BAB IV

HASIL PENELITIAN DAN PEMBAHASAN

Penelitian ini merupakan penelitian pengembangan yang bertujuan untuk menghasilkan suatu produk di bidang pendidikan. Produk yang dihasilkan adalah instrumen penilaian tes tertulis bentuk uraian non objektif untuk pembelajaran agama Islam berbasis masalah pada materi Fiqh.

Prosedur pengembangan yang digunakan dalam penelitian ini mengacu pada prosedur pengembangan instrumen tes yang dikemukakan oleh Djemari Mardapi.[1] langkah-langkah yang dipakai yakni: (1) menyusun spesifikasi tes (2) menulis soal tes (3) menelaah soal tes (4) memperbaiki tes (5) melakukan uji coba tes (6) menganalisis butir soal tes (7) menafsirkan hasil tes. Ketujuh langkah tersebut dibedakan menjadi dua tahap, yaitu tahap perancangan dan tahap uji coba.

A. Data Proses Pengembangan Instrumen Penilaian Hasil Belajar.

Data mengenai proses pengembangan instrumen hasil belajar didapatkan dari tahap perancangan. Tahapan ini terdiri dari empat langkah yaitu: menyusun spesifikasi tes, menulis soal tes, menelaah soal tes, dan memperbaiki tes. Tahap perancangan dilakukan dari tanggal 18 Agustus sampai 7 September, dengan rincian sebagai berikut:

1. Menyusun Spesifikasi Tes

Spesifikasi tes berisi uraian yang menunjukan keseluruhan karakteristik yang harus dimiliki suatu tes. Penyusunan spesifikasi tes

[1] Djemari Mardapi, *Teknik Penyusunan Instrumen Tes Dan Non Tes*, (Yogyakarta : Mitra Cendekia, 2008) 88.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

mencakup kegiatan berikut: *pertama*, menentukan tujuan tes, Tujuan tes sangat penting karena setiap tujuan memiliki penekanan yang berbeda-beda. Dalam penilitian ini tujuan tes yang digunakan adalah tes formatif, tes tersebut bertujuan untuk mengetahui, sudah sejauh manakah peserta didik telah terbentuk setelah mereka mengikuti proses pembelajaran dalam jangka waktu tertentu. Tes formatif ini biasanya dilaksanakan di tengah-tengah perjalanan program pengajaran, yaitu dilaksanakan pada setiap kali satuan pelajaran atau sub pokok bahasan berakhir atau dapat diselesaikan. Di sekolah-sekolah tes formatif ini biasa dikenal dengan istilah ulangan harian.[2] Karena tujuan tes adalah untuk tes formatif atau ulangan harian maka soal yag dibuat adalah soal uraian dengan tingkat kesukaran sedang, yaitu berkisar di angka 0,3-0,7.

Kedua, menyusun kisi-kisi tes. Setelah proses penentuan tujuan tes, kegiatan berikutnya adalah menyusun kisi-kisi tes. Kisi-kisi yang disusun berdasarkan materi tentang ibadah puasa dengan KD "Menganalisis ketentuan ibadah puasa". Kisi-kisi tes disajikan dalam bentuk matriks yang berisi komponen: kompetensi dasar, indikator, teknik penilaian dan bentuk instrumen. Tujuan penyusunan kisi-kisi soal adalah untuk menentukan ruang lingkup dan sebagai petunjuk dalam menulis soal agar kegiatan penilaian yang disusun relevan dengan

[2] Dadan Rosana, *Modul Evaluasi UT Bab I*, (Yogyakarta, 2011) 72.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

kegiatan pembelajaran.[3] Selanjutnya memilih bentuk tes yang sesuai. (terlampir)

Ketiga, Memilih bentuk tes. Spesifikasi tes berfungsi sebagai petunjuk praktis bagi penyusun tes dalam merencanakan isi materi yang akan diujikan, bentuk tes dan panjang tes. Bentuk tes yang dikembangkan dalam penelitian ini adalah tes tertulis bentuk uraian non objektif. Pemilihan bentuk tes tersebut berdasarkan pertimbangan bahwa tes tersebut memiliki keunggulan untuk mengukur tes hasil belajar yang kompleks. Sehingga bisa digunakan untuk mengukur kemampuan berfikir tingkat tinggi peserta didik yang dihasilkan dari dampak pembelajaran berbasis masalah.

2. Menulis Soal Tes.

Setelah penyusunan kisi-kisi soal, langkah selanjutnya adalah menulis butir-butir soal tes. Banyaknya butir soal tes untuk setiap indikator minimal satu butir soal, yang disesuaikan dengan model pembelajaran berbasis masalah. Tingkatan kognitif yang sesuai dengan pembelajaran berbasis masalah adalah Taksonomi Bloom edisi revisi level menganalisis, mengevaluasi, dan mengkreasi (C4, C5, C6). Namun, bukan berarti tujuan pembelajaran dengan level lebih rendah dianggap sepele dan harus ditinggalkan.[4]

Kompetensi dasar yang dipilih dalam penelitian ini didasarkan pada Kurikulum 2013, yaitu kompetensi dasar yang berbunyi

[3] Kusaeri, *Acuan dan Teknik Penilaian Proses dan Hasil Belajar dalam Kurikulum 2013*, (Yogyakarta : Ar-Ruzz Media, 2014) 45.
[4] Kusaeri, *Acuan dan Teknik Penilaian* , 36.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

"Menganalisis ketentuan ibadah puasa." Setelah itu kompetensi dasar dijabarkan kedalam beberapa indikator sesuai dengan level tujuan pembelajaran yang terdapat dalam perjenjangan Taksonomi Bloom edisi revisi. Indikator dijabarkan dalam Tabel 4.1 sebagai berikut:

Tabel 4.1 Indikator

No	Indikator	level
3.3.1	Menyebutkan pengertian puasa beserta syarat, rukun, sunah, hal-hal yang makruh dan membatalkan puasa.	C1
3.3.2	Menyebutkan ketentuan puasa Ramadhan dan puasa sunnah	C1
3.3.3	Mengelompokkan antara syarat, rukun, dan sunnah puasa.	C2
3.3.4	Membandingkan antara ketentuan puasa Ramadhan dan puasa sunnah.	C4
3.3.5	Menemukan hal-hal yang membatalkan puasa dalam kehidupan sehari-hari.	C5
3.3.6	Memberi penilaian mengenai sempurna dan tidaknya puasa, berdasarkan ketentuan yang ada.	C5
3.3.7	Merumuskan amalan yang perlu dilakukan agar puasa yang dikerjakan lebih baik.	C6

Dalam penulisan butir soal, taksonomi yang diacu adalah Taksonomi Bloom edisi revisi. Taksonomi Bloom edisi revisi terdiri dari enam tingkatan berfikir yaitu C1, C2, C3, C4, C5, dan C6. Selain memperhatikan tingkatan berfikir, Taksonomi Bloom edisi revisi juga

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

tidak bisa dilepaskan dari dimensi pengetahuan yang meliputi pengetahuan faktual, konseptual, prosedural dan metakognitif. Dalam penulisan butir soal tidak menyertakan level berpikir tingkat menerapkan (C3), karena diakomodasi oleh KD 4.3 yaitu "Mensimulasikan tata cara melaksanakan puasa dengan instrumen penilaian non tes". Sehingga level tes yang dibuat adalah sebagai berikut:

a. Mengingat pengetahuan faktual.
b. Memahami pengetahuan konseptual.
c. Menganalisa pengetahuan prosedural.
d. Menganalisa pengetahuan konseptual.
e. Mengevaluasi pengetahuan prosedural.
f. Menciptakan pengetahuan metakognitif.

Berdasarkan penjabaran indikator tersebut, maka dihasilkan enam butir soal uraian non objektif yang sesuai dengan pembelajaran agama Islam berbasis masalah beserta dengan pedoman penskorannya (pedoman penskoran terlampir). Adapun butir soal yang dihasilkan adalah sebagai berikut:

1. Setelah mempelajari ketentuan puasa, jawablah pertanyaan berikut ini sesuai dengan apa yang telah kamu pelajari!
   a. Apa pengertian puasa menurut bahasa dan istilah?
   b. Apa syarat untuk menjalankan ibadah puasa?
   c. Apa rukun puasa?
   d. Hal apa saja yang membatalkan puasa?

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Soal nomor 1, merupakan kategori proses kognitif mengingat dan pengetahuan faktual. Dalam soal ini peserta didik diminta untuk mengingat pengetahuan faktual yang berupa pengertian, syarat, rukun dan hal-hal yang membatalkan puasa. Soal ini digunakan ntuk mengukur ketercapaian indikator 3.3.1.

2. Di bulan Ramadhan Alif menjalani ibadah puasa. Pada malam hari ia berniat puasa kemudian makan sahur, dia juga menahan lapar dan memperbanyak bacaan al-Qur'an di siang hari. Ketika waktu berbuka tiba dia berbuka dengan memakan makanan yang manis.
   a. Bagaimana puasa Alif jika dia tidak berniat di malam hari? Jelaskan beserta alasannya!
   b. Bagaimana puasa Alif jika ia tidak makan sahur ? jelaskan beserta alasannya!

   Soal nomor 2, merupakan kategori proses kognitif memahami dan pengetahuan konseptual. Dalam soal ini peserta didik diminta untuk memahami pengetahuan konseptual tentang kategorisasi syarat, rukun, dan sunah puasa. Soal ini digunakan ntuk mengukur ketercapaian indikator 3.3.3.

3. Deskripsikan ketentuan puasa Ramadhan dan puasa sunnah, kemudian bandingkan dengan menyajikan perbedaan dan persamaan antara keduanya !

   Soal nomor 3, merupakan kategori proses kognitif menganalisa dan pengetahuan konseptua. Dalam soal ini peserta

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

didik diminta untuk menganalisa pengetahuan prosedural tentang ketentuan puasa ramadhan dan puasa sunah. Soal ini digunakan ntuk mengukur ketercapaian indikator 3.3.2 dan 3.3.4.

4. Berdasarkan pengalamanmu ketika bulan Ramadhan, apa saja hal-hal yang membatalkan puasa yang banyak kamu temukan di sekitarmu? Jelaskan beserta alasannya?

   Soal nomor 4, merupakan kategori proses kognitif menganalisa dan pengetahuan konseptual. Dalam soal ini peserta didik diminta untuk menemukan pengetahuan konseptual tentang hal-hal yang membetalkan puasa. Soal ini digunakan ntuk mengukur ketercapaian indikator 3.3.5.

5. Miftah dan arief adalah murid kelas VIII. Pada saat bulan Ramadhan mereka pergi ke Sidoarjo untuk membeli peralatan sekolah, di tengah jalan mereka hampir ditabrak oleh seorang tukang becak. Miftah dan arief kemudian marah-marah dan mengucapkan kata kotor kepada tukang becak tersebut.

   c. Bagaimana pendapatmu mengenai puasa Miftah dan Arief ? Jelaskan alasannya!
   d. Apa yang seharusnya dilakukan oleh Miftah dan Arief ketika kejadian tersebut?

   Soal nomor 5, merupakan kategori proses kognitif mengevaluasi dan pengetahuan prosedural. Dalam soal ini peserta didik diminta untuk mengevaluasi pengetahuan prosedural tentang

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

kesempurnaan berpuasa. Soal ini digunakan ntuk mengukur ketercapaian indikator 3.3.6.

6. Susunlah jadwal kegiatan di bulan Ramadhan, agar puasa yang kamu jalani bisa menjadi lebih sempurna. Perhatikan amalan sunnah ketika menjalankan ibadah puasa

Soal nomor 6, merupakan kategori proses kognitif mencipta dan pengetahuan prosedural. Dalam soal ini peserta didik diminta untuk menciptakan pengetahuan prosedural tentang bagaimana berpuasa dengan sempurna. Soal ini digunakan ntuk mengukur ketercapaian indikator 3.3.7.

3. Menelaah Soal Tes.

Salah satu tahapan untuk menghasilkan tes yang baik adalah dengan melakukan penelaahan tes. Tes yang telah disusun diserahkan pada ahli untuk ditelaah. Penelaahan tes dilakukan melalui teknik panel yaitu dengan cara beberapa penelaah menelaah tes ditempat terpisah, sehingga menghasilkan perbaikan dan komentar terhadap tes yang ditelaah.[5] Kegiatan tersebut melibatkan dua orang pakar pendidikan agama Islam. Kegiatan penelaahan tes secara panel dilakukan pada tanggal 29 agustus 2016 di ruang guru MTSN 4 Sidoarjo, dengan peserta Ahsan Bisri, S,Ag dan Muh. Ali Mashudi, S.Pd.I.

Kedua pakar disamakan persepsinya mengenai instrumen yang dikembangkan. Penelaahan butir tes didahului dengan penetapan level tes

[5] Kusaeri, *Acuan dan Teknik Penilaian*, 103

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

berdasarkan perjenjangan taksonomi Bloom edisi revisi dan kesesuaian antara instrumen tes tertulis bentuk uraian non objektif dengan pembelajaran agama Islam berbasis masalah. Para pakar setuju dengan model tes uraian yang diajukan. Secara lisan pakar menyatakan bahwa " tes uraian yang disusun bisa digunakan sebagai alat ukur yang sesuai dengan pembelajaran berbasis masalah."

Kemudian para pakar diminta untuk menelaah soal di tempat terpisah terhadap instrumen yang disusun beserta kisi-kisi dan pedoman penskorannya dengan menggunakan lembar validasi yang telah disediakan. Adapun hasil dari telaah para pakar berdasarkan lembar validasi, disajikan pada Tabel 4.2

Tabel 4.2 Hasil Validasi Instrumen

Aspek	Validator		Rata-Rata
	1	2	
Materi	3,78	3,75	3,76
Konstruksi	3,43	3,73	3,58
Bahasa	3,58	3,44	3,51
Rata Rata Total Validitas			3,62

Berdasarkan hasil validasi instrumen penilaian tes tertulis bentuk uraian non objektif untuk pembelajaran agama Islam berbasis masalah pada Tabel 4.2, instrumen yang telah disusun dinyatakan valid. Adapun indikatornya adalah rata-rata total validitas dari instrumen yang disusun mencapai angka 3,6. Jika dibandingkan dengan kriteria total kevalidan

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

maka instrumen yang telah disusun berada pada kategori sangat valid. Secara visual hasil validasi ahli disajikan dalam Diagram 4.1.

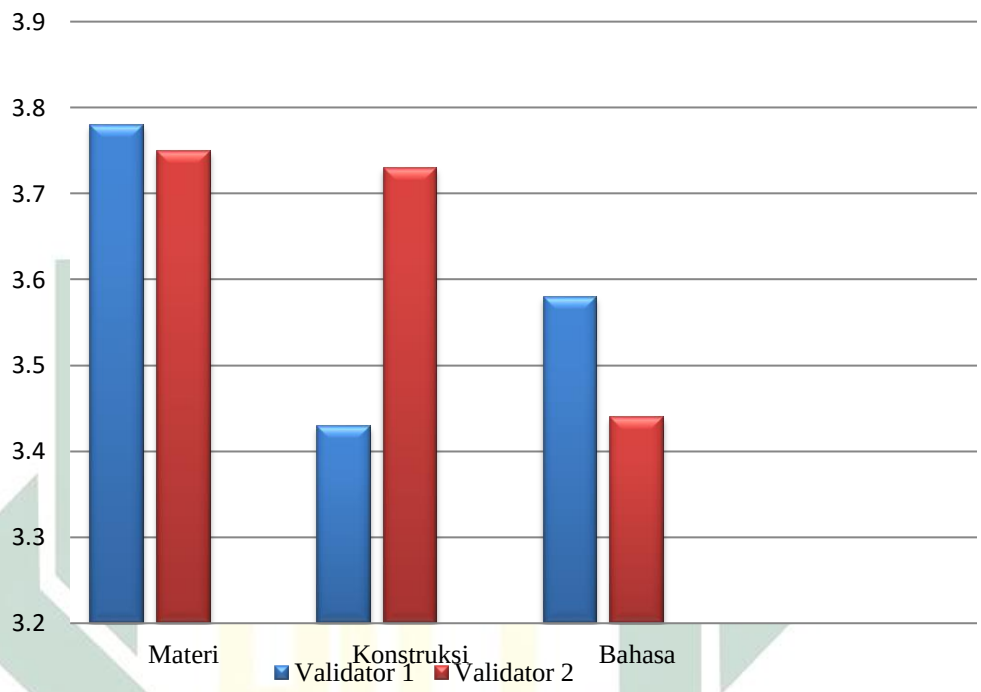

Diagram 4.1 Hasil Validasi

Dari diagram batang diatas diketahui bahwa tidak terjadi perbedaan penilaian yang signifikan antar kedua validator. Pada aspek materi, selisih skor antara validator pertama dan kedua adalah 0,03. Selisih skor aspek konstruksi adalah 0,3. Selisih skor aspek bahasa adalah 0,14. Selisih terbesar antara kedua validator adalah pada aspek konstruksi yakni 0,3. Walaupun perbedaan keduanya cukup besar, jika dibandingkan dengan tabel interpretasi rata-rata total validitas, skor keduanya berada pada kategori sangat valid. Ini artinya instrumen yang dikembangkan valid secara materi, konstruksi dan bahasa. Sehingga dapat digunakan

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

sebagai alat penilaian tes tertulis bentuk uraian objektif untuk pembelajaran agama Islam berbasis masalah .

Selain memberi penilaian menggunakan lembar validasi terhadap instrumen yang dikembangkan, para pakar juga dipersilahkan untuk memberikan beberapa masukan dan saran guna menghasilkan instrumen yang lebih baik. Masukan dan saran pakar melalui teknik panel dijadikan acuan untuk merevisi instrumen tes tertulis bentuk uraian objektif. Rincian masukan dari para pakar disajikan pada tahap perbaikan tes.

4. Memperbaiki Tes.

Setelah proses validasi oleh para pakar, maka selanjutnya instrumen direvisi berdasarkan saran yang telah diberikan. Hasil revisi tersebut diuraikan sebagai berikut:

a. Butir tes pada indikator 3.3.1 (menyebutkan pengertian puasa beserta syarat, rukun, sunah, dan hal-hal yang mebatalkan puasa). Indikator tersebut terdiri dari empat butir soal. Berikut redaksi soal pada indikator 3.3.1:

"*Setelah mempelajari ketentuan puasa, jawablah pertanyaan berikut ini sesuai dengan apa yang telah kamu pelajari!*

1. *Apa pengertian puasa menurut bahasa dan istilah?*
2. *Apa syarat untuk menjalankan ibadah puasa?*
3. *Apa rukun puasa?*
4. *Hal apa saja yang membatalkan puasa?"*

Adapun saran menurut validator untuk butir soal tersebut adalah:" kalimat tanya/perintah yang digunakan kurang jelas sehingga menimbulkan penafsiran ganda". Berdasarkan saran tersebut maka

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

harus dilakukan revisi terhadap soal. Setelah direvisi redaksi soal menjadi:

"*Setelah mempelajari ketentuan puasa, jawablah pertanyaan berikut ini sesuai dengan apa yang telah kamu pelajari!*

1. *Apa pengertian puasa menurut bahasa dan istilah?*
2. *Sebutkan syarat-syarat untuk menjalankan ibadah puasa?*
3. *Sebutkan rukun-rukun puasa?*
4. *Apa saja hal-hal yang membatalkan puasa?"*

b. Indikator 3.3.2 yang berbunyi *"menyebutkan ketentuan puasa Ramadhan dan puasa sunnah"* hampir sama dengan indikator 3.3.4 yang berbunyi *" membandingkan antara ketentuan puasa ramadhan dan puasa sunnah"*. Kedua indikator tersebut bisa dijadikan satu, Sehingga indikator ini diputuskan untuk dibuang dan mengakibatkan banyaknya indikator dalam kisi-kisi soal berkurang.

c. Butir tes pada indikator 3.3.3 (mengelompokkan antara syarat, rukun, dan sunnah puasa). Indikator ini terdiri dari dua butir soal. Berikut redaksi soal pada indikator 3.3.3:

"Di bulan Ramadhan Alif menjalani ibadah puasa. Pada malam hari ia berniat puasa kemudiann makan sahur, dia juga menahan lapar dan memperbanyak bacaan al-Qur'an di siang hari. Ketika waktu berbuka tiba dia berbuka dengan memakan makanan yang manis.

1. *Bagaimana puasa Alif jika dia tidak berniat di malam hari? Jelaskan beserta alasannya!*
2. *Bagaimana puasa Alif jika ia tidak makan sahur ? jelaskan beserta alasannya!"*

Saran dari validator untuk butir soal tersebut adalah:" pertanyaan dengan menggunakan kata bagaimana kurang tepat untuk digunakan

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

sebagai pengandaian, karena mengandung jawaban yang tidak pasti dan aspek yang diukur kurang lengkap, syarat puasa belum tercantum". Berdasarkan saran tersebut maka harus dilakukan revisi terhadap soal. Setelah direvisi redaksi butir menjadi:

"Di bulan Ramadhan Alif menjalani ibadah puasa. Pada malam hari ia berniat puasa kemudiann makan sahur, dia juga menahan lapar dan memperbanyak bacaan al-Qur'an di siang hari. Ketika waktu berbuka tiba dia berbuka dengan memakan makanan yang manis.

1. *Apakah puasa Alif sah jika ia belum baligh ! jelaskan alasannya!*
2. *Apakah puasa Alif sah seandainya dia tidak berniat di malam hari? Jelaskan alasannya!*
3. *Apakah puasa Alif sah seandainya ia tidak makan sahur ? jelaskan alasannya!"*

d. Butir tes pada indikator 3.3.4 (membandingkan ketentuan puasa Ramadhan dan puasa sunnah). Indikator ini kurang tepat, karena ketentuan puasa masih terlalu umum sehingga direvisi menjadi *"membandingkan tata cara puasa Ramadhan dan puasa sunnah."* Indikator tersebut terdiri dari satu butir soal. Berikut redaksi soal pada indikator 3.3.4:

"*Deskripsikan ketentuan puasa Ramadhan dan puasa sunnah, kemudian bandingkan dengan menyajikan perbedaan dan persamaan antara keduanya !"*

Saran dari validator untuk butir soal tersebut adalah:" kata deskripsikan kurang cocok bagi peserta didik jenjang MTS, kata yang tepat adalah jelaskan". Berdasarkan saran tersebut maka harus dilakukan revisi terhadap soal. Setelah direvisi redaksi butir menjadi:

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

"Jelaskan tata cara puasa Ramadhan dan puasa sunnah, kemudian carilah perbedaan dan persamaan antara keduanya !"

e. Butir tes pada indikator 3.3.5 (menemukan hal-hal yang membatalkan puasa dalam kehidupan sehari-hari). Butir tes sudah bagus, menurut validator kata "di sekitarmu" sebaiknya dihilangkan, hal ini dikarenakan kata tersebut sudah dapat difahami tanpa disebutkan lagi. Sehingga redaksi soal untuk indikator 3.3.5 adalah sebagai berikut:

Berdasarkan pengalamanmu ketika bulan Ramadhan, apa saja hal-hal yang membatalkan puasa yang banyak kamu temukan? Jelaskan alasannya!

f. Butir tes pada indikator 3.3.6 (memberi penilaian mengenai sempurna dan tidaknya puasa berdasarkan ketentuan yang ada). Tidak ada revisi untuk butir tes pada indikator 3.3.6.

g. Butir tes pada indikator 3.3.7 (merumuskan amalan-amalan yang perlu dilakukan agar puasa menjadi lebih baik). Indikator tersebut terdiri dari satu butir soal, adapun redaksi dari soal tersebut adalah:

"Susunlah jadwal kegiatan di bulan Ramadhan, agar puasa yang kamu jalani bisa menjadi lebih sempurna. Perhatikan amalan sunnah ketika menjalankan ibadah puasa."

Validator menyarankan bahwa kalimat "perhatikan amalan sunnah" hendaknya diganti, karena bisa menuntun peserta didik pada jawaban. Berdasarkan saran tersebut maka soal direvisi menjadi:

"Susunlah jadwal kegiatan di bulan Ramadhan, agar puasa yang kamu jalani bisa menjadi lebih sempurna. Perhatikan hal-hal yang menjadikan puasamu lebih baik !"

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

B. Data Karakteristik Instrumen Penilaian Hasil Belajar.

Data karakteristik intrumen penilaia hasil belajar dapat diperoleh pada tahap uji coba. Tahapan ini terdiri dari tiga langkah yaitu: melakukan uji coba tes, menganalisis butir soal tes, dan menafsirkan hasil tes. Uji coba dilakukan di MTSN 4 Sidoarjo pada tanggal 8 September 2016.

Kegiatan uji coba diawali dengan pemilihan kelas yang akan dijadikan subyek coba. Kelas yang terpilih adalah kelas VIII B. Pemilihan kelas didasarkan pada pertimbangan bahwa kelas VIII B merupakan kelas dengan kriteria keagamaan. Sehingga uji coba yang dilakukan mampu menginterpretasikan hasil yang maksimal dan akurat.

Uji coba dalam penelitian ini dilakukan sekali atau disebut dengan *single test method*. Sebelum dilakukan uji coba, terlebih dahulu dilakukan uji keterbacaan oleh 5 orang peserta didik dan 1 orang guru Fiqh. Pelaksanaan kegiatan uji keterbacaan dimaksudkan untuk mengetahui "Apakah peserta didik mampu memahami maksud dari soal atau belum."

Hasil uji keterbacaan diperoleh kesimpulan bahwa soal nomor 4 perlu direvisi. Revisi tersebut adalah sebagai berikut:

Berdasarkan pengalamanmu ketika bulan Ramadhan, apa saja hal-hal yang membatalkan puasa yang banyak kamu temukan?jelaskan alasannya!

Direvisi menjadi

Berdasarkan pengalamanmu ketika bulan Ramadhan, apa saja hal-hal yang membatalkan puasa yang banyak kamu temukan?

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Redaksi tersebut direvisi dengan membuang kalimat "jelaskan alasannya." Dikarenakan pada saat uji keterbacaan peserta didik tidak bisa memahami dengan baik maksud dari perintah tersebut. Sehingga diputuskan untuk dihapus untuk menghindari kesalahan dalam memahami soal.

Setelah melakukan uji keterbacaan, langkah selanjutnya adalah melakukan uji coba tes. Hasil dari uji coba kemudian diskor dengan menggunakan rubrik penskoran yang telah dibuat dan divalidasi oleh pakar.

Setelah hasil uji coba tes diketahui, tes dianalisis agar bisa mendapatkan jawaban dari "Apakah tes layak digunakan atau tidak." Alasan diperlukannya analisis butir soal adalah: (1) untuk mengetahui kekuatan dan kelemahan soal (2) untuk menyediakan informasi tentang spesifikasi butir soal secara lengkap (3) untuk mengetahui masalah yang ada dalam butir soal.[6]

Analisis yang digunakan adalah teknik klasik. Menurut Kusaeri dan Suprananto, aspek yang diperhatikan dalam teori klasik adalah tingkat kesukaran soal, daya pembeda dan reliabilitas.[7]

1. **Tingkat Kesukaran Soal.**

   Tingkat kesukaran soal adalah peluang menjawab benar suatu soal pada tingkat tertentu, yang besarnya berkisar dari 0 sampai 1. Semakin besar indeks tingkat kesukaran suatu soal maka semakin mudah soal itu. Untuk mendapatkan hasil tingkat kesukaran soal, dapat

[6] Kunandar, *Penilaian Autentik Penilaian Hasil Belajar Peserta Didik Berdasarkan Kurikulum 2013,* (Jakarta : Raja Grafindo Persada, 2014) 238.
[7]Kusairi dan Suprananto, *Pengukuran dan Penilaian Pendidikan,* (Yogyakarta: Graha Ilmu, 2012) 173.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

dilakukan dengan langkah-langkah sebagai berikut:[8] (a), menghitung jumlah skor setiap butir yang diperoleh setiap peserta didik (b) menghitung total skor maksimal setiap butir yang seharusnya diperoleh peserta tes (c) membagi perhitungan langkah pertama dengan langkah kedua. Hasil dari analisis tingkat kesukaran soal disajikan dalam Tabel 4.3.

Tabel 4.3 Tingkat Kesukaran Soal

No Soal	Tingkat Kesukaran Soal
1	0,70
2	0,63
3	0,65
4	0,64
5	0,53
6	0,65

Dari tabel di atas, terlihat bahwa butir-butir soal didominasi oleh soal dengan tingkat kesukaran mendekati atau sama dengan 0,60. Butir-butir tes yang memiliki tingkat kesukaran mendekati atau sama dengan 0,60 adalah butir tes nomor 2, 3, 4 dan 6. Dengan hasil tersebut maka diketahui bahwa instrumen tes yag dikembangkan berada pada kategori sedang atau diterima.

[8]Kusaeri, *Acuan dan Teknik Penilaian*, 106.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Untuk memperjelas hasil analisis tingkat kesukaran soal, maka secara visual hasil perhitungan tingkat kesukaran soal untuk isntrumen yang dikembangkan disajikan dalam Diagram 4.2.

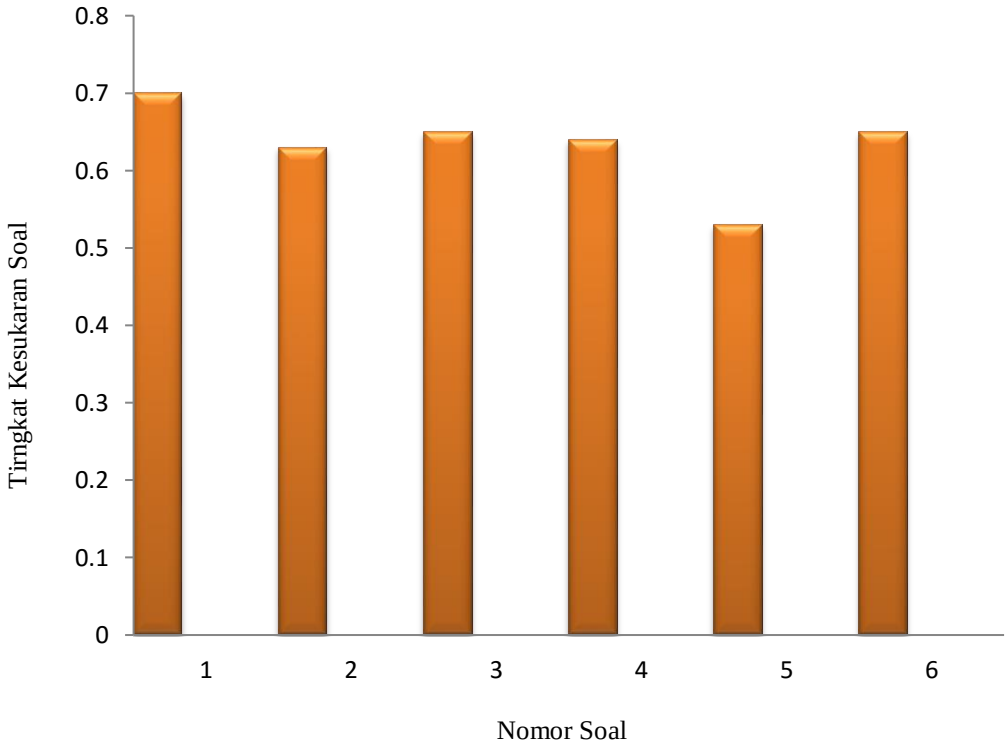

Diagram 4.2 Tingkat Kesukaran Soal

Berdasarkan diagram batang diatas, butir soal nomor 1 memiliki tingkat kesukaran soal sebesar 0,70. Hal ini menandakan bahwa soal nomor satu adalah soal paling mudah. Butir soal dengan indeks kesukaran soal terendah adalah butir soal nomor 5 dengan rentang sebesar 0,54. Hal ini menandakan bahwa soal nomor 5 adalah soal paling sulit diantara soal-soal yang dikembangkan. Secara keseluruhan, butir soal yang dikembangkan berada pada kategori sedang. Sehingga butir

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

soal yang dikembangkan diterima dan bisa digunakan untuk melakukan tes dengan tujuan formatif atau ulangan.

2. Daya Pembeda Soal.

Daya pembeda soal adalah kemampuan sebuah soal membedakan antara peserta didik yang pandai dan kurang. Untuk dapat menghitung daya pembeda soal dilakukan langkah-langkah sebagai berikut:[9] (a) mengurutkan peserta tes berdasarkan skor dari yang teratas sampai terendah (b) menentukan kelompok atas dengan mengambil sebanyak 27% peserta dengan urutan teratas dan kelompok bawah sebanyak 27% peserta dengan urutan terbawah (c) menghitung tingkat kesukaran soal kelompok atas dan bawah (d) menghitung selisih tingkat kesukaran soal kelompok atas dan bawah. Setelah menentukan kelompok atas dan bawah beserta tingkat kesukaran tiap-tiap kelompok, maka selanjutnya mencari daya pembeda soal. Hasil dari analisis daya pembeda soal disajikan dalam Tabel 4.4.

Tabel 4.4 Daya Pembeda Soal

Soal	Tingkat Kesukaran Soal Kelompok Atas	Tingkat Kesukaran Soal Kelompok Bawah	Daya Pembeda Soal (Dp)
1	0,84	0,60	0,24
2	0,80	0,44	0,35
3	0,81	0,52	0,29
4	0,86	0,56	0,30
5	0,71	0,40	0,31
6	0,87	0,51	0,36

[9] Kusaeri, *Acuan dan Teknik Penilaian*, 113.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Pada Tabel 4.4 diketahui bahwa butir-butir soal memiliki indeks daya pembeda soal yang bervariasi. Indeks daya pembeda terendah adalah soal nomor 1 dengan daya pembeda sebesar 0,24. Indeks daya pembeda tertinggi adalah soal nomor 6 dengan daya pembeda sebesar 0,36. Secara visual hasil analisis daya pembeda soal dapat dilihat dalam Diagram 4.3.

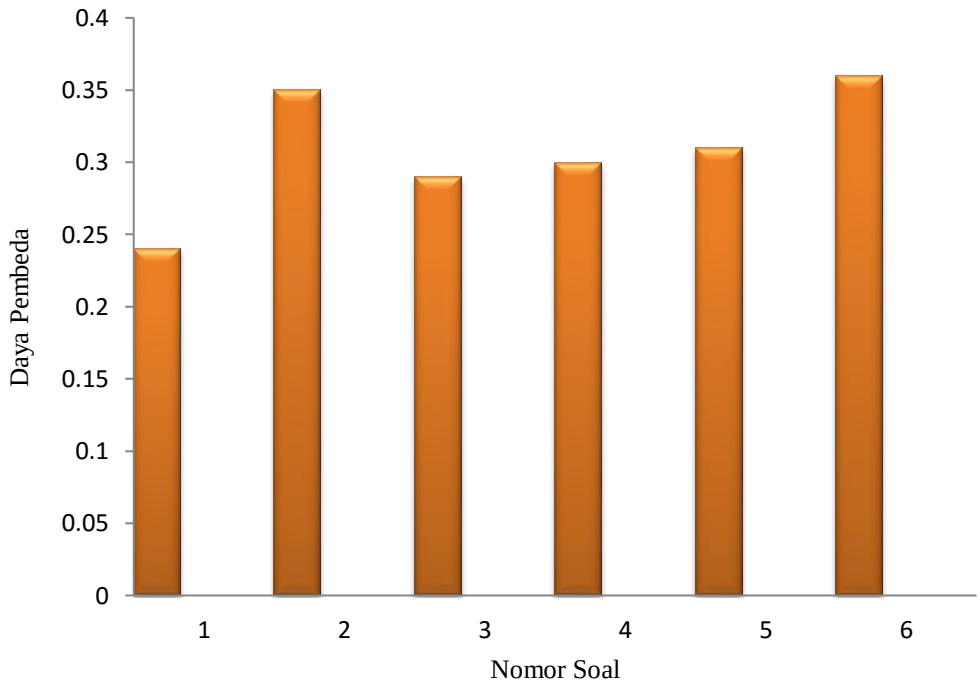

Diagram 4.3 Daya Pembeda

Pada Diagram 4.3 diketahui bahwa Indeks daya pembeda soal nomor 1 termasuk kategori kurang memuaskan. Artinya soal tidak dapat membedakan dengan baik antara peserta yang pandai dan kurang pandai. Hal ini dikarenakan soal nomor 1 memang dirancang agar mudah dikerjakan untuk mengurangi tingkat kepanikan peserta didik dalam

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

menjawab soal tes selanjutnya. Sehingga banyak peserta didik yang mampu menjawab, baik dari kelompok atas maupun bawah.

Indeks daya pembeda soal nomor 2, 3, 4, 5, dan 6 berada pada rentangan 0,3-0,4. Hal ini menandakan bahwa butir soal berfungsi sebagaimana mestinya. Secara keseluruhan butir tes yang disusun berada dalam kategori memuaskan dan diterima.

3. Reliabilitas Soal.

Reliabilitas soal pada instrumen tes ini menggunakan model belah dua *(split half methods).* Metode belah dua dilaksanakan dengan cara satu kali pelaksanaan tes. Metode ini sering disebut dengan *single test- single trial method*. Metode ini memungkinkan untuk mengestimasi reliabilitas tanpa harus menyelenggarakan tes dua kali. Untuk mengukur tingkat reliabilitas soal yang dikembangkan digunakan persamaan *Flanagan.*[10]

Untuk menghitung reliabilitas soal menggunakan persamaan *Flanagan,* dilaksanakan Langkah-langkah sebagai berikut: (a) memilah soal menjadi dua bagian (b) menentukan jumlah skor dan kuadrat total semua belahan (c) menentukan jumlah varian total (d) menentukan jumlah varian belahan pertama (e) menentukan jumlah varian belahan kedua (f) menentukan reliabilitas dengan memasukkan angka-angka yang diperoleh pada langkah ketiga sampai kelima kedalam rumus persamaan *Flanagan.*

[10] Sumarna Surapranata, *Analisis, Validitas, Reliabilitas Dan Interpretasi Hasil Tes*, (Bandung: Remaja Rosda Karya, 2009) 100.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Menurut Sumarna Surapranata faktor yang mempengaruhi ketidakajegan reliabilitas dengan penggunaan metode belah dua adalah karena dalam merespon tes itu sendiri. Sehingga yang paling berpengaruh pada reliabilitas tes adalah tingkat kesukaran soal. Ada banyak opsi untuk membagi tes menjadi dua bagian.[11] Dengan mempertimbangkan tingkat kesukaran tes dari masing-masing butir soal, diputuskan bahwa bagian pertama terdiri dari soal nomor 1, 4, 5. Bagian kedua terdiri dari soal nomor 2, 3, 6.

Berdasarkan perhitungan reliabilitas terhadap instrumen yang dikembangkan, didapatkan hasil 0,819 atau dibulatkan menjadi 0,82. Jika dibandingkan dengan tabel interpretasi indeks reliabilitas, maka instrumen yang dikembangkan memiliki indeks reliabilitas yang sangat tinggi yaitu 0,82. Dengan demikian disimpulkan bahwa instrumen yang dibuat reliabel atau bisa digunakan untuk mengukur kemampuan peserta didik dalam pembelajaran agama Islam dengan menggunakan metode pembelajaran berbasis masalah.

C. Pembahasan.

Penyusunan butir soal tes dalam penelitian ini dimulai dari merumuskan tujuan pembelajaran yang jelas. Tujuan pembelajaran yang jelas akan sangat membantu agar penilaian yang dilakukan benar-benar mengukur apa yang telah diajarkan kepada peserta didik.[12] Para ahli pendidikan telah sepakat bahwa untuk merumuskan tujuan pembelajaran yang baik, hendaknya

[11] Sumarna Surapranata, *Analisis, Validitas*, 109.
[12] Kusaeri, *Acuan dan Teknik Penilaian*, 30.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

mengacu pada salah-satu klasifikasi (taksonomi) tujuan pembelajaran.[13] Dalam penelitian ini taksonomi tujuan pembelajaran yang digunakan adalah Taksonomi Bloom edisi revisi.

Menurut Brookhart, kemampuan berpikir tingkat tinggi adalah (1) berpikir tingkat tinggi berada pada bagian atas taksonomi kognitif Bloom, meliputi kemampuan menganalisis, mengevaluasi, dan mengkreasikan (2) tujuan pengajaran di balik taksonomi kognitif yang dapat membekali peserta didik untuk melakukan transfer pengetahuan, (3) mampu berpikir artinya peserta didik mampu menerapkan pengetahuan dan keterampilan yang mereka kembangkan selama belajar pada konteks yang baru.[14]

Pembelajaran berbasis masalah atau disingkat dengan PBM adalah suatu model pembelajaran yang melibatkan peserta didik untuk memecahkan suatu masalah.[15] Menurut Taufiq Amir Pembelajaran berbasis masalah mampu untuk menunjang pembangunan kecakapan diri sendiri, kolaboratif dan kemampuan berpikir analisis, evaluasi dan mengkreasi.[16] Menurut Masidjo, untuk menunjang keberhasilan pembelajaran berbasis masalah, hendaknya menggunakan teknik penilaian yang tepat, agar kemampuan peserta didik dapat terukur. [17]

---

[13] Dadan Rosana, *Modul Evaluasi*, 46.
[14] Brookhart, "*How to Assess Higher Order Thinking Skills in Your Classroom,"* (Alexandria: ASCD, 2010) 5.
[15] Sudarman, *Problem Based Learning : Suatu Model Pembelajaran untuk Mengembangkan dan Meningkatkan Kemampuan Memecahkan Masalah,* (Jakarta, 2007, Dalam jurnal pendidikan inovatif).
[16] Taufiq Amir, *Inovasi Pendidikan Melalui Problem Based Learning*, (Prenada Media Group : Jakarta, 2010) 22.
[17] Masidjo, *Penilaian Pencapaian Hasil Belajar Siswa Disekolah*, (Kanisius: Yogyakarta, 1995) 46.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Jadi untuk mengukur tingkat keberhasilan pembelajaran berbasis masalah pada peserta didik, hendaknya digunakan instrumen penilaian yang tepat. Karena model penilaian akan sangat berpengaruh pada peserta didik. Menurut Van den Berg seperti dikutip Sa'dun Akbar, bahwa kurikulum memiliki potensi yang kaya untuk mengembangkan keterampilan berpikir tingkat tinggi peserta didik.[18]

Tes tertulis bentuk uraian non objektif dapat digunakan untuk mengukur kemampuan hasil belajar peserta didik pada tingkat C4, C5, C6. Karena tes tertulis bentuk uraian non objektif dapat menilai berbagai jenis kemampuan seperti: mengemukakan pendapat, berpikir kritis, berpikir kreatif dan pemecahan masalah.[19] Sehingga butir soal uraian non objektif dalam penelitian ini didesain berdasarkan perjenjangan Taksonomi Bloom edisi revisi dengan memperhatikan kemampuan menganalisis, mengevaluasi, dan mengkreasi yang dimiliki oleh peserta didik.

Dalam penulisan butir soal, materi awal cukup melibatkan aspek pengetahuan dan pemahaman (C1 dan C2). Karena sangat tidak mungkin untuk menguasai materi pembelajaran yang lebih tinggi tanpa menguasai materi pembelajaran yang lebih rendah. Selanjutnya untuk topik inti disusun berdasarkan level yang lebih kompleks yaitu level C4, C5, C6. Kategori pada Taksonomi Bloom edisi revisi disusun menjadi sebuah hierarki kumulatif.

---

[18] Sa'dun Akbar, *Instrumen Perangkat Pembelajaran*, (Bandung : PT Remaja Rosda Karya, 2013) 8.
[19] Kunandar, *Penilaian Autentik*, 209

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Artinya, penguasaan kategori yang lebih kompleks mensyaratkan penguasaan semua kategori yang dibawahnya.[20]

Berdasarkan penelitian yang dilakukan oleh Darodjat, Darmiyati Zuchdi dan Zamroni tentang " Model Evaluasi Pembelajaran Akidah Dan Akhlak Di Madrasah Tsanawiyah (MTs)." Diketahui, butir soal yang disusun mulai dari soal yang mudah mampu mengurangi rasa panik peserta tes. Sehingga peserta tes mampu merespon butir soal dengan baik.[21] Samritin dalam disertasinya menyatakan, tes yang dirakit dimulai dari butir tes yang mudah dapat mengurangi kecemasan peserta tes.[22]

Oleh karena itu, dalam penyusunan butir soal diperlukan soal dengan level mengingat dan memahami. Dengan demikian peserta didik mampu untuk mengorganisasikan penguasaan materi tingkat rendah sampai pada penguasaan materi yang lebih kompleks yaitu pada tingkatan berpikir level menganalisa, mengevaluasi dan mengkreasi. Sehingga peserta didik tidak panik dan mampu merespon soal dengan baik.

Untuk menjamin keakuratan soal tes tertulis bentuk uraian non objektif, maka soal harus memenuhi syarat-syarat sebagai berikut: (a) membatasi ruang lingkup dengan memilih materi atau bahan pelajaran yang esensial (b) menggunakan bahasa yang baik dan benar sehingga mudah difahami dengan baik oleh peserta didik (c) jangan mengulang pertanyaan

---

[20] Kusaeri, *Acuan dan Teknik Penilaian*, 33.

[21] Darodjat, Darmiyati Zuchdi dan Zamroni, *"Model Evaluasi Pembelajaran Akidah Dan Akhlak Di Madrasah Tsanawiyah (MTs),"* Jurnal Penelitian dan Evaluasi Pendidikan UNY, Volume 20, No 1, Juni 2016..

[22] Samritin, *"Pengembangan Instrumen Penilaian Kemampuan Higher Order Thinking Siswa SMP dalam Mata Pelajaran Matematika" (Disertasi, -- UNY Yogyakarta, 2014)* 123.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

pada materi yang sama (d) tuliskan rubrik penskoran sebelum menulis soal (e) menuliskan skor untuk masing-masing soal (f) rumusan soal harus jelas dan tegas (g) rumusan soal tidak boleh menggunakan kata yang menimbulkan penafsiran ganda (h) memiliki kriteria parameter butir soal yang baik (i) memiliki reliabilitas yang tinggi. [23]

Pembuatan instrumen tes tertulis bentuk uraian harus disertai pedoman penskoran yang disebut rubrik. Rubrik penskoran yang digunakan dalam penilitian ini adalah rubrik penskoran analitik. Penggunaan rubrik penskoran analitik dimaksudkan agar penyekoran yang dilakukan lebih teliti. [24]

Instrumen penilaian yang telah dihasilkan kemudian direvisi. Revisi terhadap instrumen dilakukan dua kali. Revisi pertama dilakukan berdasarkan saran para pakar. Pada revisi pertama menghasilkan perubahan besar terhadap instrumen. Revisi tersebut terdiri dari penghapusan salah satu indikator dan revisi konstruksi atau redaksi soal. Revisi kedua dilakukan pada saat melakukan uji keterbacaan. Revisi tersebut hanya menghasilkan perubahan redaksi dari soal.

Setelah soal disusun maka soal diuji cobakan terhadap subjek coba. Kemudian dianalisis berdasarkan kevalidan, tingkat kesukaran, daya pembeda dan reliabilitas soal.

1. Validitas.

Menurut Anas Sudijono, salah satu ciri tes hasil belajar yang baik adalah memiliki validitas. Suatu tes hasil belajar dengan validitas yang

[23] Kunandar, *Penilaian Autentik* . 212.
[24] Kusaeri, *Acuan dan Teknik Penilaian* , 94.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

tinggi dapat dikatakan handal dan tidak perlu diragukan ketepatan dalam mengukur hasil belajar peserta didik.[25] *Wainer & Braun* juga berpendapat bahwa tes yang baik harus memiliki karakteristik validitas agar dapat menyajikan informasi yang tepat tentang kondisi siswa yang mengikuti tes.[26] Keakuratan soal tes tertulis bentuk uraian non objektif untuk pembelajaran agama Islam berbasis masalah yang dikembangkan dalam penelitian ini telah terjamin kevalidannya.

Kevalidan instumen tes yang dikembangkan dibuktikan dengan penilaian para pakar. Berdasarkan penilaian pakar mengenai kesesuaian instrumen dengan level Taksonomi Bloom, dinyatakan bahwa tingkatan level instrumen yang dikembangkan telah sesuai dengan level Taksonomi Bloom edisi revisi.

Berdasarkan hasil penilaian pakar dengan menggunakan lembar validasi, dihasilkan rata-rata total validitas sebesar 3,62. Jika dibandingkan dengan tabel kriteria rata-rata total validitas, maka instrumen yang dikembangkan berada pada kategori sangat valid. Karena instrumen yang dikembangkan dinyatakan valid maka instrumen dapat digunakan untuk tahap selanjutnya, yaitu untuk tahap uji coba dalam mengukur kemampuan peserta didik.

[25] Anas Sudijono, *Pengantar Evaluasi Pendidikan* (Jakarta: Raja Granfindo Persada, 2011) 163.
[26] Kusairi dan Suprananto, *Pengukuran dan Penilaian Pendidikan,* (Yogyakarta: Graha Ilmu, 2012) 74.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

2. Tingkat Kesukaran.

Tingkat kesukaran butir soal adalah peluang menjawab benar suatu soal pada tingkat kemampuan tertentu. Dari hasil analisis tingkat kesukaran soal, didapatkan indeks tingkat kesukaran soal yang rata-rata berada pada rentang 0,3-0,7. Menurut Djemari mardapi Mardapi, butir soal yang baik memiliki kisaran indeks kesulitan 0,3 – 0,7. Butir soal yang memiliki tingkat kesulitan di bawah 0,3 dianggap terlalu sulit dan butir soal yang memiliki tingkat kesulitan di atas 0,7 dianggap terlalu mudah. Kriteria indeks daya beda butir soal yang boleh digunakan adalah ≥ 0,3.[27] Dengan demikian, instrumen yang dikembangkan berada pada kategori baik.

Butir soal nomor 1 memiliki tingkat kesukaran soal sebesar 0,70. Ini menandakan bahwa soal tersebut tergolong dalam soal yang mudah. Hal ini dapat dilihat dari banyaknya jumlah peserta yang mampu menjawab soal tersebut. Butir soal nomor 1 tidak direvisi, karena soal dirancang agar mudah untuk dikerjakan. Supaya peserta didik tidak panik saat mengerjakan soal selanjutnya.

Butir soal dengan indeks kesukaran soal terendah adalah butir soal nomor 5 dengan hasil sebesar 0,54. Hal ini menandakan bahwa soal tersebut adalah soal paling sulit diantara soal-soal yang lain. Secara keseluruhan, butir soal yang dikembangkan berada pada kategori sedang yaitu pada rentang 0,60. Rentang 0,60 adalah indeks tingkat kesukaran

[27] Djemari Mardapi, *Teknik Penyusunan Instrumen Tes*, 143.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

soal yang sedang tetapi mendekati mudah. Hal ini dipengaruhi oleh kualitas subjek coba. Subjek coba yang digunakan dalam penelitian ini adalah kelas dengan kriteria keagamaan, yang artinya subjek coba menguasai materi pembelajaran. Apabila instrumen diuji cobakan di kelas dengan kriteria selain agama, maka dipastikan tingkat kesukaran soal berada pada kategori sedang atau ideal.

Fungsi tingkat kesukaran butir soal biasanya dikaitkan dengan tujuan tes. Menurut Sukiman "Butir soal yang digunakan untuk keperluan ulangan atau ujian semester memiliki tingkat kesukaran yang sedang". Indeks tingkat kesukaran butir soal yang baik antara 0,3- 0,7. Jadi dapat disimpulkan bahwa intrumen tes tertulis bentuk uraian non objektif untuk pembelajaran agama Islam berbasis masalah yang dikembangkan dalam penelitian ini bisa digunakan untuk keperluan tes formatif atau ulangan harian. [28]

3. Daya Pembeda.

Daya pembeda soal adalah kemampuan butir soal dalam membedakan siswa yang pandai dan siswa yang kurang pandai.[29] berdasarkan hasil analisis daya pembeda soal, instrumen yang dikembangkan memiliki dua kriteria utama. Soal dengan daya pembeda dibawah 0,3 dan soal diatas 0,3. Soal dengan indeks daya pembeda dibawah 0,3 adalah soal nomor 1 dan 3 dengan indeks daya pembeda

[28] Sukiman, *Pengembangan Sistem Evaluasi.* (Yogyakarta: Insan Madani, 2012) 201.
[29] Kunandar, *Penilaian Autentik* . 240.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

0,24 dan 0,29 atau 33%. Soal dengan indeks daya pembeda diatas 0,3 adalah soal nomor 2, 4, 5, 6 atau 77%.

Tindak lanjut butir soal sesudah dianalisis daya pembedanya sebagai berikut:[30]

a. Butir soal yang memiliki daya pembeda baik disimpan.
b. Butir soal dengan daya pembeda rendah, ada dua kemungkinan tidak lanjut yaitu: (1) ditelusuri untuk kemudian diperbaiki dan selanjutnya digunakan kembali dalam tes hasil belajar mendatang guna mengetahui daya pembedanya meningkat atau tidak. (2) Dibuang.
c. Butir item yang angka indeks diskriminasinya bertanda negatif, sebaiknya dibuang karena kualitas butir soalnya sangat jelek.

Jika ditafsirkan dengan menggunakan tabel kriteria indeks daya pembeda soal, maka soal nomor 1 dan 3 berada pada kategori kurang baik. Soal nomor satu memiliki indeks daya pembeda soal yang paling rendah, hal ini dikarenakan soal nomor 1 memang dirancang agar mudah untuk dikerjakan, sehingga memiliki indeks daya pembeda dibawah 0,3. Namun soal nomor satu tetap dipertahankan karena tujuannya adalah untuk mengurangi rasa cemas peserta tes.

Hasil penelusuran soal nomor tiga yang memiliki indeks daya pembeda soal dibawah 0,3 ditemukan, bahwa ada dua orang peserta tes dari kelompok atas yang mendapatkan skor rendah yaitu 4. Oleh karena

[30] Anas Sudijono, *Pengantar Evaluasi*, 408.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

itu diputuskan bahwa soal nomor 3 tetap dipertahankan, karena faktor yang mempengaruhi indeks daya pembeda soal nomor 3 tidak terlalu signifikan.

Jadi disimpulkan bahwa instrumen tes tertulis bentuk uraian non objektif untuk pembelajaran agama Islam berbasis masalah dilihat dari segi daya pembeda soal memiliki kualitas baik. Sehingga tes berfungsi sebagaimana mestinya, yaitu mampu membedakan antara peserta didik yang pandai dan kurang pandai. Hal ini dibuktikan bahwa rata-rata soal memiliki indeks daya pembeda pada rentang 0,3-0,4 atau memuaskan.

4. Reliabilitas.

Reliabilitas adalah derajat konsistensi pengukur seperangkat soal. Reliabilitas soal pada instrumen tes ini menggunakan model belah dua *(split half methods).* Menurut Sumarna Surapranata metode belah dua dapat mengatasi kelemahan yang terdapat pada metode tes ulang dan tes paralel. Metode ini memungkinkan untuk mengestimasi reliablitas tanpa harus menyelenggarakan tes dua kali. Dengan demikian beberapa kelamahan seperti *reactivity effect* dan khususnya pengaruh waktu terhadap perolehan skor sebenarnya dapat diminimalisasi. Dengan demikian ketidakajegan perolehan skor bukan karena penyelenggaraan tes tetapi karena dalam merespon tes itu sendiri.[31]

Berdasarkan perhitungan indeks reliabilitas yang dilakukan dengan menggunakan persamaan *flanagan* didapatkan hasil 0,819 atau

[31] Sumarna Surapranata, *Analisis Validitas*, 100.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

dibulatkan menjadi 0,82. Jika dibandingkan dengan tabel interpretasi indeks reliabilitas, maka instrumen yang dikembangkan memiliki indeks reliabilitas yang sangat tinggi yaitu 0,82. Disimpulkan bahwa indeks reliabilitas intrumen yang dikembangkan berada pada kategori sangat tinggi atau instrumen yang dikembangkan ajeg bila digunakan untuk mengukur kemampuan peserta didik.

D. Temuan Penelitian.

Penelitian ini adalah penelitian pengembangan yang menghasilkan produk dalam bidang pendidikan, berupa instrumen tes tertulis bentuk uraian non objektif untuk pembelajaran agama Islam berbasis masalah. Intrumen yang dihasilkan dalam penelitian telah memenuhi kriteria baik dari segi validitas, parameter butir maupun reliabilitas. Instrumen tes tertulis bentuk uraian non objektif terdiri dari satul soal level C1, satu soal level C2, satu soal level C4, dua soal level C5 dan satul soal level C6.

Instrumen tes tertulis bentuk uraian non objektif sangat cocok untuk digabungkan dengan pembelajaran berbasis masalah. Keduanya memiliki karakteristik yang hampir sama, yaitu untuk meningkatkan kemampuan berfikir level C4, C5, dan C6. Dengan menggabungkan keduanya maka diharapkan pembelajaran yang telah dilakukan bisa mendapatkan hasil yang maksimal. Pada dasarnya semua bentuk tes bisa digunakan, namun yang paling tepat digunakan dengan pembelajaran berbasis masalah adalah bentuk uraian non objektif.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Butir-butir instrumen tes tertulis bentuk uraian non objektif yang dihasilkan dalam penelitian ini memiliki tingkat kesukaran soal pada rentang 0,3 sampai 0,7. Tingkat kesukaran soal berada pada kategori sedang atau dapat diterima. Hal ini karena pengembangan tes telah melalui prosedur yang sistematis dan dilakukan dengan baik. Proses pengembangan tes dimulai dari penyusunan spesifikasi tes kemudian dilanjutkan dengan penulisan butir-butir soal dengan mengacu pada perjenjangan tingkat berfikir menurut Taksonomi Bloom edisi revisi. Selain itu ada berbagai aspek yang mempengaruhi pembuatan soal seperti aspek materi, konstruksi dan bahasa.

Butir-butir tes memiliki indeks reliabilitas sebesar 0,82. Dengan indeks reliabilitas itu, maka butir tes dinyatakan sangat valid dan mampu digunakan untuk mengukur kemampuan peserta didik.

Butir tes disusun dengan dimulai dari soal yang mudah, sehingga peserta didik tidak panik ketika mengerjakan soal. Ketenangan dalam menjawab tes mempengaruhi peserta didik dalam menunjukkan kemampuannya dalam menjawab soal. Hasil dari peserta tes akan berpengaruh terhadap tingkat kesukaran soal yang dikembangkan. Secara tidak langsung akan berpengaruh terhadap semua aspek yang akan dianalisis yaitu daya pembeda dan reliabilitas. Oleh karena itu instrumen disusun dengan dimulai dari yang paling mudah.

Butir tes yang disusun memiliki rata-rata total validitas 3,6 yang berarti instrumen yang dikembangkan valid berdasarkan penilaian validator. Para validator adalah praktisi pendidikan, sehingga mereka paham mengenai

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

kemampuan peserta didik. Pemahaman tersebut menjadi pertimbangan dalam menyusun sebuah instrumen penilaian yang baik. Hal ini sejalan dengan hasil analisis secara empirik dilapangan yang menyatakan bahwa instrumen yang dibuat valid berdasarkan kriteria parameter butir dan reliabilitas.

Ketika berdikusi dengan guru mata pelajaran Fiqh, dihasilkan fakta bahwa guru-guru setuju untuk menggunakan soal-soal model uraian bentuk uraian dengan mengacu pada perjenjangan Taksonomi Bloom edisi revisi. Guru-guru beralasan bahwa jika model pembelajarannya sudah bagus namun penilaiannya kurang, maka tidak bisa mengukur kemampuan peserta didik yang sesungguhnya. Menyusun soal berdasarkan perjenjangan Taksonomi Bloom edisi revisi juga tidak begitu sulit, karena sudah banyak panduan untuk menyusunnya. Yang diperlukan hanya latihan dan pembiasaan untuk menyusun soal yang baik sesuai dengan perjenjangan Taksonomi Bloom edisi revisi.[32]

Temuan di lapangan menunjukkan bahwa peserta didik dipaksa berfikir lebih keras untuk menjawab soal-soal. Karena soal-soal yang dikembangkan tidak hanya bersifat mengingat dan memahami materi. Soal-soal juga dirancang agar peserta didik mampu menganalisa, mengevaluasi dan mengkreasikan hal baru dari apa yang telah mereka pelajari. Dengan proses tersebut maka kemampuan peserta didik yang dihasilkan dari dampak pembelajaran berbasis masalah mampu tersalurkan dengan baik dan tidak sia-sia.

[32] Ahsan Bisri dan Muh. Ali Mashudi, Wawancara, sidoarjo, 29 agustus 2016.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Jadi secara keseluruhan instrumen penilaian yang dikembangkan dalam penelitian ini dikatakan baik karena telah memenuhi kriteria sebagai berikut: (1) valid berdasarkan penilaian para ahli (2) memiliki reliabilitas minimal 0,7 (3) butir-butir tes memiliki parameter tingkat kesulitan dan daya beda pada rentang 0,3-0,7 .

E. Keterbatasan Penelitian.

Beberapa keterbatasan yang ada dalam penelitian adalah sebagai berikut:

1. Dalam penelitian ini uji coba instrumen didasarkan pada asumsi bahwa proses pembelajaran berbasis masalah yang dilakukan oleh guru sudah baik. Peneliti tidak langsung mengajar peserta didik, namun hanya memberi masukan kepada guru mata pelajaran Fiqh.
2. Subyek coba dalam penelitian ini hanya kelas VIII B di MTSN 4 Sidoarjo. Peneliti berasumsi bahwa subyek coba dengan karakteristik keagamaan mampu mewakili subyek yang lain. Untuk penelitian lebih lanjut dapat menggunakan subyek coba yang lebih luas untuk meningkatkan keakuratan butir tes yang dikembangkan.
3. Instrumen tes tertulis bentuk uraian non objektif yang dikembangkan terbatas pada materi Fiqh. Untuk materi agama Islam yang lain bisa dikembangkan sendiri instrumen penilaiannya, dengan mengikuti tahap pengembangan yang telah dilakukan. Karena tahap-tahap pengembangan yang telah dilakukan terbukti valid dan teruji.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id