

## **ANALISIS BUTIR SOAL EVALUASI PEMBELAJARAN PENDIDIKAN AGAMA ISLAM MENGGUNAKAN PROGRAM ANATES**

**Elviana<sup>1)</sup>**

**<sup>1)</sup>Universitas Islam Negeri Ar- Raniry Banda Aceh, Indonesia**  
email: elviana.baharuddin@ar-raniry.ac.id

### ***Abstract***

*The purpose of this study is to determine the characteristics of PAI Learning Evaluation items using ANA Test program. This research uses a quantitative descriptive approach and documentation technique in data collection. The document was in form of final semester exam questions which were 25 items of multiple choice test. The number of documents were 26 similar to the number of students who responded to the problem. Data analysis used ANA Test version 4 program. The results showed that the power difference was low, only 20% items could distinguish between upper and lower groups, the level of difficulty was moderate, there were 64%, the test validity was 0.24 in low category, only 24% was significant, the reliability was low and the effectiveness of the distractor was only 32 % functioning properly.*

***Keywords:*** *Item Characteristics, PAI Learning Evaluation Questions, Ana Test Program*

### **Abstrak**

Tujuan penelitian ini untuk mengetahui karakteristik butir-butir soal Evaluasi Pembelajaran PAI dengan menggunakan program anates. Penelitian ini menggunakan pendekatan deskriptif kuantitatif dan teknik pengumpulan data yang dilakukan dengan teknik dokumen. Dokumen berupa soal ujian Akhir Semester berjumlah 25 butir, bentuk tes yaitu pilihan ganda. Jumlah dokumen ada 26 sejumlah mahasiswa yang merespon soal tersebut. Analisis data menggunakan program anates versi 4. Hasil penelitian menunjukkan bahwa daya beda rendah, hanya 20% butir yang bisa membedakan kelompok atas dan bawah, taraf kesukaran sedang, ada 64%, validitas tes sebesar 0,24

kategori rendah hanya 24 % yang signifikan, reliabilitas rendah dan efektifitas distraktor hanya 32 % yang berfungsi dengan baik.

**Kata Kunci:** *Karakteristik Butir, Soal Evaluasi Pembelajaran PAI, Program Anates*

## **PENDAHULUAN**

Evaluasi merupakan proses mengukur dan menilai, dua hal ini sangat erat kaitannya dengan sistem evaluasi. Pengukuran merupakan sebuah proses perbandingan antara objek ukur dengan alat ukur tertentu yang dilakukan secara sistematis. Sedangkan penilaian merupakan interpretasi dari hasil mengukur. Hasil mengukur diperoleh dari instrumen yang digunakan atau berupa tes. Evaluasi dapat dilakukan dengan teknik tes dan non Tes, teknik tes sering dilakukan oleh dosen yaitu quiz, Ujian Tengah Semester (UTS), dan Ujian Akhir semester (UAS). Bentuk tes yang diberikan dapat berupa tes objektif dan tes subjektif. Untuk mengukur tes yang berkualitas maka butir soal harus dianalisis secara tepat. Tes yang berkualitas menurut Suharsimi Arikunto yaitu mempunyai ciri-ciri tes yang baik yaitu harus memenuhi persyaratan tes yaitu validitas, reliabilitas, objektivitas, kepraktisan, dan ekonomis.<sup>1</sup>

Mutu pendidikan memiliki keterkaitan dengan sistem pengajaran dan sistem evaluasi dalam rangka memantau. Jadi harus ada usaha yang terus menerus berkelanjutan untuk meningkatkan mutu pendidikan, karena tuntutan kualitas pendidikan selalu berubah sesuai dengan tuntutan masyarakat dan perkembangan ilmu pengetahuan dan teknologi. Evaluasi merupakan salah satu komponen penting dan tahap yang harus ditempuh oleh dosen untuk mengetahui keefektifan pembelajaran. Hasil yang

---

<sup>1</sup> Arikunto, Suharsimi. *Dasar-dasar Evaluasi Pendidikan*. (Jakarta: Bumi Aksara, 2013), h. 57-58

diperoleh dapat dijadikan *feed-back* bagi guru dalam memperbaiki dan menyempurnakan program dan kegiatan pembelajaran<sup>2</sup>. Anas sudijono juga menyatakana bahwa evalusai merupakan proses/kegiatan penentu kemajuan Pendidikan dan suatu usaha untuk memperoleh informasi berupa *feed back* atau umpan bali bagi penyempurnaan pendidikan<sup>3</sup>. Hal ini dapat diperoleh dari hasil evaluasi dan pengujian kualitas sistem pembelajaran secara menyeluruh. Oleh karena itu dalam prakteknya masalah pengukuran mempunyai kedudukan yang sangat penting dalam proses evaluasi. Baik buruknya hasil evaluasi tergantung pada hasil pengukuran.

Analisis butir soal perlu dilakukan untuk menguji mutu setiap butir soal dan seperangkat soal dalam berbagai aspek. Analisis butir soal dapat dilakukan secara kualitatif maupun secara kuantitaif. Tujuan utama analisis butir soal adalah untuk mendapatkan informasi tentang karakteristik setiap butir soal, baik melalui telaah butir soal maupun analisis empiris. Hasilnya dapat digunakan untuk mengetahui mutu soal dan mutu belajar peserta didik dari analisis hasil ujian.

Dosen sebagai pengajar dan pendidik harus membuat instrumen tes yang baik sesuai kriteria tersebut di atas. Namun masih ditemukan dosen dalam mengevaluasi domain kognitif mahasiswa hanya melakukan analisis butir secara kualitatif, sementara analisis secara kuantitatif masih sangat jarang dilakukan. Kondisi demikian membuat dosen tidak mengukur mahasiswa dengan tepat dan teliti, sehingga berpengaruh pada pemberian nilai yang tidak sesuai dengan kemampuan kognitif

---

<sup>2</sup> Arifin, Zainal. *Evaluasi Pembelajaran*, (Bandung: Remaja Rosdakarya, 2012), h. 1

<sup>3</sup> Sudijono, Anas. *Pengantar Evaluasi Pendidikan*, (Jakarta: Raja Grafindo Persada, 2012), h. 2

mahasiswa. Kompetensi dosen merupakan suatu keharusan yang harus dimiliki oleh seorang pendidik dalam melaksanakan tugasnya. Tugas dosen bukan hanya sebatas merancang tetapi lebih pada pelaksanaan pembelajaran dan evaluasi. Soal tes hendaknya memiliki daya pembeda antara mahasiswa yang pandai dengan yang kurang pandai. Selain itu juga memiliki tingkat kesukaran karena hal tersebut merupakan tujuan tes atau penilaian<sup>4</sup>. Perangkat dalam menganalisis butir soal sudah banyak dikembangkan diantaranya SPSS, RASCH Model, Anates, IteMan, Bilog.

Evaluasi Pembelajaran PAI merupakan salah satu matakuliah pada semester V dengan bobot 4 SKS. Matakuliah ini termasuk matakuliah keahlian membahas tentang konsep evaluasi, teknik penilaian, penyusunan instrumen penilaian, analisis butir soal, kriteria/standar penilaian, aplikatif, dan sistematis. Matakuliah ini dianggap penting karena akan mempersiapkan mahasiswa sebagai calon guru yang profesional. Kemampuan ini bisa dilakukan oleh calon guru, jika mereka dapat menguasai materi Evaluasi Pembelajaran PAI. Mengingat pentingnya matakuliah ini maka sistem pengajaran dan pembelajaran harus ditangani dengan sebaik-baiknya. Dosen dituntut untuk mengukur kemampuan mahasiswa melalui instrumen yang valid. Untuk itu penelitian ini akan melihat butir tes yang sudah dibuat oleh dosen dalam mengevaluasi mahasiswa menggunakan Anates untuk mengetahui validitas, reliabilitas, taraf kesukaran, daya beda, dan efektifitas pengecoh.

---

<sup>4</sup> Wiguna, Satria, Syaokani, dan Ananda, Rusydi. *Kemampuan Guru PAI dalam Merancang Tes (Analisis Aplikasi Anates Ganda di Sekolah SMA Negeri 1 Hinai, EDU Riligia*, 2 (1) (2018) : 1-10

Berdasarkan uraian di atas maka rumusan masalah pada penelitian ini adalah Bagaimana karakteristik butir-butir soal ujian Mata Kuliah Evaluasi Pembelajaran PAI dengan menggunakan Program Anates ?. Adapun tujuan dari penelitian ini adalah untuk mengetahui karakteristik butir-butir soal ujian Mata Kuliah Evaluasi Pembelajaran PAI dengan menggunakan Program Anates. Kategori dan jenis penelitian adalah penelitian deskriptif kuantitatif. Penelitian deskriptif kuantitatif dimaksud hanya untuk menggambarkan, menjelaskan atau meringkaskan berbagai kondisi, situasi, fenomena, atau berbagai variabel penelitian menurut kejadian sebagaimana adanya yang dapat dipotret, diwawancara, diobservasi, serta yang dapat diungkapkan melalui bahan-bahan dokumenter<sup>5</sup>. Metode deskriptif adalah metode penelitian untuk membuat gambaran mengenai situasi atau kejadian, atau memberi gambaran terhadap fenomena-fenomena.<sup>6</sup> Metode penelitian yang digunakan pada penelitian ini dengan pendekatan kajian lapangan atau kancah (*field research*). Data yang diperoleh pada penelitian diperoleh dari lapangan yaitu di prodi PAI UIN Ar-Raniry Banda Aceh.

Lokasi penelitian ini dilakukan pada Prodi PAI Fakultas Tarbiyah dan Keguruan UIN Ar-Raniry Banda Aceh. Lokasi ini dipilih karena memiliki karakteristik yang sesuai dengan tujuan penelitian. Sumber Data merupakan data yang diperoleh dari subjek yang akan diteliti. Sumber data pada penelitian ini diperoleh dari dokumen yaitu soal final yang telah dibuat oleh dosen pengampu matakuliah Evaluasi Pembelajaran PAI. Jumlah soal terdiri dari 25 butir dan bentuk tes yang dibuat adalah bentuk

---

<sup>5</sup> Burhan Bungin, *Metodologi Penelitian Kuantitatif*, Ed. 1., Cet. 3., (Jakarta: Kencana, 2008), h. 36

<sup>6</sup> Nazir, Moh. *Metode Penelitian*, cet. 6 (Bogor: Ghalia Indonesia, 2005), h. 55

pilihan ganda, dengan opsi atau pilihan jawaban sebanyak 4. Jumlah dokumen yang ada sesuai dengan jumlah mahasiswa yang mengikuti Ujian Akhir Semester sebanyak 26 mahasiswa. Data pada penelitian ini berasal hasil respon soal Ujian Akhir Semester matakuliah Evaluasi Pembelajaran PAI. Teknik pengumpulan data pada penelitian ini adalah teknik dokumentasi. Dokumentasi yang digunakan pada penelitian ini berupa soal ujian akhir semester ganjil, kunci jawaban, dan lembar jawaban dari mahasiswa prodi PAI FTK UIN Ar-Raniry pada matakuliah Evaluasi Pembelajaran PAI. Analisis data yang digunakan dalam penelitian ini melakukan analisis butir soal secara kuantitatif dengan menggunakan program Anates versi 4. Data yang diperoleh untuk mengetahui validitas, reliabilitas, taraf kesukaran, daya beda dan keefektifan distraktor pada soal Evaluasi Pembelajaran PAI.

## **PEMBAHASAN**

Analisis butir soal dilakukan untuk mengidentifikasi mana butir-butir soal yang baik dan yang buruk. Dari butir soal tersebut akan diperoleh informasi yang dapat digunakan untuk menghasilkan soal-soal yang lebih berkualitas, sehingga dapat diperoleh informasi yang baik terhadap peserta tes. Analisis soal dapat dilakukan dengan dua cara yaitu analisis kualitatif dan analisis kuantitatif. Penelitian ini menggunakan analisis kuantitatif untuk melihat butir keberfungsian butir soal dalam suatu tes. Salah satu cara untuk menganalisis soal yaitu dengan menggunakan program Anates versi 4, yang merupakan sebuah aplikasi komputer sederhana yang mudah diterapkan, cepat dan akurat. Aplikasi ini dirancang oleh Karno dan Yudi Wibisono yang merupakan perangkat lunak yang khusus dikembangkan untuk menganalisis tes pilihan ganda dan tes uraian.

Manfaat dari Anates yaitu dapat menganalisis butir soal secara otomatis memeriksa jawaban benar dan salah secara cepat dan praktis. Kelebihan program ini mudah dipahami karena petunjuk cara menjalankan program menggunakan Bahasa Indonesia dan hasil analisis dapat ditransfer ke Ms Excel untuk dihitung nilainya. Kemampuan anates untuk mengetahui analisis butir soal yang meliputi: reliabilitas, kelompok unggul dan asor, daya pembeda, tingkat kesukaran, korelasi skor butir dengan skor total dan kualitas pengecoh<sup>7</sup>.

Hasil dari analisis butir soal dengan menggunakan program Anates versi 4 diperoleh sebagai berikut:

### **1. Daya Pembeda**

Daya beda butir soal merupakan suatu indeks yang menunjukkan tingkat kemampuan butir soal membedakan kelompok atas (peserta tes yang berprestasi tinggi) dari kelompok bawah (peserta tes yang berprestasi rendah). Pernyataan ini sesuai dengan pendapat Sundayana bahwa daya pembeda merupakan kemampuan soal untuk dapat membedakan siswa yang berkemampuan tinggi dengan siswa yang berkemampuan rendah<sup>8</sup>. Menurut Anas daya pembeda butir adalah kemampuan suatu butir tes untuk dapat membedakan antara peserta tes yang berkemampuan tinggi dengan peserta tes yang berkemampuan rendah, sehingga sebagian peserta yang berkemampuan tinggi menjawab butir tes lebih banyak menjawab betul, sementara peserta yang berkemampuan rendah sebagian besar tidak menjawab betul pada tes tersebut<sup>9</sup>. Tujuan mencari daya pembeda

---

<sup>7</sup> Wiguna, Satria, Syauckani dan Ananda Rusydi. Kemampuan Guru PAI dalam Merancang Tes (Analisis Aplikasi Anates Ganda di Sekolah SMA Negeri Hinai) EDU Riligia, 2 (1): 1-10

<sup>8</sup> Sundayana. R. *Statistika Penelitian Pendidikan*, (Bandung: Alfabeta, 2016), h. 76

<sup>9</sup> Sudijono, Anas. *Pengantar Evaluasi...*, h. 385-386

adalah untuk mengukur keefektifan butir atau untuk menentukan apakah butir soal tersebut termasuk kategori skor tinggi atau skor rendah dalam keseluruhan tes serta butir soal tersebut memiliki kemampuan membedakan kelompok dalam aspek yang diukur sesuai dengan perbedaan kedua kelompok tersebut. Dalam analisis ini digunakan nilai koefisien korelasi biserial menunjukkan hubungan antara dua skor, yaitu skor butir soal dan skor keseluruhan dari peserta tes yang sama. Koefisien daya beda berkisar antara  $-1,00$  sampai dengan  $+1,00$ . Daya beda  $+1,00$  berarti bahwa semua anggota kelompok atas menjawab benar terhadap butir soal itu, sedangkan kelompok bawah seluruhnya menjawab salah satu terhadap butir soal itu. Sebaliknya daya beda  $-1,00$  berarti bahwa semua anggota kelompok atas menjawab salah butir soal itu, sedangkan kelompok bawah seluruhnya menjawab benar terhadap soal itu.

Daya pembeda soal atau daya beda butir soal merupakan indeks yang menunjukkan tingkat kemampuan butir soal membedakan kelompok atas (unggul) dari kelompok bawah (asor). Daya pembeda suatu butir berfungsi untuk menentukan apakah suatu butir soal dapat membedakan kelompok dalam aspek yang diukur sesuai dengan perbedaan pada kelompok tersebut. Sehingga butir soal tersebut dapat membedakan mahasiswa yang mempunyai kemampuan tinggi dan mahasiswa yang mempunyai kemampuan rendah. Hasil analisis dengan menggunakan anates daya beda soal diperoleh butir dengan kriteria sangat baik adalah butir 1, 8, 12, 17 dan 21, kriteria baik adalah butir 9 dan 20, kriteria agak baik atau cukup adalah

butir 6, 7, 10, 13, 14, 22, 24 dan 25, kriteria buruk adalah butir 3, 4, 11, 16, 18, 19, sebaiknya butir ini dibuang, kriteria sangat buruk adalah butir 2, 5, 15 dan 23 butir harus dibuang. Hal ini diklasifikasi berdasarkan tabel di bawah ini :

Tabel 1  
Kriteria Indeks Daya Pembeda Soal

Indeks Daya Pembeda	Kriteria Daya Pembeda
Negatif - 9%	Sangat buruk (harus dibuang)
10% - 19%	Buruk (sebaiknya dibuang)
20% - 29%	Agak Baik atau Cukup
30% - 49%	Baik
50% ke atas	Sangat Baik

Koefisien daya beda berkisar antara -1,00 sampai dengan +1,00, nilai positif saja yang mempunyai arti semakin besar indeks daya beda, maka semakin mendekati 1 maka butir tersebut mampu membedakan kelompok atas dan kelompok bawah. Butir yang dianggap mempunyai daya beda yang baik apabila butir tersebut  $> 0,5$ .

## 2. Tingkat Kesukaran

Tingkat Kesukaran (*Difficulty level*) disebut juga sebagai indeks kesukaran butir soal yang dilambangkan dengan huruf p yaitu *proportion*. Tingkat kesukaran adalah peluang untuk menjawab benar pada suatu soal pada tingkat kemampuan tertentu yang biasanya dinyatakan dalam bentuk indeks. Tingkat kesukaran menunjukkan proporsi peserta didik yang dapat mengerjakan soal secara benar dari suatu tes/ujian<sup>10</sup>. Dapat dikatakan

---

<sup>10</sup> Sumintono, Bambang dan Widhiarso, Wahyu. *Aplikasi Pemodelan Rasch pada Assessment Pendidikan*, (Bandung: Trim Komunikata, 2015), h.13

bahwa taraf kesukaran merupakan proporsi peserta menjawab benar terhadap butir soal. Suatu tes dikatakan baik apabila butir soal pada tes tersebut tidak terlalu sukar dan juga tidak terlalu mudah. Soal yang terlalu mudah tidak merangsang mahasiswa untuk mempertinggi usaha memecahkannya. Sebaliknya soal yang terlalu sukar akan menyebabkan mahasiswa menjadi putus asa dan tidak mempunyai semangat untuk mencoba lagi karena di luar jangkauannya<sup>11</sup>.

Angka indeks yang menunjukkan mudah sukarnya suatu butir soal dapat diketahui dari besaran indeks kesukaran yaitu antara 0,00 sampai dengan 1,00. Aiken juga menyatakan bahwa indeks kesukaran dinyatakan dalam bentuk proporsi yang besarnya berkisar 0,00 - 1,00 <sup>12</sup>. Taraf kesukaran soal berkorelasi dengan jumlah siswa yang dapat mengerjakan dengan benar terhadap butir pada suatu perangkat tes. Kriteria butir dikategorikan menjadi tiga yaitu mudah, sedang dan sukar dapat dilihat pada tabel di bawah ini :

Tabel. 2  
Kriteria Tingkat Kesukaran Butir Soal

Kriteria	Kategori
$P \geq 0,76$	Mudah
$0,25 \leq p \leq 0,75$	Sedang
$P \leq 0,24$	Sukar

Kriteria soal sangat mudah berjumlah 2 butir, kriteria Mudah berjumlah 1 butir, kriteria Sedang berjumlah 15 butir, kriteria Sukar berjumlah 4 butir, dan kriteria sangat sukar berjumlah 3 butir. Butir yang baik dengan kriteria sedang, yaitu tidak terlalu

---

<sup>11</sup> Arikunto, Suharsimi. *Dasar-dasar Evaluasi Pendidikan*, (Jakarta: Bumi Aksara, 2008), h. 207

<sup>12</sup> Aiken, Lewis R. *Psychological Testing and Assessment*, Eight Edition, (Boston: Allyn and Bacon, 1994), h. 66

mudah atau tidak terlalu sukar, karena butir yang mudah tidak merangsang mahasiswa untuk mempertinggi usaha memecahkannya, sedangkan butir yang terlalu sukar akan menyebabkan mahasiswa tidak bersemangat untuk mencoba lagi karena diluar kemampuannya.

### **3. Validitas**

Validitas merupakan kecermatan tes dalam menjalankan fungsi pengukurannya. Suatu tes dikatakan valid apabila tes tersebut mengukur apa yang hendak diukur. Menurut Arikunto validitas sebuah tes dapat diketahui dari hasil pengalaman. Validitas tes tersebut dibagi menjadi dua kelompok, yaitu validitas logis dan validitas empiris. Validitas logis meliputi validitas isi (*content validity*) dan validitas konstruk (*construct validity*), sedangkan validitas empiris meliputi validitas “ada sekarang” atau konkruen (*concurrent validity*) dan validitas prediksi (*predictive validity*)<sup>13</sup>. Menurut Sukardi bahwa suatu tes valid apabila koefisien 0,5 dapat diterima, jika hanya satu-satunya, sebaliknya jika ternyata ada tes prediksi lain yang sejenis dan mempunyai koefisien lebih tinggi maka koefisien 0,5 tidak diterima<sup>14</sup>. Maka suatu tes valid minimal koefiennya 0,5. Menurut Azwar butir soal dikatakan valid apabila suatu koefisien validitas dianggap memuaskan apabila koefisien diperoleh berkisar antara 0,30 sampai 0,50<sup>15</sup>. Validitas atau (kesahihan) tes dapat diartikan sebagai ketetapan dan

---

<sup>13</sup> Arikunto, Suharsimi, *Dasar-dasar...*, h. 65-69

<sup>14</sup> Sukardi. *Evaluasi Pendidikan Prinsip & Operasionalnya*, (Jakarta : Bumi Aksara, 2008), h. 38

<sup>15</sup> Azwar, Saifuddin, *Reliabilitas dan Validitas*, (Yogyakarta: Pustaka Pelajar, 2008), h. 158

kecermatan tes dalam menjalankan fungsi pengukurannya. Semakin tinggi koefisien maka semakin cermat suatu tes. Pada penelitian ini berdasarkan hasil analisis data yang diperoleh dengan menggunakan anates diperoleh koefisien validitas sebesar 0,24 termasuk klasifikasi sangat rendah. Dari hasil analisis data menggunakan program anates diketahui bahwa dari 25 soal yang dianalisis diperoleh soal yang signifikan hanya 6 butir yaitu butir 1, 8, 9, 12, 17 dan 21. Butir tidak signifikan berjumlah 19 butir yaitu butir 2,3,4,5,6,7,10, 11, 13, 14, 15, 16, 18, 19, 20, 22, 23, 24, dan 25.

Suatu tes mempunyai validitas tinggi apabila tes tersebut menjalankan fungsi ukurnya, atau memberikan hasil ukur yang sesuai dengan maksud dan tujuan diadakannya tes tersebut. Sebaliknya, suatu tes yang menghasilkan data tidak relevan dengan tujuan pengukuran dikatakan sebagai tes yang memiliki validitas rendah.

#### **4. Reliabilitas**

Reliabilitas suatu tes merupakan konsistensi dari suatu tes dalam mengukur apa yang seharusnya diukur sehingga pengukuran itu memberikan informasi yang dapat dipercaya. Hasil pengukuran dikatakan mempunyai reliabilitas yang tinggi apabila pengukuran pertama dan kedua menunjukkan hasil yang hampir sama, sebaliknya apabila pengukuran pertama dan kedua jauh berbeda maka reliabilitas suatu tes mempunyai reliabilitas rendah. Inti pokok dari reliabilitas adalah sejauhmana hasil dari suatu pengukuran dapat dipercaya. Tujuan utama mengestimasi reliabilitas adalah untuk menentukan seberapa besar variabilitas

yang terjadi karena adanya kesalahan pengukuran dan seberapa besar variabilitas tes yang sebenarnya<sup>16</sup>.

Reliabilitas suatu tes dapat diperoleh dengan beberapa cara yaitu:

- a) metode dua tes, disebut juga tes parallel atau setara (*equivalen*) yaitu dua tes yang diberikan kepada sekelompok peserta tes. Dua hasil tes tersebut dicari korelasinya.
- b) metode satu tes, sebuah tes diberikan dua kali kepada sekelompok peserta tes, tetapi dalam waktu yang berbeda, kemudian kedua hasil tes tersebut dicari korelasinya.
- c) Metode "*split-half*" (satu tes), suatu tes dibagi menjadi dua bagian yang sama tingkat kesukarannya, sama isi dan bentuknya. Boleh dibagi ganjil-genap atau atas- bawah. Kemudian dilihat skor dari masing-masing bagian paruhan tes tersebut dan dicari korelasinya<sup>17</sup>

Berbeda dengan metode pertama dan kedua yang setelah diketemukan koefisien korelasi langsung ditafsirkan itulah koefisien realibilitas, maka dengan metode ketiga ini tidak dapat demikian. Pada waktu membelah dua dan mengkorelasikan dua belahan, baru diketahui reliabilitas separo tes. Berdasarkan hasil analisis diperoleh reliabilitas tes diperoleh sebesar 0,24 dengan menggunakan teknik belah yaitu ganjil dan genap. Hasil koefisien yang diperoleh dapat dilihat pada klasifikasi pada tabel berikut ini:

Ta  
b  
e  
1  
3

---

<sup>16</sup> Supranata, Sumarna, Analisis, Validitas..., h. 89

<sup>17</sup> Purwanto, M. Ngalim. *Prinsip-prinsip dan Teknik Evaluasi Pengajaran*, (Bandung: Remaja Rosdakarya, 2010), h. 139-140

**Klasifikasi Tingkat  
Reliabilitas**

Tes Kategori Reliabilitas	Koefisien Korelasi
Sangat Tinggi	0,800 - 1,000
Tinggi	0,600 - 0,799
Cukup	0,400 - 0,599
Rendah	0,200 - 0,399
Sangat Rendah	0,000- 0,199

Dari hasil yang diperoleh melalui Anates reliabilitas soal Evaluasi Pembelajaran PAI yaitu sebesar 0,24, sesuai dengan klasifikasi pada tabel di atas masuk dalam kategori rendah karena dalam rentang 0,200 - 0,399. Berarti tes Evaluasi Pembelajaran PAI ini mempunyai koefisien yang kurang baik dan kurang handal dalam mengukur kemampuan mahasiswa. Hasil pengukuran harus reliabel artinya harus memiliki tingkat konsistensi dan kemampuan. Suatu alat tes dikatakan reliabel apabila alat tes tersebut dapat dipercaya, konsisten, atau tetap. Untuk membuktikan apakah suatu alat tes memiliki sifat tetap, perlu diadakan uji coba terhadap alat tes yang akan digunakan tersebut. Menurut Azwar hasil pengukuran dapat dipercaya hanya apabila dalam beberapa kali pelaksanaan pengukuran terhadap kelompok subjek yang sama diperoleh hasil yang relatif sama, selama aspek yang diukur dalam diri subjek memang belum berubah

**5. Efektifitas Pengecoh**

Distraktor disebut juga dengan pengecoh, pada soal pilihan ganda kemungkinan jawaban terbagi dua yaitu kunci jawaban dan

distraktor. Dari sekian banyak alternatif jawaban hanya satu yang benar yaitu kunci jawaban dan kemungkinan jawaban yang tidak benar disebut dengan pengecoh. Pengecoh berfungsi untuk mengidentifikasi peserta tes yang mempunyai kemampuan tinggi<sup>18</sup>.

Tujuan analisis distraktor adalah untuk mengetahui seberapa banyak siswa menjawab benar sesuai dengan kunci jawaban dan seberapa banyak yang memilih distraktor atau pengecoh. Efektifitas distraktor perlu diperhatikan untuk mengetahui banyaknya mahasiswa menjawab benar sesuai dengan kunci jawaban dan berapa banyak mahasiswa yang memilih distraktor atau pengecoh. Pengecoh berfungsi untuk mengidentifikasi peserta tes yang berkemampuan tinggi. Pengecoh dikatakan efektif jika dipilih oleh peserta tes yang berasal dari kelompok bawah yaitu mahasiswa yang mempunyai kemampuan rendah. Dan jika dipilih oleh mahasiswa yang berkemampuan tinggi maka pengecoh tersebut tidak berfungsi dengan baik. Dari hasil analisis data menggunakan program anates diperoleh hasil dikatakan distraktor yang baik apabila dipilih paling sedikit 5% dari peserta tes. Dari butir soal terdapat beberapa pengecoh dengan kriteria sangat baik, Baik, Buruk dan sangat buruk. Butir soal yang mempunyai pengecoh yang baik adalah pada butir soal 3, 4, 10, 11, 16, 22, 23 dan 25. Menurut Sumarna suatu pengecoh dikatakan berfungsi dengan baik apabila dipilih oleh 5% peserta tes. Jika pengecoh dipilih secara merata, maka pengecoh tersebut termasuk sangat baik. Membuat pengecoh yang baik pada tes pilihan ganda sulit,

---

<sup>18</sup> Surapranata, Sumarna, *Analisis, Validitas, Reliabilitas dan Interpretasi Hasil Tes*, (Bandung: Remaja Rosdakarya, 2004), h. 43

karena pengecoh yang kurang baik akan mengakibatkan rendahnya daya pembeda, dan apabila satu atau dua pengecoh tidak berfungsi baik maka mengakibatkan tingkat kesukaran juga rendah.

## **PENUTUP**

Berdasarkan hasil analisis yang diperoleh secara keseluruhan, maka dapat ditarik kesimpulan bahwa :

Karakteristik butir soal ujian akhir semester mata kuliah Evaluasi Pembelajaran PAI yaitu daya beda soal diperoleh butir dengan kriteria sangat baik adalah butir 1, 8, 12, 17 dan 21, kriteria baik adalah butir 9 dan 20, kriteria agak baik atau cukup adalah butir 6, 7, 10, 13, 14, 22, 24 dan 25, kriteria buruk adalah butir 3, 4, 11, 16, 18, 19, sebaiknya butir ini dibuang, kriteria sangat buruk adalah butir 2, 5, 15 dan 23 butir harus dibuang. Tingkat kesukaran kriteria soal sangat mudah berjumlah 2 butir, kriteria Mudah berjumlah 1 butir, kriteria Sedang berjumlah 15 butir, kriteria Sukar berjumlah 4 butir, dan kriteria sangat sukar berjumlah 3 butir. Butir yang baik dengan kriteria sedang, yaitu tidak terlalu mudah atau tidak terlalu sukar. Validitas yang diperoleh dengan menggunakan anates sebesar 0,24 termasuk klasifikasi sangat rendah, dari 25 soal yang dianalisis diperoleh soal yang signifikan hanya 6 butir yaitu butir 1, 8, 9, 12, 17 dan 21. Reliabilitas tes diperoleh sebesar 0,39 dengan menggunakan teknik belah yaitu ganjil dan genap. Distraktor yang baik apabila dipilih paling sedikit 5% dari peserta tes, pengecoh pada butir soal sebagian besar sangat baik, baik, buruk dan beberapa pengecoh sangat buruk. Butir soal yang

mempunyai pengecoh yang baik adalah pada butir soal 3, 4, 10, 11, 16, 22, 23 dan 25.

## **DAFTAR PUSTAKA**

- Aiken, Lewis R. (1994). *Psychological Testing and Assessment*, (Eighth Edition). Boston: Allyn and Bacon.
- Arifin, Zainal. (2011). *Evaluasi Pembelajaran*, Bandung: Remaja Rosdakarya.
- Arikunto, Suharsimi. (2013). *Dasar-dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.
- Azwar, Saifuddin. (2008). *Reliabilitas dan Validitas*, Yogyakarta: Pustaka Pelajar
- Bungin, Burhan. (2008). *Metodologi Penelitian Kuantitatif*, Ed.1., Cet. 3., Jakarta: Kencana.
- Nazir, Moh.. (2005). *Metode Penelitian*. Cet. 6., Bogor: Ghalia Indonesia
- Purwanto, M. Ngalim. (2010). *Prinsip-prinsip dan Teknik Evaluasi Pengajaran*. Bandung: Remaja Rosdakarya.
- Sudijono, Anas. (2011). *Pengantar Evaluasi Pendidikan*, Jakarta: Raja Grafindo Persada.
- Anastasi dan Urbina. (1997). *Psychological Testing*, New York: Macmillan Publishing Company.
- Sukardi, (2008). *Evaluasi Pendidikan Prinsip & Operasionalnya*, Jakarta: Bumi Aksara
- Sumintono, Bambang dan Widhiarso, Wahyu. (2015). *Aplikasi Pemodelan Rasch Pada Assessment Pendidikan*, Cimahi: Trim Komunikata.
- Sundayana, R. (2016). *Statistika Penelitian Pendidikan*, Bandung: Alfabeta
- Supranata, Sumarna. (2004). *Analisis, Validitas, Reliabilitas dan Interpretasi Hasil Tes*, Bandung: Remaja Rosdakarya.
- Wiguna, Satria, Syaukani, dan Ananda, Rusydi. (2018). Kemampuan Guru PAI dalam Merancang Tes (Analisis Aplikasi Anates Ganda Di Sekolah SMA Negeri 1 Hinai), *EDU Riligia*, 2 (1) : 1-10.