

BAB IV

LAPORAN HASIL PENELITIAN

A. Gambaran Umum Tempat Penelitian

Penelitian ini mengambil tempat penelitian pada 3 tempat, yaitu di TPA (taman pendidikan al qur'an) YABABERMA (Yayasan Banua Berkah Mandiri), di TPA Darul Muthaharah Sungai Gardu Banjarmasin. Penjelasan lebih lengkap tentang kedua tempat tersebut akan dijelaskan sebagai berikut:

1. TPA YABABERMA (Taman Pendidikan Al Qur'an Yayasan Banua Berkah Mandiri) Banjarmasin

TPA (Taman Pendidikan Al Qur'an) YABABERMA (Yayasan Banua Berkah Mandiri) merupakan tempat pembelajaran membaca dan menghafal al Qur'an dan al Hadits dengan menggunakan metode Azza di bawah naungan Yayasan Rumah Yatim Bahagia yang beralamat Jl. Simpang Bali, Gang Telerama No. 13, Antasan Besar, Banjarmasin Tengah, Kalimantan Selatan untuk asrama putra, dan di Jl. Dahlia Raya No. 3, Telawang, Banjarmasin Barat, Kota Banjarmasin, Kalimantan Selatan untuk asrama putri dan balita.¹

TPA (Taman Pendidikan Al Qur'an) YABABERMA (Yayasan Banua Berkah Mandiri) merupakan lembaga pendidikan Al Qur'an dan Hadist yang mengikuti Yayasan Darud Dakwah di Kuala kapuas yaitu dengan menggunakan metode Azza yang mana lembaga pendidikan TPA YABABERMA ini di bawah naungan Yayasan Banua Mandiri yang merupakan salah satu adalah lembaga

¹ Dokumentasi Yayasan Banua Berkah Foundation Banjarmasin 2024

rumah yatim yang menerima dan menyalurkan zakat, infaq, sedekah, wakaf kepada yatim dan dhuafa di Banjarmasin yang berada dalam pengawasan dan bantuan dari pemerintah dan sudah terdaftar di MENKUMHAM.

TPA YABABERMA memulai mendirikan lembaga pendidikan Al Qur'an dan Hadist dengan menggunakan metode Azza pada awal tahun 2018 yang dimulai dengan awal pertemuan pendiri sekaligus pengawas yayasan berkah mandiri Bapak Rafiq dengan Ustadz Ahmad Zaki Zamani. Dari awal pertemuan tersebut terlintas sebuah ide untuk mendirikan TPA yang akan menjadi awal pergerakan yayasan banua mandiri dalam bidang pendidikan yang akan direncanakan akan terus berkembang sampai tingkat perguruan tinggi.²

Tanggapan masyarakat serta antusias anak didaerah sekitar sangat baik terhadap yayasan dalam bidang pendidikan Al Qur'an dan Hadist dengan metode pembelajaran Azza, sehingga perkembangan lembaga pendidikan dapat terlihat dari jumlah santri yang makin hari terus bertambah hingga saat ini, adapun jumlah santri pada saat ini ada 152 santri yang terbagi menjadi beberapa halaqoh, yaitu:

- a. Mubtadi
- b. Ibtida 1
- c. Ibtida 2
- d. Ibtida 3
- e. Wushto

² Wawancara dengan Ibu Ignes, Sekretaris Yayasan Banua Mandiri Banjarmasin, (28 Juli 2025)

f. Ulya 1

g. Ulya 2

2. TPA (taman pendidikan al Qur'an) Darul Muthaharah Banjarmasin

TPA (taman pendidikan al Qur'an) Darul Muthaharah Banjarmasin merupakan tempat pembelajaran membaca dan menghafal al Qur'an dan al Hadits dengan menggunakan metode Azza. TPA Darul Muthaharah beralamat di Jl. Veteran Km5, Sungai Gardu, RT 13, RW 02, Gang H. Arsyad, Kelurahan Pengambangan, Kecamatan Banjarmasin Timur, Kota Banjarmasin.³

TPA Darul Muthaharah merupakan cabang dari TPA Darud Dakwah di Kuala Kapuas, didirikan oleh Muhammad Nor yang merupakan murid dari pendiri Yayasan Darud Dakwah Azza Kuala Kapuas yaitu Ustadz Ahmad Zaki Zamani. Muhammad Noor yang sebelumnya merupakan salah pengajar di TPA Darud Dakwah di Kapuas, tetapi karena beliau berpindah tempat tinggal sementara yang sebelum di Kapuas kemudian pindah ke Banjarmasin maka beliau pun mendirikan TPA di Mushalla dekat dengan tempat tinggal beliau di Banjarmasin atas izin dan bantuan oleh Ustadz Ahmad Zakki Zamani pada bulan Agustus 2019.⁴

Peminat masyarakat terhadap TPA Darul Muthaharah cukup banyak, sehingga saat ini jumlah santri di TPA Darul Muthaharah berjumlah 55 santri yang terbagi ke dalam beberapa halaqah, yaitu:

a. Mubtadi

³ Dokumentasi TPA Darul Muthaharah Banjarmasin

⁴ Wawancara dengan Ustadah Ainun, Ketua TPA Darul Muthaharah Banjarmasin, (25 Juli 2025)

- b. Ibtida 1
- c. Ibtida 2
- d. Ibtida 3
- e. Wushto
- f. Ulya 1
- g. Ulya 2
- h. Ulya 3

3. TPA Yayasan Generasi Yatim Tahfidz (YGYT) Banjarbaru

TPA (Taman Pendidikan Al Qur'an) YGYT (Yayasan Generasi Yatim Tahfidz) merupakan tempat pembelajaran yang mempelajari cara membaca yang baik dan menghafal al Qur'an dan Hadits dengan menggunakan metode Azza di bawah naungan Yayasan Rumah Yatim Bahagia serupa dengan YABABERMA yang beralamat Jl. P. M. Noor, komplek Permata Hijau, Blok. D, No. 29, RT. 1, RW. 2, Kelurahan Sungai Ulin, Kecamatan Banjarbaru Utara, Kota Banjarbaru, Kalimantan Selatan.

TPA Yayasan Generasi Yatim Tahfidz ini juga merupakan lembaga pendidikan yang dinaungi oleh Yayasan Rumah Yatim Bahagia seperti TPA YABABERMA di kota Banjarmasin yang didirikan pada tahun 2021 di kota Banjarbaru dengan berfokus pada program tahfidz al Qur'an dan Hadits yang merupakan lanjutan dari TPA YABABERMA di kota Banjarmasin, oleh karena itu, lembaga tersebut juga menggunakan metode Azza sebagai metode pembelajaran untuk mengajarkan para peserta didik dalam pembelajaran membaca dan menghafal

al Qur'an dan Hadits.⁵

Adapun jumlah santri yang berada saat adalah berjumlah 63 santri, yang terbagi menjadi beberapa halaqoh, yaitu:

- a. Muhtadi
- b. Ibtida 1
- c. Ibtida 2
- d. Ibtida 3
- e. Wusho
- f. Ulya 1

B. Hasil Penelitian

Data hasil penelitian terhadap objek penelitian yaitu “penyusunan instrumen tes hasil belajar pendidikan al Qur'an dan Hadits metode azza pada kelas Wusho di TPA kota Banjarmasin diuraikan berdasarkan desain metode penelitian, yaitu *Exploratori Sequential Design*, yang mana mencakup pada beberapa langkah seperti yang telah peneliti buat pada bab 3 pada karya ilmiah ini, sebagai berikut:

1. Menentukan Tujuan Tes

Tujuan tes ini diadakan untuk kelas Wusho pada pembelajaran al Qur'an dan Hadist dengan metode Azza adalah sebagai berikut:

- 1) Untuk mengetahui apakah strategi guru dalam menggunakan metode Azza dalam proses pembelajaran sudah tepat atau belum pada anak yang berada di

⁵ Wawancara dengan Ustad Murtadha, Ketua TPA Darul Muthhaharah Banjarmasin, (12 Juli 2025)

kelas Wustho

- 2) Untuk mengetahui penguasaan materi al Qur`an dalam segi bacaan dan teori pada beberapa materi sebagai berikut:
 - a) Mempelajari dasar-dasar tempat keluarnya bunyi huruf hijaiyyah pada mulut dari huruf Alif -Ya`
 - b) Mempelajari sifat-sifat tempat keluarnya huruf hijaiyyah pada mulut dari huruf Alif-Ya`
 - c) 11 Tanda-tanda wadaf, yaitu:
 - (1) Waqaf lazim
 - (2) Waqaf laa washal
 - (3) Waqaf jaiz
 - (4) Waqaf al washlu aula
 - (5) Waqaf al waqfu aula
 - (6) Waqaf mu`annaqah
 - (7) Waqaf mustahab
 - (8) Waqaf muthlaq
 - (9) Waqaf murokh-khas
 - (10) Waqaf mujawwaz
 - (11) Qila `alaihi waqaf
 - d) bacaan pada Mad Lazim Musyabba` Harfi dan Mad Lazim Mukhoffaf Harfi pada setiap awal surah
- 3) Untuk mengetahui penguasaan materi al Hadist dalam segi kehidupan sehari-

hari, dalam hal ini materi yang harus dikuasai pada kelas Wustho dimulai dari Hadist bagian 1 sampai dengan Hadist bagian 5 yang berjumlah 20 Hadist

- 4) Untuk mengetahui hafalan surah-surah pendek di al Qur'an dari surah An-Naas sampai dengan surah Adh-Dhuha dan al Hadist

Setelah mengetahui tujuan dari diadakannya tes untuk kelas Wustho, mengetahui indikator juga menjadi landasan dalam menyusun kisi-kisi sebagai salah satu langkah dalam penyusunan instrumen tes ini setelah menemukan tujuan tes tersebut. Indikator tersebut adalah sebagai berikut:

- 1) Siswa dapat menyebutkan arti waqaf lazim dan menandai letaknya dalam potongan ayat.
- 2) Siswa dapat menjelaskan makna waqaf laa washal dan menunjukkan contohnya dalam ayat.
- 3) Siswa dapat menjelaskan pengertian waqaf jaiz dan membedakannya dari waqaf lainnya.
- 4) Siswa dapat memilih bacaan yang tepat saat bertemu waqaf al washlu aula.
- 5) Siswa dapat menentukan kapan harus berhenti saat ada tanda waqaf al waqfu aula.
- 6) Siswa dapat menyebutkan ciri-ciri waqaf mu'annaqah dan menjelaskan aturannya.
- 7) Siswa dapat menyebutkan arti waqaf mustahab dan memilih tempat berhenti yang sesuai.
- 8) Siswa dapat mengenali tanda waqaf muthlaq dalam Al-Qur'an.

- 9) Siswa dapat menyebutkan maksud waqaf murokh-khas dan contohnya.
- 10) Siswa dapat menjelaskan kapan boleh berhenti di waqaf mujawwaz.
- 11) Siswa dapat menyebutkan pengertian dan penggunaan qila 'alaihi waqaf.
- 12) Siswa dapat mengenali huruf qalqalah shugro dan cara membacanya dalam kata.
- 13) Siswa dapat membaca huruf-huruf muqatta'ah dengan mad lazim mutsaqqal harfi secara benar.
- 14) Siswa dapat membaca mad lazim mukhoffaf harfi dengan panjang yang sesuai.
- 15) Siswa dapat menghafal dan melafalkan Surah An Nas sampai Ad Dhuha dengan lancar dan tajwid yang benar.
- 16) Siswa dapat menghafal dan menyebutkan isi Hadits 1 sampai 5 secara lengkap.
- 17) Siswa dapat menghafal dan menyebutkan isi Hadits 6 sampai 10 secara lengkap.
- 18) Siswa dapat menghafal dan menyebutkan isi Hadits 11 sampai 15 secara lengkap.
- 19) Siswa dapat menghafal dan menyebutkan isi Hadits 16 sampai 20 secara lengkap.

2. Menyusun Kisi-Kisi

Setelah mengetahui tujuan dan indikator yang terdapat dalam sebuah tes, maka penyusunan kisi-kisi merupakan langkah awal yang harus dilakukan pada setiap penyusunan instrumen tes hasil belajar dan menulis soal.

Berikut ini adalah kisi-kisi instrumen tes berdasarkan indikator pembelajaran yang sebelumnya telah peneliti susun. Kisi-kisi ini mencakup:

- a. Kompetensi dasar
- b. Materi Pokok
- c. Indikator Soal
- d. Level kognitif

Tabel 4.1

Daftar Kisi-kisi Untuk Kelas Wustho

No	Kompetensi Dasar	Materi Pokok	Indikator Soal	Level Kognitif
1	Memahami tanda waqaf	Waqaf lazim	Menyebutkan arti waqaf lazim dan letaknya dalam ayat	C1
2	Memahami tanda waqaf	Waqaf mustahab	Menyebutkan arti dan penggunaannya	C1
3	Memahami tanda waqaf	Waqaf muthlaq	Mengenali tanda dalam ayat	C1
4	Memahami tanda waqaf	Qila ‘alaihi waqaf	Menjelaskan pengertian dan penggunaannya	C1
5	Menghafal hadits	Hadits 1–5	Menyebutkan isi hadits dengan benar	C1
6	Menghafal hadits	Hadits 6–10	Menyebutkan isi hadits dengan benar	C1
7	Menghafal hadits	Hadits 11–15	Menyebutkan isi hadits dengan benar	C1
8	Menghafal hadits	Hadits 16–20	Menyebutkan isi hadits dengan benar	C1
9	Memahami tanda waqaf	Waqaf Lazim	Menjelaskan makna waqaf Lazim dan contohnya	C2

10	Memahami tanda waqaf	Waqaf jaiz	Menjelaskan pengertian dan membedakan dengan waqaf lain	C2
11	Memahami tanda waqaf	Waqaf mu'annaqah	Menyebutkan ciri-ciri dan aturan bacanya	C2
12	Memahami tanda waqaf	Waqaf murokh-khas	Menyebutkan maksud dan contoh	C2
13	Memahami tanda waqaf	Waqaf mujawwaz	Menjelaskan kapan boleh berhenti	C2
14	Membaca tajwid	Qalqalah Kubro	Membedakan qalqalah kubro dan shugro	C2
15	Memahami tanda waqaf	Waqaf al washlu aula	Memilih bacaan yang tepat saat ada waqaf ini	C3
16	Memahami tanda waqaf	Waqaf al waqfu aula	Menentukan kapan harus berhenti saat ada waqaf ini	C3
17	Membaca dan memahami tajwid	Mad Lazim Mutsaqal Harfi	Membaca huruf muqatta'ah secara benar	C3
18	Membaca dan memahami tajwid	Mad Lazim Mukhoffaf Harfi	Membaca sesuai panjangnya	C3
19	Menghafal surah	Surah An Nas s.d. Ad Dhuha	Menghafal dan melafalkan dengan benar	C3

3. Menentukan Bentuk Instrumen Tes

Setelah mendapat kisi-kisi yang tersusun dari indikator dan di urutkan menurut level kognitif anak. Maka penulis menentukan bentuk instrumen tes tersebut menjadi 3 bentuk, yaitu:

- a. Pilihan ganda
- b. Uraian

4. Penulisan Instrumen tes

Setelah mengetahui bentuk soal dari susunan kisi-kisi yang di urutkan menurut level kognitif, peneliti melanjutkan pada penulisan instrumen tes dalam bentuk pilihan ganda, dan uraian sebagai berikut:

A. Jawab pertanyaan di bawah ini dengan benar!

1. Tempat keluarnya huruf ba' (ب) adalah dari ...
 - a. Tenggorokan
 - b. Pangkal Lidah
 - c. Kedua bibir
 - d. Ujung Lidah
2. Huruf 'ain (ع) keluar dari bagian ...
 - a. Pangkal Lidah
 - b. Ujung Tenggorokan
 - c. Tengah Tenggorokan
 - d. Rongga Mulut
3. Huruf yang penyebutannya di tengah tenggorokan yaitu . . .
 - a. ح
 - b. خ
 - c. هـ
 - d. غ
4. Huruf “ض” keluar dari ...
 - a. Ujung lidah dan gigi seri atas
 - b. Tengah lidah dan langit-langit
 - c. Samping lidah dan gigi geraham
 - d. Pangkal lidah dan langit-langit
5. Sifat huruf ر adalah . . .

d. Keutamaan al Qur'an

10. “jangan suka marah”

Judul Hadits di atas adalah Hadits bagian?..

a. Bagian 3 nomor 1

c. Bagian 3 nomor 3

b. Bagian 3 nomor 2

d. Bagian 5 nomor 1

11. Apa bunyi ayat sebelum ayat berikut ini ?..

وَأَنَّهُ ۖ عَلَىٰ ذَٰلِكَ لَشَهِيدٌ

a. أَفَلَا يَعْلَمُ إِذَا بُعْثِرَ مَا فِي الْقُبُورِ

b. إِنَّ الْإِنْسَانَ لِرَبِّهِ ۖ لَكَنُودٌ

c. وَأَنَّهُ ۖ حُبِّ الْخَيْرِ لَشَدِيدٌ

d. فَاتَّزَنَ بِهِ ۖ نَقْعًا

12. apa nama surah dari ayat berikut ini

يَوْمَ يَكُونُ النَّاسُ كَالْفَرَاشِ الْمَبْثُوثِ

a. At Takasur

c. Al Alaq

b. Al Qori'ah

d. At Tin

13. Dalam metode azza, apa saja nama surah pada bagian 5?

a. Al – Adiyat, Al – Zalzalah, Al – Qoriah

b. Al – Asr, Al – Takasur, Al – Qoriah

c. Al – Quraishy, Al – Fiil, Al-Humazah

d. Al – Qadr, Al – Alaq, At – Tin

14. Apa bunyi surah Al – Alaq ayat 4?

- a. عِلْمَ الْإِنْسَانِ مَا لَمْ يَعْلَمْ c. أَنْ رَأَاهُ اسْتَغْنَى
b. الَّذِي عَلَّمَ بِالْقَلَمِ d. كَلَّا إِنَّ الْإِنْسَانَ لَيْطَغَى

15. Apa arti dari nama surah Al-Humazah ?

- a. Pertolongan c. Pemberian yang banyak
b. pengumpat d. Api yang bergejolak

16. Salah satu tanda waqaf yang berarti "berhenti lebih utama" adalah ...

- a. لا c. صلى
b. قللى d. ج

17. Berapa harakat panjang bacaan "اَلَمْ" (alif-laam-miim)?

- a. 2 harakat c. 6 harakat
b. 4 harakat d. 8 harakat

18. Tanda waqaf “ج” dalam Al-Qur'an disebut dengan ..

- a. Waqaf laa washal c. Waqaf lazim
b. Waqaf Jaiz d. Waqaf mustahab

19. Arti dari tanda waqaf (•.....•) adalah?..

- a. Lebih baik berhenti
b. Lebih baik diteruskan
c. Boleh berhenti, tapi lebih baik diteruskan
d. Boleh berhenti disalah satu tanda

20. Baris fathatain, kasratain dan dhommatain dikatakan juga...

- a. Tanwin c. Ghorib

b. Tasydid

d. Jahr

B. Isilah titik-titik di bawah ini!

1. Jelaskan tempat keluarnya huruf “fa (ف)” dan sifat utamanya dalam ilmu tajwid!
2. Apa perbedaan antara Waqaf Muthlaq, Waqaf Lazim dan Waqaf Jaiz?
3. Apa arti dari Judul Hadits yang artinya “berbuat kebaikan” pada Hadits bagian 2 nomor 2?
4. Berapa jumlah ayat dari surah Adh-Dhuha?
5. Ada berapa jumlah surah dari Surah Al-Fiil sampai dengan Surah An-Naas?

Adapun kunci jawaban terdapat pada lampiran yang telah dilampirkan oleh peneliti beserta dengan poin skor tiap jawaban yang dijawab oleh peserta didik dengan jawaban yang benar.

5. Pelaksanaan Tes

Agar memperoleh instrumen tes yang baik maka instrumen tes tersebut harus di uji coba terlebih dahulu dan hasilnya akan dianalisis sehingga memenuhi syarat-syarat tes yang baik. Peserta uji coba adalah siswa atau santri TPA YABABERMA dan TPA Darul Muthaharah yang berada pada kelas atau halaqoh Wustho berjumlah 100 siswa yang mana instrumen tes ini disusun untuk menjadi tes evaluasi hasil belajar kenaikan kelas atau halaqoh dari Wustho menuju halaqoh

Ulya yang memiliki pembelajaran di atasnya.

Adapun hasil uji coba tes yang diadakan untuk santri kelas Wustho pada TPA YABABERMA, TPA Darul Muthaharah dan TPA Yayasan Generasi Yatim Tahfidz terdapat pada halaman lampiran.

Berikut hasil gambar tangkap kamera kegiatan pelaksanaan uji coba tes:

Gambar 4.1

Pelaksanaan Uji Coba Tes



Dengan jawaban yang telah diperoleh oleh peneliti dari peserta didik yang mengikuti uji coba tes ini, maka peneliti dapat melakukan analisis terhadap tiap butir soal yang meliputi:

- a) Validitas
- b) Reliabilitas
- c) Tingkat kesukaran
- d) Daya pembeda
- e) Efektivitas pengecoh

6. Analisis Instrumen Tes

a. Menghitung Validitas Butir (*r butir*)

Setiap item soal dianalisis validitas butirnya dengan cara sebagai berikut:

Yang pertama Tabel skor diurutkan dari skor tertinggi ke terendah.

Selanjutnya setiap butir soal dihitung nilai validitasnya dengan menggunakan rumus *korelasi Product Moment Pearson*:

$$r_{xy} = \frac{n \sum XY - (\sum X)(\sum Y)}{\sqrt{\{n \sum X^2 - (\sum X)^2\} \{n \sum Y^2 - (\sum Y)^2\}}}$$

Keterangan:

r_{xy} = Koefisien korelasi antara variabel X dengan variabel Y.

N = Jumlah Responden.

$\sum X$ = Jumlah skor item variabel X

$\sum Y$ = Jumlah skor variabel Y

$\sum XY$ = Jumlah skor dalam sebaran X dan Y

$\sum X^2$ = Jumlah kuadrat dari jumlah skor dalam sebaran X

$\sum Y^2$ = Jumlah kuadrat dari jumlah skor dalam sebaran Y

Nilai koefisien korelasi (*r*) dikonfirmasi dengan tabel kritik *Product*

moment pada taraf signikansi 5% ($\alpha = 0,05$) dan dk n-1. Kriteria nilai koefisien korelasi tersebut dikategorikan sebagai berikut:

$0,8 \leq r \leq 1$ sangat tinggi

$0,6 \leq r \leq 0,79$ tinggi

$0,4 \leq r \leq 0,59$ cukup

$0,2 \leq r \leq 0,39$ rendah

$0,0 \leq r \leq 0,19$ rendah sekali⁶

Apabila hasil perhitungan validitas butir untuk soal tertentu diperoleh r-hitung lebih besar jika dibandingkan dengan r-tabel pada taraf signifikasi 5% ($\alpha = 0,05$) dan dk n-1 berarti soal tersebut valid. Perhitungan nilai koefisien korelasi r , Widoyoko dan Sugiyono mengatakan, bahwa harga kritik untuk validitas instrumen adalah 0,3. Artinya bahwa butir soal dikatakan valid jika nilai koefisien korelasi 0,3 atau lebih (paling kecil 0,3) dan sebaliknya dikategorikan butir soal tidak valid.⁷ Dari sejumlah soal yang divalidasi, maka kemungkinan terdapat beberapa soal yang tidak valid.

Dengan bantuan software aplikasi SPSS, maka penelitian melakukan analisis validitas tes hasil belajar dan mendapatkan hasil analisis sebagai berikut:

Tabel. 3.1

⁶ Arikunto Suharsimi, *Dasar-Dasar Evaluasi Pendidikan*, hal. 253

⁷ Widoyoko, E.S. *Teknik Penyusunan Instrumen Penelitian*. Yogyakarta: Pustaka Pelajar, 2012, hal. 149

Hasil Analisis Validitas Instrumen Tes Hasil Belajar

Bentuk Soal Pilihan Ganda

Pilihan Ganda	Pearson Corelation	Validitas
Soal 1	0,286**	Valid
Soal 2	0,441**	Valid
Soal 3	0,494**	Valid
Soal 4	0,440**	Valid
Soal 5	0,710**	Valid
Soal 6	0,286**	Valid
Soal 7	0,447**	Valid
Soal 8	0,751**	Valid
Soal 9	0,417**	Valid
Soal 10	0,635**	Valid
Soal 11	0,710**	Valid
Soal 12	0,558**	Valid
Soal 13	0,499**	Valid
Soal 14	0,447**	Valid
Soal 15	0,711**	Valid
Soal 16	0,558**	Valid
Soal 17	0,635**	Valid
Soal 18	0,384**	Valid
Soal 19	0,284**	Valid
Soal 20	0,710**	Valid

Berikut hasil dari perhitung analisis validitas tes hasil belajar Qur'an hadits metode Azza kelas Wustho pada 20 butir soal dalam bentuk pilihan ganda dengan bantuan aplikasi software SPSS, yang mana hasilnya pada 20 butir soal dinyatakan valid, artinya nilai r-hitung lebih besar daripada r-tabel pada taraf signifikan 5% ($\alpha = 0,05$).

Tabel. 3.2

Hasil Analisis Validitas Bentuk Soal Esai

Esai	Pearson Corelation	Validitas
Soal 1	0,742**	Valid
Soal 2	0,793**	Valid
Soal 3	0,344**	Valid
Soal 4	0,471**	Valid
Soal 5	0,798**	Valid

Dari penghitungan validitas menggunakan aplikasi bantuan SPSS dalam menganalisis validitas pada tiap butir soal, maka di diperolehlah hasil seperti tabel di atas. Tabel di atas menunjukkan bahwa nilai tiap butir soal bernilai valid, artinya nilai r-hitung lebih besar daripada r-tabel pada taraf signifikan 5% ($\alpha = 0,05$).

Dengan pernyataan di atas yang menyatakan bahwa semua butir soal dinyatakan valid untuk dapat digunakan sebagai instrumen tes hasil belajar dalam pembelajaran pendidikan agama Islam dalam mata pelajaran al Qur'an dan Hadits di TPA kota Banjarmasin dan Banjarbaru pada kelas Wustho dengan metode pembelajaran menggunakan metode Azza.

b. Reliabilitas

Tujuan utama menghitung reliabilitas skor tes adalah untuk mengetahui tingkat ketepatan (*Precision*) dan keajegan (*consistency*) skor tes. Suatu instrumen atau tes dikatakan memiliki nilai reliabilitas yang tinggi apabila tes yang dibuat memiliki hasil yang konsisten dalam mengukur yang hendak diukur.⁸ Indeks

⁸ Sukardi. *Metodologi Pendlitian Pendidikan: Kompetensi dan Praktiknya*. Jakarta: Bumi Aksara, 2012, hal. 127

reliabilitas berkisar antara 0-1. Semakin tinggi koefisien reliabilitas pada suatu tes (mendekati 1), maka semakin tinggi pula keajegan/ketepatannya.⁹

Tes yang memiliki konsistensi reliabilitas yang tinggi adalah akurat, *resproducibel*, dan *generalized* terhadap kesempatan *testing* dan instrumen tes lainnya. Secara rinci faktor yang dapat mempengaruhi reliabilitas sebuah tes di antaranya:

- 1) Semakin banyak jumlah butir soal, semakin ajek suatu tes
- 2) Semakin lama waktu tes, semakin ajek
- 3) Semakin sempit range kesukaran butir soal, semakin besar keajekan
- 4) Soal-soal yang saling berhubungan akan mengurangi keajekan
- 5) Semakin objektif pemberian skor, semakin besar keajekan
- 6) Tidak tepatnya pemberian skor
- 7) Menjawab besar soal dengan cara menebak
- 8) Semakin homogen materi semakin besar keajekan
- 9) Pengalaman peserta ujian
- 10) Salah penafsiran terhadap butir soal
- 11) Menjawab soal dengan buru-buru/cepat
- 12) Kesiapan mental peserta ujian
- 13) Adanya gangguan dalam pelaksanaan tes
- 14) Jarak antara tes pertama dengan tes kedua

⁹ Depdiknas, *Juknis Penulisan Analisis Butir Soal di SMA*. Jakarta: Direktorat Pembinaan SMA, 2010, hal. 121.

15) Posisi individu dalam belajar

16) Kondisi fisik peserta tes/ujian

Selanjutnya ada 3 cara yang dapat dilakukan untuk menentukan reliabilitas tes, yaitu:

- 1) Keajekan pengukur ulang: kesesuaian hasil pengukuran pertama dan kedua dari suatu alat ukur terhadap kelompok yang sama.
- 2) Keajekan pengukur setara: kesesuaian hasil pengukuran ada 3 atau lebih alat ukur berdasarkan kompetensi kisi-kisi yang lama
- 3) Keajekan belah dua: kesesuaian antara hasil pengukuran belahan pertama dan belahan kedua dari alat ukur yang sama.

Untuk mengetahui koefisien reliabilitas tes/soal bentuk pilihan ganda digunakan rumus Kuder Richadson 21 (KR21) dengan persamaan sebagai berikut:

$$r_{11} = \frac{n}{n-1} \left[1 - \frac{\sum \bar{x}(n - \bar{x})}{n(SD)^2} \right]$$

Keterangan:

r_{11} = koefisien reliabilitas

n = Banyaknya butir soal

$\bar{x}$ = Rata-rata skor total

$(SD)^2$ = Varians skor total

Untuk menghitung reliabilitas tes bentuk uraian dapat dilakukan dengan

menggunakan rumus Alpha-Cronbach, yaitu:

$$r_{11} = \frac{n}{n-1} \left[1 - \frac{\sum_{i=1}^n s_i^2}{s_t^2} \right]$$

Keterangan:

r_{11} = koefisien reliabilitas

n = Banyaknya butir soal

s_i^2 = varian skor soal ke-i

s_t^2 = varian skor total

Nilai korelasi (r) yang diperoleh kemudian dikonsultasikan dengan nilai tabel koefisien korelasi dengan kategori sebagai berikut:

- 1) $0,90 \leq r \leq 1,00$ = korelasi sangat tinggi (sempurna)
- 2) $0,70 \leq r \leq 0,90$ = korelasi tinggi
- 3) $0,40 \leq r \leq 0,70$ = korelasi sedang
- 4) $0,20 \leq r \leq 0,40$ = korelasi rendah
- 5) $0,00 \leq r \leq 0,20$ = korelasi sangat rendah ¹⁰

Jika hasil perhitungan diperoleh besarnya $r = 0,76$, maka berarti 0,76

¹⁰ Depdiknas, *Juknis Penulisan Analisis Butir Soal di SMA*. Jakarta: Direktorat Pembinaan SMA, 2010, hal. 129

terletak antara $0,6 \leq r \leq 0,8$, berarti koefisien reliabilitas tes tergolong kategori tinggi.

Dengan menggunakan aplikasi bantuan SPSS, maka peneliti melakukan analisis reliabilitas pada instrumen tes hasil belajar dan mendapatkan hasil analisis reliabilitas sebagai berikut:

Tabel. 6.1

Hasil Analisis Reliabilitas Instrumen Tes Bentuk Soal Pilihan Ganda

Cronbach's Alpha	N Of Items	Reliabilitas
0,859	20	Reliabel

Berdasarkan Tabel 6.1 diketahui bahwa instrumen tes bentuk pilihan ganda yang terdiri atas 20 butir soal memperoleh nilai Cronbach's Alpha sebesar 0,859. Nilai tersebut menunjukkan bahwa instrumen memiliki tingkat reliabilitas yang tinggi sehingga termasuk dalam kategori reliabel. Dengan demikian, instrumen tes pilihan ganda memiliki konsistensi internal yang baik, sehingga setiap butir soal mampu mengukur konstruk yang sama secara konsisten.

Koefisien reliabilitas sebesar 0,859 menunjukkan bahwa variasi skor yang diperoleh peserta didik lebih banyak dipengaruhi oleh perbedaan kemampuan yang dimiliki daripada oleh kesalahan pengukuran (*measurement error*). Oleh karena itu, instrumen tes pilihan ganda layak digunakan sebagai alat evaluasi hasil belajar Al-Qur'an Hadis pada kelas Wustho dalam pembelajaran berbasis Metode Azza.

Tabel. 6.1

Hasil Analisis Reliabilitas Instrumen Tes Bentuk Soal Esai

Cronbach's Alpha	N Of Items	Reliabilitas
0,647	5	Tidak Reliabel

Berdasar hasil perhitungan analisis menggunakan aplikasi SPSS, maka tes yang di analisis reliabilitasnya adalah tes yang sah untuk instrumen tes hasil belajar dalam bentuk soal pilihan ganda, yaitu tes yang sudah valid atau memenuhi ke 3 kriteria analisis butir (r butir, indeks kesukaran, daya beda). Namun pada instrumen tes hasil belajar dalam bentuk soal esai belum menyatakan reliabilitas setelah dilakukan perhitungan dengan aplikasi bantuan yaitu SPSS, hal ini dikarenakan ada beberapa dari 3 aspek kriteria analisis butir soal yang kurang dalam memenuhi persyaratan untuk menjadi instrumen tes hasil belajar yang reliabel. Berdasarkan hal tersebut maka tidak semua soal dapat dipakai, untuk itu perlu diperbaiki kembali ulang pada beberapa soal.

c. Tingkat Kesukaran

Tingkat kesukaran soal adalah peluang untuk menjawab benar suatu soal pada tingkat kemampuan tertentu yang biasanya dinyatakan dalam bentuk suatu indeks. Indeks tingkat kesukaran ini pada umumnya dinyatakan dalam bentuk proporsi yang berkisar sebagai berikut:

- 1) Indeks kesukaran 0,00-0,30 tergolong sukar,
- 2) Indeks kesukaran 0,31-0,70 tergolong sedang, dan
- 3) Indeks kesukaran 0,71-1,00 tergolong mudah.

Semakin besar indeks tingkat kesukaran yang diperoleh dari hasil hitungan, maka semakin mudah soal itu. Sebaliknya apabila indeks tingkat

kesukaran kecil, maka artinya banyak siswa yang tidak dapat menjawab dengan benar. Perhitungan indeks tingkat kesukaran ini dilakukan untuk setiap nomor soal. Pada prinsipnya, skor rata-rata yang diperoleh peserta didik pada butir soal yang bersangkutan dinamakan tingkat kesukaran butir soal itu.

Fungsi analisis tingkat kesukaran butir soal ini berkaitan dengan tujuan diadakannya tes. Misalnya untuk keperluan ujian semester digunakan butir soal yang memiliki tingkat kesukaran sedang, untuk keperluan seleksi digunakan butir soal yang memiliki tingkat kesukaran tinggi, dan untuk keperluan diagnostik biasanya digunakan butir soal yang memiliki tingkat kesukaran rendah/mudah. Menganalisis tingkat kesukaran soal adalah untuk menentukan kualitas soal yang baik, mengetahui klasifikasi soal mudah, sedang, dan sukar/sulit. Rumus yang digunakan untuk soal yang berbentuk pilihan ganda adalah sebagai berikut:

$$p = \frac{B}{N}$$

Keterangan:

P = indeks kesukaran

B = jumlah siswa yang menjawab benar suatu soal

N = jumlah siswa yang mengikuti tes/mengerjakan soal tes

Selanjutnya untuk mengetahui tingkat kesukaran soal dengan bentuk uraian isian digunakan rumus yang diawali dengan mencari rata-rata suatu skor:

$$rata - rata = \frac{jumlah\ skor\ siswa\ pada\ suatu\ soal}{jumlah\ siswa\ yang\ mengikuti\ tes}$$

Dilanjutkan dengan rumus yang mencari tingkat kesukaran sebagai berikut:

$$tingkat\ kesukaran = \frac{rata - rata}{skor\ maksimum\ yang\ di\ tetapkan}$$

Butir soal dengan indeks tingkat kesukaran yang mudah dan terlalu sukar akan dibuang. Berdasarkan hasil perhitungan indeks kesukaran, maka kemungkinan tidak semua soal dapat memenuhi kriteria pengujian. Soal yang mempunyai indeks kesukaran sedang/cukup yang dapat diambil.

Dengan bantuan software aplikasi SPSS, maka penelitian melakukan analisis tingkat kesukaran instrumen tes hasil belajar dan mendapatkan hasil analisis sebagai berikut:

Tabel. 4.1

Hasil Analisis Tingkat Kesukaran Instrumen Tes Bentuk Pilihan Ganda

Pilihan Ganda	Mean	Kesukaran
Soal 1	0,75	Mudah
Soal 2	0,77	Mudah
Soal 3	0,20	Sukar
Soal 4	0,34	Sedang
Soal 5	0,58	Sedang
Soal 6	0,75	Mudah
Soal 7	0,80	Mudah

Soal 8	0,79	Mudah
Soal 9	0,88	Mudah
Soal 10	0,76	Mudah
Soal 11	0,58	Sedang
Soal 12	0,76	Mudah
Soal 13	0,59	Sedang
Soal 14	0,80	Mudah
Soal 15	0,70	Sedang
Soal 16	0,76	Mudah
Soal 17	0,76	Mudah
Soal 18	0,47	Sedang
Soal 19	0,74	Mudah
Soal 20	0,58	Sedang

Berdasarkan Tabel 4.1 diketahui bahwa dari 20 butir soal pilihan ganda, terdapat 12 butir soal (60%) yang termasuk dalam kategori mudah, yaitu soal nomor 1, 2, 6, 7, 8, 9, 10, 12, 14, 16, 17, dan 19. Selanjutnya terdapat 7 butir soal (35%) yang berada pada kategori sedang, yaitu soal nomor 4, 5, 11, 13, 15, 18, dan 20. Adapun 1 butir soal (5%), yaitu soal nomor 3, berada pada kategori sukar.

Nilai rata-rata (mean) tingkat kesukaran pada soal pilihan ganda berkisar antara 0,20 hingga 0,88. Soal nomor 9 memiliki nilai tingkat kesukaran tertinggi, yaitu sebesar 0,88, yang menunjukkan bahwa sebagian besar peserta didik mampu menjawab soal tersebut dengan benar sehingga tergolong sangat mudah. Sebaliknya, soal nomor 3 memperoleh nilai mean sebesar 0,20, yang menunjukkan bahwa hanya sebagian kecil peserta didik yang dapat menjawab dengan benar sehingga soal tersebut termasuk dalam kategori sukar.

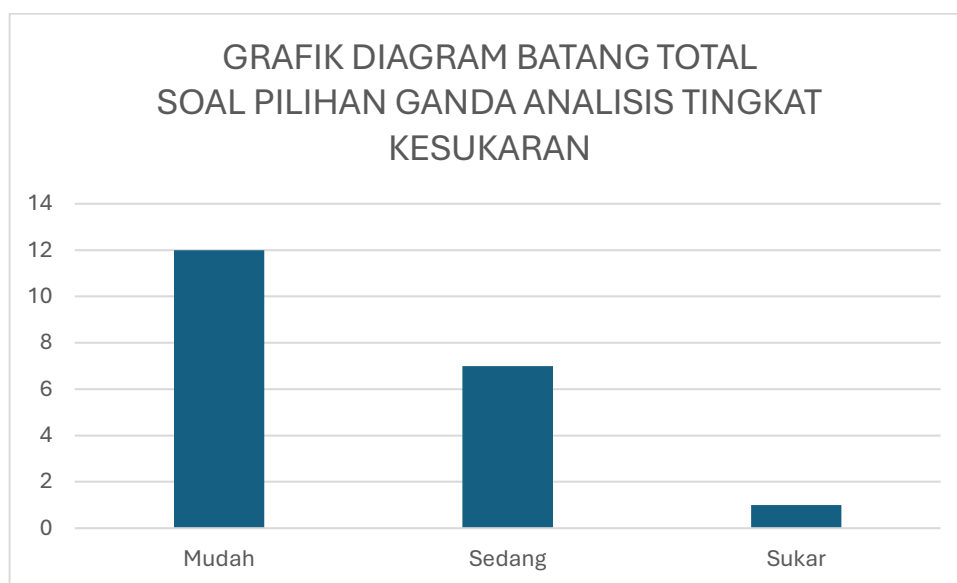
Secara keseluruhan, distribusi tingkat kesukaran pada instrumen tes pilihan ganda didominasi oleh soal berkategori mudah, diikuti oleh soal berkategori sedang, sedangkan jumlah soal berkategori sukar relatif sedikit. Komposisi tersebut

menunjukkan bahwa instrumen tes lebih banyak mengukur kemampuan dasar peserta didik dibandingkan kemampuan berpikir yang lebih tinggi.

Gambar. 4.1

Grafik Diagram Batang Total Hasil Analisis Tingkat Kesukaran Soal Pilihan

Ganda



Tabel. 4.2

Hasil Analisis Tingkat Kesukaran Instrumen Tes Bentuk Esai

Esai	Mean	Kesukaran
Soal 1	1,51	Sedang
Soal 2	2,19	Sedang
Soal 3	3,37	Mudah
Soal 4	3,52	Mudah
Soal 5	2,22	Sedang

Berdasarkan Tabel 4.2 diketahui bahwa dari 5 butir soal esai, terdapat 2 butir soal (40%) yang termasuk dalam kategori mudah, yaitu soal nomor 3 dan 4.

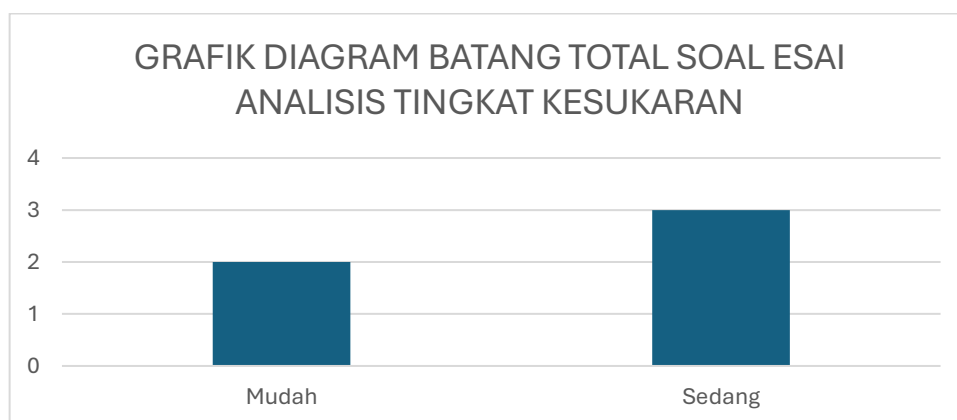
Sementara itu, 3 butir soal (60%), yaitu soal nomor 1, 2, dan 5, berada pada kategori sedang. Tidak terdapat soal esai yang termasuk dalam kategori sukar.

Nilai rata-rata (mean) tingkat kesukaran soal esai berada pada rentang 1,51 hingga 3,52. Soal nomor 4 memperoleh nilai mean tertinggi sebesar 3,52, sehingga termasuk kategori mudah karena sebagian besar peserta didik mampu memperoleh skor yang tinggi pada butir tersebut. Sebaliknya, soal nomor 1 memperoleh nilai mean sebesar 1,51, yang meskipun merupakan nilai terendah di antara soal esai, masih berada dalam kategori sedang.

Secara keseluruhan, hasil analisis menunjukkan bahwa instrumen tes bentuk esai didominasi oleh soal dengan tingkat kesukaran sedang, disertai beberapa soal yang tergolong mudah. Tidak ditemukannya soal berkategori sukar mengindikasikan bahwa seluruh soal esai masih dapat dijawab oleh sebagian besar peserta didik sesuai dengan kompetensi yang telah dipelajari selama proses pembelajaran menggunakan Metode Azza.

Gambar. 4.2

Grafik Diagram Batang Total Hasil Analisis Tingkat Kesukaran Soal Esai



Hasil analisis butir soal yang telah disusun oleh peneliti dengan menggunakan bantuan aplikasi SPSS diperoleh hasil bahwa dari keseluruhan 20 soal dalam bentuk pilihan ganda dan 5 soal bentuk esai yang dianalisis, menunjukkan sebanyak 12 soal atau 60% soal bentuk pilihan ganda dalam kategori mudah, 7 soal atau 35% soal bentuk pilihan ganda dalam kategori sedang, dan 1 soal atau 5% soal bentuk pilihan ganda dalam kategori sukar atau sulit. Adapun yang lainnya, yaitu 2 soal atau 40% soal bentuk esai dalam kategori mudah, dan 3 soal atau 60% soal bentuk esai dalam kategori sedang. Jika didistribusikan dalam indeks tingkat kesukaran, maka dapat dilihat hasilnya sebagai berikut:

Tabel 4.3

Kategori Butir Soal Dalam Indeks Kesukaran

No.	kategori	Nomor Butir		Jumlah		Persentase	
		pilihan ganda	esai	pilihan ganda	esai	pilihan ganda	esai
1	sukar	3	-	1	-	5%	-
2	sedang	4,5,11,13,15,18,20	1,2,5	7	3	35%	60%
3	mudah	1,2,6,7,8,9,10,12,14,16,17,19	3,4	12	2	60%	40%

d. Daya Pembeda

Daya pembeda soal adalah kemampuan suatu butir soal yang dapat membedakan antara siswa yang telah menguasai materi yang ditanyakan pada

instrumen tes dan siswa yang tidak/kurang/belum menguasai materi yang ditanyakan. Manfaat daya pembeda butir soal adalah sebagai berikut:

- 1) untuk meningkatkan mutu setiap butir soal melalui data empiriknya.
Berdasarkan dengan indeks daya pembeda, setiap butir soal dapat diketahui apakah butir soal itu baik, direvisi, atau ditolak.
- 2) untuk mengetahui seberapa jauh setiap butir soal dapat mendeteksi/membedakan kemampuan siswa, yaitu siswa yang telah memahami atau belum memahami materi yang telah diajarkan guru.

Apabila suatu butir soal tidak dapat membedakan kedua kemampuan siswa tersebut, maka butir soal dapat diduga bahwa kemungkinan bahwa kunci jawaban butir soal itu tidak tepat, butir soal itu memiliki 2 atau lebih kunci jawaban yang benar, kompetensi yang diukur tidak jelas, pengecoh tidak berfungsi, materi yang diajarkan terlalu sulit, sehingga banyak siswa yang menebak atau menjawab dengan asal, atau bisa juga sebagian siswa yang memahami materi yang ditanyakan berpikir ada yang salah informasi dalam butir soalnya.

Adapun rumus yang digunakan untuk mengetahui daya pembeda pada tes yang berbentuk pilihan ganda adalah sebagai berikut:

$$DP = \frac{BA - BB}{\frac{1}{2}N}$$

Atau

$$DP = \frac{2(BA - BB)}{N}$$

Keterangan:

DP = Daya pembeda soal

BA = Jumlah jawaban soal benar pada kelompok atas

BB = Jumlah jawaban soal benar pada kelompok bawah

N = Jumlah siswa yang mengerjakan tes.

Untuk mengetahui daya pembeda tes dengan bentuk uraian adalah dengan menggunakan rumus berikut:

$$DP = \frac{\text{mean kelompok atas} - \text{mean kelompok bawah}}{\text{skor maksimum soal}}$$

Hasil perhitungan dengan menggunakan rumus di atas dapat menggambarkan tingkat kemampuan soal dalam membedakan antara peserta didik yang sudah memahami materi yang diujikan dengan peserta didik yang belum/tidak memahami materi yang diujikan.

Kriteria indeks daya pembeda soal dibuat klasifikasinya dengan sebagai berikut:

- 1) 0,40 – 1,00 soal diterima dengan baik
- 2) 0,30 – 0,39 soal diterima tetapi perlu diperbaiki

- 3) 0,20 – 0,29 soal diperbaiki
 4) 0,19 – 0,00 soal tidak dipakai/dibuang.

Setelah diadakan perhitungan daya beda maka sejumlah soal yang disusun kemungkinan tidak semuanya dapat terpakai. Soal yang dapat terpakai adalah soal yang mempunyai daya beda cukup (0,21 – 0,40), baik (0,41 – 0,71), dan baik sekali (0,71 – 1,00).¹¹

Menggunakan aplikasi SPSS untuk melakukan analisis daya pembeda, maka peneliti mendapatkan hasil analisis sebagai berikut:

Tabel. 5.1

Hasil Analisis Daya Pembeda Instrumen Tes Bentuk Pilihan Ganda

Pilihan Ganda	Corrected Item-Total Correlation	Status
Soal 1	0,197	Diperbaiki
Soal 2	0,363	Diterima Perbaikan
Soal 3	0,424	Diterima
Soal 4	0,351	Diterima Perbaikan
Soal 5	0,650	Diterima
Soal 6	0,197	Diperbaiki
Soal 7	0,373	Diterima Perbaikan
Soal 8	0,708	Diterima
Soal 9	0,356	Diterima Perbaikan
Soal 10	0,574	Diterima
Soal 11	0,650	Diterima
Soal 12	0,489	Diterima
Soal 13	0,412	Diterima
Soal 14	0,373	Diterima Perbaikan
Soal 15	0,656	Diterima
Soal 16	0,489	Diterima

¹¹ Crocker, L. & Algina, J. *Introduction to Classical and Modern Test Theory*. New York: Holt, Rinehart and Winston, Inc. 1986, hal.315

Soal 17	0,574	Diterima
Soal 18	0,286	Diperbaiki
Soal 19	0,193	Diperbaiki
Soal 20	0,650	Diterima

Berdasarkan Tabel 5.1 diketahui bahwa dari 20 butir soal pilihan ganda, terdapat 11 butir soal (55%) yang termasuk dalam kategori diterima, yaitu soal nomor 3, 5, 8, 10, 11, 12, 13, 15, 16, 17, dan 20. Butir-butir soal tersebut memiliki nilai Corrected Item-Total Correlation berkisar antara 0,412 hingga 0,708, sehingga telah memenuhi kriteria daya pembeda yang baik dan dapat digunakan dalam tes hasil belajar tanpa memerlukan revisi yang berarti.

Selanjutnya terdapat 5 butir soal (25%) yang termasuk dalam kategori diterima dengan perbaikan, yaitu soal nomor 2, 4, 7, 9, dan 14. Nilai Corrected Item-Total Correlation pada kelompok ini berada pada rentang 0,351 sampai 0,373. Meskipun masih memenuhi syarat untuk digunakan, butir-butir soal tersebut memerlukan penyempurnaan, misalnya pada aspek redaksi soal, kejelasan alternatif jawaban, atau kesesuaian dengan indikator pembelajaran agar daya pembedanya menjadi lebih optimal.

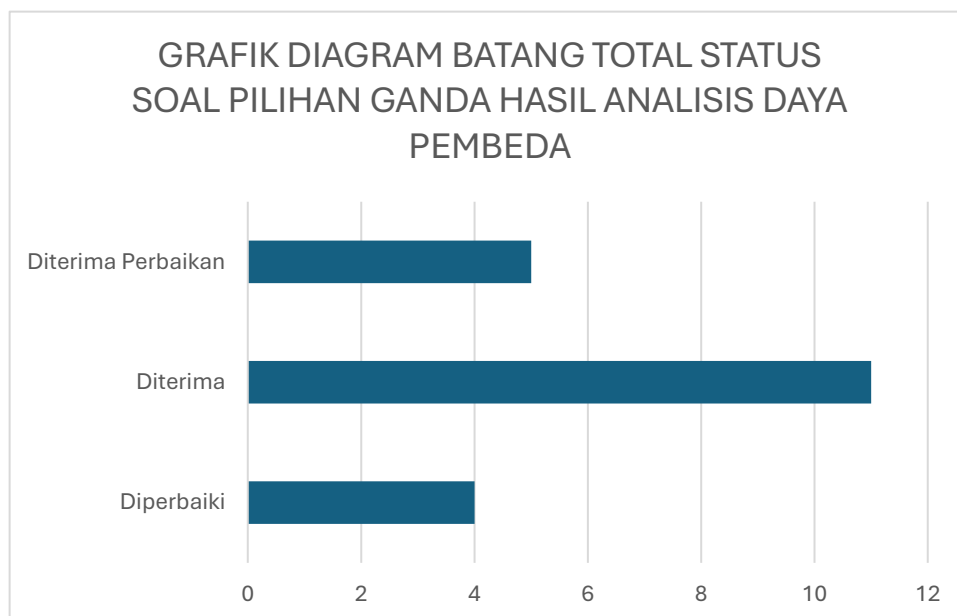
Adapun 4 butir soal (20%), yaitu soal nomor 1, 6, 18, dan 19, berada pada kategori diperbaiki. Nilai Corrected Item-Total Correlation pada keempat soal tersebut berkisar antara 0,193 hingga 0,286, yang menunjukkan bahwa kemampuan soal dalam membedakan peserta didik berkemampuan tinggi dan rendah masih rendah. Oleh karena itu, butir-butir soal tersebut memerlukan revisi sebelum digunakan kembali sebagai instrumen evaluasi.

Nilai Corrected Item-Total Correlation tertinggi pada instrumen pilihan ganda diperoleh oleh soal nomor 8, yaitu sebesar 0,708, yang menunjukkan bahwa butir soal tersebut memiliki kemampuan membedakan peserta didik secara sangat baik. Sebaliknya, nilai terendah diperoleh oleh soal nomor 19, yaitu sebesar 0,193, sehingga butir soal tersebut memerlukan perbaikan karena belum mampu memberikan daya pembeda yang optimal.

Secara keseluruhan, hasil analisis menunjukkan bahwa sebagian besar butir soal pilihan ganda telah memiliki daya pembeda yang baik, ditunjukkan dengan 55% soal berada pada kategori diterima, 25% diterima dengan perbaikan, dan hanya 20% yang masih memerlukan perbaikan.

Gambar. 5.1

Grafik Diagram Batang Total Hasil Analisis Daya Pembeda Soal Pilihan Ganda



Tabel. 5.2

Hasil Analisis Daya Pembeda Instrumen Tes Bentuk Esai

Esai	Corrected Item-Total Correlation	Daya
Soal 1	0,592	Diterima
Soal 2	0,567	Diterima
Soal 3	0,108	Diperbaiki
Soal 4	0,218	Diperbaiki
Soal 5	0,565	Diterima

Berdasarkan Tabel 5.2 diketahui bahwa dari 5 butir soal esai, terdapat 3 butir soal (60%) yang berada pada kategori diterima, yaitu soal nomor 1, 2, dan 5. Ketiga butir soal tersebut memiliki nilai Corrected Item-Total Correlation masing-masing sebesar 0,592, 0,567, dan 0,565, yang menunjukkan bahwa soal mampu membedakan peserta didik berdasarkan tingkat penguasaan materi secara baik.

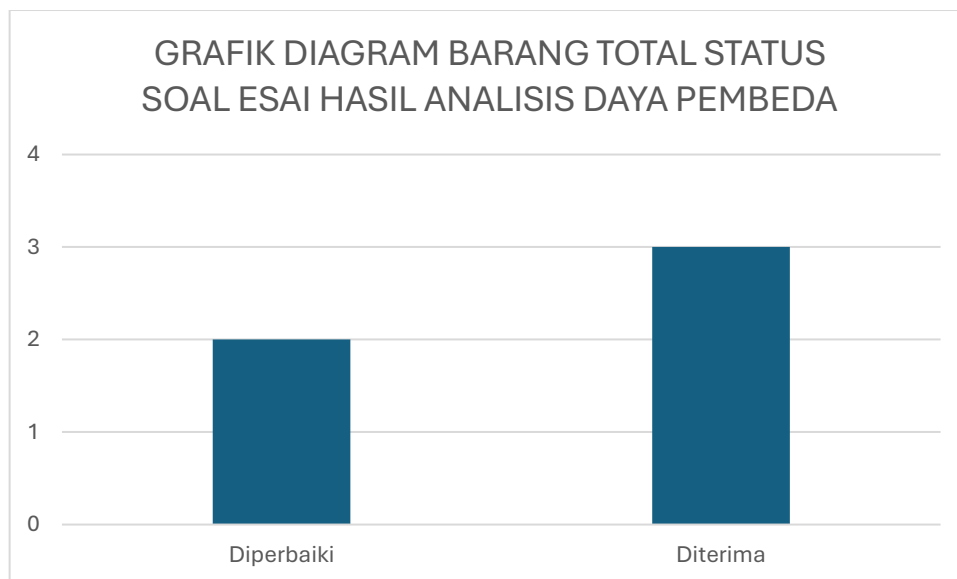
Sementara itu, terdapat 2 butir soal (40%), yaitu soal nomor 3 dan 4, yang termasuk dalam kategori diperbaiki. Kedua soal tersebut memiliki nilai Corrected Item-Total Correlation sebesar 0,108 dan 0,218. Nilai tersebut menunjukkan bahwa kemampuan kedua butir soal dalam membedakan peserta didik yang memiliki kemampuan tinggi dan rendah masih belum optimal, sehingga diperlukan perbaikan terhadap isi, redaksi, atau pedoman penskorannya.

Nilai Corrected Item-Total Correlation tertinggi pada instrumen esai diperoleh oleh soal nomor 1, yaitu sebesar 0,592, sedangkan nilai terendah terdapat pada soal nomor 3, yaitu sebesar 0,108. Perbedaan tersebut menunjukkan adanya variasi kemampuan masing-masing butir soal dalam mengukur perbedaan

kemampuan peserta didik.

Gambar. 5.2

Grafik Diagram Batang Total Hasil Analisis Daya Pembeda Soal Esai



Berdasarkan hasil perhitungan analisis dengan bantuan aplikasi SPSS terhadap butir soal yang telah disusun oleh peneliti untuk mata pelajaran al Qur'an dan Hadits pada halaqoh Wustho di TPA kota Banjarmasin dan kota Banjarbaru dengan menggunakan metode pembelajaran Azza, yang dianalisis dari segi daya pembeda maka diperoleh hasil dari keseluruhan 20 butir soal dalam bentuk pilihan ganda dan 5 soal dalam bentuk esai yang dianalisis menunjukkan sebanyak 4 soal atau 20% soal dalam bentuk pilihan ganda dalam kategori harus menerima perbaikan, sebanyak 5 soal atau 25% soal bentuk pilihan ganda dalam kategori diterima namun setelah diperbaiki, sebanyak 11 soal atau 55% soal bentuk pilihan ganda dalam kategori diterima dan dapat digunakan dalam tes hasil belajar,

sebanyak 2 soal atau 40% soal bentuk esai dalam kategori harus diperbaiki dan yang terakhir sebanyak 3 soal atau 60% soal bentuk esai dalam kategori diterima dan dalam digunakan untuk sebagai instrumen tes hasil belajar pendidikan agama Islam dalam mata pelajaran al Qur'an dan Hadits pada kelas Wustho di TPA kota Banjarmasin dan Banjarbaru yang menggunakan metode pembelajarannya metode Azza.

e. Efektivitas Pengecoh

Analisis distraktor atau pengecoh diperlukan hanya untuk pembuat soal. Selain menghitung indeks kesukaran dan daya beda dalam analisis butir soal juga perlu diketahui apakah distraktor atau pengecoh yang disediakan tepat atau tidak benar. Apakah semua pilihan yang disediakan dipilih semua karena dianggap betul, jawaban terkumpul pada pilihan tertentu atau pilihan yang sama sekali tidak ada pemilihnya. Dengan memeriksa pola pilihan jawaban, dapat ditentukan dalam beberapa hal, yaitu:

- 1) Berapa jumlah subjek yang menjawab betul
- 2) Distraktor atau pengecoh mana yang terlalu jelas atau mencolok sehingga sangat sedikit yang terkecoh untuk memilihnya
- 3) Pengecoh mana yang justru menyesatkan subjek yang termasuk kelompok tinggi yang seharusnya tidak terkecoh
- 4) Pengecoh mana yang dapat menarik bagi subjek kelompok rendah, tetapi tidak cukup menarik bagi subjek kelompok tinggi.

Dengan bantuan software excel, peneliti melakukan penghitungan analisis

efektivitas pengecoh dan mendapatkan hasil sebagai berikut:

Tabel. 6.1

Jumlah Siswa Yang Memilih Jawaban

ITEM		A	B	C	D
Soal Pilihan Ganda	Soal 1	10	8	75	7
	Soal 2	7	9	77	7
	Soal 3	20	29	19	32
	Soal 4	34	20	34	12
	Soal 5	58	28	6	8
	Soal 6	11	75	6	8
	Soal 7	7	5	8	80
	Soal 8	8	8	79	5
	Soal 9	7	7	78	8
	Soal 10	7	9	76	8
	Soal 11	18	58	17	7
	Soal 12	11	76	6	7
	Soal 13	8	59	19	14
	Soal 14	9	80	6	5
	Soal 15	14	70	10	6
	Soal 16	9	9	76	6
	Soal 17	5	76	12	7
	Soal 18	8	47	39	6
	Soal 19	10	6	10	74
	Soal 20	58	29	7	6

Dari hasil rekapitulasi jawaban siswa berikut, dapat dilihat bahwa pilihan jawaban yang disediakan oleh peneliti terdapat perbedaan pada jumlah. Maka selanjutnya peneliti membuat tabel presentase agar untuk mengetahui berapa persen tiap pengecoh yang efektif dalam suatu tiap butir soal.

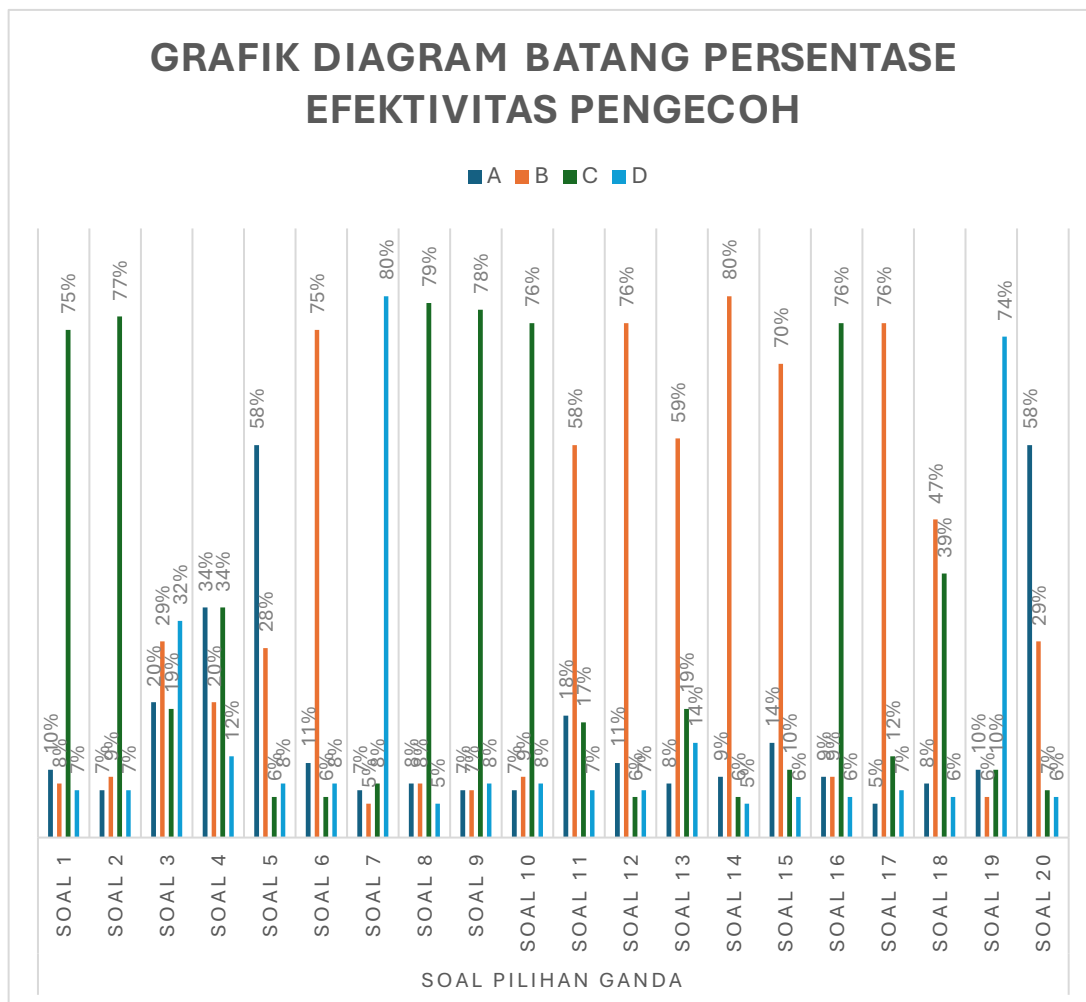
Tabel. 6.2

Persentase Efektivitas Pengecoh

ITEM		A	B	C	D
Soal Pilihan Ganda	Soal 1	10%	8%	75%	7%
	Soal 2	7%	9%	77%	7%
	Soal 3	20%	29%	19%	32%
	Soal 4	34%	20%	34%	12%
	Soal 5	58%	28%	6%	8%
	Soal 6	11%	75%	6%	8%
	Soal 7	7%	5%	8%	80%
	Soal 8	8%	8%	79%	5%
	Soal 9	7%	7%	78%	8%
	Soal 10	7%	9%	76%	8%
	Soal 11	18%	58%	17%	7%
	Soal 12	11%	76%	6%	7%
	Soal 13	8%	59%	19%	14%
	Soal 14	9%	80%	6%	5%
	Soal 15	14%	70%	10%	6%
	Soal 16	9%	9%	76%	6%
	Soal 17	5%	76%	12%	7%
	Soal 18	8%	47%	39%	6%
	Soal 19	10%	6%	10%	74%
	Soal 20	58%	29%	7%	6%

Gambar. 6.1

Grafik Diagram Batang Persentase Efektivitas Pengecoh



Dari tabel gambar grafik di atas yang telah di buat oleh peneliti, dapat dilihat seberapa persen efektivitas suatu pengecoh tiap butir soal, dan peneliti membuat keterangan tingkat kesukaran dalam bentuk tabel sebagai berikut:

Tabel. 6.3

Keterangan Efektivitas Pengecoh

ITEM		A	B	C	D
Soal pilihan ganda	Soal 1	Efektif	Efektif	Efektif	Efektif
	Soal 2	Efektif	Efektif	Efektif	Efektif
	Soal 3	Efektif	Efektif	Efektif	Efektif
	Soal 4	Efektif	Efektif	Efektif	Efektif
	Soal 5	Efektif	Efektif	Efektif	Efektif
	Soal 6	Efektif	Efektif	Efektif	Efektif
	Soal 7	Efektif	Efektif	Efektif	Efektif
	Soal 8	Efektif	Efektif	Efektif	Efektif
	Soal 9	Efektif	Efektif	Efektif	Efektif
	Soal 10	Efektif	Efektif	Efektif	Efektif
	Soal 11	Efektif	Efektif	Efektif	Efektif
	Soal 12	Efektif	Efektif	Efektif	Efektif
	Soal 13	Efektif	Efektif	Efektif	Efektif
	Soal 14	Efektif	Efektif	Efektif	Efektif
	Soal 15	Efektif	Efektif	Efektif	Efektif
	Soal 16	Efektif	Efektif	Efektif	Efektif
	Soal 17	Efektif	Efektif	Efektif	Efektif
	Soal 18	Efektif	Efektif	Efektif	Efektif
	Soal 19	Efektif	Efektif	Efektif	Efektif
	Soal 20	Efektif	Efektif	Efektif	Efektif

Dengan mengetahui hasil dari analisis efektivitas pengecoh di atas yang telah di ubah menjadi bentuk persen, maka dapat dilihat bahwa efektivitas pengecoh tiap butir soal terdapat bekerja dengan baik, maka dalam setiap butir soal yang terdapat pengecoh pada setiap alternatif jawaban yang telah disediakan efektif serta digunakan dalam instrumen tes hasil belajar al qur'an dan hadits dengan metode Azza.

C. Pembahasan Hasil Penelitian

1. Penyusunan Instrumen Tes Hasil Belajar

Dalam hal penyusunan dan pengembangan Instrumen tes hasil belajar sebagai alat ukur, Sumadi Suryabrata memberikan perhatian besar yang tertulis

dalam bukunya yang menyatakan bahwa kegiatan penyusunan dan pengembangan spesifikasi alat ukur atau yang dimaksud adalah instrumen tes pada dasarnya adalah proses pengambilan keputusan. Setiap keputusan harus diambil berdasarkan atas pertimbangan mengenai beberapa hal. Pertimbangan-pertimbangan itu pada umumnya mengenai hal-hal berikut:¹²

- a) Menentukan wilayah yang akan dikenai pengukuran atau menentukan wilayah yang akan di laksanakan tes.
- b) Menentukan dasar konseptual atau dasar teoretis yang akan digunakan sebagai landasan
- c) Menentukan subjek yang akan dikenai pengukuran
- d) Menentukan tujuan pengukuran
- e) Menentukan materi alat ukur
- f) Menentukan taraf kesukaran alat ukur dan masing-masing bagiannya
- g) Menentukan taraf kesukaran soal dan distribusinya
- h) Menyusun kisi-kisi atau "*test blue-print*"
- i) Merencanakan tugas-tugas untuk para penulis soal
- j) Merencanakan perakitan soal
- k) Merencanakan jadwal penerbitan soal

Adapun dalam hal pengembangan tes hasil belajar, pada umumnya menurut Sumadi Suryabrata meliputi langkah-langkah sebagai berikut:¹³

¹² Sumadi Suryabrata, *Pengembangan Alat Ukur Psikologis*, Yogyakarta: CV. Andi Offset, 1998, hal. 48

¹³ *Ibid...* hal. 68

- a) Pengembangan spesifikasi tes
- b) Penulisan tes
- c) Penelaahan tes
- d) Perakitan tes
- e) Uji coba tes
- f) Analisis butir soal
- g) Seleksi dan perakitan soal kembali
- h) Pencetakan tes
- i) Administrasi tes bentuk akhir
- j) Penyusunan skala dan norma

Berdasarkan pendapat di atas, peneliti telah melakukan beberapa hal yang telah dipaparkan di atas, namun tetap masih jauh ada beberapa langkah yang belum terlaksanakan dan juga ada yang masih perlu perbaikan. Terkhusus pada langkah seleksi dan perakitan kembali butir soal yang telah di analisis, sehingga mendapatkan hasil yang baik dan dapat butir soal dapat di gunakan dengan tujuan yang semestinya.

Adapun hasil analisis tiap butir soal instrumen tes hasil belajar yang telah disusun oleh peneliti, yang dianalisis dari segi validitas, tingkat kesukaran, daya pembeda, efektivitas pengecoh dan reliabilitas melalui bantuan software SPSS maka diperoleh hasil dari kelas Wustho tiap TPA dalam pendidikan Al Qur'an dan Hadits dari keseluruhan 20 butir soal pilihan ganda dan 5 soal esai yang dianalisis.

2. Validitas

Instrumen tes hasil belajar Al-Qur'an dan Hadits dalam penelitian ini terdiri atas 20 butir soal pilihan ganda dan 5 butir soal esai. Pengujian validitas dilakukan menggunakan aplikasi SPSS dengan taraf signifikansi 5% ($\alpha = 0,05$). Kriteria validitas ditentukan berdasarkan nilai r tabel sebesar 0,30, yang digunakan sebagai batas minimal untuk menyatakan butir valid.

Berdasarkan hasil analisis, diketahui bahwa seluruh butir soal, baik bentuk pilihan ganda maupun esai, dinyatakan valid, dengan nilai r hitung tertinggi 0,751 dan terendah 0,284. Meskipun terdapat satu butir soal dengan nilai r hitung di bawah batas ideal, yaitu 0,284, namun karena perbedaannya relatif kecil dan butir tersebut memiliki relevansi isi yang kuat terhadap indikator pembelajaran, maka masih dianggap layak untuk dipertahankan dengan perbaikan redaksional.

Hasil ini menunjukkan bahwa setiap butir soal mampu mengukur kompetensi yang diharapkan sesuai dengan indikator hasil belajar pada mata pelajaran Al-Qur'an dan Hadits. Dengan demikian, instrumen tes ini dapat dikatakan memiliki validitas yang baik dan dapat digunakan untuk mengukur hasil belajar peserta didik secara representatif.

Keterkaitan dengan Penelitian Sebelumnya hasil validitas yang diperoleh sejalan dengan penelitian Sarea, Dewantara, dan Sabbar, yang menegaskan bahwa kualitas butir merupakan penentu utama dari validitas dan reliabilitas instrumen tes. Mereka menulis *"Item quality is a critical factor in determining the validity and*

reliability of test instruments".¹⁴ Makna kutipan tersebut adalah bahwa setiap butir soal yang dirancang dengan kualitas isi dan struktur yang baik akan menghasilkan alat ukur yang sahih dan konsisten. Dalam penelitian Sarea et al., dari 20 butir soal yang dikembangkan, 18 dinyatakan valid setelah melalui proses validasi ahli, sementara dua butir direvisi karena masalah redaksi dan distraktor yang tidak efektif.

Kondisi ini selaras dengan hasil penelitian ini, di mana seluruh butir soal dinyatakan valid, meskipun terdapat satu butir dengan nilai korelasi mendekati batas minimal (0,284) yang tetap dipertahankan karena relevansi kontennya.

Penelitian Mawangir juga memberikan dukungan yang kuat terhadap temuan ini. Ia menjelaskan bahwa "*The results prove that overall, the items have a good level of difficulty since they are in the range of 0.3–0.7*".¹⁵

Meskipun pernyataan ini menyinggung tingkat kesukaran, Mawangir mengaitkannya dengan validitas empiris, yaitu bahwa butir dengan tingkat kesukaran ideal umumnya juga memiliki korelasi tinggi terhadap skor total tes. Dalam penelitiannya, semua butir yang memiliki korelasi di atas r tabel (0,197) dinyatakan valid secara empiris. Dengan demikian, validitas yang tinggi dalam penelitian ini menunjukkan bahwa butir soal telah sesuai dengan prinsip

¹⁴ Sarea, M. S., Dewantara, A. H., & Sabbar, S. D.. "*Development of a Comprehensive Instrument for the Islamic Education Subject Examination. Tafkir: Interdisciplinary*" Journal of Islamic Education, 6(3), 2025, hal. 758

¹⁵ Mawangir, M.. "*The Development of Higher Order Thinking Skill (HOTS) Test Instrument on the Fiqh Subject: The Case of a State Islamic Senior High School in West Bangka Regency.*" Journal of Higher Education Theory and Practice, 22(7), 2022, hal. 167,

penyusunan tes berbasis teori pengukuran klasik sebagaimana dijelaskan oleh Mawangir.

Selain itu, hasil penelitian ini juga sejalan dengan penelitian Reza, Fakhruddin, dan Isnaeni yang menekankan pentingnya validitas konstruk dan isi dalam pengembangan instrumen berbasis nilai-nilai Islam. Mereka menyatakan *“The use of the Rasch Model as an analytical tool in the development of this test is very important because it is able to ensure the validity and reliability of the test according to the student's ability”*.¹⁶ Maknanya, validitas instrumen tidak hanya ditentukan oleh korelasi statistik, tetapi juga oleh kesesuaian antara butir soal dan tingkat kemampuan peserta didik. Pernyataan ini memperkuat hasil penelitian ini bahwa instrumen yang dikembangkan telah sesuai dengan tingkat kompetensi siswa dalam memahami materi Al-Qur'an dan Hadits, baik dari segi substansi isi maupun kesesuaian tingkat kesulitan.

Secara keseluruhan, hasil validitas penelitian ini memperlihatkan bahwa seluruh butir soal telah memenuhi kriteria empiris dan konseptual untuk dinyatakan valid. Temuan ini konsisten dengan teori yang dikemukakan oleh Downing dan Haladyna bahwa validitas merupakan aspek paling fundamental dalam pengukuran hasil belajar, karena mencerminkan sejauh mana suatu instrumen benar-benar mengukur apa yang seharusnya diukur.¹⁷

¹⁶ Reza, M., Fakhruddin, & Isnaeni, W. *“Integration of Islamic Religious Education in the Development of Reading and Numerical Literacy Tests: An Analysis of the Rasch Model.”* *Tafkir: Interdisciplinary Journal of Islamic Education*, 6(4),2024, hal. 1457

¹⁷ Downing, S. M., & Haladyna, T. M.. *Test item development: Validity evidence from quality items. Applied Measurement in Education*, 10(1), 1997, hal. 61

Data hasil analisis yang peneliti lakukan menunjukkan bahwa seluruh soal dalam bentuk pilihan ganda dan esai valid secara validitas. Hal ini sesuai dengan yang telah disampaikan oleh para ahli seperti salah satunya Ata Nayla dan Ani Widayati dalam tulisan jurnalnya yang menyatakan bahwa “suatu tes dikatakan memiliki validitas yang baik apabila tes tersebut dapat mengukur objek yang seharusnya diukur dan sesuai dengan kriteria tertentu”. Dengan memperhatikan beberapa hal yang telah dipaparkan oleh Zainal Arifin dalam bukunya yang telah peneliti kutip pada bab 2 yang menjadi salah satu faktor bahwa instrumen tes hasil belajar dinyatakan valid.

Validitas yang tinggi juga menunjukkan bahwa instrumen yang dikembangkan dalam penelitian ini dapat digunakan sebagai alat evaluasi pembelajaran yang sah, mendukung peningkatan kualitas assesmen PAI, dan dapat dijadikan model untuk pengembangan tes hasil belajar di TPA yang bermetode Azza kota Banjarmasin dan Banjarbaru. Namun tetap harus dilakukan pengembangan pada instrumen tes hasil belajar ini, karena melihat pada aspek lainnya tidak hanya pada validitas saja.

3. Tingkat Kesukaran

tingkat kesukaran butir soal juga merupakan salah satu indikator penting dalam analisis kualitas instrumen tes. Parameter ini digunakan untuk melihat sejauh mana suatu butir soal dapat dikerjakan oleh peserta tes, serta untuk memastikan bahwa distribusi kesukaran mencerminkan karakteristik peserta didik yang heterogen. Dalam penelitian ini, analisis tingkat kesukaran dilakukan terhadap 20

butir soal dalam bentuk pilihan ganda dan 5 soal dalam bentuk esai menggunakan hasil pengolahan data melalui aplikasi SPSS.

Hasil analisis menunjukkan bahwa mayoritas soal berada pada kategori sedang, disusul kategori mudah dan sukar mengingat bahwa tujuan instrumen tes hasil belajar ini untuk mengetahui sejauh mana pemahaman dan ingatan para peserta didik terhadap materi yang telah diajarkan. Distribusi ini menunjukkan bahwa instrumen tes hasil belajar telah memenuhi prinsip penyebaran tingkat kesukaran yang proporsional sesuai dengan tujuannya, yaitu tidak didominasi oleh soal yang terlalu mudah maupun terlalu sukar.

Temuan penelitian ini sejalan dengan standar yang digunakan dalam berbagai penelitian pengembangan instrumen tes hasil belajar PAI. Misalnya, dalam penelitian Sarea et al. (2025) menunjukkan bahwa instrumen tes hasil belajar yang baik umumnya memiliki distribusi tingkat kesukaran yang seimbang, dengan komposisi 60% butir soal dalam kategori sedang, 25% sukar, dan 15% mudah pada instrumen tes komprehensif PAI yang mereka kembangkan. *“The item difficulty index showed a more balanced distribution: 25% difficult, 60% moderate, and 15% easy.”*¹⁸

Komposisi kesukaran tersebut menjadi rujukan bahwa suatu instrumen yang baik tidak hanya berfokus pada soal mudah atau sukar, akan tetapi menempatkan butir kategori sedang sebagai mayoritas karena lebih representatif

¹⁸ Sarea, M. S., Dewantara, A. H., & Sabbar, S. D.. ”Development of a Comprehensive Instrument for the Islamic Education Subject Examination. *Tafkir: Interdisciplinary*” Journal of Islamic Education, 6(3), 2025, hal. 758

dalam membedakan variasi kemampuan peserta tes yang telah disesuaikan oleh peneliti dengan tujuan penyusunan instrumen tes hasil belajar tersebut.

Selain itu, menurut Reza et al. (2024) yang telah melakukan suatu pengembangan pada tes literasi numerasi dan membaca berbasis PAI melalui analisis Rasch Model menunjukkan bahwa instrumen yang baik harus memiliki variasi tingkat kesukaran yang berlapis, mulai dari sangat sukar hingga sangat mudah.¹⁹ Penelitian tersebut melaporkan hasil temuannya, adanya penyebaran 2 item sangat sukar, 7 sukar, 31 sedang, 29 mudah, dan 12 sangat mudah. Temuan tersebut memperkuat suatu prinsip bahwa instrumen yang baik tidak bersifat homogen, tetapi harus mampu mencakup spektrum kemampuan peserta didik secara luas.

Dengan begitu, hasil tingkat kesukaran instrumen ini, yang telah didominasi oleh butir soal dengan kategori sedang dengan proporsi tertentu pada kategori mudah dan sukar, dapat dikatakan bahwa konsisten dengan temuan Reza, yang menekankan bahwa variasi kesukaran merupakan elemen penting dalam memastikan instrumen mampu mengukur kemampuan peserta tes secara komprehensif.

Oleh karena itu, untuk pengembangan lebih lanjut, instrumen tes hasil belajar ini perlu dirancang ulang dengan menambahkan butir soal dengan tingkat kesukarannya dalam kategori sedang dan sukar, agar lebih sesuai dengan standar

¹⁹ Reza, M., Fakhruddin, & Isnaeni, W. “*Integration of Islamic Religious Education in the Development of Reading and Numerical Literacy Tests: An Analysis of the Rasch Model.*” *Tafkir: Interdisciplinary Journal of Islamic Education*, 6(4),2024, hal. 1476

kualitas instrumen tes hasil belajar PAI, meski tingkat kesukaran instrumen tes hasil belajar yang telah disusun oleh peneliti dapat dikategorikan cukup baik.

Adapun pendapat lain tentang analisis tingkat kesukaran pada tiap butir soal, Sumadi Suryabrata menyatakan dalam bukunya bahwa, rentang atau distribusi kesukaran soal dipengaruhi oleh tujuan testing. Apabila tes yang memiliki tujuan sebagai tes diagnostik, maka perlu mempunyai taraf kesukaran yang luas, karena dengan taraf kesukaran yang luas diharapkan dapat diidentifikasi bagian-bagian tes mana yang dikuasai secara baik oleh peserta didik, dan bagian-bagian mana yang merupakan kesukaran bagi peserta didik, baik secara kelompok maupun individual.²⁰ Hal ini selaras dengan tujuan peneliti dalam penyusunan tes hasil belajar, yang mana tujuan dalam penyusunan instrumen tes hasil belajar ini untuk mengetahui bagian mana peserta didik yang dikuasai dengan baik dan bagian mana yang merupakan kesukaran bagi peserta didik dengan berbekal bahan materi yang telah diajarkan dengan menggunakan metode Azza. Sehingga menghasilkan tingkat kesukaran soal yang tidak banyak memiliki tingkat sukar atau sulit.

4. Daya Pembeda

Hasil analisis daya pembeda yang dilakukan peneliti pada tiap butir soal terdapat 4 soal atau 20% soal dalam bentuk pilihan ganda dalam kategori harus menerima perbaikan, sebanyak 5 soal atau 25% soal bentuk pilihan ganda dalam kategori diterima namun setelah diperbaiki ulang, sebanyak 11 soal atau 55% soal

²⁰ Sumadi Suryabrata, *Pengembangan Alat Ukur Psikologis*, Yogyakarta: CV. Andi Offset, 1998, hal. 70

bentuk pilihan ganda dalam kategori diterima dan dapat digunakan dalam tes hasil belajar, sebanyak 2 soal atau 40% soal bentuk esai dalam kategori harus diperbaiki dan yang terakhir sebanyak 3 soal atau 60% soal bentuk esai dalam kategori diterima. Dengan masih banyaknya butir soal dengan tingkat daya beda yang kurang, maka perlu adanya perbaikan agar dapat digunakan dengan sesuai dengan tujuannya yaitu dapat membedakan peserta didik yang memiliki pemahaman materi yang baik dan yang memiliki pemahaman materi yang kurang baik. Hal ini selaras pendapat Kunandar yang telah peneliti kutip pada bab 2 di atas, bahwa tingkat daya soal adalah kemampuan suatu soal untuk membedakan antara peserta didik yang sudah menguasai materi. Sama seperti halnya Sudijono yang menyatakan, bahwa suatu daya pembeda item butir soal adalah kemampuan suatu butir item tes hasil belajar yang dapat membedakan antara peserta didik yang berkemampuan tinggi dan yang berkemampuan rendah.

5. Efektivitas Pengecoh

Hasil analisis efektivitas pengecoh yang dilakukan peneliti terhadap tiap butir soal menunjukkan bahwa terdapat beberapa pengecoh yang tidak efektif, atau masih kurangnya pengecoh jawaban tersebut terpilih oleh peserta didik yang menggunakan pendapat dari Sumadi Suryabrata dalam menganalisis efektivitas pengecoh tiap butir soal tersebut. Hal ini disebabkan kurangnya keahlian peneliti dalam menyusun pengecoh yang kemungkinan peserta didik akan memilih tersebut. Banyak alternatif jawaban bukan berarti banyak kemungkinan pengecoh efektif terlaksanakan, seperti halnya yang dinyatakan oleh Sumadi Suryabrata bahwa

mencari pengecoh keempat, kelima sering kali sukar, menimbulkan frustrasi, dan menghasilkan pengecoh yang kurang berfungsi.²¹

6. Reliabilitas

Selain hasil analisis yang telah dibahas di atas, terdapat juga hasil analisis reliabilitas yang merupakan indikator penting untuk menentukan sebuah instrumen tes hasil belajar konsistensi internal dalam mengukur konstruk yang sama secara stabil.

Hasil analisis yang terakhir peneliti lakukan adalah analisis reliabilitas terhadap instrumen tes hasil belajar dengan bantuan software SPSS yang mana hasilnya 20 butir soal dalam bentuk pilihan ganda dinyatakan reliabel dan 5 butir soal dalam bentuk esai tidak reliabel. Hal ini disadari oleh peneliti setelah mengetahui suatu pernyataan yang menyatakan bahwa secara teori bahwa banyaknya pengecoh maka makin tinggi reliabilitas.²² beliau juga menganalog kan dengan makin banyaknya jumlah soal suatu tes, maka makin tinggi reliabilitasnya.

Hasil analisis perhitungan reliabilitas pada instrumen tes dalam bentuk pilihan ganda pada penelitian ini menunjukkan nilai Cronbach's Alpha sebesar 0,859 seperti yang telah peneliti paparkan sebelumnya, yang berada pada kategori sangat tinggi. Nilai ini mengindikasikan bahwa butir-butir soal pada instrumen pilihan ganda memiliki konsistensi internal yang kuat dan saling mendukung dalam

²¹ Sumadi Suryabrata, *Pengembangan Alat Ukur Psikologis*, Yogyakarta: CV. Andi Offset, 1998, hal.93

²² *Ibid...* hal.92

mengukur kompetensi yang sama.

Temuan ini sesuai dengan standar yang digunakan dalam penelitian Reza et al. (2024), yang melaporkan bahwa instrumen dengan Cronbach's Alpha di atas 0,80 termasuk dalam kategori "*very good reliability*".²³

Adapun instrumen tes hasil belajar dalam bentuk esai memperoleh nilai reliabilitas 0,674 yang berada sedikit di bawah batas idealnya yaitu 0,70. Secara umum, nilai ini dapat dikategorikan cukup, namun belum optimal. Kondisi ini dapat dipahami mengingatkan bentuk soal esai cenderung memiliki tingkat subjektivitas penilaian yang lebih tinggi dibandingkan soal dalam bentuk pilihan ganda, sehingga kemungkinan terjadinya variasi skor antar penilaian juga lebih besar.

²³ Reza, M., Fakhruddin, & Isnaeni, W. "*Integration of Islamic Religious Education in the Development of Reading and Numerical Literacy Tests: An Analysis of the Rasch Model.*" *Tafkir: Interdisciplinary Journal of Islamic Education*, 6(4), 2024, hal. 1476