

PENELITIAN INDIVIDUAL

**UJI KUALITAS INSTRUMEN PENILAIAN HASIL BELAJAR
MATA PELAJARAN PAI SD DI PURWOKERTO**



LAPORAN PENELITIAN

Diajukan kepada LPPM IAIN Purwokerto
Untuk mendapatkan proyek dosen penelitian 2016

Oleh:

Dwi Priyanto, M.Pd

**INSTITUT AGAMA ISLAM NEGERI (IAIN)
PURWOKERTO
2016**

BAB I

PENDAHULUAN

A. Latar Belakang Masalah

Kegiatan penilaian dalam sebuah lembaga pendidikan merupakan satu rutinitas yang dilakukan dalam rangka untuk mengetahui perkembangan peserta didik dalam memahami, mensikapi dan melaksanakan nilai yang harus diejawantahkan dalam kehidupan sehari-hari. Penilaian akan memberikan informasi atas peserta didik dalam menindaklanjuti materi yang telah diberikan juga akan memberikan informasi ketercapaian tujuan pembelajaran. Dua hal ini sangat penting dalam dunia pendidikan untuk bisa terdeteksi sebagai masukan dalam kegiatan pembelajaran berikutnya agar lebih baik dan bermutu sehingga bisa menghasilkan lulusan yang berkualitas. Mencermati hal ini maka kita bisa melihat dalam Undang-Undang Sistem Pendidikan Nasional (UUSPN) Nomor 20 Tahun 2003 disebutkan bahwa pemerintah akan mengupayakan pengendalian mutu pendidikan nasional melalui sistem penilaian. UUSPN merupakan pedoman dan *role* pemerintah untuk menjamin peningkatan kualitas pendidikan nasional. Sebagaimana termaktub dalam pasal 57 ayat 1 bahwa penilaian dilakukan dalam rangka pengendalian kualitas pendidikan secara nasional sebagai bentuk akuntabilitas penyelenggaraan pendidikan kepada pihak-pihak yang berkepentingan.

Penilaian dalam penyelenggaraan pendidikan mempunyai peranan yang sangat penting sehingga harus mendapatkan perhatian yang serius. Karenanya penilaian merupakan serangkaian aktivitas untuk mengetahui keberhasilan proses pembelajaran, baik keberhasilan siswa (prestasi belajar) dan keberhasilan guru mengajar maupun keberhasilan proses pembelajaran. Instrumen merupakan alat yang akan digunakan dalam kegiatan penilaian agar sampai pada hakikat penilaian dilakukan. Proses instrumen untuk mampu mengungkap semua itu perlu disiapkan sebagai sebuah alat ukur yang memiliki daya ukur yang tepat. Instrumen yang memiliki tingkat validitas yang tinggi dan reliabilitas yang handal akan menghantarkan kita mencapai kegiatan penilaian mencapai apa yang diinginkan.

Sehingga informasi atau data secara kuantitatif dan kualitatif dari peserta didik akan bisa diketahui (Jahja Umar, et al., 1997: 2).

Penilaian juga sebagai proses mengumpulkan informasi untuk mengetahui pencapaian belajar kelas atau kelompok. Hasil penilaian 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 penilaian dapat digunakan sebagai pedoman dalam meningkatkan kualitas proses belajar mengajar. Penilaian pada dasarnya melakukan *judgment* terhadap hasil proses pembelajaran, maka kesalahan pada penilaian dan pengukuran diupayakan sekecil mungkin (Djemari Mardapi, 2008: 9).

Pernyataan penilaian 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 penilaian 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 penilaian 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. Untuk mendapatkan penilaian 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 dijalankannya.

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 Sekolah Dasar (SD) 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. Valid maksudnya betul-betul mengukur apa yang hendak diukur secara tepat. Alat ukur hanya mengukur satu dimensi atau aspek tertentu saja. Instrumen Pendidikan Agama Islam (PAI) hanya mengukur materi PAI bukan mengukur keterampilan materi yang lain. Handal maksudnya jika ia mempunyai hasil yang taat asas (*consistent*), 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, kondisi di lapangan masih banyak guru yang belum menerapkan penyusunan instrumen yang sesuai norma penilaian. 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 PAI di SD. Hasil survey awal yang peneliti peroleh bahwa soal ujian SD untuk ulangan tengah semester dibuat belum memperhatikan prinsip-prinsip evaluasi yang baik. Soal ujian sekolah dibuat oleh tim yang pemilihan tim tersebut tidak memperhatikan kualitas dari penyusun soal dalam tim yang ada. Prinsip penilaian yang ada tidak dilakukan dalam proses penyusunan soal ujian. Hal ini menjadikan peneliti merasa tertarik untuk meneliti lebih jauh tentang proses penyusunan butir soal yang ada di SD Purwokerto pada khususnya. Penelitian lebih mengarah kepada proses uji kualitas soal yang telah digunakan sebagai alat ukur pada siswa SD sehingga akan bisa diketahui kualitas soal tersebut, layak atau tidak soal ujian sebagai instrumen untuk mengukur kemampuan siswa. Penelitian ini akan mengambil setting SD di Purwokerto.

B. Rumusan Masalah

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

1. Apa kendala guru dalam proses penyusunan instrumen penilaian hasil belajar PAI siswa SD di Purwokerto?
2. Bagaimana pedoman penyusunan instrumen penilaian hasil belajar PAI siswa SD di Purwokerto?
3. Bagaimana proses penyusunan instrumen penilaian hasil belajar PAI siswa SD di Purwokerto?
4. Bagaimana kualitas instrumen penilaian hasil belajar PAI siswa SD di Purwokerto?
5. Bagaimana implikasi uji kualitas instrumen penilaian hasil belajar PAI siswa SD di Purwokerto?

C. Tujuan Penelitian dan

Tujuan penelitian adalah untuk; (1) mengetahui kendala guru dalam proses penyusunan instrumen penilaian hasil belajar PAI siswa SD di Purwokerto, 2) mengetahui pedoman penyusunan instrumen penilaian hasil belajar PAI siswa SD

di Purwokerto, 3) mengetahui proses penyusunan instrumen penilaian hasil belajar PAI siswa SD di Purwokerto, 4) mengetahui kualitas instrumen penilaian hasil belajar PAI siswa SD di Purwokerto, dan 5) mengetahui implikasi uji kualitas instrumen penilaian hasil belajar PAI siswa SD di Purwokerto? Penelitian ini sebagai bentuk penguatan keilmuan program studi (prodi) PAI. Mahasiswa prodi PAI akan sangat mungkin bisa terlibat dalam penelitian seperti ini, bahkan mahasiswa PAI juga akan dibekali untuk kemampuan dalam proses penyusunan instrumen penilaian hasil belajar yang baik (memiliki tingkat validitas dan reliabilitas yang tinggi). Sehingga instrumen penilaian benar-benar berfungsi sebagai alat ukur yang tepat, mampu mengukur apa yang seharusnya diukur. Mata kuliah yang akan menghantarkan penguasaan mahasiswa pada kompetensi ini adalah mata kuliah evaluasi pembelajaran PAI dan statistik pendidikan.

D. Signifikansi Penelitian

Signifikansi penelitian ini akan memberikan manfaat kepada mahasiswa prodi PAI, para guru pada umumnya dan guru mata pelajaran PAI SD pada khususnya di Purwokerto dalam penyusunan instrumen penilaian sebagai alat ukur yang valid. Sehingga benar-benar akan mengukur sebagai mana mestinya, untuk bisa mengetahui peserta didik yang memiliki kemampuan tinggi akan bisa kelihatan dalam penguasaan kompetensi dan sebaliknya yang belum maksimal penguasaan kompetensi juga akan bisa dilihat. Sehingga manfaat selanjutnya dari instrumen penilaian hasil belajar yang valid akan bisa untuk menentukan proses pembelajaran berikutnya untuk lebih baik lagi karena ada *feed back* hasil tes dari peserta didik. Dengan kata lain dengan penelitian ini guru PAI khususnya akan berusaha untuk bisa menyusun instrumen penilaian hasil belajar yang lebih baik. Bagi mahasiswa prodi PAI sebagai calon guru agama Islam juga sudah berusaha mempersiapkan agar bisa menjadi guru yang profesional khususnya dari sisi proses penilaian telah mempersiapkan bekal dalam menyusun instrumen penilaian yang baik karena ini sebagai faktor pendukung keberhasilan pendidikan agar lebih bermutu dan berkualitas.

E. Telaah Pustaka Penelitian Terkait

Penelitian Intan Rezki Kurniasari, Sunarmi, Nugraningsih yang berjudul Pengembangan Instrumen Penilaian Kognitif Materi Gerak Tumbuhan Dan Hama Penyakit Tumbuhan Kelas VIII Semester Genap Tahun Ajaran 2013-2014. Penelitian pengembangan instrumen penilaian kognitif materi gerak tumbuhan dan hama penyakit tumbuhan kelas VIII semester genap tahun ajaran 2013-2014 ini dilakukan bertujuan untuk : (1) mengembangkan soal evaluasi Biologi untuk mengukur prestasi belajar siswa kelas VIII SMP, (2) mengetahui tingkat kesukaran soal Biologi, (3) mengetahui daya beda hasil tes Biologi, (4) mengetahui reliabilitas instrumen tes Biologi, (5) mengetahui validitas konstruk instrumen tes Biologi, (6) mengetahui validitas isi instrumen tes Biologi. Penelitian pengembangan ini menggunakan teori pengembangan Thiagarajan yang terdiri dari 4 tahap yaitu: *define, design, develop, dan disseminate*. Pengumpulan data diperoleh dari hasil validasi ahli instrumen, materi dan lapangan serta hasil uji coba produk, kemudian dianalisis dengan deskriptif kuantitatif.

Hasil penelitian pengembangan instrumen adalah (1) instrumen penilaian kognitif sebanyak 40 soal dengan bentuk soal obyektif pilihan ganda dengan 38 soal valid dan 2 kurang valid. Butir soal yang kurang valid telah melalui revisi. (2) instrumen penilaian kognitif mempunyai tingkat kesukaran soal mudah 67,7%, soal sedang 22,5% dan soal sukar 10%, (3) instrumen penilaian kognitif mempunyai daya beda jelek 62,5%, cukup 25%, baik 10%, sangat baik 0%, negatif 2,5%, (4) instrumen penilaian kognitif mempunyai nilai reliabilitas 0,90, (5) instrumen penilaian kognitif mempunyai tingkat validitas konstruk yang valid, (6) instrumen penilaian kognitif mempunyai tingkat validitas isi yang valid. Jadi instrumen ini bisa digunakan namun sebelumnya melalui tahap revisi.

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. Hasil uji validitas konstruk dengan analisis faktor yang menghasilkan 6 subskala spiritualitas.

Penelitian Zulkifli 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 esai. 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.

Kemudian penelitian 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 konstruknya 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.

Sementara penelitian yang akan peneliti lakukan ini lebih menekankan untuk *men-create* instrumen penilaian dari sisi kualitas soal tes yang *nota bene* sebagai alat ukur untuk mengetahui kemampuan peserta dalam mencapai tujuan pembelajaran mata pelajaran PAI SD Adapun produk yang akan dicapai dalam penelitian ini adalah bisa diketahuinya kualitas instrumen penilaian hasil belajar mata pelajaran PAI SD di Purwokerto. Hasil penelitian ini akan memiliki kontribusi terlihatnya instrumen penilaian hasil belajar yang digunakan untuk evaluasi apakah memiliki tingkat validitas dan reliabilitas yang baik atau tidak sehingga bisa digunakan sebagai pedoman guru SD yang lain dalam kegiatan penilaian yang sudah semestinya harus menggunakan alat ukur 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, sehingga tidak salah dalam melakukan proses pengukuran dan penilaian peserta didiknya. Kesalahan dalam proses penilaian akan berdampak tidak bisa diketahuinya perkembangan kemampuan peserta didik yang

sebenarnya sehingga tidak mudah untuk menentukan langkah dan strategi pembelajaran berikutnya. Jika ini terjadi maka output atau lulusan tidak bisa mencapai hasil yang maksimal. Di sinilah pentingnya seorang guru dalam menyusun instrumen penilaian harus benar dan baik.

F. Kerangka Teori

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

1. 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*” (Gilbert, 1978: 621). Maksudnya adalah bahwa seberapa 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 tes menurut Anatasi dalam Howard Wainer dan Henry I. Braun (1988: 25) dapat dibedakan atas tiga macam yaitu:

- a. Validitas isi (*content validity*)
- b. Validitas konstruk (*construct validity*)
- c. Validitas kriteria- relasi (*criterion-related validity*)

Untuk mendapatkan gambaran masing-masing validitas tersebut akan diterangkan secara detail.

a. Validitas Isi (*content validity*)

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.

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 menvalidasi 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.*

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.

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).

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.

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. Dalam penelitian ini akan lebih menfokuskan pada validitas konstruk, karenanya akan tetap melihat validitas yang lain. Validitas konstruk di sisi akan digunakan untuk mengukur instrumen penilaian hasil belajar PAI pada

tingkat SD Aliyah. Aspek yang akan diukur meliputi tiga ranah yaitu; kognitif, afektif, dan psikomotor.

2. Instrumen Penilaian Hasil Belajar PAI

a. 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 pertanyaan/ 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.

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 SD, maka sampelnya juga peserta didik SD. Sampel yang diperlukan minimal 30 peserta didik, bisa berasal dari satu SD atau lebih.

Analisis hasil ujicoba, analisis hasil ujicoba meliputi variasi jawaban tiap butir pertanyaan/ pernyataan. Jika menggunakan skala instrumen 1 sampai 4, dan jawaban responden bervariasi dari 1 sampai 4, 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. 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.

b. Hasil Belajar PAI

Hasil belajar PAI sebaiknya mencakup pada tiga ranah dalam pembelajaran, yaitu ranah kognitif, afektif, dan psikomotor (Anderson & Krathwohl, 2001). Ranah kognitif meliputi kemampuan menghafal, memahami, menerapkan, menganalisis, mensintesis, dan kemampuan mengpenilaian. 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.

Adapun mata pelajaran PAI di SD bertujuan untuk:

1. menumbuhkembangkan akidah melalui pemberian, pemupukan, dan pengembangan pengetahuan, penghayatan, pengamalan, pembiasaan, serta pengalaman peserta didik tentang agama Islam sehingga menjadi manusia muslim yang terus berkembang keimanan dan ketakwaannya kepada Allah SWT,
2. mewujudkan manusia Indonesia yang taat beragama dan berakhlak mulia yaitu manusia yang berpengetahuan, rajin beribadah, cerdas, produktif, jujur, adil, etis, berdisiplin, bertoleransi (tasamuh), menjaga keharmonisan secara personal dan sosial serta mengembangkan budaya agama dalam komunitas sekolah.

c. Instrumen Penilaian Hasil Belajar PAI

Instrumen penilaian yang akan digunakan dalam penelitian akan sangat bergantung dari materi pelajaran, dalam hal ini adalah PAI. Mata pelajaran PAI 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.

3. Penilaian

Berbicara penilaian tidak bisa lepas dari pembahasan penilaian 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 penilaian, (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 Raynolds (1956) yang mengemukakan bahwa penilaian pendidikan adalah penaksiran terhadap pertumbuhan dan kemajuan siswa ke arah tujuan atau nilai-nilai yang telah ditetapkan dalam kurikulum. Penilaian juga dapat diartikan sebagai proses menilai sesuatu berdasarkan kriteria atau tujuan yang telah ditetapkan, yang selanjutnya diikuti dengan pengambilan keputusan terhadap focus pembahasan yang dipenilaian.

4. Mata Pelajaran PAI di SD

Pendidikan Agama Islam di SD/MI bertujuan untuk:

1. menumbuhkembangkan akidah melalui pemberian, pemupukan, dan pengembangan pengetahuan, penghayatan, pengamalan, pembiasaan, serta pengalaman peserta didik tentang agama Islam sehingga menjadi manusia muslim yang terus berkembang keimanan dan ketakwaannya kepada Allah SWT,
2. mewujudkan manuasia Indonesia yang taat beragama dan berakhlak mulia yaitu manusia yang berpengetahuan, rajin beribadah, cerdas, produktif, jujur, adil, etis, berdisiplin, bertoleransi (tasamuh), menjaga keharmonisan secara personal dan sosial serta mengembangkan budaya agama dalam komunitas sekolah.

Sedangkan Ruang lingkup Pendidikan Agama Islam meliputi aspek-aspek sebagai berikut.

1. Al-Qur'an dan Hadits

2. Aqidah
3. Akhlak
4. Fiqih
5. Tarikh dan Kebudayaan Islam

Kemudian untuk kompetensi yang akan dicapai dalam PAI SD khususnya kelas V adalah sebagai berikut:

Kompetensi Inti

1. Menerima, menjalankan dan menghargai ajaran agama yang dianutnya
2. Menunjukkan perilaku jujur, disiplin, tanggung jawab, santun, peduli dan percaya diri dalam berinteraksi dengan keluarga, teman, guru dan tetangganya serta cinta tanah air
3. Memahami pengetahuan faktual dan konseptual dengan cara mengamati menanya dan mencoba berdasarkan rasa ingin tahu tentang dirinya, makhluk ciptaan Tuhan dan kegiatannya dan benda-benda yang dijumpainya di rumah, di sekolah dan tempat bermain
4. Menyajikan pengetahuan faktual dan konseptual dalam bahasa yang jelas, sistematis, logis dan kritis, dalam karya yang estetis, dalam gerakan yang mencerminkan anak sehat dan dalam tindakan yang mencerminkan perilaku anak beriman dan berakhlak mulia

Kompetensi Dasar

- 1.1 Meyakini Al-Qur'an sebagai kitab suci terakhir dan menjadikannya sebagai pedoman hidup
- 1.2 Mengamalkan isi kandungan Q.S. At-Tin dan Al- Insyirah dalam kehidupan
- 1.3 Melaksanakan kewajiban puasa Ramadhan sebagai implementasi dari pemahaman rukun Islam
- 1.4 Melaksanakan shalat tarawih dan tadarus Al-Qur'an di bulan Ramadhan sebagai wujud ketaatan kepada Allah dan rasul-Nya
- 2.1 Menunjukkan perilaku suka menolong sebagai implementasi dari pemahaman Q.S. Al- Insyirah

- 2.2 Menunjukkan sikap saling mengingatkan dalam kebajikan sebagai implementasi dari pemahaman Q.S. At-Tin
- 2.3 Menunjukkan sikap jujur sebagai implementasi dari pemahaman kisah masa kanak-kanak Nabi Muhammad SAW
- 2.4 Menunjukkan sikap sabar dan pengendalian diri sebagai implementasi dari pemahaman puasa ramadhan
- 2.5 Mengenal nama-nama Rosul Allah SWT dan rasul ulul azmi sebagai implementasi pemahaman rukun iman
- 3.1 Memahami makna diturunkannya kitab-kitab suci melalui rosul-rosul-Nya sebagai implementasi rukun iman
- 3.2 Mengetahui makna Q.S. At-Tin dan Al- Insyirah dengan benar
- 3.3 Mengetahui hikmah puasa ramadhan yang dapat membentuk akhlak mulia
- 3.4 Mengetahui kisah kelahiran dan masa kanak-kanak Nabi Muhammad SAW
- 4.1 Membaca Q.S. At-Tin dan Al- Insyirah dengan benar
- 4.2 4.2 Menulis kalimat-kalimat dalam Q.S. At-Tin dan Al- Insyirah dengan baik dan benar
- 4.3 Menunjukkan hafalan Q.S. At-Tin dan Al- Insyirah dengan baik dan benar
- 4.4 Mencontohkan sikap saling mengingatkan dalam hal kebajikan sebagai implementasi dari pemahaman Q.S. At-Tin
- 4.5 Mencontohkan perilaku suka menolong sebagai implementasi Q.S. Al- Insyirah
- 4.6 Menceritakan kisah kelahiran dan masa kanak-kanak Nabi Muhammad SAW

G. Sistematika Pelaporan

Penelitian ini akan meliputi; BAB I Pendahuluan yang meliputi latar belakang, rumusan masalah, tujuan dan signifikansi penelitian, telaah pustaka penelitian terkait, kerangka teori dan metode penelitian serta sistematika pelaporan.

BAB II Instrumen Evaluasi Hasil Belajar yang meliputi instrumen evaluasi, persyaratan instrumen evaluasi yang baik, pengembangan instrumen pembelajaran, ciri-ciri tes hasil belajar yang baik dan pengujian validitas reliabilitas.

BAB III Analisis Butir Soal meliputi penilaian acuan normatif, penilaian acuan patokan, penilaian acuan gabungan, analisis butir soal dan analisis butir soal kuantitatif PAI SD di Purwokerto.



BAB II

INSTRUMEN EVALUASI

A. Persyaratan Instrumen Evaluasi

Sebuah instrumen evaluasi hendaknya memenuhi syarat sebelum di gunakan untuk mengevaluasi atau mengadakan penilaian agar terhindar dari kesalahan dan hasil yang tidak valid (tidak sesuai kenyataan sebenarnya). Alat evaluasi yang kurang baik dapat mengakibatkan hasil penilaian menjadi bias atau tidak sesuainya hasil penilaian dengan kenyataan yang sebenarnya, seperti contoh anak yang pintar dinilai tidak mampu atau sebaliknya. Jika terjadi demikian perlu ditanyakan apakah persyaratan instrumen yang digunakan menilai sudah sesuai dengan kaidah-kaidah penyusunan instrumen.

Menurut Sukardi mengemukakan bahwa, suatu evaluasi memenuhi syarat-syarat sebelum diterapkan kepada siswa yang kemudian direfleksikan dalam bentuk tingkah laku. Instrumen evaluasi yang baik, harus mempunyai syarat seperti berikut: 1) valid, 2) andal, 3) objektif, 4) seimbang, 5) membedakan, 6) norma, 7) fair, dan 8) praktis. Sedangkan Wina Sanjaya mengatakan bahwa syarat-syarat alat evaluasi yang baik harus:

1 Memberikan motivasi

Memberikan penilaian evaluasi diarahkan untuk meningkatkan motivasi belajar bagi siswa melalui upaya pemahaman akan kekuatan dan kelemahan yang dimiliki baik oleh guru maupun siswa. Siswa perlu memahami makna dari hasil penilaian.

2 Validitas

Penilaian diarahkan bukan semata-mata untuk melengkapi syarat administrasi saja, akan tetapi diarahkan untuk memperoleh informasi tentang ketercapaian kompetensi seperti yang terumuskan dalam kurikulum. Oleh sebab itu,

penilaian tidak menyimpang dari kompetensi yang ingin dicapai. Dengan kata lain penilaian harus menjamin validitas.

3 Adil

Setiap siswa memiliki kesempatan yang sama dalam proses pembelajaran tanpa memandang perbedaan sosial-ekonomi, latar belakang budaya dan kemampuan. Dalam penilaian, siswa disejajarkan untuk mendapatkan perlakuan yang sama.

4 Terbuka

Alat penilaian yang baik adalah alat penilaian yang dipahami baik oleh penilai maupun yang dinilai. Siswa perlu memahami jenis atau prosedur penilaian yang akan dilakukan beserta kriteria penilaian. Keterbukaan ini bukan hanya akan mendorong siswa untuk memperoleh hasil yang baik sehingga motivasi belajar mereka akan bertambah juga, akan tetapi sekaligus mereka akan memahami posisi mereka sendiri dalam pencapaian kompetensi.

5 Berkesinambungan

Penilaian tidak pernah mengenal waktu kapan penilaian seharusnya dilakukan. Penilaian dilakukan secara terus-menerus dan berkesinambungan.

6 Bermakna

Penilaian tersusun dan terarah akan memberikan makna kepada semua pihak khususnya siswa untuk mengetahui posisi mereka dalam memperoleh kompetensi dan memahami kesulitan yang dihadapi dalam mencapai kompetensi. Dengan demikian, hasil penilaian itu juga bermakna bagi guru juga termasuk bagi orang tua dalam memberika bimbingan kepada siswa dalam upaya memperoleh kompetensi sesuai dengan target kurikulum.

7 Menyeluruh

Kurikulum diarahkan untuk perkembangan siswa secara utuh, baik perkembangan afektif, kognitif maupun psikomotorik. Oleh sebab itu, guru dalam melaksanakan penilaian harus menggunakan ragam penilaian, misalnya tes, penilaian produk, skala sikap, penampilan, dan sebagainya. Hal ini sangat penting, sebab hasil penilaian harus memberikan informasi secara utuh tentang perkembangan setiap aspek.

8 Edukatif

Penilaian kelas tidak semata-mata diarahkan untuk memperoleh gambaran kemampuan siswa dalam pencapaian kompetensi melalui angka yang diperoleh, akan tetapi hasil penilaian harus memberikan umpan balik untuk memperbaiki proses pembelajaran, baik yang dilakukan oleh guru maupun siswa, sehingga hasil belajar lebih optimal. Dengan demikian, proses penilaian tidak semata-mata tanggung jawab guru akan tetapi juga merupakan tanggung jawab siswa. Artinya siswa harus ikut terlibat dalam proses penilaian, sehingga mereka menyadari, bahwa penilaian adalah bagian dari proses pembelajaran.¹ Sedangkan Daryanto membagi syarat-syarat evaluasi menjadi 5 (lima) bagian, diantaranya:²

1) Keterpaduan

Evaluasi merupakan komponen integral dalam program pengajaran disamping tujuan serta metode. Tujuan instruksional, materi dan metode, serta evaluasi merupakan tiga keterpaduan yang tidak boleh dipisahkan.

2) Koherensi

Dengan prinsip koherensi diharapkan evaluasi harus berkualitas dengan materi pengajaran yang sudah disajikan dan sesuai dengan ranah kemampuan yang hendak diukur.

¹ Wina Sanjaya. Kurikulum dan Pembelajaran. Bandung: Kencana Prenada Media Group. 2008.h. 352-354

² Daryanto. Evaluasi Pendidikan. Solo: Rineka Cipto. 2009. h. 19-28

3) Pedagogis

Evaluasi perlu diterapkan sebagai upaya perbaikan sikap dan tingkah laku ditinjau dari segi pedagogis. Evaluasi dan hasilnya hendaknya dapat dipakai sebagai alat motivasi untuk siswa dalam kegiatan belajarnya.

4) Akuntabilitas

Sejauh mana keberhasilan program pengajaran perlu disampaikan kepada pihak-pihak yang berkepentingan dengan pendidikan sebagai laporan pertanggungjawaban (*accountability*). Artinya bahwa seorang guru dalam merencanakan kegiatan evaluasi haruslah mempertimbangkan hal yang mendukung pada terciptanya evaluasi yang tepat agar mendapatkan hasil yang tepat pula.

B. Pengembangan Instrumen Pembelajaran

Instrumen atau alat pengumpul data adalah alat yang digunakan untuk mengumpulkan data dalam satu kegiatan penelitian. Dalam konteks pembelajaran instrumen akan dijadikan sebagai alat untuk mengumpulkan informasi tentang perkembangan peserta didik. Instrumen sangat penting dalam upaya mendapatkan informasi atau data peserta didik sehingga informasi atas peserta didik kepada pihak yang membutuhkan bisa diberikan dengan data yang benar. Keakuratan informasi ditentukan baik tidaknya instrumen, semakin baik instrumen akan semakin akurat informasi yang diperoleh, sebaliknya dengan instrumen yang kurang baik maka informasi yang diperoleh akan semakin kurang akurat.

Instrumen dalam pembelajaran lebih diorientasikan pada upaya menghasilkan data mengenai perkembangan peserta didik dalam memahami dan menguasai materi yang telah diberikan. Kegiatan yang biasa dilakukan adalah pada evaluasi sebagai moment untuk mengetahui progres peserta didik maupun keberhasilan proses pembelajaran dalam menghantarkan peserta didik mencapai tujuan pembelajaran. Yang kedua sekaligus sebagai parameter keberhasilan pembelajaran dalam mencapai target sebagaimana telah ditetapkan dalam perencanaan.

1. Langkah-langkah Penyusunan dan Pengembangan Instrumen Tes

Penyusunan dan pengembangan pada prinsipnya kegiatan yang akan menghasilkan item-item instrumen yang berupa tes, adapun langkah-langkahnya adalah sebagai berikut:³

a. Menyusun spesifikasi tes

Uraian yang menunjukkan keseluruhan karakteristik yang harus dimiliki suatu tes. Spesifikasi yang jelas akan mempermudah dalam menulis soal, dan siapa saja yang menulis akan mendapatkan soal yang relatif sama tingkat kesukarannya. Penyusunan spesifikasi meliputi kegiatan berikut:

- 1) menentukan tujuan tes; ada empat macam tes dilihat dari tujuannya yaitu tes penempatan, tes diagnostik, tes formatif dan tes sumatif.
- 2) menyusun kisi-kisi tes; dalam mengembangkan kisi-kisi tes kita perlu mempertimbangkan empat hal di antaranya membuat tujuan umum pembelajaran, membuat daftar pokok dan subpokok bahasan yang akan diujikan, menentukan indikator, dan menentukan jumlah soal setiap pokok bahasan dan sub pokok bahasan
- 3) memilih bentuk tes; pemilihan bentuk tes yang tepat ditentukan oleh tujuan tes, jumlah peserta tes, waktu yang tersedia untuk memeriksa lembar jawaban tes, cakupan materi tes, dan karakteristik mata pelajaran yang diujikan, dan
- 4) menentukan panjang tes; penentuan panjang tes berdasarkan cakupan materi ujian dan tingkat kesulitan dari soal tes yang dibuat. Pada umumnya tes tertulis menggunakan waktu 90 sampai 150 menit, untuk soal tes praktik bisa lebih dari itu. Soal pilihan ganda biasanya membutuhkan waktu kurang lebih 2 menit setiap soalnya, namun tetap ada pengaruh dari tingkat kesulitan soal yang ada, sehingga waktu mungkin ada yang kurang dari 2 menit dan ada yang lebih dari 2 menit.

³ Djemari Mardapi. Teknik Penyusunan Instrumen Tes dan Nontes. Yogyakarta: Mitra Cendekia Press. 2008. h. 88-97

b. Menulis soal tes

Penulisan soal merupakan langkah menjabarkan indikator menjadi pertanyaan-pertanyaan yang karakteristiknya sesuai dengan perincian pada kisi-kisi yang telah dibuat. Setiap pertanyaan perlu disusun sedemikian sehingga jelas yang ditanyakan dan jelas pula jawaban yang diinginkan atau diharapkan.

c. Menelaah soal tes

Proses telaah soal dilakukan setelah soal selesai dibuat untuk menghindari kekeliruan yang akan berakibat pada peserta didik tidak bisa memahami maksud soal. Sebaiknya penelaah orang lain yang bukan penyusun agar lebih obyektif. Maksud telaah soal agar soal yang dihasilkan adalah soal yang benar-benar berkualitas.

d. Melakukan ujicoba tes

Tujuan soal diujicobakan agar diperoleh informasi mengenai soal tersebut dari sisi reliabilitas, validitas, tingkat kesukaran, daya beda, pola jawaban, efektifitas distraktor. Hasil ujicoba akan menjadi dasar dalam memperbaiki soal jika diketahui soal belum sesuai dengan parameter soal berkualitas. Ujicoba dilakukan kepada peserta didik yang bukan akan mendapatkan soal tersebut sebagai evaluasi. Misal jika soal hanya untuk lingkup satu sekolah maka bisa diujicobakan kepada peserta didik sekolah lain meskipun dekat jaraknya, namun jika soal diperuntukan untuk peserta didik dalam satu kecamatan maka harus diujicobakan pada peserta didik di kecamatan lain.

e. Menganalisis butir soal

Analisis butir soal dilakukan dengan mendasarkan soal yang telah diujicobakan. Melalui analisis ini akan bisa diketahui tingkat kesukaran soal, daya beda soal maupun keberfungsiaan distraktor dalam rangka untuk mendapatkan soal yang berkualitas. Soal-soal yang telah diujicobakan akan dianalisis satu per satu butir sehingga masing-masing butir soal akan diketahui kualitasnya.

f. Memperbaiki tes

Soal yang tidak sesuai dengan yang diharapkan berarti soal tidak berkualitas, pada soal ini perlu ada upaya perbaikan agar sampai pada soal yang masuk kategori soal yang sesuai dengan patokan yang telah distandarkan oleh para *expert* evaluasi.

g. Merakit tes

Soal yang telah diperbaiki selanjutnya penyusun soal akan merakitnya tentunya harus berhati-hati dalam merakit menjadi kesatuan soal tes yang terpadu. Proses ini perlu memperhatikan hal yang dapat mempengaruhi validitas soal di antaranya nomor urut soal, pengelompokan bentuk soal, *lay out*, dan yang lainnya.

h. Melaksanakan tes

Ketika proses merakit atau menyusun selesai dan telah dilakukan revisi pasca ujicoba maka soal tes bisa digunakan dalam pelaksanaan tes kepada para testee. Pelaksanaan tes memerlukan pemantauan atau pengawasan agar tes benar-benar dikerjakan oleh testee masing-masing dengan jujur dan sesuai dengan ketentuan yang ada. Meskipun demikian pemantauan diusahakan tidak boleh mengganggu ketenangan testee dalam mengerjakan soal tes, karena dapat mengakibatkan hasil tes yang diperoleh kurang akurat.

b. Menafsirkan hasil tes

Hasil tes dari para testee berupa data kuantitatif yang masih perlu ditafsirkan untuk menjadi nilai yang bisa memiliki arti. Oleh karenanya harus ditafsirkan melalui proses penilaian dan pada akhirnya dievaluasi yang sampai keputusan untuk masing-masing testee siapa yang masuk dalam kategori rendah, sedang ataupun tinggi. Tinggi rendahnya nilai selalu dikaitkan dengan acuan penilaian. Ada dua acuan penilaian yang sering digunakan dalam bidang psikologi pendidikan, yaitu acuan norma dan acuan kriteria. Jadi tinggi rendahnya suatu nilai dibandingkan dengan kelompoknya atau dengankriteria yang harus dicapai. Nilai bisa memberikan informasi keberhasilan peserta didik

dalam mengikuti pembelajaran, guru dalam memberikan pelajaran, dosen dalam melakukan proses perkuliahan. Bahan untuk memotivasi peserta didik sebagai bahan untuk mempertimbangkan siapa yang akan mendapat *reward*. Sehingga pada akhirnya nilai ini pula yang akan bisa sebagai *feed back* digunakan pendidik dalam melakukan proses pembelajaran berikutnya untuk semakin baik.

C. Ciri-ciri Tes Hasil Belajar yang Baik

Adapun tes di katakan baik sebagai alat pengukur harus memenuhi persyaratan tes sebagai berikut:⁴

1. Alat evaluasi harus Objektivitas, maksudnya tidak adanya unsur pribadi yang mempengaruhi lawan dari objektif adalah subyektif, artinya terdapat unsur pribadi yang masuk mempengaruhinya, untuk menghindari masuknya unsur subyekifitas maka penilaian harus di laksanakan dengan mengingat pedoman, pedoman tersebut, antara lain :
 - a. mencakup keseluruhan materi.
 - b. mencakup berbagai aspek berpikir (ingatan, pemahaman, aplikasi, dan sebagainya.
 - c. melalui berbagai cara yaitu tes tertulis, tes lisan, tes perbuatan, pengamatan insidental, dan sebagainya.
2. Evaluasi harus dilaksanakan secara kontinu (terus menerus), dengan demikian akan di peroleh gambaran yang lebih jelas tentang keadaan siswa.
3. Evaluasi harus dilaksanakan secara komperhensif (menyeluruh), yang dimaksud dengan komperhensif di sini adalah atas berbagai segi peninjauan, yaitu;
 - a. mencakup keseluruhan materi.
 - b. mencakup berbagai aspek berpikir (ingatan, pemahaman, aplikasi, dan sebagainya.
 - c. melalui berbagai cara yaitu tes tertulis, tes lisan, tes perbuatan, pengamatan insidental, dan sebagainya.
4. Alat evaluasi harus sah (valid), maksudnya tes betul-betul secara tepat, benar, shahih atau absah telah dapat mengungkap atau mengukur apa yang seharusnya diungkap atau diukur lewat tes tersebut.⁵
5. Alat evaluasi harus andal (reliable), suatu tes dapat di katakan andal jika mempunyai hasil yang taat asas (consistent). dengan kata lain, jika suatu tes

⁴ Suharsimi Arikunto dan Jabar. *Evaluasi Program Pendidikan*. Jakarta: Bumi Aksara. 2004. h. 51-61

⁵ Anas Sudijono. *Pengantar Evaluasi Pendidikan*. Jakarta: raja Gravindo Persada. 2008. h. 93

diberikan kepada sekelompok subjek saat ini, kemudian di berikan lagi kepada sekelompok subjek yang sama pada saat yang akan datang, dan hasilnya ternyata sama atau mendekati sama, maka dapat di katakan tes itu mempunyai keandalan yang tinggi.

6. Alat evaluasi harus Diskriminatif, maksudnya alat evaluasi harus di susun sedemikian rupa sehingga dapat menunjukkan perbedaan-perbedaan yang kecil sekalipun. Semakin baik suatu alat evaluasi, maka semakin mampu tes itu menunjukkan perbedaan secara teliti.
7. Alat evaluasi harus Praktibilitas, sebuah tes di katakan memiliki praktibilitas yang tinggi apabila tes tersebut bersifat praktis dan mudah pengadministrasiannya, tes yang praktis adalah tes yang :
 - a. Mudah dilaksanakan, misalnya tidak menuntut peralatan yang banyak dan memberi kebebasan kepada siswa untuk mengerjakan terlebih dahulu bagian yang dianggap mudah oleh siswa
 - b. Mudah pemeriksaanya, artinya bahwa tes itu di lengkapi dengan kunci jawaban maupun pedoman skoringnya.
 - c. Di lengkapi dengan petunjuk yang jelas sehingga dapat di berikan kepadaorang lain.
 - d. Tes harus ekonomis, maksudnya adalah pelaksanaan tes itu tidak membutuhkan biaya yang mahal, tenaga banyak dan waktu yang lama

D. Pengujian Validitas dan Reliabilitas Instrumen

Pengujian suatu hipotesa tentang hubungan variabel penelitian sangat tergantung pada kualitas data yang dipakai dalam pengujian tersebut. Hal-hal yang menyebabkan data yang dikumpulkan tidak valid dan tidak reliabel, merupakan prasyarat agar hasil penelitian yang dicapai dapat dipertanggungjawabkan secara ilmiah. Validitas menunjukkan sejauh mana alat pengukur itu mengukur apa yang ingin diukur. Sedangkan reliabilitas adalah istilah yang dipakai untuk menunjukkan sejauhmana suatu hasil pengukuran relatif konsisten apabila pengukuran diulangi dua kali atau lebih. Contoh: apa bila kita ingin mengukur panjang suatu ruangan, alat yang diperlukan adalah penggaris (meteran), berarti

alat tersebut valid karena sesuai dengan fungsinya untuk mengukur panjang, dan sekalipun diukur lebih dari dua kali, hasilnya juga tetap sama artinya reliabilitasnya teruji. Tetapi apabila diukur dengan langkah kaki, maka hasilnya bila diukur lebih dari dua kali maka hasilnya akan berbeda, berarti tidak reliabel.

1. Validitas

Ada beberapa macam validitas antara lain:

a. Validitas konstruk

Konstruk adalah kerangka dalam suatu konsep, misalkan seorang peneliti ingin mengukur konsep 'relegiusitas'. Konsep relegiusitas, harus dijabarkan dalam kerangka konsep yang dapat dijabarkan dalam tolak ukur operasional. Ada tiga cara untuk mencapai kerangka konsep dalam suatu penelitian, antara lain:

- 1) Mencari definisi-definisi konsep yang dikemukakan para ahli yang tertulis dalam literatur. Apabila definisi tersebut sudah mengandung kerangka konsep (ada tolok ukurnya), maka peneliti langsung bisa menggunakannya, namun apabila belum maka perlu dioperasionalkan sehingga ada tolok ukur yang jelas.
- 2) Kalau sekiranya di dalam literatur tidak dapat diperoleh definisi konsep yang ingin diukur, peneliti harus mendefinisikan konsep tersebut, dengan cara mendiskusikan dengan para ahli yang kompeten dibidangnya.
- 3) Menyamakan definisi yang akan diukur kepada calon responden, atau orang-orang yang memiliki karakteristik yang sama dengan responden.

Misalnya untuk mengukur relegiusitas, dapat menggunakan pendapat dari Nienke yang menyatakan bahwa untuk mengetahui kadar relegiusitas individu dapat dipakai kerangka berikut:⁶

⁶ Nienke Nieveen. Prototyping to Reach Product Quality. In J. van den Akker, R. Branch, K. Gustafson, N. Nieveen and Tj. Plomp (Eds). *Design Approaches and Tools in Education and Training*. Dordrecht : Kluwer Academic Publisher, 1999. p. 125-136).

- 1) Keterlibatan ritual (*ritual involvement*), yaitu sejauhmana seorang mengerjakan kewajiban ritual di dalam agama mereka (sholat, zakat, puasa, membayar zakat).
- 2) Keterlibatan ideologi (*ideological involvement*), yaitu tingkatan sejauh mana orang menerima hal-hal yang dogmatis di dalam agama mereka (misal: apakah seorang percaya pada malaikat, hari kiamat)
- 3) Keterlibatan intelektual (*intellectual involvement*), yang menggambarkan seberapa jauh seseorang mengetahui tentang ajaran agama. Seberapa jauh aktivitasnya di dalam menambah pengetahuan agama (misal: ikut pengajian, membaca buku agama).
- 4) Keterlibatan pengalaman (*experiential involvement*), yaitu menunjukkan apakah seseorang pernah mengalami pengalaman spektakuler yang merupakan keajaiban dari Tuhan (misal: merasakan do'anya terkabul).
- 5) Keterlibatan secara konsekuen (*consequential involvement*), yaitu tingkatan sejauhmana perilaku seorang konsekuen dengan ajaran agama. (misal: berjudi, berzina).

b. Validitas isi

Validitas isi suatu alat pengukur ditentukan oleh sejauh mana isi alat tersebut mewakili sebagai aspek kerangka konsep. Contoh seorang peneliti ingin meneliti tingkat religiusitas suatu masyarakat, maka seluruh aspek (5 aspek religiusitas: Keterlibatan ritual, Keterlibatan ideologi, Keterlibatan intelektual, Keterlibatan pengalaman, Keterlibatan secara konsekuen) harus dimasukkan dalam kerangka konsep yang disusun dalam teknik dan instrumen pengumpulan data.

c. Validitas Eksternal

Validitas eksternal berkaitan dengan hasil yang dicapai dari instrumen yang digunakan sesuai dengan data atau informasi lain mengenai variabel penelitian yang diteliti. Contoh: seorang peneliti ingin mengetahui validitas tes

IPS. Caranya adalah mencoba tes tersebut kepada siswa yang diambil sebagai subyek uji coba. Hasil yang diperoleh kemudian dikorelasikan dengan nilai IPS anak-anak tersebut, misal dari nilai rapor, sebagai ukuran atau kriterium.

d. Validitas Prediktif

Alat pengukur yang dibuat oleh peneliti dimaksudkan untuk memprediksi apa yang akan terjadi dimasa yang akan datang. Contoh tes masuk perguruan tinggi bagi siswa yang lulus diprediksikan mampu mengikuti pelajaran di perguruan tinggi dengan sukses.

2. Reliabilitas

Kata reliabilitas dalam bahasa Indonesia yang digunakan saat ini, sebenarnya diambil dari kata *reliability* dalam bahasa Inggris dan berasal dari kata *reliable* yang artinya dapat dipercaya, keajegan, konsisten, keandalan, kestabilan. Suatu tes dapat dikatakan reliabel jika tes tersebut menunjukkan hasil yang dapat dipercaya dan tidak bertentangan. Menurut Sugiono Reliabilitas adalah serangkaian pengukuran atau serangkaian alat ukur yang memiliki konsistensi bila pengukuran yang dilakukan dengan alat ukur itu dilakukan secara berulang.⁷ Reliabilitas tes adalah tingkat keajegan (konsistensi) suatu tes, yakni sejauh mana suatu tes dapat dipercaya untuk menghasilkan skor yang ajeg, relatif tidak berubah walaupun diteskan pada situasi yang berbeda-beda. Sedangkan Sukadji (2000) mengatakan bahwa reliabilitas suatu tes adalah seberapa besar derajat tes mengukur secara konsisten sasaran yang diukur.⁸ Reliabilitas dinyatakan dalam bentuk angka, biasanya sebagai koefisien, jika nilai koefisien tinggi berarti memiliki tingkat reliabilitas yang tinggi.

Menurut Nursalam reliabilitas adalah kesamaan hasil pengukuran atau pengamatan bila fakta atau kenyataan hidup tadi diukur atau diamati berkali-kali

⁷ Sugiyono. Metodologi Penelitian Kualitatif. Bandung: Alfabeta. 2005. h.34

⁸ Sukaji. Menyusun dan mengevaluasi laporan penelitian. Jakarta: UI Press. 2000. h, 32

dalam waktu yang berlainan. Alat dan cara mengukur atau mengamati sama–sama memegang peranan penting dalam waktu yang bersamaan.⁹ Menurut Arifin suatu tes dapat dikatakan andal (*reliable*) jika tes tersebut mempunyai hasil yang taat asas (konsisten).¹⁰ Sedangkan Sudjana mengatakan bahwa reliabilitas suatu tes adalah ketepatan atau kejegan tes tersebut dalam menilai apa adanya, artinya kapan pun tes tersebut digunakan akan memberikan hasil yang sama atau relatif sama.¹¹

Berdasarkan beberapa pendapat tentang reliabilitas di atas, maka dapat diambil kesimpulan bahwa reliabilitas adalah suatu pengukuran terhadap suatu tes yang melihat apakah tes tersebut dapat mengukur apa yang seharusnya di ukur.

a. Jenis- Jenis Reliabilitas

Salah satu syarat agar hasil suatu tes dapat dipercaya adalah tes tersebut harus mempunyai reliabilitas yang memadai. Oleh karena itu Djaali dan Pudji membedakan reliabilitas menjadi 2 macam, yaitu; 1) Reliabilitas Konsistensi tanggapan, dan 2) Reliabilitas konsistensi gabungan item¹²

1) Reliabilitas Konsistensi Tanggapan

Reliabilitas ini selalu mempersoalkan mengenai tanggapa responden atau objek terhadap tes tersebut apakah sudah baik atau konsisten. Dalam artian apabila tes yang telah dicobakan tersebut dilakukan pengukuran kembali terhadap obyek yang sama, apakah hasilnya masih tetap sama dengan pengukuran sebelumnya. Jika hasil pengukuran kedua menunjukkan ketidakonsistenan, maka hasil pengukuran tersebut tidak menggambarkan keadaan obyek yang sesungguhnya. Untuk mengetahui apakah suatu tes atau instrument tersebut sudah

⁹ Nursalam. Konsep dan Penerapan Metodologi Penelitian Ilmu Keperawatan; pedoman skripsi, tesis dan instrumen penelitian keperawatan. Jakarta: Salemba Medika. 2003. h, 43

¹⁰ Zainal Arifin. Evaluasi Instruksional. Badung: Remaja Rosdakarya. 1991. 21

¹¹ D. Sudjana. Manajemen Program Pendidikan untuk Lembaga Nonformal dan Pengembangan Sumber Daya Manusia. Bandung: Falah Production. 2004. h, 54

¹² Djaali dan Puji Mulyono. Pengukuran dalam Bidang Pendidikan. Jakarta: Grasindo. 2008. h. 56

mantap atau konsisten, maka tes/instrument tersebut harus diuji kepada obyek ukur yang sama secara berulang-ulang.

Ada tiga mekanisme untuk memeriksa reliabilitas tanggapan responden terhadap tes yaitu :

- a) Teknik test-retest ialah pengtesan dua kali dengan menggunakan suatu tes yang sama pada waktu yang berbeda.
- b) Teknik belah dua ialah pengtesan (pengukuran) yang dilakukan dengan dua kelompok item yang setara pada saat yang sama.
- c) Bentuk ekivalen ialah pengtesan (pengukuran) yang dilakukan dengan menggunakan dua tes yang dibuat setara kemudian diberikan kepada responden atau obyek tes dalam waktu yang bersamaan.

2) Reliabilitas Konsistensi Gabungan Item

Reliabilitas ini terkait dengan konsistensi antara item-item suatu tes atau instrument.. Apabila terhadap bagian obyek ukur yang sama, hasil pengukuran melalui item yang satu kontradiksi atau tidak konsisten dengan hasil ukur melalui item yang lain maka pengukuran dengan tes (alat ukur) sebagai suatu kesatuan itu tidak dapat dipercaya. Untuk itu jika terjadi hal demikian maka kita tidak bisa menyalahkan obyek ukur, melainkan alat ukur (tes) yang dipersalahkan, dengan mengatakan bahwa tes tersebut tidak reliabel atau memiliki reliabilitas yang rendah.

Koefisien reliabilitas konsistensi gabungan item dapat dihitung dengan menggunakan 3 rumus yakni :

- a) Rumus Kuder-Richardson, yang dikenal dengan nama KR-20 dan KR-21.
- b) Rumus koefisien Alpha atau Alpha Cronbach.
- c) Rumus reliabilitas Hoyt, yang menggunakan analisis varian.¹³

¹³ Djaali dan Puji Mulyono. Pengukuran dalam Bidang Pendidikan. Jakarta: Grasindo. 2008. h. 58

3. Kepraktisan Suatu Instrumen

Dalam kamus besar bahasa Indonesia kepraktisan diartikan sebagai suatu yang bersifat praktis atau efisien. Arikunto (2010) mengartikan kepraktisan dalam evaluasi pendidikan merupakan kemudahan-kemudahan yang ada pada instrument evaluasi baik dalam mempersiapkan, menggunakan, menginterpretasi/memperoleh hasil, maupun kemudahan dalam menyimpannya.¹⁴

Kepraktisan juga merupakan salah satu ukuran suatu instrumen evaluasi dikatakan baik atau tidak. Bila guru menggunakan esay tes untuk mengukur tanggapan siswa terhadap suatu produk pembelajaran, dan jumlah siswa yang dibimbingnya mencapai dua ratus orang, maka upaya ini cenderung tidak praktis. Diperlukan cara lain untuk menilai tanggapan siswa tersebut, misalnya dengan tes lisan terhadap hasil diskusi kelompok. Kepraktisan diartikan pula sebagai kemudahan dalam penyelenggaraan, membuat instrumen, dan dalam pemeriksaan atau penentuan keputusan yang objektif, sehingga keputusan tidak menjadi bias dan meragukan. Kepraktisan dihubungkan pula dengan efisien dan efektifitas waktu dan dana. Sebuah tes dikatakan baik bila tidak memerlukan waktu yang banyak dalam pelaksanaannya, dan tidak memerlukan dana yang besar atau mahal.

Kepraktisan sebuah alat evaluasi lebih menekankan pada tingkat efisiensi dan efektivitas alat evaluasi tersebut, beberapa kriteria dalam mengukur tingkat kepraktisan, diantaranya adalah:

- a) Waktu yang diperlukan untuk menyusun tes tersebut
- b) Biaya yang diperlukan untuk menyelenggarakan tes tersebut
- c) Waktu yang diperlukan untuk melaksanakan tes
- d) Tingkat kesulitan menyusun tes
- e) Tingkat kesulitan dalam proses pemeriksaan tes
- f) Tingkat kesulitan melakukan intrepetasi terhadap hasil tes

¹⁴ Suharsimi Arikunto. Dasar-dasar Evaluasi pendidikan. Jakarta: Bumi Aksara. 1995. h. 51

Kepraktisan alat evaluasi akan memberikan manfaat yang besar bagi pelaksanaan maupun bagi peserta didik karena dirancang sedemikian sistematis terutama materi instrumen tersebut. Berkaitan kepraktisan dalam penelitian pengembangan Van den Akker menyatakan :

“Practically refers to the extent that user (or other expert) consider the intervention as appealing and usable in ‘normal’ conditions”¹⁵

Artinya, kepraktisan mengacu pada tingkat bahwa pengguna (atau pakar-pakar lainnya) mempertimbangkan intervensi dapat digunakan dan disukai dalam kondisi normal.

Untuk mengukur tingkat kepraktisan yang berkaitan dengan pengembangan instrument berupa materi pembelajaran, Nieveen berpendapat bahwa untuk mengukur kepraktisannya dengan melihat apakah guru (dan pakar-pakar lainnya) mempertimbangkan bahwa materi mudah dan dapat digunakan oleh guru dan siswa. Khusus untuk pengembangan model yang dikembangkan dalam penelitian pengembangan, model tersebut dikatakan praktis jika para ahli dan praktisi menyatakan bahwa secara teoritis bahwa model dapat diterapkan di lapangan dan tingkat keterlaksanaannya model tersebut termasuk kategori “baik”. Istilah “baik” ini masih memerlukan indikator-indikator yang diperlukan untuk menentukan tingkat “kebaikan” dari keterlaksanaan model yang di kembangkan.

Berkaitan dengan kepraktisan di tinjau dari apakah guru dapat melaksanakan pembelajaran di kelas. Biasanya peneliti dan observer mengamati aktivitas yang dilakukan guru dalam pelaksanaan pembelajaran. Misalnya, melihat kegiatan guru dalam mempersiapkan siswa untuk belajar, memeriksa pekerjaan siswa, dll.

¹⁵ JV. Akker. Principles and Methods of Development Research. In. J. Van den Akker, R. Branch, K. Gustafon, N Nieveen and Tj.Plomp (Eds). *Design Approaches and Tools in Education and Training* (hlm. 1-14). Dodrecht : Kluwer Academic Publisher. 1999. p.10

4. Efek Potensial (Efektivitas)

Aspek penting dalam keefektifan (efek potensial) dari suatu instrumen, teori, atau model adalah mengetahui tingkat/derajat dari penerapan teori, atau model dalam suatu situasi tertentu. Tingkat keefektifan ini biasanya dinyatakan dengan suatu skala numerik yang didasarkan pada kriteria tertentu. Berkaitan dengan keefektifan pengembangan instrument, model, teori dalam dunia pendidikan, Van den Akker menyatakan :

“Effectiveness refer to the extent that the experiences and outcomes with the intervention are consistent with the intended aims”¹⁶

Artinya, keefektifan mengacu pada tingkatan bahwa pengalaman dan hasil intervensi konsisten dengan tujuan yang dimaksud.

Keefektifan suatu bahan ajar biasanya dilihat dari potensial efek berupa kualitas hasil belajar, sikap., dan motivasi peserta didik. Menurut Akker sebagaimana dikemukakan Yazid ada dua aspek keefektifan yang harus dipenuhi oleh suatu bahan ajar, yaitu:

1. Ahli dan praktisi berdasarkan pengalamannya menyatakan bahwa bahan ajar tersebut efektif.
2. Secara operasional bahan ajar tersebut memberikan hasil sesuai yang diharapkan.¹⁷

Menurut Suryadi dalam Yazid bahan ajar dapat dikatakan efektif apabila:¹⁸

1. Rata-rata siswa aktif dalam aktivitas pembelajaran.

¹⁶ JV. Akker. Principles and Methods of Development Research. In. J. Van den Akker, R. Branch, K. Gustafon, N Nieveen and Tj.Plomp (Eds). *Design Approaches and Tools in Education and Training* (hlm. 1-14). Dodrecht : Kluwer Academic Publisher. 1999. p.10

¹⁷ A Yazid. Kevalidan, kepraktisan dan efek Potensial Bahan Ajar. Pascasarjana Pendidikan Matematika Universitas Sriwijaya.
<http://aisyahyazid.blogspot.com/2011/12/kevalidan-kepraktisan-dan-efek.html>

¹⁸ A Yazid. Kevalidan, kepraktisan dan efek Potensial Bahan Ajar. Pascasarjana Pendidikan Matematika Universitas Sriwijaya.
<http://aisyahyazid.blogspot.com/2011/12/kevalidan-kepraktisan-dan-efek.html>

2. Rata-rata siswa aktif dalam mengerjakan tugas.
3. Rata-rata siswa efektif dalam keefektifan relatif penguasaan bahan pengajaran.
4. Respon siswa terhadap pembelajaran yang dilaksanakan baik/positif
5. Respon guru terhadap pembelajaran yang dilaksanakan baik/positif

5. Analisis Butir Soal

a. Pengertian

Kegiatan menganalisis butir soal merupakan suatu kegiatan yang harus dilakukan guru untuk meningkatkan mutu soal yang telah ditulis. Kegiatan ini merupakan proses pengumpulan, peringkasan, dan penggunaan informasi dari jawaban siswa untuk membuat keputusan tentang setiap penilaian.¹⁹ Tujuan penelaahan adalah untuk mengkaji dan menelaah setiap butir soal agar diperoleh soal yang bermutu sebelum soal digunakan. Di samping itu, tujuan analisis butir soal juga untuk membantu meningkatkan tes melalui revisi atau membuang soal yang tidak efektif, serta untuk mengetahui informasi diagnostik pada siswa apakah mereka sudah/belum memahami materi yang telah diajarkan.²⁰ Soal yang bermutu adalah soal yang dapat memberikan informasi setepat-tepatnya sesuai dengan tujuannya di antaranya dapat menentukan peserta didik mana yang sudah atau belum menguasai materi yang diajarkan guru.

Dalam melaksanakan analisis butir soal, para penulis soal dapat menganalisis secara kualitatif, dalam kaitan dengan isi dan bentuknya, dan kuantitatif dalam kaitan dengan ciri-ciri statistiknya²¹ atau prosedur peningkatan secara judgment dan prosedur peningkatan secara empirik.²² Analisis kualitatif mencakup pertimbangan validitas isi dan konstruk, sedangkan analisis kuantitatif mencakup pengukuran kesulitan butir soal dan

¹⁹ Nitko, Anthony J. (1996). *Educational Assessment of Students, Second Edition*. Ohio: Merrill an imprint of Prentice Hall Englewood Cliffs. 1996. p.308

²⁰ Aiken, Lewis R. *Psychological Testing and Assessment, (Eight Edition)*, Boston: Allyn and Bacon. 1994.p.63

²¹ Anastasi. Anne and Urbina, Susana. (1997). *Psicohological Testing. (Seventh Edition)*. New Jersey: Prentice-Hall, Inc. 1997. p.172

²² Popham, James W. (1995). *Classroom Assessment: What Teachers Need to Know*. Boston: Allyn and Bacon. 1995. P 195

diskriminasi soal yang termasuk validitas soal dan reliabilitasnya.

Jadi, ada dua cara yang dapat digunakan dalam penelaahan butir soal yaitu penelaahan soal secara kualitatif dan kuantitatif. Kedua teknik ini masing-masing memiliki keunggulan dan kelemahan. Oleh karena itu teknik terbaik adalah menggunakan keduanya (penggabungan). Kedua cara ini diuraikan secara rinci dalam buku ini.

b. Manfaat Soal yang Telah Ditelaah

Tujuan utama analisis butir soal dalam sebuah tes yang dibuat guru adalah untuk mengidentifikasi kekurangan-kekurangan dalam tes atau dalam pembelajaran.²³ Berdasarkan tujuan ini, maka kegiatan analisis butir soal memiliki banyak manfaat, di antaranya adalah: (1) dapat membantu para pengguna tes dalam evaluasi atas tes yang digunakan, (2) sangat relevan bagi penyusunan tes informal dan lokal seperti tes yang disiapkan guru untuk siswa di kelas, (3) mendukung penulisan butir soal yang efektif, (4) secara materi dapat memperbaiki tes di kelas, (5) meningkatkan validitas soal dan reliabilitas. Di samping itu, manfaat lainnya adalah: (1) menentukan apakah suatu fungsi butir soal sesuai dengan yang diharapkan, (2) memberi masukan kepada siswa tentang kemampuan dan sebagai dasar untuk bahan diskusi di kelas, (3) memberi masukan kepada guru tentang kesulitan siswa, (4) memberi masukan pada aspek tertentu untuk pengembangan kurikulum, (5) merevisi materi yang dinilai atau diukur, (6) meningkatkan keterampilan penulisan soal.

Linn dan Gronlund juga menambahkan tentang pelaksanaan kegiatan analisis butir soal yang biasanya didesain untuk menjawab pertanyaan-pertanyaan berikut ini; (1) Apakah fungsi soal sudah tepat? (2) Apakah soal ini memiliki tingkat kesukaran yang tepat? (3) Apakah soal bebas dari hal-hal yang tidak relevan? (4) Apakah pilihan jawabannya efektif? Lebih lanjut Linn dan Gronlund menyatakan bahwa kegunaan analisis butir soal bukan hanya terbatas

²³ Anastasi, Anne and Urbina, Susana. (1997). *Psychological Testing. (Seventh Edition)*. New Jersey: Prentice-Hall, Inc. 1997. p.184

untuk meningkatkan butir soal, tetapi ada beberapa hal, yaitu bahwa data analisis butir soal bermanfaat sebagai dasar: (1) diskusi kelas efisien tentang hasil tes, (2) untuk kerja remedial, (3) untuk peningkatan secara umum pembelajaran di kelas, dan (3) untuk peningkatan keterampilan pada konstruksi tes.²⁴

Berbagai uraian di atas menunjukkan bahwa analisis butir soal adalah: (1) untuk menentukan soal-soal yang cacat atau tidak berfungsi penggunaannya; (2) untuk meningkatkan butir soal melalui tiga komponen analisis yaitu tingkat kesukaran, daya pembeda, dan pengecoh soal, serta meningkatkan pembelajaran melalui ambiguitas soal dan keterampilan tertentu yang menyebabkan peserta didik sulit.



IAIN PURWOKERTO

²⁴ Linn, Robert L. and Gronlund, Norman E. (1995). *Measurement and Assessment in teaching (Seventh Edition)*. Ohio: Merrill, an imprint of Prentice Hall. 1995. p.318

BAB III

METODE PENELITIAN

Penelitian yang akan dilakukan adalah menyusun instrumen penilaian hasil belajar untuk mata pelajaran PAI di SD, 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 develop and validate educational products... 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 found in the field testing stage.(1983: 772)

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.

a. Jenis dan Pendekatan Penelitian

Jenis penelitian ini adalah penelitian lapangan (*field research*) karena data yang akan menjadi data utama akan diperoleh dari lapangan tempat penelitian. Pendekatan yang digunakan dalam penelitian ini adalah pendekatan kuantitatif, yaitu pendekatan yang menemukan pengetahuan menggunakan angka sebagai alat menemukan keterangan mengenai apa yang ingin kita ketahui. Kuantitatif dilihat dari jenis datanya adalah “penelitian yang datanya berupa angka-angka atau gejala yang diangkakan. Sedangkan metode penelitian yang digunakan dalam penelitian ini adalah metode survei, yaitu metode yang digunakan untuk mengumpulkan

informasi mengenai pemahaman, sikap dan perilaku, pengalaman, opini, dan karakteristik serta penerapan. Metode survei sering digunakan dalam jumlah banyak pada penelitian sosial dan penelitian lapangan lainnya. Dalam survei ini, informasi yang diinginkan dikumpulkan dari responden dengan menggunakan alat pengumpul data berupa angket dan atau instrumen tes.

b. Tempat dan Waktu Penelitian

Dalam mencari dan mengumpulkan data untuk menyusun laporan penelitian ini diambil tempat dan waktu penelitian, sebagai berikut:

1. Tempat Penelitian

Penelitian ini bertempat di SD di Purwokerto dan dilaksanakan pada peserta didik kelas V tahun ajaran 2015/2016.

2. Waktu Penelitian

Penelitian ini dilaksanakan selama kurang lebih 5 bulan mulai bulan Maret sampai Juli 2016 dengan menyesuaikan tahun pelajaran yang sedang berjalan.

c. Subyekdan Obyek Penelitian

1. Subyek Penelitian

Subyek penelitian dalam penelitian ini para guru PAI di SD Purwokerto. Sejumlah guru PAI yang ada peneliti mengambil beberapa sebagai subyek yang akan mewakili.

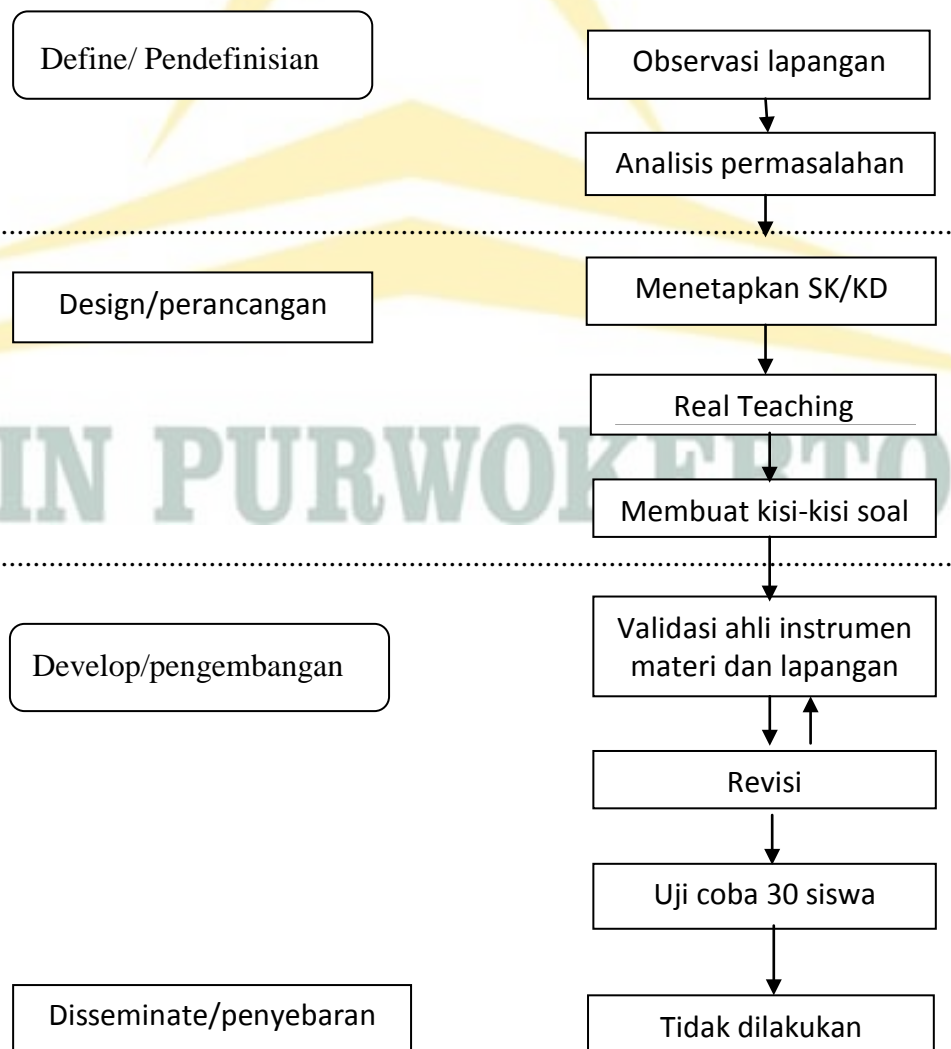
2. Obyek Penelitian

Obyek Penelitian dalam penelitian ini adalah instrumen evaluasi PAI. Khususnya adalah pada materi kelas V pada tahun pelajaran 2015/2016. Instrumen ini disusun oleh para guru PAI yang ada di wilayah Purwokerto.

d. Teknik Pengumpulan Data

Instrumen pengumpulan data dalam penelitian ini adalah tes, observasi dan interview. (Sugiyono, 2006). Tes adalah suatu alat atau prosedur yang

sistematis untuk mengukur suatu suasana aturan yang ditetapkan. Dalam penelitian ini digunakan tes untuk mengukur kemampuan siswa dalam mengerjakan tes penilaian hasil belajar pada mata pelajaran PAI di SD. Penilaian yang digunakan untuk mengukur kemampuan siswa dalam mengerjakan instrumen penilaian hasil belajar dengan menggunakan tes prestasi. Dalam tahap ini yang dilakukan juga penyusunan kisi-kisi penelaahan, perbaikan, perakitan kisi-kisi, penelaahan kisi-kisi dilakukan oleh pakar psikometri, pakar pendidikan, dosen, guru mapel PAI dalam *Focus Group Discuss* (FGD). 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 FGD. Langkah-langkah pengembangan sebagaimana bagan berikut:



e. Uji Coba Produk

Setelah uji ahli tentang instrumen penilaian hasil belajar mapel PAI di SD yang dilengkapi dengan kurikulum dan silabi PAI SD, peneliti merevisi masukan-masukan dari para ahli untuk melengkapi produk tersebut. Desain uji coba dapat dilihat pada tahapan 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 dikonsepsikan.
2. Subyek Uji Coba, subyek ujicoba dalam penelitian ini adalah siswa-siswa SD di Purwokerto.
3. Jenis Data, jenis data dalam penelitian ini adalah kuantitatif yaitu instrumen penilaian hasil belajar mata pelajaran PAI di SD, 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. Teknis Analisis Data

Metode penelitian ini akan menggunakan perhitungan analisis butir soal untuk melihat *difficulty level test* dan *discriminating power test*. Program Itecan bisa digunakan dalam proses validasi instrumen evaluasi dari sisi *difficulty level test* dan *discriminating power test* serta keberfungsian distraktor. Instrumen mata pelajaran fiqh akan di uji kualitasnya menggunakan program Itecan sehingga bisa diketahui mengenai kualitas dari instrumen yang digunakan untuk mengevaluasi peserta didik.

BAB IV

HASIL PENELITIAN DAN PEMBAHASAN

Peneliti mengadakan kunjungan ke beberapa SD di Purwokerto untuk mennggali data melalui observasi dan interview. Data yang peneliti akses adalah data proses penyusunan kisi-kisi sebagai awal dalam menyusun instrumen evaluasi.

1. Kendala guru dalam proses penyusunan instrumen penilaian hasil belajar PAI siswa SD di Purwokerto.

Proses penyusunan instrumen penilaian hasil belajar PAI di SD Purwokerto berdasarkan data dari hasil interview dengan salah satu guru adalah bahwa mereka masih banyak yang belum mengenai instrumen itu sendiri sebagai alat ukur untuk evaluasi. Sebagaimana statemen dari guru PAI SD di Purwokerto Bapak Isdi Atmanto, S.Ag berikut ini;

“Sebenarnya banyak dari para guru bahkan mungkin umumnya guru belum begitu memahami betul sebenarnya apa perbedaan antara; jenis, bentuk, model, dan instrument penilaian/evaluasi. Kalau menurut saya sendiri instrument evaluasi adalah alat penilaian seperti soal-soal test, lembar observasi, ceck list, dan sebagainya yang digunakan oleh guru untuk menilai sejauh mana kemampuan siswa menguasai materi pelajaran”.

Dari pernyataan tersebut di atas ternyata masih banyak guru yang belum memahami instrumen evaluasi, padahal semetinya guru harus sudah paham instrumen evaluasi karena sangat penting dalam proses evaluasi. Banyak wadah yang bisa dijadikan sebagai tempat untuk sharing mengenai instrumen evaluasi, biasanya para guru ada kegiatan KKG dalam hal ini PAI sangat mungkin untuk menjadi sarana sharing pengetahuan mengenai instrumen evaluasi. Namun demikian sarana ini juga belum semua guru memanfaatkan dengan maksimal, bahkan masih ada beberapa sekolah yang belum mengikuti kegiatan KKG PAI ini. Sebagaimana tutur Pak Isdi berikut ini;

“Penyusunan instrument bagi guru PAI di Sekolah Dasar khususnya pada sekolah kami tidak dilakukan bersama-sama dalam forum KKG maple PAI sehingga saya dan guru PAI lain di sini mempersiapkan sendiri penyusunan instrumen evaluasinya dengan mengacu kepada silabus dan RPP yang ada kemudian setelah dibuat instrumennya maka bisa digunakan untuk berkali-kali dengan menggunakan instrumen yang sama setiap tahun bisa dikatakan permanen begitu lah, paling jika dianggap perlu dilakukan peninjauan lagi dan dilakukan perbaikan dan penyesuaian seperlunya”.

Hal ini menunjukkan bahwa kesadaran untuk membahas bersamaapa yang menjadi permasalahan dalam pembelajaran belum berjalan dengan baik. Perlu ada pihak yang memfasilitasi agar kegiatan KKG PAI berjalan dengan baik dengan melibatkan semua guru PAI. Yang terpenting KKG PAI sebagai wahana untuk meningkatkan profesionalisme guru PAI bukan sekedar agenda bulanan biasa.

2. Pedoman penyusunan instrumen penilaian hasil belajar PAI siswa SD di Purwokerto.

Pedoman penyusunan instrumen penilaian PAI di SD Purwokerto berdasarkan informan yang dapat peneliti kemukakan dengan mendaarkan pada kurikulum sebagaimana dalam Standar Kompetensi, Kompetensi Dasar, indikator sebagaimana terjabarkan dalam silabus dan RPP PAI di SD. Mereka para guru dalam menyusun instrumen penilaian juga mendasarkan pada kemampuan siswa meskipun yang ini belum begitu diperhatikan sekali karena para guru PAI SD memiliki asumsi kemampuan mereka relatif sama. Sebagaimana pernyataan guru PAI di SD Purwokerto;

“Saya lebih dominan mempertimbangkan kesesuaian instrumen dengan materi yang sudah disampaikan kepada siswa, karena kalau tidak disesuaikan dengan materi maka siswa akan merasa kesulitan atas materi yang belum/tidak diberikan kepada mereka, selain itu kita juga tidak mengetahui apakah siswa sudah menguasai tujuan pembelajaran yang ada di SK-KD silabus atau belum. Sebenarnya pertimbangan lain adalah latar belakang siswa, tapi menurut saya karena materinya relatif sama di berbagai sekolah dasar yang bersumber dari silabus dan kurikulum yang sama maka saya kira pertimbangan siswa tidak terlalu diperhatikan karena kemampuan siswa juga relatif tidak ada perberbedaan”.

Pendapat ini diperkuat oleh guru PAI SD yang lain yaitu Ibu Rohmah Nurhidayah, S.Pd.I dia memaparkan mengenai pertimbangan dalam proses penyusunan instrumen evaluasi yaitu;

“Idealnya pertimbangannya didasarkan kepada jenis tujuan kompetensi yang ingin dicapai, aspek materi pembelajaran yang diajarkan, dan situasi kondisi latar belakang siswa yang diajar. Namun realitasnya yang terjadi adalah guru pada umumnya lebih dominan mempertimbangkan pada aspek materi yang sudah disampaikan kepada siswa, karena kalau tidak disesuaikan dengan materi maka siswa akan merasa kesulitan dan kita juga tidak mengetahui apakah siswa sudah menguasai tujuan pembelajaran yang ada di SK-KD-indikator silabus/ RPP atau belum”.

Pedoman dalam menyusun instrumen pada setiap mata pelajaran akan mendasarkan pada kurikulum dengan menyesuaikan mulai dengan mencermati SK kemudian KD yang dijabarkan dengan beberapa indikator sebagai rujukan dalam menyusun instrumen evaluasi. Guru-guru PAI di SD Purwokerto juga telah melakukan penyusunan instrumen dengan mempertimbangkan apa yang seharusnya.

3. Proses penyusunan instrumen penilaian hasil belajar PAI siswa SD di Purwokerto.

Proses penyusunan instrumen penilaian hasil belajar PAI di SD Purwokerto dengan mendasarkan pada informasi dari interview dengan beberapa guru yang mengampu mata pelajaran PAI adalah bahwa

“Secara urut proses penyusunan instrumen penilaiannya adalah diawali dengan melihat dan menganalisis SK-dan KD yang ada dalam silabus, kemudian SK dan KD tersebut dijabarkan menjadi rincian indikator kompetensi yang tercantum juga dalam RPP, kemudian dari indikator tersebut dipilih mana indikator yang penting untuk dijadikan sebagai bahan merumuskan instrumen penilaian baik penilaian kognitif, sikap maupun psikomotorik, demikian juga untuk penilaian kognitif berupa soal-soal tes mengacu kepada indikator-indikator tersebut”.

Juga pernyataan dari guru yang lain yang senada dengan pernyataan pertama adalah;

“Proses penyusunan instrument penilaian dimulai dengan melihat dan menganalisis Standar Kompetensi dan Kompetensi Dasar yang ada dalam silabus, kemudian Standar Kompetensi dan Kompetensi Dasar tersebut dijabarkan menjadi indikator-indikator kompetensi, kemudian dari keseluruhan indikator yang ada guru menentukan indikator-indikator yang dinilai penting dan representatif untuk dijadikan sebagai bahan merumuskan instrumen penilaian. Dari indikator-indikator tersebut dibuatlah kisi-kisi soal, terakhir kisi-kisi tersebut dijadikan acuan dalam menyusun instrumen tes maupun nontes”.

Artinya bahwa untuk sebagian besar guru PAI dalam proses penyusunan instrumen sudah lebih baik pemahamannya. Namun demikian masih banyak juga guru yang belum menggunakan tahapan tersebut dalam proses penyusunan instrumen evaluasi, di antara mereka ada yang masih mengambil dari lembar kerja siswa (LKS) sebagai jalan pintas. Ini adalah kesalahan besar yang guru lakukan karena telah keluar dari role yang seharusnya dilakukan dalam proses penyusunan instrumen. Namun demikian masih banyak guru yang mengambil jalan pintas ini. Beberapa guru menuturkan karena mereka sudah tidak memiliki waktu yang cukup untuk melakukan hal tersebut karena banyak beban yang harus diselesaikan, di antaranya beban administrasi yang semestinya menjadi tugas tenaga kependidikan. Tetapi untuk guru PAI di Purwokerto sudah menjalankan prosedur proses penyusunan instrumen evaluasi yang baik hal ini bisa lihat pada dokumen kisi-kisi yang dirumuskan sebagai pedoman penyusunan instrumen.

Ada enam langkah yang harus dilewati ketika kita akan menyusun instrumen evaluasi. Keenam langkah tersebut adalah:

a. Menentukan tujuan dalam mengadakan evaluasi.

Tujuan di sini berorientasi pada materi. Dan materi ini bergantung pada luasnya evaluasi yang dikehendaki. Seperti misalnya: “ingin mengetahui seberapa jauh siswa telah memahami sejarah pendidikan pada masa klasik”. Dengan demikian tujuan ini mengarah kepada Standar Kompetensi.

- b. Membatasi materi yang akan diteskan. Hal ini dilakukan agar dalam instrumen tes tidak terdapat materi-materi di luar tujuan tes. Pembatasan ini mengarah pada Kompetensi Dasar dari bab tertentu.
- c. Merumuskan Kompetensi Dasar. Sesuai dengan Kompetensi Dasar dari setiap pembahasan (dari tiap-tiap bahan).
- d. Menderetkan semua indikator dalam tabel persiapan yang juga memuat aspek tingkah laku yang terkandung dalam indikator.
- e. Menyusun tabel spesifikasi yang memuat pokok materi, aspek berpikir yang diukur beserta imbalan antara keduanya. Tabel ini disebut juga kisi-kisi, *blue print, lay-out*.
- f. Menuliskan butir-butir soal, didasarkan atas indikator yang telah dituliskan sebagai bentuk operasional dari KD. Kisi-kisi tersebut adalah sebagai berikut:

Kisi-kisi Penulisan Soal Ulangan Akhir Semester II Tahun Pelajaran 2015 / 2016

Satuan Pendidikan : Sekolah Dasar
 Kelas : V (lima)
 Mata Pelajaran : Pendidikan Agama Islam
 Tahun Pelajaran : (2015/2016)
 Kurikulum : KTSP / 2006
 Alokasi Waktu : 60 menit
 Jumlah soal : 50 soal
 A. Pilihan Ganda = 35 soal
 B. Isian = 10 soal
 C. Uraian = 5 soal

No	Standar Kompetensi	Kompetensi Dasar	Jumlah Soal	Materi	Indikator	Bentuk Soal	No Urut Soal
1	2	3	4	5	6	7	8
	6. Mengartikan Al Quran surat pendek	6.1 Membaca surat Al Ma'un dan Al Fiil 6.2	12	Surat Al Ma'un Surat Al fiil	Disajikan dalam kalimat siswa dapat menyebutkan jumlah ayat- ayat dalam surat Al Fiil. Di sajikan dalam ayat	PG	1

	pilihan	Mengartikan surat Al Ma'un dan Al Fiil		siswa dapat menyebutkan ayat surat Al Fiil.	PG	2
				Disajikan dalam ayat surat Al Fiil yang mengandung bacaan izhar.	PG	3
				Disajikan dalam kalimat siswa dapat menyebutkan lahir nabi Muhammad SAW	PG	4
				Disajikan dalam kalimat siswa dapat menyebutkan nama raja Yaman yang menyerang Ka'bah	PG	5
				Disajikan dalam kalimat siswa dapat menyebutkan jumlah ayat- ayat dalam surat Al Ma'un.	PG	6
				Disajikan dalam kalimat siswa dapat menyebutkan arti Al Ma'un.	PG	7
				Disajikan dalam ayat siswa dapat menyebutkan ayat surat Al Ma'un.	PG	8
				Disajikan dalam potongan ayat siswa dapat meneruskan ayat dalam surat Al Ma'un.	PG	9
				Disajikan dalam ayat surat Al Ma'un yang mengandung bacaan Ihfa.	PG	10
				Disajikan dalam kalimat siswa dapat mengartikan kata Al Fiil	PG	11
				Disajikan dalam potongan ayat siswa dapat meneruskan ayat dalam	PG	12

7. Mengenal Rasul-rasul Allah SWT	7.1 Menyebutkan nama-nama Rosul Allah SWT	11	Nama-nama Rosul Allah Rosul Ulul Azmi. Perbedaan antara nabi dan rosul	surat Al Fiil		
				Disajikan dalam kalimat siswa dapat menyebutkan pengertian Nabi.	PG	13
				Disajikan dengan kalimat siswa dapat menyebutkan pengertian Rasul	PG	14
				Disajikan dalam kalimat siswa dapat menyebutkan salah satu nabi Ulul Azmi	PG	15
				Disajikan dalam kalimat siswa dapat mengartikan sifat fatonah	PG	16
				Disajikan dalam kalimat siswa dapat menyebutkan arti sifat wajib bagi rosul	PG	17
				Disajikan dalam kalimat siswa dapat menyebutkan salah satu nama Nabi / Rasul	PG	18
				Disajikan dalam kalimat siswa dapat menyebutkan pengertian sifat mustahil bagi Rasul	PG	19
				Disajikan dalam kalimat siswa dapat mengartikan salah satu sifat mustahil (baladah)	PG	20
				Disajikan dengan kalimat siswa dapat menyebutkan mu'jizat nabi Ibrahim a.s	PG	21
8. Menceritakan kisah sahabat Nabi	8.1 Menceritakan kisah khalifah Abu Bakar	2	Kisah khalifah	Disajikan dalam kalimat siswa dapat mengartikan zaman Jahiliyah.	PG	22
				Disajikan dalam kalimat siswa dapat menyebutkan	PG	23

9. Membiasakan perilaku terpuji.	RA	1	Abu Bakar RA	nama anak nabi Nuh yang tidak taat pada orang tuanya.	PG	24			
	8.2 Menceritakan kisah Umar Bin khattab. RA		1	Abu Bakar RA			Disajikan dalam kalimat siswa dapat menyebutkan nama asli Abu Bakar.	PG	25
							Disajikan dalam kalimat siswa dapat menyebutkan salah satu khulafahurosyidin	PG	26
							Disajikan dalam kalimat siswa dapat menyebutkan tgl terjadinya Isra Mi'raj.	PG	27
							Disajikan dengan kalimat siswa dapat menyebutkan salah satu usaha yang dilakukan oleh Abu Bakar dalam menegakkan Islam.	PG	28
	9.1 Meneladani perilaku khalifah Abu Bakar RA		2	Kisah Umar Bin khattab. RA			Disajikan dalam kalimat siswa dapat mengartikan kata Al Kazab.	PG	29
							Disajikan dalam kalimat siswa dapat menyebutkan pembunuh khalifah Umar bin Khattab.	PG	30
	9.2 Meneladani perilaku Umar bin Khattab.RA		4	Keteladanan Abu Bakar RA			Disajikan dalam kalimat siswa dapat menyebutkan julukan Umar bin khattab karena kecerdasannya.	PG	31
							Keladanan Umar bin Khattab.RA	Disajikan dalam kalimat sederhana siswa dapat menyebutkan julukan yang diberikan oleh nabi Muhammad SAW kepada Abu Bakar.	PG
								Disajikan dengan kalimat sederhana siswa dapat	

					menyebutkan surat dalam Al Quran yang di baca oleh Fatimah sehingga Umar sadar dan masuk Islam.	PG	33
					Disajikan dalam kalimat siswa dapat menunjukan ayat Al Quran yang memerintahkan puasa		
					Disajikan dalam bentuk kalimat siswa dapat menyebutkan salah satu rukun puasa.	PG	34
					Disajikan potongan ayat siswa dapat melengkapi (surat al Ma'un ayat 2)	PG	35
10. Mengenal Puasa Wajib	10.1 Menyebutkan ketentuan – ketentuan puasa Ramadhan	3	Ketentuan – ketentuan puasa Ramadhan	Disajikan dengan kalimat siswa dapat menyebutkan lafal yang mengakhiri surat Al ma'un.	Isian		36
	10.2 Menyebutkan hikmah puasa		Hikmah puasa Ramadhan	Disajikan dalam kalimat siswa dapat meneruskan potongan ayat surat Al Fiil ayat 3	Isian		37
6. Mengartikan Al Qur'an Surat pendek pilihan	6.1 Membaca surat Al Ma'un dan Al Fiil	3	Surat Al Ma'un Surat Al fiil	Disajikan dalam kalimat siswa dapat menyebutkan orang yang mendustakan agama.	Isian		38
	6.2 Mengartikan suarat Al Ma'un dan Al			Disajikan dalam kalimat siswa dapat mengartikan salah satu sifat mustahil bagi rosul.	Isian		39
7. Mengenal Rasul-Rasul			Nama –	Disajikan dalam kalimat			40

	Allah SWT	Fiil	2	nama Rasul Allah. Nama – nama Rasul Ulul Azmi	siswa dapat menyebutkan mu'jizat nabi Muhammad SAW yang paling besar. Disajikan dalam kalimat siswa dapat menyebutkan kholifah yang pertama Disajikan dalam kalimat siswa dapat menyebutkan berapa lama Umar Khattab memerintah.	Isian	41
	8. Menceritakan kisah sahabat Nabi	7.1 Menyebutkan nama – nama Rasul Allah.	2	Kholifah Abu bakar RA	Disajikan dalam kalimat siswa dapat menyebutkan pengertian puasa	Isian	42
		7.2 Menyebutkan nama – nama Rasul Ulul Azmi.		Kholifah Umar bin Khattab ra	Disajikan dalam kalimat siswa dapat menyebutkan lamanya puasa di bulan Ramadhan.	Isian	43
	10 Mengenal Puasa wajib	8.1 Menceritakan kisah Kholifah Abu bakar RA	3	Ketentuan – ketentuan puasa Ramadhan	Disajikan dalam kalimat siswa dapat menyebutkan hukum puasa nazar. Disajikan dalam kalimat siswa dapat mengartikan salah satu ayat dalam surat Al Ma'un.	Isian	44
		8.2 Menceritakan kisah Umar bin Khattab RA			Disajikan dalam kalimat siswa dapat menyebutkan isi kandungan suarat Al Fiil	Isian	45
	6. Mengartikan Al Qur'an surat pendek pilihan	10.1 Menyebutkan Ketentuan – ketentuan puasa Ramadhan	2	Surat Al Ma'un Surat Al fiil	Disajikan dalam kalimat siswa dapat menyebutkan 3 Rasul ulul azmi. Disajikan dalam kalimat siswa dapat menyebutkan 3 khulafahurosidin	Uraian	46
	7. Mengenal Rasul-rasul Allah SWT.		2	Nama – nama Rasul Ulul Azmi.	Disajikan dalam kalimat siswa dapat menyebutkan 3 puasa sunah.	Uraian	47

	8. Menceritakan kisah sahabat Nabi	6.2 Mengartikan surat Al Ma'un dan Al Fil		Kholifah Abu bakar r.a. dan Umar bin Khattab r.a.		Uraian	48
	10. Mengenal Puasa Wajib	7.2 Menyebutkan nama – nama Rasul Ulul Azmi.	1	Ketentuan – ketentuan puasa Ramadha		Uraian	49
		8.1 Menceritakan kisah Kholifah Abu bakar r.a.				Uraian	50
		8.2 Menceritakan kisah kholifah Umar bin Khotob r.a					
		10.1 Menyebutkan Ketentuan – ketentuan puasa Ramadhan					

4. Kualitas instrumen penilaian hasil belajar PAI siswa SD di Purwokerto.

Dengan melihat kisi-kisi dan soal yang tersusun menurut hemat peneliti masih banyak yang harus dibenahi dalam rangka untuk mendapatkan soal yang berkualitas. Para guru menurut informan guru PAI di Purwokerto tidak pernah melakukan uji coba instrumen. Bahkan mereka belum tahu seperti apa dan bagaimana untuk melakukan uji coba instrumen evaluasi. Mereka tidak mau merepotkan diri dengan melakukan uji coba instrumen, karena selama ini sudah dianggap cukup. Mereka menyadari bahwa instrumen yang berkualitas memang memerlukan adanya ujicoba namun karena belum ada pengarah dari pengawas sehingga mereka sudah merasa cukup sebagaimana pernyataan guru PAI di Purwokerto berikut;

“Uji kualitas dalam penyusunan instrumen kok sepertinya belum lazim dilakukan oleh para guru khususnya di tingkat sekolah dasar. Para pengawas dan kepala sekolah juga selama ini tidak menekankan hal tersebut, karena untuk menyusun instrumen saja ternyata sudah cukup repot dan memerlukan waktu yang tidak singkat. Sehingga para guru khususnya saya jarang atau belum pernah melakukannya, mungkin nanti jika sudah ada ketentuan seperti itu dan mendapat pelatihan penerapannya saya kira banyak guru yang mau melakukannya”.

Juga pendapat guru yang lain yang menyatakan bahwa kualitas instrumen dengan mengujicobakan adalah sebagai berikut;

“Nah, saya yakin untuk melakukan uji kualitas umumnya para guru menurut saya termasuk saya yang mengajar PAI betul-betul tidak memiliki waktu untuk melakukannya karena kesibukan guru yang memiliki banyak pekerjaan teknis administratif yang sangat menyibukkan di samping kewajiban mengajar dengan jam mengajar yang tinggi, sehingga tidak cukup waktu untuk melakukan uji kualitas tersebut. Selain itu faktor lainnya adalah tidak adanya pembinaan dan tuntutan khusus bagi guru untuk melakukannya, sehingga selain alasan waktu juga alasan keterbatasan kemampuan untuk melakukannya yang belum memadai”.

Informasi tersebut memberikan informasi bahwa mereka masih belum terbiasa dengan menyusun instrumen yang kemudian diujicobakan meskipun mereka memahami bahwa itu adalah tahapan yang mestinya dilalui untuk mendapatkan instrumen yang berkualitas.

Secara umum untuk soal yang disusun oleh guru PAI di SD Purwokerto dengan mengkroscek antara kisi-kisi dengan soal sudah cukup sesuai artinya jika ini bisa diartikan berkualitas maka soal sudah cukup baik. Hanya jika kita mendasarkan pada prinsip bagaimana untuk mendapatkan soal yang berkualitas tidak cukup sampai di situ. Karena instrumen yang berkualitas adalah soal yang dapat memberikan informasi setepat-tepatnya sesuai dengan tujuannya di antaranya dapat menentukan peserta didik mana yang sudah atau belum menguasai materi yang diajarkan guru. Kemudian untuk dipahami para guru bahwa teknik untuk mendapatkan instrumen yang berkualitas bisa menggunakan teknik moderator yaitu teknik berdiskusi dengan satu orang sebagai penengah, setiap butir soal didiskusikan secara bersama dengan beberapa ahli; guru yang mengajarkan materi, ahli materi, penyusun/pengembang kurikulum, ahli penilaian, ahli bahasa, berlatar belakang psikologi. Teknik ini sangat baik karena setiap butir soal dilihat secara bersama-sama berdasarkan kaidah penulisannya. Di samping itu, para penelaah dipersilakan mengomentari/ memperbaiki berdasarkan ilmu yang dimilikinya. Setiap komentar/masukan dari peserta diskusi dicatat oleh notulis. Setiap butir soal dapat dituntaskan secara bersama-sama, perbaikannya seperti apa.

Teknik yang lain yang bisa digunakan adalah teknik panel merupakan suatu teknik menelaah butir soal yang setiap butir soalnya ditelaah berdasarkan kaidah penulisan butir soal, yaitu ditelaah dari segi materi, konstruksi, bahasa/budaya, kebenaran kunci jawaban/pedoman penskorannya yang dilakukan oleh beberapa penelaah. Caranya adalah beberapa penelaah diberikan: butir-butir soal yang akan ditelaah, format penelaahan, dan pedoman penilaian/ penelaahannya. Pada tahap awal para penelaah diberikan pengarahan, kemudian tahap berikutnya para penelaah berkerja sendiri-sendiri di tempat yang tidak sama. Para penelaah dipersilakan memperbaiki langsung pada teks soal dan memberikan komentarnya serta memberikan nilai pada setiap butir soalnya yang kriterianya adalah: baik, diperbaiki, atau diganti.

Barangkali sebagai bahan perbaikan ke depan guru PAI dalam menyusun instrumen perlu mempertimbangkan ke dua teknik tersebut. Sehingga diharapkan akan mendapatkan instrumen evaluasi yang berkualitas. Dengan mendapatkan instrumen evaluasi yang berkualitas tentunya akan banyak berimplikasi kepada siswa. Implikasi yang dimaksud adalah akan bisa mengukur kemampuan siswa yang sebenarnya sehingga informasi yang diterima adalah informasi yang mendekati kebenaran. Dengan demikian guru akan tepat dalam memberikan tindakan selanjutnya baik dalam remediasi maupun dalam program pengayaan. Artinya bisa memberikan terapi yang semestinya kepada siswa jika instrumen yang digunakan dalam evaluasi adalah instrumen yang berkualitas.



IAIN PURWOKERTO

BAB V

KESIMPULAN DAN SARAN

A. Kesimpulan

Berdasarkan data yang diperoleh dan dianalisis dalam bab sebelumnya maka kesimpulan dari penelitian ini adalah:

1. Kendala guru dalam proses penyusunan instrumen penilaian hasil belajar PAI siswa SD di Purwokerto adalah karena masih banyak guru PAI yang belum memahami instrumen itu sendiri. Kegiatan KKG PAI juga tidak diikuti oleh semua guru PAI sehingga saran dimana guru dapat sharing belum dimanfaatkan semaksimal mungkin
2. Pedoman penyusunan instrumen penilaian hasil belajar PAI siswa SD di Purwokerto dengan menyesuaikan mulai dengan mencermati SK kemudian KD yang dijabarkan dengan beberapa indikator sebagai rujukan dalam menyusun instrumen evaluasi
3. Proses penyusunan instrumen penilaian hasil belajar PAI siswa SD di Purwokerto diawali dengan menganalisis SKKD yang ada dalam silabus, kemudian SKKD tersebut dijabarkan menjadi rincian indikator kompetensi yang tercantum juga dalam RPP. Tetapi masih belum memperhatikan enam langkah dalam menyusun instrumen evaluasi, bahkan masih ada beberapa guru yang mendasarkan pada LKS dalam mendapatkan instrumen evaluasi.
4. Kualitas instrumen penilaian hasil belajar PAI siswa SD di Purwokerto belum terermin karena instrumen setelah disusun belum pernah dilakukan uji coba, hanya sebatas berkualitas jika dilihat dari keseuaian antara SKKD dengan Indikator sebagai rujukan menyusun instrumen.

5. Implikasi instrumen tersebut belum begitu mampu untuk memberikan informasi kemampuan yang sebenarnya atas siswa, hal ini menjadikan guru tidak mudah untuk melakukan kegiatan berikutnya setelah evaluasi.

B. Saran-saran

1. Ada kegiatan pengayaan bagi guru untuk lebih bisa memahami instrumen evaluasi.
2. Penyusunan instrumen perlu memahami pedoman yang digariskan dalam evaluasi.
3. Proses penyusunan instrumen evaluasi perlu ada tahapan yang mengerahkan seperti pada enam tahapan dan perlu menyusun tabel spesifikasi.
4. Perlu ada uji coba instrumen agar mendapatkan instrumen yang berkualitas



IAIN PURWOKERTO

Daftar Pustaka

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.

Anderson, L., & Krathwohl, D. (2001). *A taxonomy for learning, teaching and assessing: A revision of Bloom's taxonomy of educational objectives*. New York: Longman.

Borg, Walter.R and Gall, Meredith, D. (1983). *Educational research in introduction*. New York: Longman Inc

Danim, Sudarwan, (2007). *Metode Penelitian untuk Ilmu-ilmu Perilaku*, Jakarta: Bumi Aksara.

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*.

Djemari dkk. (1999b: 79). *Survey kegiatan guru dalam melakukan penilaian di kelas*. Laporan Penelitian. Yogyakarta: IKIP Yogyakarta.

Purwanti, Endang, (2008). *Dimensi-dimensi Riset Ilmiah*

Gronlund, N.E. (1976). *Measurement and evaluation in teaching*. New York: Macmillan.

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

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 penilaian, Nomor 7, Tahun VI, 2004*

Mardapi, Djemari, Program Pascasarjana UNY. (2002). *Kumpulan seminar dan loka karya*. Yogyakarta: Program pascasarjana.

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

Matondang, Zulkifli, jurnal tabularasa PPs UNIMED Vol.6 No.1, Juni 2009

Nield, Anthony F dan Wintre, Maxine Galander. Multiple choice questions with an option to comment : student attitudes and use. *Teaching of psychology*. Vol.13. No.4, 1986.

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>)

Raynolds, Cecil R. dkk ((2010). *Measurement and assessment in education*. Pearson Education. Upper River New Jersey

Sax, Gilbert. (1980). *Principal of education and psychological measurement and evaluation*. Belmont, California. Wadsworth Publishing Company.

Scouller, Karen. (2000). *The influence of assessment on student learning*. Paper

Sugiyono. (2006). *Metode Penelitian Kuantitatif, Kualitatif dan R & D*. Bandung: Alfabeta.

Sugiyono. (2013). *Metode Penelitian Pendidikan Pendekatan Kuantitatif, Kualitatif, dan R&D*, Bandung: Alfabeta

Umar, Jahja. (1985). Berbagai permasalahan penggunaan bentuk soal uraian dan pilihan ganda dalam ujian. *Buletin Pengujian dan Penilaian*, 6-10.

Wahyuningsih, Hepi, *Jurnal Psikologi Universitas Diponegoro* Vol.3 No. 1, Juni 2009

Wainer, Howard & Henry I. Braun. (1990). *Tes validity*. New Jersey: Hillsdale 07642.

Yusrizal, *jurnal tabularasa PPs UNIMED* Vol.5 No.1, Juni 2008



IAIN PURWOKERTO