizquna 2021

ISBN : 978-623-6018-37-8

Penulis : DWI PRIYANTO, S.Ag., M.Pd

Editor : Kang Emha

Perancang Sampul: Tim Rizquna

Layout : Faishol

Penerbit Rizquna

Anggota IKAPI No. 199/JTE/2020

Jl. KS Tubun Gang Camar RT 05/04, Karangsalam Kidul, Kedungbanteng, Banyumas, Jawa Tengah

Email: cv.rizqunaa@gmail.com

Layanan SMS: 085257288761

Cetakan I, Mei 2021

Penerbit dan Agency

CV. Rizquna

Karangsalam Kidul, Kedungbanteng, Banyumas, Jawa Tengah

Email: cv.rizqunaa@gmail.com

Hak cipta dilindungi undang-undang

Diterbitkan pertama kali oleh

Hak Cipta dilindungi Undang-undang

All Right Reserved

Dilarang memperbanyak sebagian atau seluruh isi buku ini
dalam bentuk apapun tanpa izin tertulis
dari Penerbit Rizquna



Kata Pengantar

Puji terindah bagi Allah SWT, atas segala anugerah dan berkah yang Allah berikan kedalam hidup ini sehingga kami mampu untuk menyusun buku “PENGEMBANGAN INSTRUMEN PENILAIAN HASIL BELAJAR MATA PELAJARAN FIQH DI MADRASAH ALIYAH” yang diharapkan dari kami yaitu, semoga buku ini bisa memberikan banyak pembelajaran dan pengetahuan yang pastinya bermanfaat untuk kemaslahatan semuanya. Tidak lupa shalwaat beserta salam kita haturkan kepada junjungan kita Nabi Muhammad SAW yang menjadi panutan umat kita semua.

Kami sadar bahwa penulisan buku ini bukan hanya dari hasil kinerja kami sendiri. Ada banyak pihak yang sangat berjasa dalam menyelesaikan buku ini. Seperti pengambilan materi, editor, dan lain-lain. Maka dari itu, kami mengucapkan banyak terimakasih kepada semua pihak yang telah membntu memberikan wawasan dan bimbingan kepada kami sebelum maupun ketika menulis buku ini.

Kami juga sangat mengetahui buku ini banyak kekurangan sebab karya tulis manusia tidak ada yang sempurna dan yang sempurna hanyalah

milik Allah semata. Apabila banyak kekurangan atau ketidaksesuaian mohon pemaklumanya. Dan kami sangat menerima sekali kritik, saran yang membangun untuk dijadikan susunan buku yang baik.

Atas tersusunnya buku ini kami haturkan terimakasih yang sebanyak-banyaknya kepada semua pihak yang terlibat membantu terbitnya buku ini. Semoga buku ini dapat memberikan banyak manfaat bagi orang-orang yang membaca dan mengamalkan ilmunya.

Penulis

Daftar Isi

Kata Pengantar iii

Daftar Isi v

BAB I PENDAHULUAN 1

- A. Latar Belakang Masalah 1
- B. Rumusan Masalah 9
- C. Tujuan dan Signifikansi 10
- D. Telaah Pustaka/Review Penelitian Terkait
11

BAB II KAJIAN TEORI 19

- A. Validitas 19
- B. Instrumen Penilaian Hasil Belajar Fiqh 38
- C. Tes 55
- D. Penilaian 65
- E. Mata Pelajaran Fiqh di MA 68

BAB III METODE PENELITIAN 85

- A. Model Pengembangan 85

BAB IV PENGEMBANGAN INSTRUMEN PENILAIAN HASIL BELAJAR MATA PELAJARAN FIQH MAN 89

- A. Pedoman penyusunan instrumen hasil
belajar Fiqh di MAN 1 dan MAN 2
Purwokerto 89

- B. Penyusunan Instrumen Penilaian Hasil Belajar Fiqh di MAN 1 Purwokerto dan MAN 2 Purwokerto 98
- C. Pengembangan instrumen penilaian hasil belajar Fiqh di MAN 1 dan MAN 2 Purwokerto 104
- D. Prosedur Pengembangan 114
- E. Uji Coba Produk 116
- F. Teknik Pengumpulan Data 117
- G. Teknik Analisis Data 117

BAB V PENUTUP 119

DAFTAR PUSTAKA 123

BAB I

PENDAHULUAN

A. Latar Belakang Masalah

Dalam Undang-Undang Sistem Pendidikan Nasional (UUSPN) Nomor 20 Tahun 2003 disebutkan bahwa pemerintah akan mengupayakan pengendalian mutu pendidikan nasional melalui sistem evaluasi. UUSPN merupakan pedoman dan role pemerintah untuk menjamin peningkatan kualitas pendidikan nasional. Sebagaimana termaktub dalam pasal 57 ayat 1 bahwa evaluasi dilakukan dalam rangka pengendalian kualitas pendidikan secara nasional sebagai bentuk akuntabilitas penyelenggaraan pendidikan kepada pihak-pihak yang berkepentingan.

Evaluasi pendidikan mempunyai peranan yang sangat penting karena dapat untuk mengetahui pemahaman dan daya serap siswa atas materi yang telah diberikan. Evaluasi pendidikan dilaksanakan dalam rangka mendapatkan informasi mengenai aspek yang berkaitan dengan pendidikan. Karenanya evaluasi merupakan serangkaian aktivitas

untuk mengetahui keberhasilan proses pembelajaran, baik keberhasilan siswa (prestasi belajar) dan keberhasilan guru mengajar maupun keberhasilan proses pembelajaran. Keberhasilan yang disebut pertama akan lebih merupakan pembahasan dalam tulisan ini. Membahas mengenai evaluasi maka di dalam kita juga membicarakan pengukuran karena di sini merupakan kegiatan untuk mendapatkan informasi atau data secara kuantitatif dan kualitatif (Jahja Umar, et al., 1997: 2).

Evaluasi juga didefinisikan sebagai proses mengumpulkan informasi untuk mengetahui pencapaian belajar kelas atau kelompok. Hasil evaluasi akan bisa digunakan sebagai acuan guru untuk memberikan masukan lebih baik lagi dalam mengajar dan siswa lebih baik lagi dalam belajar. Yang pada gilirannya hasil evaluasi dapat digunakan sebagai pedoman dalam meningkatkan kualitas proses belajar mengajar. Evaluasi pada dasarnya melakukan judgment terhadap hasil penilaian, maka kesalahan pada penilaian dan pengukuran diupayakan sekecil mungkin (Djemari Mardapi, 2008: 9).

Pernyataan evaluasi juga ditegaskan dalam Peraturan Pemerintah Nomor 19 Tahun 2005 tentang Standar Nasional Pendidikan, bahwa dalam rangka pencapaian Standar Nasional Pendidikan, salah satu hal penting yang harus diupayakan adalah adanya standar penilaian, yaitu standar nasional pendidikan berkaitan dengan mekanisme, prosedur, dan instrumen penilaian hasil belajar siswa. Hal ini mengindikasikan bahwa kegiatan evaluasi dan penilaian hasil belajar melalui ujian, baik ujian nasional, ujian regional, ataupun ujian yang diselenggarakan oleh pendidik, satuan pendidikan dan pemerintah memerlukan mekanisme, prosedur, dan instrumen yang dapat dipertanggungjawabkan, untuk memenuhi akuntabilitas pendidikan dalam bentuk kualitas pendidikan nasional yang semakin meningkat.

Djemari Mardapi dkk. (1999b: 79) menemukan beberapa hal yang membuat sistem evaluasi hasil belajar yang dilakukan di sekolah maupun di daerah belum mendukung kualitas pendidikan, antara lain (1) kualitas tes buatan guru masih masih kurang memadai; (2) jaringan pengujian di daerah belum

dimanfaatkan dengan baik; (3) pelaporan hasil pelaksanaan ujian oleh guru kepada kepala sekolah belum terlaksana secara rutin; (4) hasil-hasil ujian belum dimanfaatkan secara optimal untuk perbaikan proses pembelajaran di kelas.

Evaluasi di samping untuk mengetahui pencapaian belajar siswa pada tingkat kelas juga untuk menelaah suatu program dan dampak pada program pendidikan. Istilah lain dari tersebut di atas, bahwa yang pertama biasa disebut evaluasi yang bersifat mikro dan kedua evaluasi yang bersifat makro. Pencapaian belajar yang ada pada level mikro tidak hanya yang bersifat kognitif saja, tetapi juga mencakup untuk mengungkap ranah yang lain yaitu afektif dan psikomotor. Atau tiga ranah dalam istilah lain adalah kemampuan berpikir, ketrampilan melakukan pekerjaan, dan perilaku. Setiap siswa mempunyai potensi pada dua ranah yaitu; kemampuan berpikir dan keterampilan, hanya saja tingkatannya antara satu siswa dengan siswa yang berbeda (Djemari Mardapi, 2008: 99).

Untuk mendapatkan evaluasi yang objektif diperlukan suatu instrumen penilaian

yang baik yaitu instrumen yang memenuhi kriteria *valid* dan reliabel. Instrumen yang baik semestinya akan bisa dijadikan sebagai alat untuk mengungkap keterampilan siswa dalam memahami materi pembelajaran. Keterampilan yang dikehendaki di sini tidak sekedar memahami, tetapi juga keterampilan melakukan pekerjaan dan perilaku harus sesuai dengan apa yang dijelankannya.

Kemampuan berpikir atau kognisi menjadi penting untuk bisa terungkap, karena kognisi kunci dalam mengetahui perilaku dan keterampilan melakukan pekerjaan atau pengamalan dalam hal ini adalah nilai-nilai agama oleh siswa. Tinjauan dari aspek psikologi dalam memahami anak seusia siswa Madrasah Aliyah (MA) akan menjadi penekanan dalam rangka membantu mengetahui karakteristik siswa. Oleh karena itu instrumen penilaian hasil belajar harus sesuai dengan apa yang akan dinilai sehingga menghasilkan data yang sah dan akurat.

Instrumen penilaian yang bermutu dapat membantu pendidik meningkatkan pembelajaran dan memberikan informasi dengan tepat tentang peserta didik mana yang

belum atau sudah mencapai kompetensi. Salah satu ciri instrumen penilaian yang bermutu adalah bahwa instrumen itu dapat membedakan setiap kemampuan peserta didik. Semakin tinggi kemampuan peserta didik dalam memahami materi pembelajaran, semakin tinggi pula peluang menjawab benar soal atau mencapai kompetensi yang ditetapkan dan kebalikannya.

Syarat instrumen penilaian yang bermutu adalah bahwa instrumen harus sah (*valid*), dan handal. Sah maksudnya bahwa setiap alat ukur hanya mengukur satu dimensi atau aspek saja. Instrumen Fiqh hanya mengukur materi Fiqh bukan mengukur keterampilan materi yang lain. Handal maksudnya bahwa setiap alat ukur harus dapat memberikan hasil pengukuran yang tepat, cermat, dan ajeg. Untuk dapat menghasilkan instrumen yang sah dan handal, penulis harus merumuskan kisi-kisi dan menulis instrumen berdasarkan kaidah penulisan yang baik.

Namun demikian, masih banyak guru yang belum menerapkan penyusunan instrumen yang sesuai norma evaluasi. Bahkan banyak guru yang belum paham dengan pengetahuan

proses penulisan soal atau instrumen yang baik. Satu hal yang perlu kita upayakan untuk diperbaiki adalah pembuatan instrumen penilaian hasil belajar khususnya pada mata pelajaran fiqh di MA. Hasil survey awal yang peneliti peroleh bahwa soal ujian Madrasah Aliyah untuk ulangan tengah semester dibuat oleh guru mapel Fiqh madrasah sendiri, sedangkan soal ujian semester madrasah menerima dari karesidenan. Pembuat soal ujian karesidenan adalah 2 guru fiqh, dari 2 madrasah yang kemudian diedit oleh guru fiqh madrasah yang lain, sehinggadihasilkan soal ujian untuk seluruh madrasah satu karesidenan. Informasi yang peneliti dapatkan soal ujian tersebut belum melalui uji validitas.

Pertimbangan peneliti memilih mata pelajaran fiqh karena berkenaan dengan pelaksanaan nilai agama dalam keseharian, benar dan tidanya pengamalan ajaran agama sangat ditentukan oleh pengetahuan fiqh siswa. Karenanya menjadi penting untuk dikaji lebih detail khususnya dalam hal proses penyusunan instrumen penilaian hasil belajar. Instrumen akan digunakan untuk menilai tiga ranah yaitu kognitif, afektif, dan psikomotor .

Sehingga tidak cukup jika hanya menggunakan instrumen tes, tetapi juga non-tes.

Mengenai penggunaan bentuk tes dalam ujian menurut Jahja Umar (1995: 9) bahwa bentuk tes yang digunakan dalam ujian tergantung dari jenis ujiannya. Jika jenis ujiannya adalah ulangan harian sebaiknya tes yang digunakan adalah tes bentuk uraian yang lebih mendetail. Jika jenis ujiannya adalah ujian akhir maka tes yang digunakan adalah tes pilihan ganda atau campuran antara pilihan ganda dan uraian sesuai dengan tuntutan kebutuhan yang ada. Baik pilihan ganda maupun uraian adalah tes yang memenuhi kriteria tes yang baik. Bentuk tes uraian bisa digunakan untuk melihat pola berpikir dari siswa (*testee*).

Dengan diberlakukannya otonomi daerah sedikit banyak akan mempengaruhi otonomi pendidikan masing-masing daerah. Sehingga sangat mungkin terjadi satu daerah dengan daerah yang lain ketika melakukan penilaian dari hasil prestasi bidang pendidikan tertentu akan mengalami perbedaan. Hal ini menjadi kurang baik karena perbedaan cara penilaian tersebut untuk menilai variabel yang sama.

Ketika hal ini terjadi maka perlu ada upaya-upaya yang lebih mengarahkan kepada perbaikan untuk tidak terjadi perbedaan yang sangat menyolok. Dengan demikian begitu penting proses mengenai bagaimana membuat instrumen penilaian yang baik yaitu instrumen penilaian yang memiliki validitas dan reliabilitas yang tinggi. Oleh karena itu peneliti ingin berusaha menyusun instrumen penilaian yang secara validitas konstruk bisa dipertanggungjawabkan kualitasnya mata pelajaran fiqh di Madrasah Aliyah. Lokasi ujicoba skala kecil dan besar penelitian ini akan mengambil di Madrasah Aliyah Negeri (MAN) 1 Purwokerto dan Madrasah Aliyah Negeri (MAN) 2 Purwokerto. Kedua MAN tersebut berlokasi di kabupaten Banyumas sehingga akan memiliki cukup variasi karakter yang mungkin berbeda sehingga akan bisa menjadi catatan yang bisa memperkaya pertimbangan dalam penyusunan penelitian.

B. Rumusan Masalah

Mendasarkan pada latar belakang masalah tersebut di atas, maka rumusan masalahnya adalah :

1. Bagaimana pedoman penyusunan instrumen hasil belajar Fiqh di MAN 1 dan MAN 2 Purwokerto?
2. Bagaimana penyusunan instrumen penilaian hasil belajar Fiqh di MAN 1 Purwokerto dan MAN 2 Purwokerto?
3. Bagaimana pengembangan instrumen penilaian hasil belajar Fiqh di MAN 1 Purwokerto dan MAN 2 Purwokerto?

C. Tujuan dan Signifikansi

1. Tujuan Penelitian

Tujuan penelitian ini adalah sebagai berikut:

- a. Menggali informasi pedoman penyusunan instrumen penilaian hasil belajar di Madrasah Aliyah
- b. Mengetahui proses penyusunan instrumen penilaian hasil belajar Fiqh di MAN 1 Purwokerto dan MAN 2 Purwokerto.
- c. Menemukan pengembangan instrumen penilaian hasil belajar Fiqh di Madrasah Aliyah.

2. Signifikansi Penelitian

Penelitian akan melakukan pengembangan instrumen penilaian hasil belajar untuk mata pelajaran Fiqh di Madrasah Aliyah. Penelitian ini merupakan penelitian yang akan berusaha menghasilkan instrumen yang *valid* dan reliabel dalam rangka mendapatkan instrumen tes yang bisa menjadi pedoman dalam melakukan evaluasi hasil belajar. Lebih jauh dari hal tersebut penelitian bisa menjadi acuan madrasah atau departemen terkait dalam proses penyusunan instrumen penilaian yang memiliki tingkat validitas yang baik.

D. Telaah Pustaka/Review Penelitian Terkait

Penyusunan instrumen penilaian hasil belajar yang baik adalah penyusunan instrumen yang disertai bukti empiris. Oleh karenanya dalam penelitian ini dilakukan ujicoba terlebih dahulu untuk memperoleh pemenuhan bukti empiris. Setelah itu perangkat tes disusun dengan menggunakan item tes yang telah memiliki bukti empiris hasil kegiatan ujicoba. Perangkat tes itulah yang digunakan untuk melakukan pengukuran yang

sesungguhnya dalam tahap pengukuran.

Penyusunan instrumen pengukuran proses pemahaman fiqh sampai mengamalkan terdiri dari tahap perancangan, tahap ujicoba, dan tahap pengukuran. Langkah pertama yang terpenting dalam tahap perancangan adalah menyusun *learning continuum* (LC). Rumusan LC yg diperoleh kemudian menjadi pegangan untuk menyusun kisi-kisi dan penulisan item. Rumusan LC yang memenuhi validitas isi akan dapat diperoleh dengan memperhatikan referensi baik dalam literatur maupun keterampilan proses memahami dan mengamalkan fiqh yang tersurat dalam Standar Kompetensi (SK) dan Kompetensi Dasar (KD) dalam Standar Isi menurut Permendiknas Nomor 22 Tahun 2006 dan telah ditelaah melalui focus group discussion (FGD). Dengan demikian rumusan LC yg dihasilkan dapat dijadikan sebagai rujukan kompetensi dalam memahami fiqh yang harus dikuasai oleh siswa MA. Rumusan LC yg dihasilkan dapat dijadikan tolok ukur target pembelajaran. Oleh karena itu, hasil pengukuran pemahaman dan pengamalan Fiqh nantinya dipakai sebagai acuan untuk menerapkan *assessment of learning*.

Penelitian Yusrizal (2009) tentang Pengujian Validitas Konstruk dengan menggunakan analisis faktor, penelitian ini melakukan uji validitas konstruk dengan menggunakan analisis faktor pada instrumen penilaian kinerja dosen yang dikembangkan. Instrumen ini dilakukan dua kali ujicoba pada 770 mahasiswa FKIP Universitas Syiah Kuala Banda Aceh. Pada ujicoba pertama validitas konstruk diuji dengan analisis faktor eksploratori, berhasil diekstraksi 7 faktor, yang sesuai dengan jumlah faktor yang diestimasi. Koefisien reliabilitas yang ditunjukkan oleh konsistensi internal alpha sebesar 0,931. Pada ujicoba kedua validitas konstraknya dianalisis dengan menggunakan analisis faktor konfirmatori. Hasil komputasi juga berhasil diekstraksi 7 faktor yang sesuai dengan kajian teoritis. Koefisien reliabilitas konsistensi internal alpha diperoleh sebesar 0,934. Dapat disimpulkan bahwa instrumen penilaian kinerja dosen yang dikembangkan memiliki validitas konstruk yang baik dan memiliki koefisien reliabilitas konsistensi internal yang sangat tinggi.

Penelitian Hepi Wahyuningsi (2009) dengan judul Validitas Konstruk Alat Ukur Spirituality Orientation Inventory (SOI), hasil uji validitas isi dengan *professional judgement*, hasil uji validitas konstruk dengan bukti homogenitas dan bukti adanya perbedaan skor pada 2 kelompok yang berbeda pada alat ukur *spiritual orientation inventory* (SOI) yang dikembangkan ini menunjukkan tidak ada dimensi spiritualitas yang gugur. Hal ini menunjukkan bahwa aitem-aitem yang dibuat berdasarkan *preliminary* mampu mencerminkan kesembilan dimensi spiritualitas. Selain itu, hasil ini juga menunjukkan bahwa spiritualitas merupakan konsep yang multidimensional seperti yang dikemukakan oleh Johnstone & Yoon (2009) dan Ho & Ho (2007). Hasil uji validitas konstruk dengan analisis faktor yang menghasilkan 6 subskala spiritualitas, kemungkinan menunjukkan adanya overlapping antara sembilan dimensi yang dikemukakan oleh Elkins et al. (1988).

Penelitian Zulkifi Matondang (2009) dalam Validitas Dan Reliabilitas Suatu Instrumen Penelitian. Hasil penelitian adalah instrumen merupakan suatu alat yang karena

memenuhi persyaratan akademis maka dapat dipergunakan sebagai alat untuk mengukur suatu obyek ukur atau mengumpulkan data mengenai suatu variabel. Validitas isi mempermasalahkan sejauh mana tes mengukur tingkat penguasaan terhadap isi atau materi yang seharusnya dikuasai sesuai dengan tujuan pengajaran dan validitas isi tidak mempunyai besaran. Validitas konstruk mempermasalahkan seberapa jauh butir-butir tes mampu mengukur apa yang benar-benar hendak diukur sesuai dengan definisi konseptual yang telah ditetapkan. Validitas empiris (validitas kriteria) yang berarti bahwa validitas ditentukan berdasarkan kriteria, baik kriteria internal maupun kriteria eksternal.

Penelitian Scouller (2008) mengenai pengaruh metode penilaian terhadap belajar siswa dengan membandingkan antara ujian menggunakan bentuk tes pilihan ganda, jawaban singkat, dan esai bebas. Penelitian dilakukan untuk mengidentifikasi strategi persiapan dan motif, persepsi terhadap level proses intelektual yang terlibat, dan cara penilaian yang lebih disukai. Tujuan penelitian ini untuk memperbaiki kualitas belajar siswa

dengan memahami pengaruh penilaian terhadap belajar siswa dan meningkatkan metode penilaian yang lebih baik. Hasil penelitian pertama, menunjukkan adanya pendekatan dan persepsi belajar bergantung pada metode penilaiannya, kedua bahwa siswa lebih suka dinilai dengan tugas esay.

Nield & Wintre dan Masling (1986 : 196-199) melakukan penelitian mengenai bentuk tes hasil belajar. Dia membandingkan sikap siswa terhadap pertanyaan pilihan ganda dengan jenis lain seperti jawaban singkat, esai dan melengkapi. Hasilnya bahwa pertanyaan pilihan ganda paling banyak diminati dibanding yang lain.

Sementara penelitian yang akan peneliti lakukan ini lebih menekankan untuk *create* instrumen penilaian dari sisi validitas konstruk mata pelajaran Fiqh Madrasah Aliyah (MA). Dengan produk instrumen penilaian hasil belajar mata pelajaran Fiqh MA akan memiliki kontribusi instrumen yang bisa dipakai dan merupakan instrumen yang memiliki tingkat validitas konstruk yang baik. Di samping itu langkah-langkah dalam menyusun instrumen yang nantinya bisa dijadikan sebagai acuan

oleh para guru dalam menyusun dan mendapatkan instrumen yang baik, artinya yang bisa mengukur apa yang seharusnya diukur.

BAB II

KAJIAN TEORI

A. Validitas

Suatu instrumen pengukuran dikatakan memiliki validitas yang tinggi apabila alat tersebut menjalankan fungsi ukurnya, atau memberikan hasil ukur yang sesuai dengan maksud dilakukannya pengukuran tersebut. Artinya hasil ukur dari pengukuran tersebut merupakan besaran yang mencerminkan secara tepat fakta atau keadaan sesungguhnya dari apa yang diukur.

Validitas suatu instrumen atau tes mempermasalahkan apakah instrumen atau tes tersebut benar-benar mengukur apa yang hendak diukur. Seperti yang dikemukakan oleh Cureton dalam bukunya *Educational Measurement Validity*, bahwa *“The essential question of test validity is how well a test does the job it is employed to do”* (Cureton, 1978: 621). Maksudnya adalah bahwa sebarapa jauh suatu tes mampu mengungkapkan dengan tepat ciri atau keadaan yang sesungguhnya dari obyek ukur, akan tergantung dari tingkat validitas tes

yang bersangkutan.

Dengan demikian, maka tes yang *valid* untuk tujuan tertentu ialah tes yang mampu mengukur apa yang hendak diukur. Suatu tes yang *valid* untuk tujuan tertentu, atau pengambilan keputusan tertentu, mungkin tidak *valid* untuk tujuan atau pengambilan keputusan lain (Cronbach, 1975). Jadi validitas suatu tes harus selalu dikaitkan dengan tujuan atau pengambilan keputusan tertentu. Tes masuk misalnya harus selalu dikaitkan dengan seberapa jauh tes masuk tersebut dapat mencerminkan prestasi belajar para calon siswa baru setelah belajar nanti.

Konsep validitas menurut Reynolds dkk (2010: 128) adalah dengan melihat lebih kepada sejarah dan perjalanan dalam rangka melengkapi, sebagai terinci yang perlu diperhatikan, yaitu:

1. Tahun 1974 Standards Validity as Three Types yaitu; *content validity, criterion validity, and construct validity.*
2. Tahun 1985 Standards Validity as Three Interrelated Types yaitu; *content-related validity, criterion-related validity, and construct-related validity.*

3. *Tahun 1999 Standards Validity as a Unitary Construct yaitu; validity evidence based on test content, validity evidence based on related to the vapiables, validity evidence based on internal structure, validity evidence based on respons processes, and validity evidence based on consequences of testing.*

Sedangkan konsep validitas tes menurut Anatasi dalam Howard Wainer dan Henry I. Braun (1988: 25) dapat dibedakan atas tiga macam yaitu:

1. Validitas isi (*content validity*)
2. Validitas konstruk (*construct validity*)
3. Validitas kriteria- relasi (*criterion-related validity*)

a. Validitas Isi

Validitas isi suatu tes mempermasalahkan seberapa jauh suatu tes mengukur tingkat penguasaan terhadap isi suatu materi tertentu yang seharusnya dikuasai sesuai dengan tujuan pengajaran. Dengan kata lain tes yang mempunyai validitas isi yang baik ialah tes

yang benar-benar mengukur penguasaan materi yang seharusnya dikuasai sesuai dengan *content* pembelajaran yang tercantum dalam kurikulum.

Validitas isi menunjukkan seberapa jauh pertanyaan, tugas atau soal dalam suatu tes mampu mewakili secara keseluruhan dan proporsional perilaku sampel yang dikenai tes tersebut. Artinya tes mencerminkan keseluruhan *content* atau materi yang diujikan atau yang seharusnya dikuasai secara proporsional.

Untuk mengetahui apakah tes itu *valid* atau tidak harus dilakukan melalui penelaahan kisi-kisi tes untuk memastikan bahwa soal-soal tes itu sudah mewakili atau mencerminkan keseluruhan *content* atau materi yang seharusnya dikuasai secara proporsional. Oleh karena itu, validitas isi suatu tes tidak mempunyai besaran tertentu yang dihitung secara statistika, tetapi dipahami bahwa tes itu sudah *valid* berdasarkan telaah kisi-kisi tes. Oleh karena itu, Wiersma dan Jurs (1990: 35) menyatakan bahwa validitas isi sebenarnya mendasarkan pada analisis

logika, jadi tidak merupakan suatu koefisien validitas yang dihitung secara statistika.

Untuk memperbaiki tes, maka isi suatu tes harus diusahakan agar mencakup semua kompetensi dasar dan indikator yang hendak diukur. Kriteria untuk menentukan proporsi masing-masing kompetensi dasar dan indikator yang tercakup dalam suatu tes ialah berdasarkan banyaknya isi (materi) masing-masing kompetensi dasar dan indikator, yang dapat dilihat dari jumlah halaman isi (materi) dan jumlah jam pertemuan untuk masing-masing kompetensi dasar dan indikator seperti tercantum dalam kurikulum.

Selain itu, penentuan proporsi tersebut dapat pula didasarkan pendapat (*judgement*) para ahli dalam bidang yang bersangkutan. Jadi suatu tes akan mempunyai validitas isi yang baik jika tes tersebut terdiri dari item-item yang mewakili semua materi yang hendak diukur. Salah satu cara yang biasa dilakukan untuk memperbaiki validitas

isi suatu tes ialah dengan menggunakan *blue-print* untuk menentukan kisi-kisi tes.

Pengujian validitas isi yang dilakukan dengan menelaah butir soal dilakukan dengan mencermati kesesuaian isi butir yang ditulis dengan perencanaan yang dituangkan dalam kisi-kisi. Kriteria yang menjadi dasar pengujian validitas isi adalah kisi-kisi yang direncanakan. Telaah dilakukan untuk menjaga agar materi butir instrumen yang dikembangkan tidak menyimpang dari kisi-kisi. Butir-butir instrumen dinyatakan valid (*logically valid*) apabila setelah dicermati isi butir-butir yang ditulis telah menunjukkan kesesuaian dengan kisi-kisi.

Pengujian validitas isi dapat dilakukan dengan meminta pertimbangan ahli (*expert judgement*). Orang yang memiliki kompetensi dalam satu bidang dapat dimintakan pendapatnya untuk menilai ketepatan isi butir instrumen. Pertimbangan juga bisa dimintakan kepada profesional (*professional judgement*). Orang yang menekuni suatu bidang tertentu yang sesuai dengan

wilayah kajian instrumen, misalnya guru, mekanik, dokter, advokat, koreografer dan sebagainya dapat dimintakan pendapatnya untuk menilai ketepatan isi instrumen. Penilaian validitas isi juga dapat dimintakan pertimbangan kepada beberapa orang yang memiliki kompetensi untuk memberikan penilaian (*inter-rater judgement*).

Pertimbangan yang dimintakan kepada ahli, professional atau rater menyangkut isi dari butir instrumen dan kisi-kisinya. Pertimbangan yang menyangkut materi akan diukur menggunakan butir-butir instrumen. Butir-butir yang mengukur materi sebagaimana dipahami dan disepakati ahli, professional atau penilai dapat dinyatakan sebagai butir-butir yang valid. Pemberian pendapat dapat dilakukan dengan memberikan respon atas kesesuaian butir yang ditulis dengan kisi-kisinya dalam hal materi.

b. Validitas Konstruk

Validitas konstruk (*construct validity*) adalah validitas yang mempermasalahkan seberapa jauh item-item tes mampu

mengukur apa yang benar-benar hendak diukur sesuai dengan konsep khusus atau definisi konseptual yang telah ditetapkan. Validitas konstruk juga untuk mempersoalkan sejauh mana skor merefleksikan konstruk teoritik yang mendasari penyusunan alat ukur tersebut.

Validitas konstruk biasa digunakan untuk instrumen-instrumen yang dimaksudkan mengukur variable-variabel konsep, baik yang sifatnya performansi tipikal seperti instrumen untuk mengukur sikap, minat, konsep diri, fokus kontrol, gaya kepemimpinan, motivasi berprestasi, dan lain-lain, maupun yang sifatnya performansi maksimum seperti instrumen untuk mengukur bakat (tes bakat), intelegensi (kecerdasan intelektual), kecerdasan emosional dan lain-lain. Djemari Mardapi (1988: 25) menyatakan bahwa bukti validitas konstruk sangat membutuhkan instrumen yang bisa mengukur desain konstruk dari sebuah teori, seperti motivasi manusia, ketertarikan, intelegensi dan selalu fokus pada definisi konsep atribut yang diukur.

Untuk menentukan validitas konstruk suatu instrumen harus dilakukan proses penelaahan teoritis dari suatu konsep dari variabel yang hendak diukur, mulai dari perumusan konstruk, penentuan dimensi dan indikator, sampai kepada penjabaran dan penulisan soal-soal item instrumen. Perumusan konstruk harus dilakukan berdasarkan sintesis dari teori-teori mengenai konsep variabel yang hendak diukur melalui proses analisis dan komparasi yang logik dan cermat. Validitas konstruk akan lebih efektif jika dalam mengenali permasalahannya dengan menerapkan prinsip teori kognitif dalam mendesain dan memvalidasi suatu tes. Sebagaimana yang diungkapkan Wainer & Braun (1990: 21) adalah *I do not see how we can effectively pursue issues of construct validity whitout some principled aplications of cognitive theory in the design and validation of test.*

Menurut Jack R. Fraenkel validasi konstruk merupakan yang terluas cakupannya dibanding dengan validasi lainnya, karena melibatkan banyak

prosedur termasuk validasi isi dan validasi kriteria. Kemudian untuk mendapatkan validitas konstruk menurutnya ada tiga langkah di dalamnya yaitu :

- 1) Variabel yang akan diukur harus didefinisikan dengan jelas.
- 2) Hipotesis, yang mengacu pada teori yang mendasari variabel penelitian harus dapat membedakan orang dengan tingkat gradasi yang berbeda pada situasi tertentu.
- 3) Hipotesis tersebut diuji secara logis dan empiris.

Sebagaimana pendapat Ary dkk yang dikutip Muhammad Khumaidi (2006: 17) bahwa segi empiris dari validitas konstruk adalah; (a) secara internal, hubungan-hubungan di dalam tes perlunya seperti yang diramalkan oleh konstruk tersebut, dan (b) secara eksternal, hubungan-hubungan antara skor tes dengan pengamatan-pengamatan lainnya perlunya konsisten dengan konstruk tersebut.

Dimensi dan indikator dijabarkan dari konstruk yang telah dirumuskan dengan

memperhatikan hal-hal sebagai berikut:

- 1) Seberapa jauh indikator tersebut merupakan indikator yang tepat dan konstruk yang telah dirumuskan.
- 2) Indikator-indikator dari suatu konstruk harus homogen, konsisten, dan konvergen untuk mengukur konstruk dari variabel yang hendak diukur.
- 3) Indikator-indikator tersebut harus lengkap untuk mengukur suatu konstruk secara utuh.

Soal-soal instrumen yang ditulis untuk masing-masing indikator harus benar-benar dapat mengukur secara tepat indikator yang hendak diukur. Jumlah soal untuk mengukur setiap indikator harus disesuaikan dengan bobot atau pentingnya masing-masing indikator sebagai penanda konsep variabel yang hendak diukur. Karena dalam hal ini sangat mungkin instrumen dalam kategori baik dianalisis dengan menggunakan analisis faktor, tapi nilai kumulatif beberapa faktor kecil, hal ini menunjukkan bahwa soal-soal belum bekerja dengan baik dan tidak bisa mewakili faktor yang di asumsikan (Lubis

& Muhadjir, 2004: 21)

Menyimak proses telaah teoritis seperti telah dikemukakan, maka proses validasi konstruk sebuah instrumen harus dilakukan melalui penelaahan atau justifikasi pakar atau melalui penilaian sekelompok panel yang terdiri dari orang-orang yang menguasai substansi atau konten dari variabel yang hendak diukur. Setelah melalui masukan para ahli untuk kemudian disusun secara lebih baik lagi untuk mendapatkan instrumen yang lebih berkualitas dari sebelumnya. Seperti halnya validitas isi, untuk mempertinggi validitas konstruk dapat dilakukan dengan cara memerinci dan memasang setiap soal soal dengan setiap aspek. Pengujian validitas konstruk diperlukan analisis statistik yang kompleks seperti prosedur analisis faktor. Salah satu prosedur pengujian validitas konstruk yang tidak terlalu kompleks dapat dilakukan dengan pendekatan *multi-trait multi-method*. Dua atau lebih trait yang diukur melalui dua atau lebih metode dapat diuji secara serentak dengan pendekatan ini, sehingga akan

diperoleh adanya bukti adanya validitas diskriminan dan validitas konvergen. Validitas diskriminan ditunjukkan oleh rendahnya korelasi antara faktor skala atau tes yang mengukur trait yang berbeda terutama bila digunakan metode yang sama. Validitas konvergen ditunjukkan oleh tingginya korelasi skor tes-tes yang mengukur trait yang sama dengan menggunakan metode yang berbeda.

Messick membedakan ada enam (6) aspek dalam validitas konstruk untuk penerapan dalam semua tes, yaitu:

First, the content aspect concerns the relevancy and representativeness of test content to the construct. Second, the substantive aspect concerns the theoretical rationale and evidence about the processes behind test responses. Third, the structural aspect concerns the relationship of the scoring system to the structure of the construct domain. Fourth, the generalizability aspect concerns the extent to which score interpretations may be generalized to varying populations, conditions, and settings. Fifth, the external

aspect concerns the correlations of test score with criteria and other test. Sixth, the consequential aspect concerns the social consequences of test use, such as bias, fairness, and distributive justice (Leigthon & Gierl, 2007: 121-122).

Proses Validasi

Setelah suatu instrumen dianggap valid secara konseptual, maka langkah berikutnya adalah instrumen tersebut diujicobakan pada sekelompok responden yang merupakan sampel ujicoba. Dari Jawaban atau respon dari sampel ujicoba tersebut diperoleh data yang akan dianalisis untuk menguji validitas instrumen dengan menggunakan validitas.

Analisis data hasil ujicoba soal pada pokoknya dimaksudkan untuk menguji validitas soal-soal instrumen atau soal-soal tes secara empiris atau berdasarkan data empiris yang diperoleh dari ujicoba. Dalam pembahasan ini validitas yang akan diuji adalah validitas soal dengan menggunakan kriteria internal, yaitu skor total tes. Skor total instrumen atau tes dapat ditetapkan sebagai kriteria untuk

menentukan validitas soal instrumen atau soal tes karena secara konsep atau konten perangkat instrumen atau tes yang telah dibuat dan diujicobakan sudah dianggap valid. Pengujian validitas soal tes dilakukan dengan menghitung koefisien korelasi antara skor soal tes dengan skor total tes. Soal yang dianggap valid adalah soal tes yang skornya mempunyai koefisien korelasi yang signifikan dengan skor total tes.

c. Validitas Kriteria

Validitas kriteria yang berarti bahwa validitas ditentukan berdasarkan kriteria, baik kriteria internal maupun kriteria eksternal. Kriteria internal adalah tes atau instrumen itu sendiri yang menjadi kriteria, sedangkan kriteria eksternal adalah hasil ukur instrumen atau tes lain di luar instrumen itu sendiri yang menjadi kriteria. Ukuran lain yang sudah dianggap baku atau dapat dipercaya dapat pula dijadikan sebagai kriteria eksternal.

Validitas yang ditentukan berdasarkan kriteria internal disebut validitas internal,

sedangkan validitas yang ditentukan berdasarkan kriteria eksternal disebut validitas eksternal. Validitas eksternal dapat dibedakan lagi atas dua macam yaitu (a) validitas kongkuren (*concurrent validity*), dan (b) validitas prediktif (*predictive validity*).

1) Validitas Internal

Validitas internal merupakan validitas yang diukur dengan besaran yang menggunakan instrumen sebagai suatu kesatuan (keseluruhan soal) sebagai kriteria untuk menentukan validitas instrument. Dengan demikian validitas internal memperlakukan validitas soal atau item suatu instrumen dengan menggunakan hasil ukur instrumen tersebut sebagai suatu kesatuan dan sebagai kriteria, sehingga bisa juga disebut sebagai validitas soal.

Validitas soal diperlihatkan oleh seberapa jauh hasil ukur soal tersebut konsisten dengan hasil ukur instrumen secara seluruhan. Oleh karena itu, validitas soal tercermin pada besaran

toefisien korelasi antara skor soal dengan skor total instrumen. Jika koefisien korelasi antara skor soal dengan skor total instrumen positif dan signifikan, maka soal tersebut dapat dianggap *valid* berdasarkan ukuran validitas internal.

Apabila besaran koefisien korelasi antara skor soal dengan skor bernilai positif, makin besar koefisien korelasi maka validitas soal juga makin tinggi. Koefisien korelasi yang tinggi antara skor soal dengan skor total mencerminkan tingginya konsistensi antara hasil ukur keseluruhan instrumen dengan hasil ukur soal instrumen, atau dapat dikatakan bahwa soal instrumen tersebut konvergen dengan soal-soal lain dalam mengukur suatu konsep atau konstruk yang hendak diukur.

2) Validitas Eksternal

Kriteria eksternal itu dapat berupa hasil ukur instrumen baku atau instrumen yang dianggap baku dapat pula berupa hasil ukur lain yang sudah tersedia dan dapat dipercaya sebagai ukuran dari suatu konsep atau variabel

yang hendak diukur. Validitas eksternal diperlihatkan oleh suatu besaran yang merupakan hasil perhitungan statistika.

Jika kita menggunakan hasil ukur instrumen yang sudah baku sebagai kriteria eksternal, maka besaran validitas eksternal dari instrumen yang kita kembangkan didapat dengan jalan mengkorelasikan skor hasil ukur instrumen yang dikembangkan dengan skor hasil ukur instrumen baku yang dijadikan kriteria. Makin tinggi koefisien korelasi yang didapat, maka validitas instrumen yang dikembangkan juga makin baik. Kriteria yang digunakan untuk menguji validitas eksternal adalah nilai tabel r (r -tabel).

Jika koefisien korelasi antara skor hasil ukur instrumen yang dikembangkan dengan skor hasil ukur instrumen baku lebih besar daripada r -tabel, maka instrumen yang dikembangkan dapat *valid* berdasarkan kriteria eksternal yang dipilih (hasil ukur instrumen baku). Jadi keputusan uji validitas dalam hal ini adalah mengenai

valid atau tidaknya instrumen sebagai suatu kesatuan, bukan *valid* atau tidaknya soal instrumen seperti pada validitas internal.

Contoh untuk validitas prediktif adalah jika kita hendak menguji suatu tes masuk suatu perguruan tinggi dengan menggunakan indeks prestasi semester satu sebagai kriteria eksternal, karena indeks prestasi semester satu merupakan panampilan masa yang akan datang pada saat pelaksanaan tes masuk. Jika koefisien korelasi antara skor tes masuk sebagai instrumen yang akan diuji validitasnya dengan indeks prestasi semester satu (sebagai kriteria eksternal) signifikan, maka tes masuk disebut dapat dikatakan *valid* berdasarkan ukuran validitas prediktif.

Contoh untuk validitas kongruen adalah jika kita hendak menguji tes sumatif yang dimaksudkan untuk mengukur penguasaan materi pelajaran selama satu semester dengan menggunakan hasil ulangan-ulangan harian semester yang bersamaan

sebagai kriteria eksternal. Jika koefisien korelasi antara skor tes sumatif (sebagai instrumen internal) dengan nilai ulangan-ulangan harian (sebagai kriteria eksternal) signifikan, maka tes sumatif tersebut dapat dikatakan *valid* berdasarkan validitas kongkuren.

B. Instrumen Penilaian Hasil Belajar Fiqh

1. Instrumen Penilaian

Telaah instrumen kegiatan pada telaah instrumen adalah menelaah apakah; 1) butir pertanyaan atau pernyataan sesuai dengan indikator, 2) bahasa yang digunakan komunikatif dan menggunakan tata bahasa yang benar, 3) butir peranyaaan/pernyataan tidak bias, 4) format instrumen menarik untuk dibaca, 5) pedoman menjawab atau mengisi instrumen jelas, dan 6) jumlah butir dan/atau panjang kalimat pertanyaan/pernyataan sudah tepat sehingga tidak menjemukan untuk dibaca atau dijawab.

Telaah dilakukan oleh pakar dalam bidang yang diukur dan akan lebih baik

bila ada pakar penilaian. Telaah bisa juga dilakukan oleh teman sejawat bila yang diinginkan adalah masukan tentang bahasa dan format instrumen. Bahasa yang digunakan adalah yang sesuai dengan tingkat pendidikan responden. Hasil telaah selanjutnya digunakan untuk memperbaiki instrumen. Panjang instrumen berhubungan dengan masalah kebosanan, yaitu tingkat kejemuhan dalam mengisi instrumen. Lama pengisian instrumen sebaiknya tidak lebih dari 30 menit. Langkah pertama dalam menulis suatu pertanyaan/ pernyataan adalah informasi apa yang ingin diperoleh, struktur pertanyaan, dan pemilihan kata-kata. Pertanyaan yang diajukan jangan sampai bias, yaitu mengarahkan jawaban responden pada arah tertentu, positif atau negatif.

Perbaikan instrumen dilakukan terhadap butir-butir pertanyaan/ pernyataan yang tidak baik, berdasarkan analisis hasil ujicoba. Bisa saja hasil telaah instrumen baik, namun hasil ujicoba empirik tidak baik. Untuk itu

butir pertanyaan/pernyataan instrumen harus diperbaiki. Perbaikan termasuk mengakomodasi saran-saran dari responden ujicoba. Instrumen sebaiknya dilengkapi dengan pertanyaan terbuka.

Hasil telaah instrumen digunakan untuk memperbaiki instrumen. Perbaikan dilakukan terhadap konstruksi instrumen, yaitu kalimat yang digunakan, waktu yang diperlukan untuk mengisi instrumen, cara pengisian atau cara menjawab instrumen, dan pengetikan.

Merakit instrumen Setelah instrumen diperbaiki selanjutnya instrumen dirakit, yaitu menentukan format tata letak instrumen dan urutan pertanyaan/pernyataan. Format instrumen harus dibuat menarik dan tidak terlalu panjang, sehingga responden tertarik untuk membaca dan mengisinya. Setiap sepuluh pertanyaan sebaiknya dipisahkan dengan cara memberi spasi yang lebih, atau diberi batasan garis empat persegi panjang. Urutkan pertanyaan/pernyataan sesuai dengan tingkat kemudahan dalam menjawab atau mengisinya.

Ujicoba instrumen, setelah dirakit instrumen diujicobakan kepada responden, sesuai dengan tujuan penilaian apakah kepada peserta didik, kepada guru atau orang tua peserta didik. Untuk itu dipilih sampel yang karakteristiknya mewakili populasi yang ingin dinilai. Bila yang ingin dinilai adalah peserta didik MA, maka sampelnya juga peserta didik MA. Sampel yang diperlukan minimal 30 peserta didik, bisa berasal dari satu madrasah atau lebih.

Pada saat ujicoba yang perlu dicatat adalah saran-saran dari responden atas kejelasan pedoman pengisian instrumen, kejelasan kalimat yang digunakan, dan waktu yang diperlukan untuk mengisi instrumen. Waktu yang digunakan disarankan bukan waktu saat responden sudah lelah. Selain itu sebaiknya responden juga diberi minuman agar tidak lelah. Perlu diingat bahwa pengisian instrumen penilaian afektif bukan merupakan tes, sehingga walau ada batasan waktu namun tidak terlalu ketat.

Agar responden mengisi instrumen dengan akurat sesuai harapan, maka

sebaiknya instrumen dirancang sedemikian rupa sehingga waktu yang diperlukan mengisi instrumen tidak terlalu lama. Berdasarkan pengalaman, waktu yang diperlukan agar tidak jenuh adalah 30 menit atau kurang.

Analisis hasil ujicoba, analisis hasil ujicoba meliputi variasi jawaban tiap butir pertanyaan/ pernyataan. Jika menggunakan skala instrumen 1 sampai 7, dan jawaban responden bervariasi dari 1 sampai 7, maka butir pertanyaan/ pernyataan pada instrumen ini dapat dikatakan baik. Namun apabila jawabannya hanya pada satu pilihan jawaban saja, misalnya pada pilihan nomor 3, maka butir instrumen ini tergolong tidak baik. Indikator yang digunakan adalah besarnya daya beda. Bila daya beda butir instrumen lebih dari 0,30, butir instrumen tergolong baik.

Telaah dilakukan oleh pakar dalam bidang yang diukur dan akan lebih baik bila ada pakar penilaian. Telaah bisa juga dilakukan oleh teman sejawat bila yang diinginkan adalah masukan tentang

bahasa dan format instrumen. Bahasa yang digunakan adalah yang sesuai dengan tingkat pendidikan responden. Hasil telaah selanjutnya digunakan untuk memperbaiki instrumen.

Panjang instrumen berhubungan dengan masalah kebosanan, yaitu tingkat kejemuhan dalam mengisi instrumen. Lama pengisian instrumen sebaiknya tidak lebih dari 30 menit. Langkah pertama dalam menulis suatu pertanyaan/ pernyataan adalah informasi apa yang ingin diperoleh, struktur pertanyaan, dan pemilihan kata-kata. Pertanyaan yang diajukan jangan sampai bias, yaitu mengarahkan jawaban responden pada arah tertentu, positif atau negatif.

Perbaikan instrumen dilakukan terhadap butir-butir pertanyaan/ pernyataan yang tidak baik, berdasarkan analisis hasil ujicoba. Bisa saja hasil telaah instrumen baik, namun hasil ujicoba empirik tidak baik. Untuk itu butir pertanyaan/pernyataan instrumen harus diperbaiki. Perbaikan termasuk mengakomodasi saran-saran dari

responden ujicoba. Instrumen sebaiknya dilengkapi dengan pertanyaan terbuka.

Hasil telaah instrumen digunakan untuk memperbaiki instrumen. Perbaikan dilakukan terhadap konstruksi instrumen, yaitu kalimat yang digunakan, waktu yang diperlukan untuk mengisi instrumen, cara pengisian atau cara menjawab instrumen, dan pengetikan.

Merakit instrumen, setelah instrumen diperbaiki selanjutnya instrumen dirakit, yaitu menentukan format tata letak instrumen dan urutan pertanyaan/ pernyataan. Format instrumen harus dibuat menarik dan tidak terlalu panjang, sehingga responden tertarik untuk membaca dan mengisinya. Setiap sepuluh pertanyaan sebaiknya dipisahkan dengan cara memberi spasi yang lebih, atau diberi batasan garis empat persegi panjang. Urutkan pertanyaan/ pernyataan sesuai dengan tingkat kemudahan dalam menjawab atau mengisinya.

Indikator lain yang diperhatikan adalah indeks keandalan yang dikenal dengan indeks reliabilitas. Batas indeks

reliabilitas minimal 0,70. Bila indeks ini lebih kecil dari 0,70, kesalahan pengukuran akan melebihi batas. Oleh karena itu diusahakan agar indeks keandalan instrumen minimal 0,70.

Perbaikan instrumen, perbaikan dilakukan terhadap butir-butir pertanyaan/pernyataan yang tidak baik, berdasarkan analisis hasil ujicoba. Bisa saja hasil telaah instrumen baik, namun hasil ujicoba empirik tidak baik. Untuk itu butir pertanyaan/pernyataan instrumen harus diperbaiki. Perbaikan termasuk mengakomodasi saran-saran dari responden ujicoba. Instrumen sebaiknya dilengkapi dengan pertanyaan terbuka.

2. Hasil Belajar Fiqh

Hasil belajar Fiqh sebaiknya mencakup pada tiga ranah dalam pembelajaran, yaitu ranah kognitif, afektif, dan psikomotor. Ranah kognitif meliputi kemampuan menghafal, memahami, menerapkan, menganalisis, mensintesis, dan kemampuan mengevaluasi. Kemampuan yang penting pada ranah kognitif adalah kemampuan menerapkan konsep-

konsep untuk memecahkan masalah yang dijumpai di lapangan. Kemampuan ini sering disebut dengan kemampuan mentransfer pengetahuan ke berbagai situasi sesuai dengan konteksnya.

Ranah afektif mencakup perilaku seperti perasaan, minat, sikap, emosi, dan nilai. Menurut Popham (1995: 37) bahwa ranah afektif menentukan keberhasilan belajar seseorang. Sejalan dengan pendapat ini adalah Bobby De Porter dan Mike Hernacki (1992: 35), keberhasilan seorang dalam hidupnya akan sangat ditentukan oleh emosional seseorang dalam mensikapi semua permasalahan dalam hidupnya, maka penting menjaga emosional dalam diri kita. Ranah psikomotor yaitu kemampuan yang berkaitan dengan gerak menggunakan otot seperti lari, melompat, menari, melukis, membongkar dan memasang peralatan, dan sebagainya. Peringkat kemampuan psikomotor ada lima, yaitu; 1) gerakan reflek, 2) gerakan dasar, 3) kemampuan persepsual, 4) kemampuan fisik, gerakan terampil, dan 5) komunikasi nondiskursip

(Sax, 1980: 76).

Hasil belajar menurut Bloom (1976: 53) mencakup prestasi belajar, kecepatan belajar, dan hasil afektif. Andersen (1981: 37) sependapat dengan Bloom bahwa karakteristik manusia meliputi cara yang tipikal dari berpikir, berbuat, dan perasaan. Tipikal berpikir berkaitan dengan ranah kognitif, tipikal berbuat berkaitan dengan ranah psikomotor, dan tipikal perasaan berkaitan dengan ranah afektif. Ketiga ranah ini adalah merupakan hasil belajar. Kemampuan afektif merupakan bagian dari hasil belajar yang memiliki peran sangat penting. Keberhasilan pembelajaran pada ranah kognitif dan psikomotor sangat ditentukan oleh kondisi afektif peserta didik. Peserta didik yang memiliki minat belajar dan sikap positif pada pelajaran akan merasa senang mempelajari pelajaran tersebut sehingga hasil pembelajaran akan optimal.

Mata pelajaran Fiqh di Madrasah telah diberikan mulai di Madrasah Ibtidaiyah (MI), kemudian di Madrasah Tsanawiyah (MTs), dan masih dilanjutkan di Madrasah

Aliyah (MA). Mencermati Fiqh sudah diberikan sejak pada pendidikan dasar dan menengah di madrasah, maka harapannya untuk jenjang pendidikan atas di MA siswa sudah memiliki pengetahuan dan pemahaman yang baik terhadap materi-materi fiqh. Perilaku dan pengamalan materi fiqh bagi siswa MA mestinya sudah menjadi rutinitas sehari-hari sebagaimana tujuan dan ruang lingkup mata pelajaran di MA. Namun demikian perlu ada upaya untuk lebih meningkatkan lagi siswa MA agar lebih aplikatif atas materi yang sudah diberikan agar lebih *impersonalized* pada diri siswa-siswa MA.

Adapun mata pelajaran Fiqh di MA bertujuan untuk:

1. Mengetahui dan memahami prinsip-prinsip, kaidah-kaidah dan tata cara pelaksanaan hukum Islam baik yang menyangkut aspek ibadah maupun muamalah untuk dijadikan pedoman hidup dalam kehidupan pribadi dan sosial.
2. 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.

Ruang lingkup mata pelajaran Fiqih di Madrasah Aliyah meliputi; kajian tentang prinsip-prinsip ibadah dan syari'at dalam Islam, hukum Islam dan perundang-undangan tentang zakat dan haji, hikmah dan cara pengelolaannya, hikmah kurban dan akikah, ketentuan hukum Islam tentang pengurusan jenazah, hukum Islam tentang kepemilikan; konsep perekonomian dalam Islam dan hikmahnya, hukum Islam tentang pelepasan dan perubahan harta beserta hikmahnya, hukum Islam tentang *wakaalah* dan *sulhu* beserta hikmahnya, hukum Islam tentang *daman* dan *kafaalah* beserta hikmahnya, riba, bank dan asuransi, ketentuan Islam tentang *jinaayah*, *Huduud* dan hikmahnya, ketentuan Islam tentang peradilan dan hikmahnya, hukum Islam tentang

keluarga, waris, ketentuan Islam tentang *siyaasah syar'iyah*, sumber hukum Islam dan hukum taklifi, dasar-dasar *istinbaath* dalam fikih Islam; kaidah-kaidah usul fikih dan penerapannya.

3. Instrumen Penilaian Hasil Belajar Fiqh

Instrumen penilaian yang akan digunakan dalam penelitian akan sangat bergantung dari materi pelajaran, dalam hal ini adalah fiqh. Mata pelajaran fiqh syarat dengan emosi, nilai, dan perilaku yang akan menjadi amalan dalam sehari-hari. Karenanya instrumen yang akan banyak digunakan adalah instrumen yang terkait dengan nilai, dengan kata lain instrumen untuk ranah afektif (meskipun tidak mengesampingkan dua ranah yang lain yaitu kognitif dan psikomotor) akan menjadi dasar kerangka dalam penelitian ini.

Menurut Andersen (1980: 55) ada dua metode yang dapat digunakan untuk mengukur ranah afektif yaitu metode observasi dan metode laporan-diri. Penggunaan metode observasi berdasarkan pada asumsi bahwa karakteristik afektif

dapat dilihat dari perilaku atau perbuatan yang ditampilkan, reaksi psikologi atau keduanya. Metode laporan-diri berasumsi bahwa yang mengetahui keadaan afektif seseorang adalah dirinya sendiri. Menurut Lewin (dalam Andersen, 1980: 58) perilaku seseorang merupakan fungsi dari watak (kognitif, afektif, dan psikomotor) dan karakteristik lingkungan saat perilaku atau perbuatan ditampilkan. Oleh karena tindakan, perbuatan atau pengalaman seseorang ditentukan watak dirinya dan lingkungan.

Penilaian ranah afektif (Djemari Mardapi, 2002: 13) peserta didik dilakukan dengan menggunakan instrumen afektif, di antaranya adalah:

- 1) Instrumen sikap, sikap merupakan perasaan positif atau negatif terhadap suatu obyek. Pertanyaan tentang sikap meminta responden menunjukkan perasaan yang positif atau negatif terhadap suatu obyek atau suatu kebijakan. Kata yang sering digunakan pada pertanyaan sikap menyatakan arah perasaan seseorang seperti;

menerima-menolak, menyenangkan-tidak menyenangkan, baik-buruk, diingini-tidak diingini. Contoh indikator sikap terhadap mata pelajaran Fiqh adalah:

- a) Membaca buku panduan fiqh
- b) Belajar fiqh
- c) Interaksi dengan guru fiqh
- d) Mengerjakan tugas fiqh
- e) Diskusi tentang fiqh
- f) Memiliki buku fiqh

Contoh kuesionernya seperti; (1) saya senang membaca buku fiqh, (2) tidak semua orang harus belajar fiqh, (3) saya sering bertanya pada guru tentang pelajaran fiqh, (4) saya tidak senang pada tugas pelajaran fiqh, (5) saya berusaha mengerjakan soal-soal fiqh sebaik mungkin, dan (6) fiqh penting untuk semua peserta didik.

- 2) Instrumen minat, instrumen minat bertujuan untuk memperoleh informasi tentang minat peserta didik terhadap suatu pelajaran, selanjutnya digunakan untuk meningkatkan meningkatkan

minat peserta didik terhadap suatu mata pelajaran. Definisi minat adalah keingintahuan seseorang tentang keadaan suatu obyek. Contoh indikator minat terhadap mata pelajaran Fiqh adalah:

- a) Catatan pelajaran fiqh
- b) Usaha memahami fiqh
- c) Memiliki buku fiqh
- d) Kehadiran dalam pelajaran fiqh

Contoh kuesioner yaitu; (1) catatan pelajaran fiqh saya lengkap, (2) saya selalu menyiapkan pertanyaan sebelum pelajaran fiqh, (3) saya berusaha akan memahami materi fiqh, (4) saya senang mengerjakan tugas fiqh, (5) saya berusaha selalu hadir dalam pembelajaran fiqh.

- 3) Instrumen Konsep diri, instrumen ini bertujuan untuk mengetahui kekuatan dan kelemahan diri sendiri. Konsep diri adalah persepsi seseorang terhadap dirinya yang menyangkut keunggulan dan kelemahannya. Secara operasional

konsep diri adalah pernyataan tentang kemampuan diri sendiri yang menyangkut mata pelajaran fiqh.

- 4) Instrumen Nilai, nilai merupakan konsep yang penting dalam pembentukan kompetensi peserta didik. Pencapaian kemampuan kognitif dan psikomotor tidak akan memberi manfaat bagi masyarakat apabila tidak diikuti dengan kompetensi afektif. Nilai adalah keyakinan seseorang tentang keadaan suatu obyek atau kegiatan. Instrumen nilai dan keyakinan bertujuan mengungkap nilai dan keyakinan individu. Informasi yang diperoleh berupa nilai dan keyakinan yang positif dan negatif. Hal-hal yang positif akan diperkuat sedangkan yang negatif pada akhirnya akan dihilangkan.
- 5) Instrumen Moral, instrumen ini bertujuan untuk mengetahui moral peserta didik. Moral didefinisikan sebagai pendapat, tindakan yang dianggap baik dan yang dianggap tidak baik.

Mata pelajaran fiqh memuat lima hal tersebut di atas yaitu sikap, minat, konsep

diri, nilai dan moral. Sehingga untuk membuat instrumen penilaian fiqh juga akan mendasarkan pada instrumen kelima hal tersebut. Instrumen penilaian hasil belajar fiqh akan akurat jika memperhatikan kepada indikator yang ada dalam materi fiqh itu sendiri. Oleh karena penyusunan instrumen fiqh akan sangat bergantung pada penemuan indikator-indikator dalam materi fiqh.

C. Tes

a. Pengertian Tes

Secara umum tes diartikan sebagai alat yang dipergunakan untuk mengukur pengetahuan atau penguasaan obyek ukur terhadap seperangkat isi dan materi tertentu. Menurut Sudijono (1996: 52), tes adalah alat atau prosedur yang dipergunakan dalam rangka pengukuran dan penilaian. Tes dapat juga diartikan sebagai alat pengukur yang mempunyai standar objektif, sehingga dapat dipergunakan secara meluas, serta betul-betul dapat digunakan untuk mengukur dan membandingkan keadaan psikis atau tingkah laku individu (Anastasi dan Turabian, 1997: 31). Menurut

Cronbach (1984, 47), tes merupakan suatu prosedur yang sistematis untuk mengamati atau mendeskripsikan satu atau lebih karakteristik seseorang dengan menggunakan standar numerik atau sistem kategori.

Menurut Bruce (1978: 75), tes dapat digunakan untuk mengukur banyaknya pengetahuan yang diperoleh individu dari suatu bahan pelajaran yang terbatas pada tingkat tertentu. Oleh karena itu, tes merupakan alat ukur yang banyak dipergunakan dalam dunia pendidikan. Hal ini dikarenakan umumnya orang masih memandang bahwa indikator keberhasilan seseorang mengikuti pendidikan adalah dilihat dari seberapa banyak orang menguasai materi yang telah dipelajarinya dalam suatu jenjang pendidikan tertentu.

Tes juga bisa diartikan sebagai instrumen atau alat yang digunakan untuk memperoleh informasi tentang individu atau objek. Sebagai alat pengumpul informasi atau data, tes harus dirancang secara khusus. Validitas yang didasarkan pada bukti dalam suatu tes sangat penting

dimana hal ini dapat diperoleh dari analisis hubungan antara isi dan konstruk suatu tes, sehingga dapat untuk dilakukan suatu penilaian dengan baik (AERA, 1999: 11). Kekhususan tes terlihat dari bentuk soal tes yang digunakan, jenis pertanyaan, rumusan pertanyaan yang diberikan, dan pola jawabannya harus dirancang menurut kriteria yang telah ditetapkan. Demikian juga waktu yang disediakan untuk menjawab pertanyaan serta pengadministrasian tes juga dirancang secara khusus. Selain itu aspek yang ditekankanpun terbatas. Biasanya meliputi ranah kognitif, afektif, dan psikomotor. Kekhususan-kekhususan tersebut berbeda antara satu tes dengan tes yang lain. Tes ini dapat berupa pertanyaan tertulis, wawancara, pengamatan tentang unjuk kerja fisik, checklist, dan lain-lain. Tes juga merupakan beberapa konstruk atau domain yang dapat disusun menurut beberapa cara yang harus diperhatikan secara seksama, sehingga nantinya akan bisa dikembangkan (AERA, 1999: 37).

b. Fungsi Tes

Secara umum ada beberapa macam fungsi tes di dalam dunia pendidikan. Pertama, tes dapat berfungsi sebagai alat untuk mengukur prestasi belajar siswa. Sebagai alat untuk mengukur prestasi belajar siswa tes dimaksudkan untuk mengukur tingkat perkembangan atau kemajuan yang telah dicapai siswa setelah menempuh proses belajar-mengajar dalam jangka waktu tertentu. Dalam kaitan ini, tes berfungsi sebagai alat untuk mengukur keberhasilan program pengajaran. Sebagai alat untuk mengukur keberhasilan program pengajaran, tes berfungsi untuk menunjukkan seberapa jauh program pengajaran yang telah ditentukan dapat dicapai, seberapa banyak yang belum tercapai serta menentukan langkah apa yang perlu dilakukan untuk mencapainya.

Kedua, tes dapat berfungsi sebagai motivator dalam pembelajaran. Hampir semua ahli teori pembelajaran menekankan pentingnya umpan balik yang berupa nilai untuk meningkatkan intensitas kegiatan belajar. Thorndike

(1991) mengemukakan bahwa siswa akan belajar lebih giat dan berusaha lebih keras apabila mereka mengetahui bahwa di akhir program yang sedang ditempuh akan ada tes untuk mengetahui nilai dan prestasi mereka. Ebel (1979) mengemukakan bahwa tes kadang-kadang dianggap sebagai motivator ekstrinsik. Fungsi ini dapat optimal apabila nilai hasil tes yang diperoleh siswa betul-betul obyektif dan sah, baik secara internal maupun secara eksternal yang dapat dirasakan langsung oleh siswa yang diberi nilai melalui tes.

Ketiga, tes dapat berfungsi untuk upaya perbaikan kualitas pembelajaran. Dalam rangka perbaikan kualitas pembelajaran ada tiga jenis tes yang perlu dibahas, yaitu tes penempatan, tes diagnostik, dan tes formatif. Tes yang dilaksanakan untuk keperluan penempatan bertujuan agar setiap siswa yang mengikuti kegiatan pembelajaran di kelas atau pada jenjang pendidikan tertentu dapat mengikuti kegiatan pembelajaran secara efektif, karena sesuai dengan bakat dan kemampuannya masing-masing.

Mengingat bahwa faktor penentu keberhasilan kegiatan pembelajaran dari aspek subyek belajar (peserta didik) adalah pengetahuan prasyarat dan bakat siswa, maka dalam evaluasi penempatan dapat digunakan alat evaluasi berupa tes bakat dan tes pengetahuan untuk mengetahui penguasaan siswa terhadap materi atau konsep prasyarat yang diperlukan untuk mempelajari konsep atau materi pada suatu kegiatan pembelajaran.

Tes bakat sangat penting dalam evaluasi penempatan, karena keberhasilan kegiatan pembelajaran dalam suatu bidang tertentu sangat dipengaruhi oleh bakat siswa terhadap bidang yang dipelajari. Kenyataan menunjukkan bahwa seorang siswa yang gagal dalam menempuh pendidikan pada suatu program studi tertentu kemudian dapat berhasil dengan cemerlang setelah beralih menempuh pendidikan pada bidang atau program studi yang lain.

Evaluasi diagnostik dilaksanakan untuk mengidentifikasi kesulitan belajar yang dialami siswa, menentukan faktor-

faktor yang menyebabkan terjadinya kesulitan belajar, dan menetapkan cara mengatasi kesulitan belajar tersebut. Berhasil atau gagalnya suatu kegiatan pembelajaran dalam proses pendidikan pada suatu jenis dan/atau jenjang pendidikan tertentu sangat dipengaruhi oleh apakah siswa mengalami kesulitan belajar atau tidak. Makin serius kesulitan belajar yang dialami siswa, maka makin besar kemungkinan gagal dan kebalikannya.

Oleh karena itu, keberhasilan dalam mengatasi serta mengurangi kesulitan belajar siswa akan meningkatkan keberhasilan kegiatan belajar. Kesulitan belajar siswa dapat bersumber dari kurangnya penguasaan mereka terhadap materi atau konsep prasyarat dari suatu konsep dan materi yang dipelajari serta dapat pula bersumber dari ketidaksesuaian antara bidang ilmu yang dipelajari dengan bakat siswa.

Tes formatif pada dasarnya adalah tes yang bertujuan untuk mendapatkan umpan balik bagi usaha perbaikan kualitas

pembelajaran dalam konteks kelas. Kualitas pembelajaran di kelas ditentukan oleh intensitas proses belajar (proses intern) dalam diri setiap siswa sebagai subyek belajar sekaligus peserta didik. Intensitas proses belajar dalam arti intern tersebut ditentukan oleh kesesuaian antara strategi dan metode pembelajaran dengan struktur kognitif (termasuk bakat) siswa sebagai peserta didik dan karakteristik konsep atau materi yang dipelajari. Atau dapat dikatakan bahwa intensitas proses belajar dalam arti intern adalah hasil dari interaksi yang harmonis antara tiga unsur yaitu karakteristik atau struktur kognitif subyek belajar, karakteristik konsep yang dipelajari dan strategi pembelajaran yang digunakan oleh guru sebagai pengelola kegiatan pembelajaran di kelas.

Oleh karena itu tes formatif yang diselenggarakan dalam selang waktu yang relatif pendek akan memberikan masukan atau umpan balik yang dapat digunakan oleh guru sebagai pengelola kegiatan pembelajaran dalam meningkatkan intensitas proses belajar

dalam diri setiap subyek belajar melalui peningkatan kesesuaian antara tiga unsur, yaitu struktur kognitif subyek belajar, karakteristik konsep yang dipelajari dan strategi pembelajaran yang digunakan.

Keempat, tes yang dimaksudkan untuk menentukan berhasil atau tidaknya siswa sebagai syarat untuk melanjutkan pendidikan pada jenjang yang lebih tinggi. Untuk keperluan ini dikenal istilah tes sumatif. Tes sumatif yang dikenal dengan istilah *summative test* adalah tes hasil belajar yang dilaksanakan setelah sekumpulan materi pelajaran atau satuan program pengajaran selesai diberikan.

Di sekolah, tes sumatif ini dikenal dengan tes ulangan umum. Tes sumatif ini dilaksanakan dengan tujuan untuk menentukan nilai yang menjadi lambang keberhasilan siswa setelah mereka menempuh proses pembelajaran dalam jangka waktu tertentu. Hasil tes sumatif berguna untuk (a) menentukan kedudukan atau ranking masing-masing siswa dikelompoknya; (b) menentukan dapat atau tidaknya siswa melanjutkan

program pembelajaran berikutnya; dan (c) menginformasikan kemajuan siswa untuk disampaikan kepada pihak lain seperti orang tua, sekolah, masyarakat, dan lapangan kerja. Jika tes sumatif dilaksanakan pada setiap akhir semester atau cawu, maka pada setiap akhir jenjang pendidikan dilaksanakan tes akhir atau biasa disebut evaluasi belajar tahap akhir. Penyusunan tes juga perlu mempertimbangkan kemampuan kognitif sebagaimana materi yang telah diberikan. Sebagaimana pendapat David R. Krathwohl (1985: 7) dalam Taksonomi Bloom's; *knowledge, comprehension, application, analysis, synthesis, evaluation*. Taksonomi pembelajaran, pengajaran, dan penilaian; *remember, understand, apply, analyze, evaluate, create*.

Sebagai contoh taksonomi Bloom tentang tujuan pengembangan pembelajaran adalah kemampuan berpikir yang berisi; (1) pengetahuan (*knowledge*), (2) pemahaman (*comprehension*), (3) penerapan (*application*), (4) analisis (*analysis*), (5) sintesis (*synthesis*), dan

(6) evaluasi (*evaluation*). Proses untuk mempertimbangkan secara hirarki kemampuan tiga yang terakhir masuk pada wilayah berpikir tingkat tinggi. Kemudian Anderson and Krathwohl (2001) merevisi taksonomi Bloom's dengan menambah kemampuan mencipta pengetahuan baru (*create new knowledge*)

D. Penilaian

Berbicara penilaian tidak bisa lepas dari pembahasan evaluasi dan pengukuran. Ketiga istilah ini ada saling keterkaitan yang kalau tidak dipahami dengan seksama spintas tiada perbedaan. Sebagai upaya menghindari bias atas ketiga istilah tersebut, diawal perlu kami bedakan yaitu; pertama evaluasi, (Gronlund, 1985) menyatakan suatu proses yang sistematis untuk menentukan atau membuat keputusan, sampai sejauh mana tujuan program telah tercapai. Pendapat yang sama juga dikemukakan oleh Wrihstone, dkk (1956) yang mengemukakan bahwa evaluasi pendidikan adalah penaksiran terhadap pertumbuhan dan kemajuan siswa ke arah tujuan atau nilai-nilai yang telah ditetapkan dalam kurikulum. Evaluasi juga dapat diartikan sebagai proses

menilai sesuatu berdasarkan kriteria atau tujuan yang telah ditetapkan, yang selanjutnya diikuti dengan pengambilan keputusan terhadap focus pembahasan yang dievaluasi.

Uraian di atas mengingatkan kepada kita bahwa evaluasi semesti berlaku adil untuk semua pihak terkait dalam subyek evaluasi. Evaluasi bukan untuk memonopoli siswa, menghakimi siswa, menyalahkan siswa ketika terjadi sesuatu yang tidak diinginkan oleh pihak evaluator, melainkan merupakan kegiatan menganalisis semua subyek yang melakukan aktivitas, termasuk di dalamnya aktivitas itu sendiri. Sehingga tidaklah tepat ketika ada upaya mendeskritkan salah satu pihak yang ada dalam aktivitas evaluasi, karena dua-duanya mempunyai probabilitas terjadinya kekeliruan, sebagai penyebab target tidak tercapai optimal.

Sepintas tiada beda antara evaluasi dengan penilaian, namun sesungguhnya berbeda, penilaian atau assessment berarti menilai sesuatu. Proses menilai tidak bisa menafikan kriteria-kriteria yang menjadi parameter atau ukuran untuk menilai, apakah final dari menilai akan diberikan predikat baik atau tidak baik, lulus atau tidak lulus, layak atau tidak

layak. Akhir dari penilaian adalah memberikan nilai sedangkan evaluasi sampai memberikan keputusan. Penilaian merupakan tindakan atau proses menentukan nilai sesuatu objek dan memberikan keputusan tentang nilai. Penilaian dapat dilakukan berdasarkan hasil pengukuran atau dapat pula dipengaruhi oleh hasil pengukuran.

Kemudian pengukuran atau istilah lainnya *measurement* merupakan satu kegiatan yang dilakukan untuk mengukur dalam arti memberi angka atas sesuatu yang menjadi obyek pengukuran. Secara konseptual angka-angka hasil pengukuran pada dasarnya adalah *kontinum* yang bergerak dari suatu kutub ke kutub yang lain berlawanan, misal dari rendah ke tinggi atau dari positif ke negatif, atau dari demokratis ke otoriter diberi angka dari 0 sampai 100.

Berdasarkan beberapa pengertian evaluasi, penilaian dan pengukuran yang dikemukakan di atas, merupakan tiga konsep yang berbeda. Namun dalam praktek terutama dalam dunia pendidikan, ketiga konsep tersebut sering dipraktekkan dalam satu rangkaian kegiatan, biasa hal ini tercermin pada instrument tes.

Linn dan Gronlund (1995: 47) menyatakan bahwa tes yang baik harus memenuhi tiga karakteristik, yaitu: validitas, reliabilitas, dan usabilitas. Validitas artinya ketepatan interpretasi hasil prosedur pengukuran, reliabilitas artinya konsistensi hasil pengukuran, dan usabilitas artinya praktis prosedurnya. Di samping itu, Cohen dkk. (1992: 28) juga menyatakan bahwa tes yang baik adalah tes yang *valid* artinya mengukur apa yang hendak diukur. Nitko (1996 : 36) menyatakan bahwa validitas berhubungan dengan interpretasi atau makna dan penggunaan hasil pengukuran peserta didik. Messick (1993: 13) menjelaskan bahwa validitas tes merupakan suatu integrasi pertimbangan evaluatif derajat keterangan empiris yang mendasarkan pemikiran teoritis yang mendukung ketepatan dan kesimpulan berdasarkan pada skor tes. Adapun validitas dalam model Rasch adalah sesuai atau fit dengan model (Hambleton dan Swaminathan, 1985: 73).

E. Mata Pelajaran Fiqh di MA

Di dalam UU No. 20 Tahun 2003 tentang Sistem Pendidikan Nasional, dinyatakan bahwa pendidikan nasional berfungsi mengembangkan

kemampuan dan membentuk watak serta peradaban bangsa yang bermartabat dalam rangka mencerdaskan kehidupan bangsa, bertujuan untuk berkembangnya potensi peserta didik agar menjadi manusia yang beriman dan bertakwa kepada Tuhan Yang Maha Esa, berakhlak mulia, sehat, berilmu, cakap, kreatif, mandiri, dan menjadi warga negara yang demokratis serta bertanggung jawab. Untuk mencapai tujuan tersebut, salah satu bidang studi yang harus dipelajari oleh peserta didik di madrasah adalah pendidikan agama Islam, yang dimaksudkan untuk membentuk peserta didik menjadi manusia yang beriman dan bertakwa kepada Tuhan Yang Maha Esa serta berakhlak mulia.

Mengacu pada Permenag No. 2 Tahun 2008 tentang SKL dan SI pada PAI dan Bahasa Arab, maka Pendidikan Agama Islam di Madrasah Aliyah terdiri atas empat mata pelajaran, yaitu: Al-Qur'an-Hadis, Akidah-Akhlak, Fiqih, dan Sejarah Kebudayaan Islam. Masing-masing mata pelajaran tersebut pada dasarnya saling terkait, isi mengisi dan melengkapi. Al-Qur'an-Hadis merupakan sumber utama ajaran Islam, dalam arti ia merupakan sumber akidah-

akhlak, syari'ah/fikih (ibadah, muamalah), sehingga kajiannya berada di setiap unsur tersebut. Akidah (Usuluddin) atau keimanan merupakan akar atau pokok agama. Syariah/fikih (ibadah, muamalah) dan akhlak bertitik tolak dari akidah, yakni sebagai manifestasi dan konsekuensi dari akidah (keimanan dan keyakinan hidup). Syari'ah/fikih merupakan sistem norma (aturan) yang mengatur hubungan manusia dengan Allah, sesama manusia dan dengan makhluk lainnya. Akhlak merupakan aspek sikap hidup atau kepribadian hidup manusia, dalam arti bagaimana sistem norma yang mengatur hubungan manusia dengan Allah (ibadah dalam arti khas) dan hubungan manusia dengan manusia dan lainnya (muamalah) itu menjadi sikap hidup dan kepribadian hidup manusia dalam menjalankan sistem kehidupannya (politik, ekonomi, sosial, pendidikan, kekeluargaan, Kebudayaan/seni, iptek, olahraga/kesehatan, dan lain-lain) yang dilandasi oleh akidah yang kokoh. Sejarah Kebudayaan Islam merupakan perkembangan perjalanan hidup manusia muslim dari masa ke masa dalam usaha bersyariah (beribadah dan bermuamalah) dan berakhlak serta dalam mengembangkan sistem kehidupannya yang

dilandasi oleh akidah.

Empat mata pelajaran PAI di Madrasah tersebut memiliki karakteristik sendiri-sendiri. Al-Qur'an-Hadis, menekankan pada kemampuan baca tulis yang baik dan benar, memahami makna secara tekstual dan kontekstual, serta mengamalkan kandungannya dalam kehidupan sehari-hari. Aspek akidah menekankan pada kemampuan memahami dan mempertahankan keyakinan/keimanan yang benar serta menghayati dan mengamalkan nilai-nilai *al-asma' al-husna*. Aspek Akhlak menekankan pada pembiasaan untuk melaksanakan akhlak terpuji dan menjauhi akhlak tercela dalam kehidupan sehari-hari.

Aspek fikih menekankan pada kemampuan cara melaksanakan ibadah dan muamalah yang benar dan baik. Aspek sejarah Kebudayaan Islam menekankan pada kemampuan mengambil *ibrah* dari peristiwa-peristiwa bersejarah (Islam), meneladani tokoh-tokoh berprestasi, dan mengaitkannya dengan fenomena sosial, budaya, politik, ekonomi, iptek dan seni, dan lain-lain untuk mengembangkan Kebudayaan dan peradaban Islam.

Penyusunan Standar Kompetensi (SK) dan Kompetensi Dasar (KD) mata pelajaran Akidah-Akhlak di Madrasah Aliyah ini dilakukan dengan cara mempertimbangkan dan *review* Peraturan Menteri Pendidikan Nasional Nomor 22 Tahun 2006 tentang Standar Isi (SI) untuk Satuan Pendidikan Dasar dan Menengah, mata pelajaran Pendidikan Agama Islam aspek keimanan/akidah dan akhlak untuk SMA/MA, serta memperhatikan Surat Edaran Dirjen Pendidikan Islam Nomor: DJ.II.1/PP.00/ED/681/2006, tanggal 1 Agustus 2006, tentang Pelaksanaan Standar Isi, yang intinya bahwa Madrasah dapat meningkatkan kompetensi lulusan dan mengembangkan kurikulum dengan standar yang lebih tinggi.

Mata pelajaran Fiqih di Madrasah Aliyah adalah salah satu mata pelajaran Pendidikan Agama Islam yang merupakan peningkatan dari fikih yang telah dipelajari oleh peserta didik di Madrasah Tsanawiyah. Peningkatan tersebut dilakukan dengan cara mempelajari, memperdalam serta memperkaya kajian fikih baik yang menyangkut aspek ibadah maupun muamalah, yang dilandasi oleh prinsip-prinsip dan kaidah-kaidah usul fikih serta menggali

tujuan dan hikmahnya, sebagai persiapan untuk melanjutkan ke pendidikan yang lebih tinggi dan untuk hidup bermasyarakat. Secara substansial, mata pelajaran Fikih 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.

Mata pelajaran Fikih di Madrasah Aliyah bertujuan untuk:

1. Mengetahui dan memahami prinsip-prinsip, kaidah-kaidah dan tata cara pelaksanaan hukum Islam baik yang menyangkut aspek ibadah maupun muamalah untuk dijadikan pedoman hidup dalam kehidupan pribadi dan sosial.
2. 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.

Ruang lingkup mata pelajaran Fiqih di Madrasah Aliyah meliputi; kajian tentang prinsip-prinsip ibadah dan syari'at dalam Islam, hukum Islam dan perundang-undangan tentang zakat dan haji, hikmah dan cara pengelolaannya, hikmah kurban dan akikah, ketentuan hukum Islam tentang pengurusan jenazah, hukum Islam tentang kepemilikan; konsep perekonomian dalam Islam dan hikmahnya, hukum Islam tentang pelepasan dan perubahan harta beserta hikmahnya, hukum Islam tentang *wakaalah* dan *sulhu* beserta hikmahnya, hukum Islam tentang daman dan *kafaalah* beserta hikmahnya, riba, bank dan asuransi, ketentuan Islam tentang *jinaayah*, *Huduud* dan hikmahnya, ketentuan Islam tentang peradilan dan hikmahnya, hukum Islam tentang keluarga, waris, ketentuan Islam tentang *siyaasah syar'iyah*, sumber hukum Islam dan hukum *taklifi*, dasar-dasar *istinbaath* dalam fikih Islam; kaidah-kaidah usul fikih dan penerapannya.

Fiqh merupakan sesuatu amalan yang harus dijalankan oleh setiap umat manusia tanpa kecuali bagi orang yang sudah terbebani hukum (*mukallaf*). Maka sebagai konsekuensi seorang yang tidak mengamalkan fiqh ibadah mereka dalam kacamata fiqh dianggap tidak sah, oleh karenanya amalan mereka akan sia-sia saja. Jika amalan seseorang sudah tidak memenuhi kaidah fiqh sehingga tidak memenuhi syarat dan rukunnya maka ibadah yang dijalankan akan mempunyai kewajiban untuk diperbaiki sampai memenuhi syarat dan rukunnya. Dengan demikian pengetahuan fiqh dan pengamalannya menjadi sesuatu yang harus dipenuhi. Bahkan dalam kaidah fiqhiyah dikatakan bahwa sesuatu perbuatan yang akan menjadi syarat sah dan tidaknya ibadah digolongkan perbuatan tersebut menjadi wajib. Kita semua perlu memperhatikan agar ibadah kita sampai pada wilayah sah dalam kacamata fiqh maka pengetahuan dan penerapan fiqh dalam keseharian menjadi sesuatu yang harus dipenuhi.

Tabel 2. Standar Kompetensi dan Kompetensi Dasar Mata Pelajaran Fiqh

a. Kelas X, Semester 1

Standar Kompetensi		Kompetensi Dasar	
(1)		(2)	
1.	Memahami dan menerapkan prinsip-prinsip ibadah dan syari'at dalam Islam	1.1	Mengidentifikasi prinsip-prinsip ibadah dalam Islam
		1.2	Menjelaskan tujuan (<i>maqashid</i>) syari'at Islam
		1.3	Menunjukkan perilaku orang yang berpegang pada prinsip-prinsip dan tujuan ibadah dan syariah
		1.4	Menerapkan cara berpegang pada prinsip-prinsip dan tujuan ibadah dan syariah.
2.	Memahami dan menerapkan hukum Islam tentang zakat dan hikmahnya	2.1	Menjelaskan ketentuan Islam tentang zakat dan hikmahnya
		2.2	Menjelaskan ketentuan perundang-undangan tentang zakat
		2.3	Menunjukkan contoh penerapan ketentuan zakat
		2.4	Menerapkan cara pelaksanaan zakat sesuai ketentuan perundang-undangan

- | | | |
|----|--|---|
| 3. | Memahami dan mempraktekkan hukum Islam tentang haji dan hikmahnya | 3.1 Menjelaskan ketentuan Islam tentang haji dan hikmahnya |
| | | 3.2 Menjelaskan ketentuan perundang-undangan tentang haji |
| | | 3.3 Menunjukkan contoh penerapan ketentuan haji |
| | | 3.4 mempraktikkan pelaksanaan haji sesuai ketentuan perundang-undangan tentang haji |
| 4. | Memahami dan menerpkan hikmah kurban dan akikah | 4.1 Menjelaskan tata cara pelaksanaan kurban dan hikmahnya |
| | | 4.2 Menerapkan cara pelaksanaan kurban |
| | | 4.3 Menjelaskan ketentuan akikah dan hikmahnya |
| | | 4.4 Menerapkan cara pelaksanaan akikah |
| 5. | Memahami dan memperagakan ketentuan hukum Islam tentang pengurusan jenazah | 5.1 Menjelaskan tatacara pengurusan jenazah |
| | | 5.2 Memperagakan tatacara pengurusan jenazah |

b. Semester 2

6. Memahami dan memperagakan hukum Islam tentang kepemilikan	6.1	Mengidentifikasi aturan Islam tentang kepemilikan
	6.2	Menjelaskan ketentuan Islam tentang akad
	6.3	Memperagakan aturan Islam tentang kepemilikan dan akad
7. Memahami dan menerapkan konsep per-ekonomian dalam Islam dan hikmahnya	7.1	Menjelaskan aturan Islam tentang jual beli dan hikmahnya
	7.2	Menjelaskan aturan Islam tentang <i>khiyaar</i>
	7.3	Menjelaskan aturan Islam tentang <i>musaaqah</i> , <i>mu-zaara'ah</i> dan <i>mukhaabarah</i> serta hikmahnya
	7.4	Menjelaskan aturan Islam tentang <i>syirkah</i> dan hikmahnya
	7.5	Menjelaskan aturan Islam tentang <i>muraabahah</i> , <i>mud-haarabah</i> , dan <i>salam</i>
	7.6	Menerapkan cara jual beli, <i>khiyaar</i> , <i>musaaqah</i> , <i>mu-zaara'ah</i> , <i>mukhaabarah</i> , <i>syirkah</i> , <i>muraabahah</i> , <i>mud-haarabah</i> , dan <i>salam</i>

8. Memahami dan menerapkan hukum Islam tentang pelepasan dan perubahan harta beserta hikmahnya

- 8.1 Menjelaskan ketentuan Islam tentang wakaf beserta hikmah pelaksanaannya
- 8.2 Menjelaskan ketentuan Islam tentang hibah dan hikmah pelaksanaannya
- 8.3 Menjelaskan ketentuan Islam tentang sadakah beserta hikmah pelaksanaannya
- 8.4 Menjelaskan ketentuan Islam tentang hadiah beserta hikmah pelaksanaannya
- 8.5 Menerapkan cara pelaksanaan wakaf, hibah, sedekah, dan hadiah

9 Memahami dan menerapkan hukum Islam tentang *wakalah* dan *sulhu* beserta hikmahnya

- 9.1 Menjelaskan ketentuan Islam tentang *wakalah* dan hikmahnya
- 9.2 Menjelaskan ketentuan Islam tentang *sulhu* dan hikmahnya
- 9.3 Menerapkan cara *wakalah* dan *sulhu*

10 Memahami dan menerapkan hukum Islam tentang *daman* dan *kafalah* beserta hikmahnya

- 10.1 Menjelaskan ketentuan Islam tentang *daman* dan hikmahnya
- 10.2 Menjelaskan ketentuan Islam tentang *kafalah* dan hikmahnya
- 10.3 Menerapkan cara *daman* dan *kafalah*

- | | | | |
|----|--|------|---|
| 11 | Memahami dan menerapkan ketentuan Islam tentang riba, bank, dan asuransi | 11.1 | Menjelaskan hukum riba, bank, dan asuransi |
| | | 11.2 | Menerapkan ketentuan Islam tentang riba, bank, dan asuransi |

c. Kelas XI Semester I

Standar Kompetensi		Kompetensi Dasar	
(1)		(2)	
1.	Memahami dan mencotohkan ketentuan Islam tentang jinayah dan hikmahnya	1.1	Menjelaskan hukum pembunuhan dan hikmahnya
		1.2	Menjelaskan ketentuan hukum Islam tentang qishash dan hikmahnya
		1.3	Menjelaskan ketentuan hukum Islam tentang diyat dan kafaarat beserta hikmahnya
		1.4	Menunjukkan contoh-contoh qishash, diyaat dan kafaarat dalam hukum Islam
2.	Memahami ketentuan Islam tentang Huudud dan hikmahnya serta menjauhinya	2.1	Menjelaskan ketentuan hukum Islam tentang zina dan qadzaf beserta hikmahnya
		2.2	Menjelaskan ketentuan hukum Islam tentang minuman keras beserta hikmahnya
		2.3	Menjelaskan ketentuan hukum Islam tentang mencuri, menyamun dan merampok beserta hikmahnya

		2.4	Menjelaskan ketentuan hukum Islam tentang bughat beserta hikmahnya
3	Memahami dan mengidentifikasi ketentuan Islam tentang peradilan dan hikmahnya	3.1	Menjelaskan proses peradilan dalam Islam
		3.2	Mengidentifikasi ketentuan tentang hakim dan saksi dalam peradilan Islam

d. Kelas XI, Semester 2

1.	Memahami dan menerapkan hukum Islam tentang hukum keluarga	1.1	Menjelaskan ketentuan hukum perkawinan dalam Islam dan hikmahnya
		1.2	Menjelaskan ketentuan perkawinan menurut perundang-undangan di Indonesia
		1.3	Menjelaskan konsep Islam tentang talak, perceraian, iddah, ruju`, dan hikmahnya
		1.4	Menjelaskan ketentuan Islam tentang pengasuhan anak (hadhaanah)
2	Memahami dan mencontohkan hukum Islam tentang waris	2.1	Menjelaskan ketentuan hukum waris dalam Islam
		2.2	Menjelaskan keterkaitan waris dengan wasiat
		2.3	Menunjukkan contoh cara pelaksanaan waris dan wasiat

e. Kelas XII, Semester 1

Standar Kompetensi		Kompetensi Dasar	
(1)		(2)	
1.	Memahami dan mengidentifikasi ketentuan Islam tentang siyasah syar'iyah	1.1	Menjelaskan ketentuan Islam tentang pemerintahan (khilafah)
		1.2	Menjelaskan majelis syura dalam Islam
		1.3	Mengidentifikasi ketentuan Islam tentang siyasah syar'iyah
2.	Memahami dan menerapkan sumber hukum Islam	2.1	Menjelaskan sumber hukum yang disepakati dan yang tidak disepakati ulama
		2.2	Menunjukkan penerapan sumber hukum yang disepakati dan yang tidak disepakati ulama
		2.3	Menjelaskan pengertian, fungsi, dan kedudukan ijtihad

f. Kelas XII, Semester 2

- | | |
|---|--|
| 1. Memahami dan mengidentifikasi hukum-hukum syar'i | 1.1 Menjelaskan hukum taklifi dan penerapannya dalam Islam
1.2 Menjelaskan hukum wadh'i dan penerapannya dalam Islam
1.2 Menjelaskan mahkum bihi (fih) |
| 2. Memahami dan menerapkan kaidah-kaidah usul fikih | 2.4 Menjelaskan macam-macam kaidah usul fikih
2.5 Menerapkan macam-macam kaidah usul fikih |

BAB III

METODE PENELITIAN

A. Model Pengembangan

Penelitian yang akan dilakukan adalah menyusun instrumen penilaian hasil belajar untuk mata pelajaran fiqh di Madrasah Aliyah, sehingga akan diperoleh produk setelah melalui tahapan yang telah ditentukan. Oleh karenanya penelitian ini termasuk jenis penelitian pengembangan. Adapun pendekatan yang dipakai adalah pendekatan kuantitatif. Dalam hal ini menurut Borg and Gall (1983) adalah

Educational research and development (R & D) is a process used to developmen and validate educational product... The steps of this process... consist of studying research findings... The product to be developed, developing the product... field testing it in the setting where it will be used and revising it to correct the deficiencies foud in the field testing stage.(1983: 772)

Senada dengan pendapat di atas adalah pendapatnya Sugiyono. Menurutnya penelitian R and D merupakan suatu metode penelitian

yang digunakan untuk menghasilkan produk tertentu dan menguji keefektifan produk yang diinginkan (2006, 297). Pendapat lain mengenai tahapan penelitian pengembangan adalah sebagaimana pendapatnya Borg and Gall (1983):

The mayor steps in the R & D used to develop minicourses are as follows: 1. research and information collecting-includes review of literarture, class-room observations, and preparation of report of state of the art. 2. Planning- invludes defining skills, stating objectives determining, course sequence, and small scale feasibility testing.3. Develop preliminary form of product-includes preparation of instructional materials, handbooks, and evaluation devices. 4. Preliminary fields testing-Conducted in form 1 to 3 schools, using 6 to 12 subjects. Inteerview, observational and questionnaire data collected and analyzed. 5. Main product revision-Revisison of product as suggested by preliminary field-test results. 6. Main field testing_conducted in 5 to 15 schools with 30 to 100 subjects. Quantitative data on subjects' prevourse and postecourse performance

are collected. Results are evaluated with respect to course objectives and are compared with control group data, when appropriate. 7. Operational product revision-Revision of product as suggested by main field-test results. 8. Operational field testing-Conducted in 10 to 30 schools involving 40 to 200 subjects. Interview, observational and questionnaire data collected and analyzed. 9. Final product revision-Revision of product as suggested by operational field-test results. 10. Dissemination and implementation-Report on product at professional meeting and journals. Work with publisher who assumes commercial distribution. Monitor distribution to provide quality control. (775-776).

Selain Borg and Gall tersebut di atas ada pendapat lain yang lebih singkat, yaitu pendapat yang diajukan oleh Semmel and Semmel yang lebih dikenal dengan istilah Four-D model, yang terdiri dari *Define, Design, Development and Disseminate*. Dalam konteks penelitian inilah Four-D ini digunakan. Berikut digambarkan Pengembangan Instrumen penilaian mata pelajaran Fiqh Di Madrasah Aliyah (MA).

BAB IV

PENGEMBANGAN INSTRUMEN PENILAIAN HASIL BELAJAR MATA PELAJARAN FIQH MAN

A. Pedoman penyusunan instrumen hasil belajar Fiqh di MAN 1 dan MAN 2 Purwokerto

Evaluasi merupakan salah komponen penting dan tahap yang harus ditempuh oleh guru untuk mengetahui keefektifan pembelajaran. Hasil yang diperoleh dapat dijadikan balikan (*feed-back*) bagi guru dalam memperbaiki dan menyempurnakan program dan kegiatan pembelajaran. Oleh karenanya evaluasi merupakan suatu proses yang sistematis dan berkelanjutan untuk menentukan kualitas (nilai dan arti), berdasarkan pertimbangan dan kriteria tertentu untuk membuat suatu keputusan. Berdasarkan pengertian ini, ada beberapa hal yang perlu kita pahami lebih lanjut, yaitu :

1. Evaluasi adalah suatu proses bukan suatu hasil (produk).

Hasil yang diperoleh dari kegiatan evaluasi adalah kualitas daripada sesuatu, baik yang menyangkut tentang nilai maupun arti. Sedangkan kegiatan untuk sampai kepada pemberian nilai dan arti itu adalah evaluasi. Jika Anda melakukan kajian tentang evaluasi, maka yang Anda lakukan adalah mempelajari bagaimana proses pemberian pertimbangan mengenai kualitas daripada sesuatu. Gambaran kualitas yang dimaksud merupakan konsekuensi logis dari proses evaluasi yang dilakukan. Proses tersebut tentu dilakukan secara sistematis dan berkelanjutan, dalam arti terencana, sesuai dengan prosedur dan aturan, dan terus menerus.

2. Tujuan evaluasi adalah untuk menentukan kualitas sesuatu, terutama yang berkenaan dengan nilai dan arti.

Pemberian nilai dilakukan apabila seorang evaluator memberikan pertimbangannya mengenai evaluasi tanpa menghubungkannya dengan sesuatu yang bersifat dari luar. Jadi pertimbangan

yang diberikan sepenuhnya berdasarkan apa evaluasi itu sendiri. Sedangkan arti, berhubungan dengan posisi dan peranan evaluasi dalam suatu konteks tertentu. Tentu saja kegiatan evaluasi yang komprehensif adalah yang meliputi baik proses pemberian keputusan tentang nilai dan proses keputusan tentang arti, tetapi hal ini tidak berarti bahwa suatu kegiatan evaluasi harus selalu meliputi keduanya. Pemberian nilai dan arti ini dalam bahasa yang dipergunakan Scriven (1967) adalah formatif dan sumatif. Jika formatif dan sumatif merupakan fungsi evaluasi, maka nilai dan arti adalah hasil kegiatan yang dilakukan oleh evaluasi.

3. Dalam proses evaluasi harus ada pemberian pertimbangan (*judgement*).

Pemberian pertimbangan ini pada dasarnya merupakan konsep dasar evaluasi. melalui pertimbangan inilah ditentukan nilai dan arti dari suatu yang sedang dievaluasi. Tanpa pemberian pertimbangan, suatu kegiatan bukanlah termasuk kategori kegiatan evaluasi.

4. Pemberian pertimbangan tentang nilai dan arti haruslah berdasarkan kriteria tertentu.

Tanpa kriteria yang jelas, pertimbangan nilai dan arti yang diberikan bukanlah suatu proses yang dapat diklasifikasikan sebagai evaluasi. Kriteria yang digunakan dapat saja berasal dari apa yang dievaluasi itu sendiri (internal), tetapi bisa juga berasal dari luar apa yang dievaluasi (eksternal), baik yang bersifat kuantitatif maupun kualitatif. Jika yang dievaluasi itu adalah proses pembelajaran, maka kriteria yang dimaksud bisa saja dikembangkan dari karakteristik proses pembelajaran itu sendiri, tetapi dapat pula dikembangkan kriteria umum tentang proses pembelajaran. Kriteria ini penting dibuat oleh evaluator dengan pertimbangan (a) hasil evaluasi dapat dipertanggungjawabkan secara ilmiah (b) evaluator lebih percaya diri (c) menghindari adanya unsur subjektivitas (d) memungkinkan hasil evaluasi akan sama sekalipun dilakukan pada waktu dan orang yang berbeda, dan (e) memberikan

kemudahan bagi evaluator dalam melakukan penafsiran hasil evaluasi.

Kriteria yang dimaksudkan dalam kegiatan evaluasi bahwa evaluasi membutuhkan satu parameter yang jelas yang mendasarkan pada silabus dan kurikulum. Parameter ini berupa instrumen yang dikembangkan dari indikator yang merupakan derivasi dari standar kompetensi lulusan (SKL). Berdasarkan indikator inilah satuan pendidikan dalam hal guru akan menyusun kisi-kisi yang akan dikembangkan menjadi instrumen penilaian. Sebagaimana keputusan Direktur Jenderal Pendidikan Islam nomor 21 tahun 2013 tentang ketentuan pelaksanaan Ujian Akhir Madrasah Berstandar Nasional Pendidikan Agama Islam dan Bahasa Arab tingkat Madrasah Ibtidaiyah, Madrasah Tsanawiyah, dan Madrasah Aliyah. Dalam keputusan ini jelas bahwa penyusunan kisi-kisi merupakan prasyarat untuk bisa dikembangkannya instrumen penilaian yang objektif. Pengembangan instrumen penilaian yang tanpa mendasarkan pada kisi-kisi akan menjadikan instrumen sebagai alat ukur tidak bisa berfungsi sebagai alat ukur yang baik (valid), yaitu instrumen yang bisa mengukur

kemampuan peserta didik dengan tepat dan akurat.

Instrumen penilaian hasil belajar Fiqh di MAN 1 dan MAN 2 Purwokerto agar mampu mengukur dengan tepat dan akurat maka perlu adanya pedoman penyusunan instrumen penilaian mata pelajaran Fiqh. Pedoman akan didasarkan pada standar kompetensi lulusan (SKL) dan standar kompetensi (SK) maupun kompetensi dasar (KD) yang ada dalam mata pelajaran Fiqh di MA. SKL mata pelajaran fiqh di MA sebagaimana yang tercantum dalam dokumen Peraturan Menteri Agama Republik Indonesia (Permenag RI) nomor 2 Tahun 2008 tentang Standar Kompetensi Lulusan dan Standar Isi. Pendidikan Agama Islam dan Bahasa Arab di Madrasah adalah memahami dan menerapkan sumber hukum Islam dan hukum taklifi, prinsip-prinsip ibadah dan syari'at dalam Islam, fikih ibadah, *mu'amalah*, *munakahat*, *mawaris*, *jinayah*, *siyasa*, serta *dasar-dasar istinbath dan kaidah usul fikih*. Aspek fikih menekankan pada kemampuan cara melaksanakan ibadah dan muamalah yang benar dan baik.

Mata pelajaran Fikih di Madrasah Aliyah adalah salah satu mata pelajaran Pendidikan Agama Islam yang merupakan peningkatan dari fikih yang telah dipelajari oleh peserta didik di Madrasah Tsanawiyah/SMP. Peningkatan tersebut dilakukan dengan cara mempelajari, memperdalam serta memperkaya kajian fikih baik yang menyangkut aspek ibadah maupun muamalah, yang dilandasi oleh prinsip-prinsip dan kaidah-kaidah usul fikih serta menggali tujuan dan hikmahnya, sebagai persiapan untuk melanjutkan ke pendidikan yang lebih tinggi dan untuk hidup bermasyarakat. Secara substansial, mata pelajaran Fikih 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.

Mata pelajaran Fikih di Madrasah Aliyah bertujuan untuk:

1. Mengetahui dan memahami prinsip-prinsip, kaidah-kaidah dan tatacara

pelaksanaan hukum Islam baik yang menyangkut aspek ibadah maupun muamalah untuk dijadikan pedoman hidup dalam kehidupan pribadi dan sosial.

2. 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.

Ruang lingkup mata pelajaran Fiqih di Madrasah Aliyah meliputi:

- a. kajian tentang prinsip-prinsip ibadah dan syari'at dalam Islam,
- b. hukum Islam dan perundang-undangan tentang zakat dan haji, hikmah dan cara pengelolaannya,
- c. hikmah kurban dan akikah,
- d. ketentuan hukum Islam tentang pengurusan jenazah,

- e. hukum Islam tentang kepemilikan,
- f. konsep perekonomian dalam Islam dan hikmahnya,
- g. hukum Islam tentang pelepasan dan perubahan harta beserta hikmahnya,
- h. hukum Islam tentang *wakaalah* dan *sulhu* beserta hikmahnya,
- i. hukum Islam tentang *daman* dan *kafaalah* beserta hikmahnya,
- j. riba, bank dan asuransi,
- k. ketentuan Islam tentang *jinaayah*, *Huduud* dan hikmahnya,
- l. ketentuan Islam tentang peradilan dan hikmahnya,
- m. hukum Islam tentang keluarga dan waris,
- n. ketentuan Islam tentang *siyaasah syar'iyah*,
- o. sumber hukum Islam dan hukum *taklifi*,
- p. dasar-dasar *istinbaath* dalam fikih Islam,
- q. kaidah-kaidah usul fikih dan penerapannya.

B. Penyusunan Instrumen Penilaian Hasil Belajar Fiqh di MAN 1 Purwokerto dan MAN 2 Purwokerto

Proses penyusunan instrumen penilaian hasil belajar fiqh di madrasah Aliyah Negeri Purwokerto ada 3 sumber; sumber pertama instrumen penilaian dari pemerintah (Kemenag RI), sumber kedua instrumen penilaian yang disusun oleh musyawarah guru mata pelajaran (MGMP) fiqh dan sumber ketiga yang disusun dari masing-masing guru fiqh di MA tersebut. Ketiga sumber tersebut yang pertama instrumen penilaian untuk ujian ahir madrasah bersetandar nasional (UAM-BN) diperuntukkan bagi peserta didik MA jurusan non-keagamaan dan instrumen penilaian untuk UN diperuntukkan bagi peserta didik jurusan keagamaan. Kemudian sumber kedua jenis instrumen penilaian untuk ujian akhir sekolah (UAS) dan ujian kenaikan kelas (UKK) serta sumber yang terakhir yaitu jenis instrumen penilaian untuk ulangan harian dan ujian tengah semester (UTS).

Proses penyusunan instrumen penilaian mata pelajaran fiqh dari sumber yang kedua dan ketiga tersebut di atas menjadi wilayah

di masing-masing MA khususnya untuk satu Karsidenan. MAN 1 dan MAN 2 Purwokerto dalam proses penyusunan instrumen hasil belajar mata pelajaran fiqh berada di wilayah karsidenan Banyumas yang meliputi Kabupaten Banyumas, Cilacap, Purbalingga dan Banjarnegara. Penyusunan sumber kedua yang instrumennya diperuntukkan bagi peserta didik yang akan mengikuti ujian akhir semester (UAS) dan ujian kenaikan kelas (UKK) di MA, prosesnya dimulai dengan adanya musyawarah di MGMP yang akan menunjuk personil guru yang akan ditugasi menyusun kisi-kisi soal instrumen penilaian hasil belajar mata pelajaran fiqh. Personil guru yang telah ditugasi sebagai penyusun kisi-kisi akan mengerjakannya untuk kemudian hasilnya akan dibahas dan direview dalam forum MGMP. Setelah kisi-kisi disepakati dalam forum MGMP maka untuk proses selanjutnya MGMP akan menunjuk beberapa guru untuk menyusun instrumen penilaian hasil belajar mata pelajaran fiqh yang mendasarkan pada kisi-kisi yang telah disepakati.

Instrumen penilaian hasil belajar mata pelajaran fiqh yang telah tersusun oleh beberapa

guru yang ditugasi MGMP mata pelajaran fiqh maka instrumen tersebut akan disunting oleh personil guru yang menyusun kisi-kisi. Jika hasil instrumen penilaian ada yang belum sesuai dengan kisi-kisi menurut penyunting akan mendapat catatan untuk diperbaiki. Ketika instrumen sudah sesuai dengan semua kisi-kisi menurut penyunting instrumen penilaian maka selanjutnya dibahas, direview dan ditetapkan oleh forum MGMP. Setelah fix baru kemudian atas arahan Kelompok Kerja Kepala Madrasah (K3M) selanjutnya dicetak dan didistribusikan ke semua MA di satu Karsidenan Banyumas.

Proses penyusunan instrumen penilaian hasil belajar fiqh MA yang menjadi alat ukur pada penilaian UAS dan UKK di MAN 1 dan MAN 2 Purwokerto bahkan untuk karsidenan Banyumas yang terdiri dari 4 kabupaten sebagaimana tersebut di atas tidak melalui ujicoba instrumen penilaian. Alasan mereka tidak melakukan ujicoba instrumen ada beberapa di antaranya karena faktor minimnya pengetahuan tentang pengetahuan mengenai ujicoba instrumen penilaian, meskipun sebenarnya dari Kemenag Provinsi maupun yang diadakan oleh MGMP sendiri

telah memberikan upaya melalui pelatihan-pelatihan dan workshop yang membahas tentang penyusunan butir soal dengan proses yang sesuai prinsip evaluasi. Yang di dalamnya membahas kriteria tingkat kesukaran butir soal, daya beda, serta keberfungsian distraktor. Mengingat rata-rata memang guru-guru fiqh kebanyakan sudah berusia relatif sepuh dan belum banyak yang telah menempuh studi lanjut maka pasca pelatihan dan mengikuti workshop aplikasi hasil pelatihan dan workshop belum bisa dilakukan, inilah yang disampaikan kepala MAN di Purwokerto. Ujicoba instrumen penilaian tidak dilakukan karena dianggap telah mencukupi dengan membuat pedoman kisi-kisi dan uji coba instrumen relatif dianggap kurang mendesak urgensi manfaatnya. Bahkan mereka sudah memiliki pemahaman bahwa instrumen penilaian mata pelajaran fiqh di MAN yang disusun dengan berpedoman pada kisi-kisi tes sebagaimana soal tes yang dibuat oleh MGMP sangat akurat dan sesuai dengan kurikulum fiqh yang ada, karena tidak ada SK-KD-Indikator yang terlewat tanpa ada soalnya dalam tes, jadi lebih menjamin mutunya.

Instrumen penilaian yang digunakan untuk menjadi alat ukur pada ulangan harian dan UTS yang membuat guru masing-masing MA, mereka juga sudah mendapatkan pelatihan dan workshop mengenai penyusunan butir soal namun pada praktiknya selama ini guru dalam menyusun tes lebih mengacu pada kisi kisi materi yang sudah diberikan, tidak atau jarang membuat kisi-kisi soal yang dijadikan pedoman untuk membuat item-item soal tes fiqh. Alasan pertama karena praktis tidak njlimet dan yang kedua soal yang dibuat guru sendiri hanya untuk menilai ulangan harian dan ujian tengah semester yang sifatnya internal madrasah, ungkap guru mata pelajaran fiqh di MAN 2 Purwokerto. Evaluasi sebagai kegiatan untuk mengetahui progres dan kompetensi peserta didik tentang materi yang sudah diberikan mestinya harus benar-benar bisa mengetahui akan kompetensi yang akan diukur. Proses pengukuran agar mendapatkan informasi yang semestinya dituntut memiliki instrumen sebagai alat ukur yang memiliki validitas.

Validitas suatu instrumen atau tes mempermasalahkan apakah instrumen atau tes tersebut benar-benar mengukur apa yang

hendak diukur. Suatu instrumen pengukuran dikatakan memiliki validitas yang tinggi apabila alat tersebut menjalankan fungsi ukurnya, atau memberikan hasil ukur yang sesuai dengan maksud dilakukannya pengukuran tersebut. Maksudnya adalah bahwa sebarapa jauh suatu tes mampu mengungkapkan dengan tepat ciri atau keadaan yang sesungguhnya dari obyek ukur, akan tergantung dari tingkat validitas tes yang bersangkutan. Sehingga hasil ukur dari pengukuran tersebut merupakan besaran yang mencerminkan secara tepat fakta atau keadaan sesungguhnya dari apa yang diukur. Peserta didik di MAN 1 dan MAN 2 Purwokerto bisa diukur kemampuan mereka akan pengetahuan dan pemahaman tentang fiqh yang telah diberikan haruslah dievaluasi menggunakan instrumen penilaian sebagai alat ukur yang valid. Meskipun kegiatan evaluasi tersebut untuk tingkat ulangan harian maupun UTS. Hal ini dilakukan agar evaluasi mendapatkan informasi yang sebenarnya dari kondisi peserta didik, yang pada akhirnya akan berfungsi sebagai *feedback* dalam pembelajaran berikutnya. Dengan harapan kualitas pembelajaran akan semakin meningkat dan peserta didik memiliki kemampuan dan kompetensi sebagaimana yang

digariskan dalam dalam kompetensi dasar. Karena tugas guru adalah menghantarkan peserta didik mencapai kompetensi dasar.

C. Pengembangan instrumen penilaian hasil belajar Fiqh di MAN 1 dan MAN 2 Purwokerto

Telaah instrumen kegiatan pada telaah instrumen adalah menelaah apakah; 1) butir pertanyaan atau pernyataan sesuai dengan indikator, 2) bahasa yang digunakan komunikatif dan menggunakan tata bahasa yang benar, 3) butir peranyaan/ pernyataan tidak bias, 4) format instrumen menarik untuk dibaca, 5) pedoman menjawab atau mengisi instrumen jelas, dan 6) jumlah butir dan/ atau panjang kalimat pertanyaan/ pernyataan sudah tepat sehingga tidak menjemukan untuk dibaca atau dijawab.

Telaah dilakukan oleh pakar dalam bidang pengukuran atau psikometri dan akan lebih baik bila ada pakar penilaian. Telaah bisa juga dilakukan oleh teman sejawat bila yang diinginkan adalah masukan tentang bahasa dan format instrumen. Pengembangan instrumen penilain mata pelajaran fiqh di MAN 1 maupun MAN 2 Purwokerto dalam menelaah instrumen terbatas pada teman sejawat, hal

ini dilakukan dalam forum MGMP. Sehingga instrumen yang telah disusun masih perlu telaah dari pakar psikometri sebagai ahli dalam pengukuran maupun pakar penilain maupun pakar bahasa agar mendapatkan masukan sebagai upaya penyempurnaan instrumen penilaian di MAN. Bahasa yang digunakan adalah yang sesuai dengan tingkat pendidikan responden. Hasil telaah selanjutnya digunakan untuk memperbaiki instrumen. Panjang instrumen berhubungan dengan masalah kebosanan, yaitu tingkat kejemuhan dalam mengisi instrumen. Lama pengisian instrumen sebaiknya tidak lebih dari 30 menit. Langkah pertama dalam menulis suatu pertanyaan/ pernyataan adalah informasi apa yang ingin diperoleh, struktur pertanyaan, dan pemilihan kata-kata. Pertanyaan yang diajukan jangan sampai bias, yaitu mengarahkan jawaban responden pada arah tertentu, positif atau negatif.

Perbaikan instrumen dilakukan terhadap butir-butir pertanyaan/pernyataan yang tidak baik, berdasarkan analisis hasil ujicoba. Bisa saja hasil telaah instrumen baik, namun hasil ujicoba empirik tidak baik. Untuk itu

butir pertanyaan/pernyataan instrumen harus diperbaiki. Perbaikan termasuk mengakomodasi saran-saran dari responden ujicoba. Instrumen sebaiknya dilengkapi dengan pertanyaan terbuka. Inilah penting *tryout* atau ujicoba instrumen penilaian agar diperoleh informasi mengenai butir soal yang telah disusun. Baik informasi tentang kualitas instrumen penilaian dari sisi tingkat kesukaran, daya beda maupun keberfungsian distraktor. Selain itu juga informasi akan dapat diperoleh dari pertanyaan terbuka yang akan diberikan oleh para *testee* (peserta *tryout*) yang mengikuti atau yang mengerjakan butir pada *tryout*. Masukan ini akan sangat berguna dalam rangka perbaikan instrumen penilaian, namun MAN 1 maupun MAN 2 Purwokerto tidak melakukan. Instrumen hanya dilihat dari sisi validitas konstruk dengan menyesuaikan apa yang ada dalam KD dan indikator. Kesesuaian instrumen penilaian dengan KD dan indikator sudah dianggap cukup.

Hasil telaah instrumen digunakan untuk memperbaiki instrumen. Perbaikan dilakukan terhadap konstruksi instrumen, yaitu kalimat yang digunakan, waktu yang diperlukan

untuk mengisi instrumen, cara pengisian atau cara menjawab instrumen, dan pengetikan. Merakit instrumen setelah instrumen diperbaiki selanjutnya instrumen dirakit, yaitu menentukan format tata letak instrumen dan urutan pertanyaan/ Pernyataan. Format instrumen harus dibuat menarik dan tidak terlalu panjang, sehingga responden tertarik untuk membaca dan mengisinya. Setiap sepuluh pertanyaan sebaiknya dipisahkan dengan cara memberi spasi yang lebih, atau diberi batasan garis empat persegi panjang. Urutkan pertanyaan/ Pernyataan sesuai dengan tingkat kemudahan dalam menjawab atau mengisinya.

Ujicoba instrumen, setelah dirakit instrumen diujicobakan kepada responden, sesuai dengan tujuan penilaian apakah kepada peserta didik, kepada guru atau orang tua peserta didik. Untuk itu dipilih sampel yang karakteristiknya mewakili populasi yang ingin dinilai. Bila yang ingin dinilai adalah peserta didik MA, maka sampelnya juga peserta didik MA. Sampel yang diperlukan minimal 30 peserta didik, bisa berasal dari satu madrasah atau lebih.

Pada saat ujicoba yang perlu dicatat

adalah saran-saran dari responden atas kejelasan pedoman pengisian instrumen, kejelasan kalimat yang digunakan, dan waktu yang diperlukan untuk mengisi instrumen. Waktu yang digunakan disarankan bukan waktu saat responden sudah lelah. Selain itu sebaiknya responden juga diberi minuman agar tidak lelah. Perlu diingat bahwa pengisian instrumen penilaian afektif bukan merupakan tes, sehingga walau ada batasan waktu namun tidak terlalu ketat.

Agar responden mengisi instrumen dengan akurat sesuai harapan, maka sebaiknya instrumen dirancang sedemikian rupa sehingga waktu yang diperlukan mengisi instrumen tidak terlalu lama. Berdasarkan pengalaman, waktu yang diperlukan agar tidak jenuh adalah 30 menit atau kurang.

Analisis hasil ujicoba, analisis hasil ujicoba meliputi variasi jawaban tiap butir pertanyaan/ pernyataan. Jika menggunakan skala instrumen 1 sampai 7, dan jawaban responden bervariasi dari 1 sampai 7, maka butir pertanyaan/ pernyataan pada instrumen ini dapat dikatakan baik. Namun apabila jawabannya hanya pada satu pilihan jawaban saja, misalnya pada

pilihan nomor 3, maka butir instrumen ini tergolong tidak baik. Indikator yang digunakan adalah besarnya daya beda. Bila daya beda butir instrumen lebih dari 0,30 butir instrumen tergolong baik. Sedangkan untuk tingkat kesukaran butir soal yang baik adalah terletak antara 0,24 sampai 0,76, jika nilainya di bawah 0,24 maka instrumen tergolong sangat sukar dan sebaliknya jika di atas 0,76 maka tergolong instrumen yang sangat mudah sehingga kedua tidakbisa menjadi alat ukur. Artinya instrumen yang ada di bawah nilai 0,24 dan di atas 0,76 tidak bisa digunakan sebagai instrumen yang akan digunakan menjadi soal pada kegiatan evaluasi karena tidak bisa menjadi alat ukur yang tepat. Indikator lain yang diperhatikan adalah indeks keandalan yang dikenal dengan indeks reliabilitas. Batas indeks reliabilitas minimal 0,70. Bila indeks ini lebih kecil dari 0,70, kesalahan pengukuran akan melebihi batas. Oleh karena itu diusahakan agar indeks keandalan instrumen minimal 0,70.

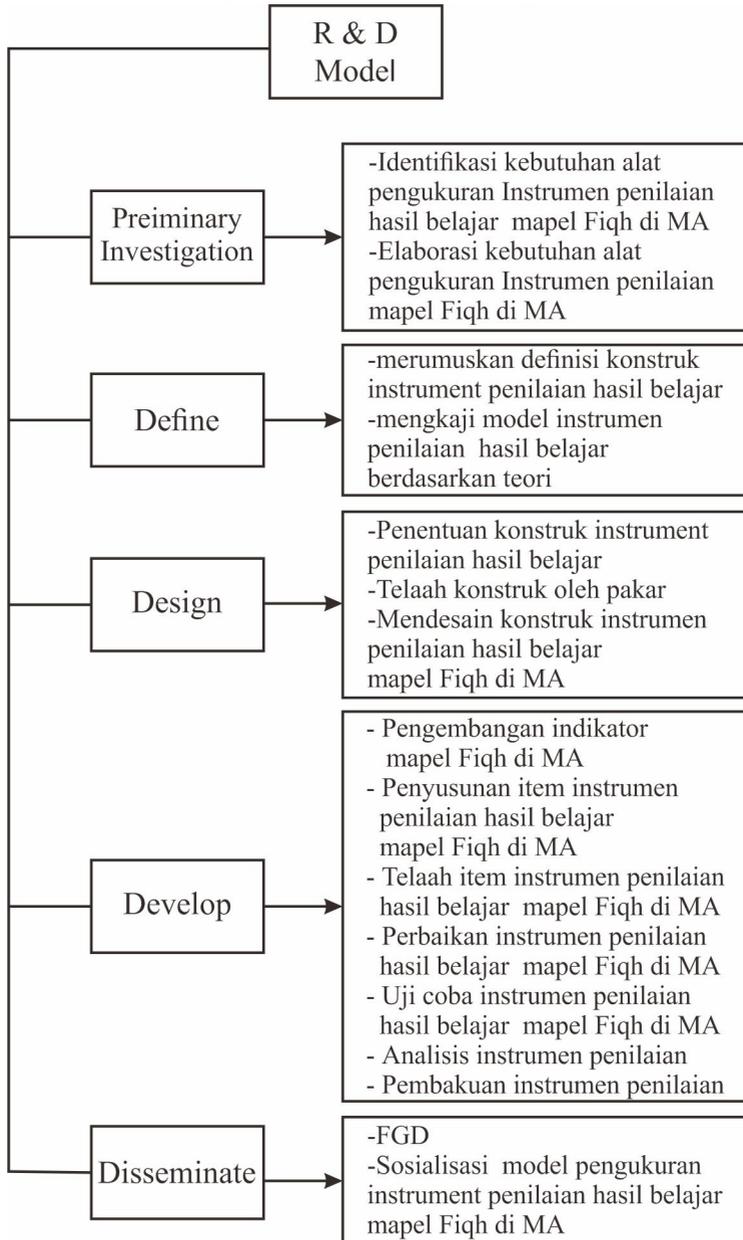
Pengembangan instrumen penilaian sebagai alat ukur pada mata pelajaran fiqh di MAN 1 maupun MAN 2 Purwokerto belum dilakukan sebagaimana dalam teori yang

harus memperhatikan proses dan langkah-langkah yang harus dilakukan. Instrumen yang dijadikan sebagai alat ukur untuk mengukur kompetensi peserta didik memerlukan instrumen penilaian yang *valid* dan reliabel, karena dalam rangka untuk mendapatkan informasi yang tepat. Instrumen yang sudah disusun berdasarkan dokumen seperti KD dan indikator belumlah cukup langsung digunakan sebagai alat ukur peserta didik. Hal yang sangat dikawatirkan yaitu jika instrumen penilaiannya tidak melalui tahapan ujicoba tidak akan bisa mendapatkan informasi yang sangat penting seperti tingkat kesukaran butir soal, daya beda maupun keberfungsian distraktor. Implikasi yang harus diperhatikan adalah ketika terjadi kesalahan pengukuran, artinya proses evaluasi tidak akan mendapatkan informasi progres peserta didik yang semestinya. Jika ini yang terjadi bisa dikatakan bahwa kegiatan evaluasi mengalami kegagalan.

Sementara proses pengembangan instrumen penilaian butir soal mata pelajaran fiqh di MAN 1 dan MAN 2 Purwokerto hanya melakukan upaya telaah dalam proses FGD dengan teman sejawat dan belum melibatkan

pakar pada bidang pengukuran, pakar penilai maupun pakar bahasa. Pengembangan instrumen yang dilakukan masih sangat sederhana. Sementara proses pengembangan instrumen membutuhkan banyak hal yang mestinya dilakukan. Dengan demikian pengembangan instrumen penilaian mata pelajaran fiqh di MAN 1 dan MAN 2 Purwokerto masih sangat membutuhkan pendampingan atau motivasi agar memandang penting melakukan upaya dan tahapan untuk mendapatkan instrumen penilaian yang memiliki akurasi yang tinggi. Sehingga tidak akan dijumpai keliru dalam memberikan penilaian kepada peserta didik.

Tabel 3. Model pengembangan instrumen penilaian mapel Fiqh di



Dari pendapat tentang langkah-langkah metode penelitian pengembangan instrumen maka secara garis besar langkah-langkah dalam penyusunan instrumen penilaian mata pelajaran fiqh untuk Madrasah Aliyah (MA) adalah sebagai berikut;

- a. melakukan kajian teoritik untuk merumuskan aspek-aspek atau indikator penilaian mata pelajaran fiqh MA,
- b. menyusun kisi-kisi instrument mapel Fiqh di MA,
- c. menyusun soal-soal instrument mapel Fiqh di MA,
- d. melakukan *expert judgment* mapel Fiqh di MA,
- e. melakukan uji coba,
- f. melakukan analisis,
- g. revisi, dan
- h. merumuskan instrumen akhir hasil penelitian

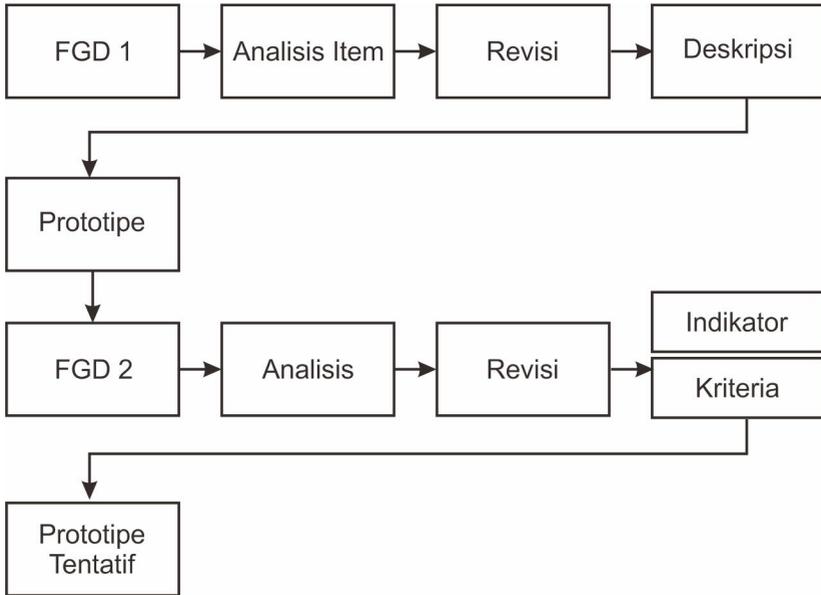
D. Prosedur Pengembangan

Dalam tahap ini yang dilakukan adalah:

- a. Penyusunan kisi-kisi penelaahan, perbaikan, perakitan kisi-kisi
- b. Penelaahan kisi-kisi dilakukan oleh pakar psikometri, pakar pendidikan, dosen, guru mapel Fiqh dalam Forum Group Discuss (FGD)
- c. Hasil FGD untuk memperbaiki kisi-kisi dan kisi-kisi dirakit menjadi acuan dalam penyusunan instrumen. Hasilnya berupa Perangkat pengukuran Instrumen penilaian. Setelah itu instrumen ini akan divalidasi oleh para ahli dalam 2 x FGD.

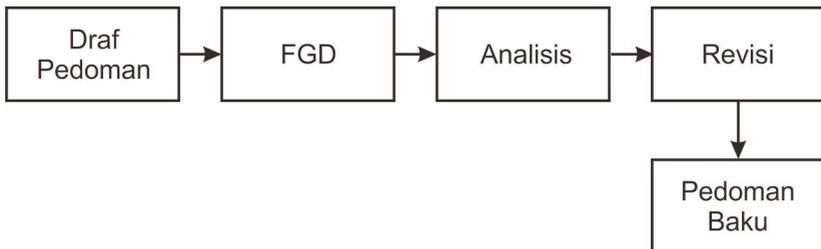
Untuk lebih jelasnya berikut ditampilkan Skema Tahapan Pengembangan model pengukuran instrumen penilaian mapel Fiqh di MA :

Tabel 4. Pengembangan



Kemudian dilakukan tahapan berikutnya sebagaimana bagan berikut untuk mendapatkan pedoman baku, yaitu;

Tabel 5. Pengembangan lanjut



E. Uji Coba Produk

Setelah uji ahli tentang instrumen penilaian hasil belajar mapel fiqh di MA yang dilengkapi dengan kurikulum dan silabi fiqh MA, peneliti merevisi masukan-masukan dari para ahli untuk melengkapi produk tersebut. Desain uji coba dapat dilihat pada gambar sebagai berikut ini.

1. Desain Uji Coba, sebagaimana dalam rencana pengembangan di atas maka desain uji coba instrumen akan dilakukan untuk mendapatkan instrumen penilaian hasil belajar yang sesuai dengan tuntutan validitas konstruk yang telah dikonsepskan.
2. Subyek Uji Coba, subyek ujicoba dalam penelitian ini adalah siswa MAN 1 Purwokerto dan MAN 2 Purwokerto. Penelitian ini menggunakan *Confirmatory Factor Analysis* (CFA).
3. Jenis Data, jenis data dalam penelitian ini adalah kuantitatif yaitu instrumen penilaian hasil belajar mata pelajaran fiqh di MA, yang tersusun setelah melalui tahapan-tahapan pengembangan sebagaimana langkah di atas. Data

yang dibutuhkan untuk mengetahui kemampuan siswa dalam mengerjakan tes adalah menggunakan tes prestasi.

F. Teknik Pengumpulan Data

Teknik pengumpulan data dalam penelitian ini adalah interview dan dokumentasi. Interview dalam penelitian ini akan dilakukan kepada guru mata pelajaran Fiqh dan kepala madrasah MAN Purwokerto dan juga kasi Mapenda Kemenag Provinsi Jawa Tengah maupun kabupaten juga siswa MAN. Dokumentasi untuk mendapatkan data mengenai butir soal yang dipakai MAN dalam kegiatan evaluasi UTS dan UKK atau UAS. Data yang akan diakses mengenai prosedur penyusunan instrumen penilaian mata pelajaran fiqh, instrumen akan dikonfirmasi kecocokannya dengan silabi dan kurikulum serta validasi instrumen yang disusun.

G. Teknik Analisis Data

Teknik analisis data yang digunakan disesuaikan dengan jenis data dikumpulkan. Beberapa hal yang perlu diperhatikan dalam analisis data:

1. Analisis data mencakup prosedur organisasi data, reduksi, dan penyajian data baik dengan tabel, bagan, atau grafik.
2. Data diklasifikasikan berdasarkan jenis dan komponen produk yang dikembangkan
3. Data dianalisis secara deskriptif maupun dalam bentuk perhitungan kuantitatif jika menggunakan data kuantitatif. Dalam penelitian ini lebih menggunakan data yang kualitatif sehingga mendeskripsikan data akan dilakukan.
4. Penyajian hasil analisis dibatasi pada hal-hal yang bersifat faktual, dengan tanpa interpretasi pengembang, sehingga sebagai dasar dalam melakukan revisi produk.

BAB V

PENUTUP

Memperhatikan data yang telah diperoleh dalam penelitian ini yaitu pengembangan instrumen penilaian hasil belajar mata pelajaran fiqh di MAN 1 dan MAN 2 Purwokerto dengan interview dan dokumentasi kemudian dibahas dengan mendiskusikan teori yang sudah peneliti bangun pada kajian teori maka kesimpulan penelitian ini adalah:

1. Pedoman penyusunan instrumen hasil belajar Fiqh di MAN 1 dan MAN 2 Purwokerto mendasarkan pada standar kompetensi lulusan (SKL) dan standar kompetensi (SK) maupun kompetensi dasar (KD) yang ada dalam mata pelajaran Fiqh di MA. SKL mata pelajaran fiqh di MA sebagaimana yang tercantum dalam dokumen Peraturan Menteri Agama Republik Indonesia (Permenag RI) nomor 2 Tahun 2008 tentang Standar Kompetensi Lulusan dan Standar Isi.
2. Penyusunan instrumen penilaian hasil belajar Fiqh di MAN 1 Purwokerto dan MAN 2 Purwokerto untuk instrumen UAS dan UKK

prosesnya dimulai dengan adanya musyawarah di MGMP yang akan menunjuk personil guru yang akan ditugasi menyusun kisi-kisi soal instrumen penilaian hasil belajar mata pelajaran fiqh. Personil guru yang telah ditugasi sebagai penyusun kisi-kisi akan mengerjakannya untuk kemudian hasilnya akan dibahas dan direview dalam forum MGMP. Setelah kisi-kisi disepakati dalam forum MGMP maka untuk proses selanjutnya MGMP akan menunjuk beberapa guru untuk menyusun instrumen penilaian hasil belajar mata pelajaran fiqh yang berdasarkan pada kisi-kisi yang telah disepakati.

Instrumen penilaian hasil belajar mata pelajaran fiqh yang telah tersusun oleh beberapa guru yang ditugasi MGMP mata pelajaran fiqh maka instrumen tersebut akan disunting oleh personil guru yang menyusun kisi-kisi. Jika hasil instrumen penilaian ada yang belum sesuai dengan kisi-kisi menurut penyunting akan mendapat catatan untuk diperbaiki. Ketika instrumen sudah sesuai dengan semua kisi-kisi menurut penyunting instrumen penilaian maka selanjutnya dibahas, direview dan ditetapkan oleh forum MGMP. Setelah fix

baru kemudian atas arahan Kelompok Kerja Kepala Madrasah (K3M) selanjutnya dicetak dan didistribusikan ke semua MA di satu Karsidenan Banyumas. Termasuk di dalamnya MAN 1 dan MAN 2 Purwokerto.

Instrumen penilaian yang digunakan untuk menjadi alat ukur pada ulangan harian dan UTS yang membuat guru masing-masing MAN akan mengacu pada kisi kisi materi yang sudah diberikan, tidak atau jarang membuat kisi-kisi soal yang dijadikan pedoman untuk membuat item-item soal tes fiqh. Alasan pertama karena praktis tidak njlimet dan yang kedua soal yang dibuat guru sendiri hanya untuk menilai ulangan harian dan ujian tengah semester yang sifatnya internal madrasah.

3. Pengembangan instrumen penilaian hasil belajar Fiqh di MAN 1 Purwokerto dan MAN 2 Purwokerto

Pengembangan instrumen penilaian butir soal mata pelajaran fiqh di MAN 1 dan MAN 2 Purwokerto hanya melakukan upaya telaah dalam proses FGD dengan teman sejawat dan belum melibatkan pakar pada bidang pengukuran, pakar penilain maupun pakar bahasa. Pengembangan instrumen yang

dilakukan masih sangat sederhana, belum dilakukan sebagaimana dalam teori yang harus memperhatikan proses dan langkah-langkah yang harus dilakukan. Instrumen yang dijadikan sebagai alat ukur untuk mengukur kompetensi peserta didik memerlukan instrumen penilaian yang *valid* dan reliabel, karena dalam rangka untuk mendapatkan informasi yang tepat. Sementara proses pengembangan instrumen membutuhkan banyak hal yang mestinya dilakukan. Dengan demikian pengembangan instrumen penilaian mata pelajaran fiqh di MAN 1 dan MAN 2 Purwokerto masih sangat membutuhkan pendampingan atau motivasi agar memandang penting melakukan upaya dan tahapan untuk mendapatkan instrumen penilaian yang memiliki akurasi yang tinggi. Sehingga tidak akan dijumpai keliru dalam memberikan penilaian kepada peserta didik.

DAFTAR PUSTAKA

- AERA. (1999). *Standart for educational and psychological testing*. Washington, DC 20036.
- Anastasi, A., & Urbina, S. (1997). *Psychological testing*. Crawfordsville: Prentice Hall
- Anderson, L., & Krathwohl, D. (2001). *A taxonomy for learning, teaching and assessing: A revision of Bloom's taxonomy of educational objectives*. New York: Longman.
- Braxton, J., & Nordvall, R. (1985). Selective Liberal Arts Colleges: Higher quality as well as Higher Prestige. *Journal of higher education*, 56 (September/October), 538-554
- Bereiter, C., and Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Lawrence Erlbaum, Publishers.
- Borg, Walter.R and Gall, Meredith, D. (1983). *Educational research in introduction*. New York: Longman Inc.
- Creswell, John W. (2008). *Educational research; planning, conducting, and evaluating quantitative and qualitative research*. New Jersey: Pearson Educational Edition.

- Depdiknas .(2003). *Undang-Undang RI Nomor 20, Tahun 2003, tentang Sistem Pendidikan Nasional*.
-(2005). *Peraturan Pemerintah RI Nomor 19, Tahun 2005, tentang Standar Nasional Pendidikan*.
-(2006). *Peraturan Menteri Pendidikan Nasional RI Nomor 22, Tahun 2006, tentang Standar isi*.
- De Porte, Bobi. & Mike Hernacki (1992). *Quantum learning: unleashing the genius in you*. Pent. Alwiyah Abdurrahman. Bandung: Kaifa
- Djaali & Muljono. (2008). *Pengukuran dalam bidang pendidikan*. Jakarta: Grasindo.
- Djemari Mardapi dkk. (1999b: 79). *Survey kegiatan guru dalam melakukan penilaian di kelas*. Laporan Penelitian. Yogyakarta: IKIP Yogyakarta.
- Embretson, Susan, & Gorir, Joanna. (2000). Improving construct validity with cognitive psychology principles. *Journal educational measurement*. Vol. 38. number 4.
- Gronlund, N.E. (1976). *Measurement and evaluation in teaching*. New York: Macmillan.

- Hambleton, R.K. (1989). Principle and selected application of IRT. Dalam Robert L. Linn (Ed). *Educational measurement* (ed., pp. 147-200). New York: Macmillan
- Hambleton, R.K., & Swaminathan, H. (1985). *Item respons theory*. Boston, MA: Kluwer Inc.
- Hambleton, R.K., Swaminathan, H., & Rogers, H.J. (1991). *Fundamental of Item respons theory*. Newbury Park, CA: Sage Publication Inc.
- Heath, T. (1984). *Method in architecture*. New York: John Wiley & Sons.
- Indrajati Sidi. (2004). *Peningkatan kualitas pendidikan melalui perbaikan system penilaian*. Media informasi Newsletter HEPI, I(2), 1.
- Joreskog, K., & Sorbum, D. (1996). *LISREL 8: User's reference guide*. Chicago: Scientific Software International Inc.
- Jurnal Psikologi Universitas Diponegoro Vol.3 No. 1, Juni 2006
- Jurnal Tabularasa PPs UNIMED Vol.5 No.1, Juni 2008
- Jurnal Tabularasa PPs UNIMED Vol.6 No.1, Juni 2009

Jurnal Psikologi Universitas Islam Indonesia, Volume 36, No. 2, Desember 2009: 116 – 129

King, A. (1995). Inquiring minds really do want to know: using questioning to teach critical thinking, *Teaching of Psychology*, 22, 13-17.

Kumaidi. (2004). *Sistem asesmen untuk menunjang kualitas pembelajaran*. Jurnal Pembelajaran, Volume 27 No.02, hal. 93-106.

Leigthon, J.P., & Gierl, M.J. (2007). *Cognitive diagnostic assessment for education: theory and application*. Cambridge University Press.

Lubis, M & Muhadjir, Noeng. (2004). *Jurnal penelitian dan evaluasi*, Nomor 7, Tahun VI, 2004)

Mardapi, J. (1988). *Practical implementation of validity generalization whit the Indonesian University selection test (sipenmaru)*. Disertasi doktor 1988.

Miles, Matthew B & Huberman, A. Michael. (2007). *Analisis data kualitatif*, (Terjemahan Tjetjep Rohidi & Mulyarto). Jakarta: Universitas Indonesia Press.

Moleong, Lexy J. (2005). *Metodologi penelitian kualitatif* (Rev. ed). Bandung: Remaja Rosda Karya

- Muhammad Khumaidi. (2007). Validitas konstruk tes kompetensi membaca gambar teknik mesin. *Jurnal Unes Semarang*.
- Muraki, E. (1997). A. Generalized partial credit model. Dalam W.J. Vander Linden., R.K. Hambleton (Eds.). *Handbook of Modern Item Respons Theory* (pp.153-163). New York: Springer-Verlag Inc.
- Muraki, E., & Bock, R.D. (1997). *PARSCALE: IRT item analysis and test scoring for rating-scale data*. Chicago, IL: Scientific Software International.
- Nield, AnthonyF dan Wintre, Maxine Galander. Multiple choice questions with an option to comment : student attitudes and use. *Teaching of psychology*. Vol.13. No.4, 1986.
- Program Pascasarjana UNY. (2002). *Kumpulan seminar dan loka karya*. Yogyakarta: Program pascasarjana.
- Paul, R.W. (tt). *Questioning to promote higher order thinking skills; the six types of socratic questions*.
- Raynolds, Cecil R. dkk ((2010). *Measurement and assessment in education*. Pearson Education. Upper River New Jersey

- Sahlan, S. & Maswan. (1988). *Multi dimensi sumber kreativitas manusia*. Bandung: Sinar Baru.
- Sax, Gilbert. (1980). *Principal of education and psychological measurement and evaluation*. Belmont, California. Wadsworth Publishing Company.
- Schon, D. (1985). *The design studios*. London: MIT and RIBA.
- Scouller, Karen. (2000). *The influence of assessment on student learning*. Paper
Presented at the annual conference of the Australian Association for research education. Sydney 4-7 December 2000. (<http://www.aare.edu.au/00pap/scoo00195.htm>)
- Sugiyono. (2006). *Metode Penelitian Kuantitatif, Kualitatif dan R & D*. Bandung: Alfabeta.
- Sukmadinata, Nana S. dkk. (2006). *Pengendalian mutu pendidikan sekolah menengah (konsep, prinsip dan instrumen)*. Bandung
- Sykes, R.C., & Yen, W.M. (2002). The scaling of mixed-item format test with the one-parameter and two-parameter partial credit model. *Journal of Educational Measurement*, 37, 221-224.

- Umar, Jahja. (1985). Berbagai permasalahan penggunaan bentuk soal uraian dan pilihan ganda dalam ujian. *Buletin Pengujian dan Penilaian*, 6-10.
- Vander Linden, W.J., & Hambleton, R.K. (1997). Item respons theory: Brief history, common model, and extention. Dalam W.J. Vander Linden., & R.K. Hambleton (Eds.). *Handbook of Modern Item Respons Theory* (pp. 1-28). New York: Springer-Verlag Inc.
- Wainer, Howard & Henry I. Braun. (1990). *Tes validity*. New Jersey: Hillsdale 07642.

