

## **ANALISIS EMPIRIK PROGRAM ANALISIS BUTIR SOAL DALAM RANGKA MENGHASILKAN SOAL YANG BAIK DAN BERMUTU SEBAGAI ALAT EVALUASI PEMBELAJARAN FISIKA**

**Muhammad Nasir**

Program Studi Pendidikan Fisika, FKIP Universitas Riau, Kampus Bina Widya Pekanbaru,  
Telp. (0761) 65804 Fax (0761) 65804  
muh\_nasir23@yahoo.com

### **ABSTRAK**

Penelitian ini bertujuan menganalisis secara empirik Program analisis butir soal untuk menghasilkan hasil soal yang baik dan bermutu. Sudah dirancang Program analisis butir soal namun untuk dapat dipergunakan perlu dilakukan analisis empirik guna memastikan bahwa program tersebut valid dan reliabel, Analisis empirik ini dilakukan dengan penilaian penggunaan program dan uji coba program. Penilaian pengguna dilakukan dengan kuesionair dan penggunaan program ini saat ujicoba. Selain itu juga diperlukan keahlian tertentu untuk menggunakan program tersebut. Program ini dirancang dengan menggunakan ADDIE models dan pembangunannya menggunakan microsoft office excell 2007 yang mudah diperoleh karena ada pada setiap komputer. Kelebihan dari program analisis butir soal ini adalah mudah menggunakannya dan tidak perlu program khusus, dengan sifat ini akan dapat membantu pengguna untuk menganalisis butir soal agar dapat menghasilkan soal Fisika yang baik dan bermutu sebagai alat evaluasi Pembelajaran Fisika. Berdasarkan hasil analisis di ujicoba dengan cara membandingkan hasil perhitungan manual dan diperoleh semua fitur dari program ini berjalan dengan baik dan memberikan hasil yang baik dan tingkat presisi yang sangat baik.

Kata Kunci :

Analisis empirik, Analisis butir soal, ADDIE Models, Presisi, evaluasi pembelajaran fisika.

### **1. PENDAHULUAN**

Kegiatan menganalisis butir soal merupakan suatu kegiatan yang harus dilakukan untuk meningkatkan mutu soal yang telah ditulis. Kegiatan ini merupakan proses pengumpulan, peringkasan, dan penggunaan informasi dari jawaban siswa untuk membuat keputusan tentang setiap penilaian (Nitko,1996: 308). Tujuan utama analisis butir soal dalam sebuah tes yang dibuat guru adalah untuk mengidentifikasi kekurangan-kekurangan dalam tes atau dalam pembelajaran (Anastasi dan Urbina, 1997:184).

Berdasarkan tujuan ini, maka kegiatan analisis butir soal memiliki banyak manfaat, di antaranya adalah : Pertama agar dapat membantu para pengguna tes dalam evaluasi atas tes yang digunakan. Kedua, sangat relevan bagi penyusunan tes informal dan lokal seperti tes yang disiapkan guru untuk siswa di kelas. Ketiga, untuk mendukung penulisan butir soal yang efektif. Keempat, secara materi dapat memperbaiki tes di kelas, dan terakhir adalah untuk meningkatkan validitas soal dan reliabilitas (Anastasi and Urbina, 1997:172). Linn dan Gronlund (1995: 315) juga menambahkan tentang pelaksanaan kegiatan analisis butir soal yang biasanya didesain untuk menjawab pertanyaan-pertanyaan berikut ini.: 1) Apakah fungsi soal sudah tepat ? 2) Apakah soal

ini memiliki tingkat kesukaran yang tepat ? 3) Apakah soal bebas dari hal-hal yang tidak relevan? 4) Apakah pilihan jawabannya efektif ?

Berbagai uraian di atas menunjukkan bahwa analisis butir soal adalah: untuk menentukan soal- soal yang cacat atau tidak berfungsi penggunaannya, dan untuk meningkatkan kualitas butir soal melalui tiga komponen analisis yaitu tingkat kesukaran, daya pembeda, dan pengecoh soal, serta meningkatkan pembelajaran melalui ambiguitas soal.

Guru selain sebagai seorang pengajar juga berperan sebagai evaluator. Pada proses evaluasi pendidikan sangat dibutuhkan kemampuan untuk menganalisis soal, sehingga soal yang digunakan dapat mencerminkan kemampuan siswa. Soal yang baik akan mampu mengevaluasi sejauh mana peserta didik menguasai indikator yang sudah ditentukan oleh pengajar. Untuk itu, kemampuan menganalisis soal setelah melakukan tes sangatlah dibutuhkan oleh pendidik untuk melakukan evaluasi apakah alat ukur yang digunakan tersebut sesuai apa tidak dengan apa yang diinginkan antara lain dapat menentukan peserta didik mana yang sudah atau belum menguasai materi yang diajarkan guru dan juga bisa membantu meningkatkan tes melalui revisi atau membuang soal yang tidak efektif, serta untuk mengetahui informasi diagnostik pada siswa apakah mereka sudah/belum memahami materi yang telah diajarkan (Aiken, 1994: 63).

Dampak hasil evaluasi terhadap motivasi belajar peserta didik bervariasi ada yang meningkat, tetap, bahkan ada yang turun. Tiap peserta didik memiliki harapan terhadap hasil ulangan suatu pelajaran, yaitu besarnya prestasi yang dinyatakan dalam skor hasil tes. Harapan ini ada yang terpenuhi dan ada yang tidak terpenuhi. Sesuai dengan karakteristik peserta didik, ada yang motivasi belajarnya naik, ada yang tetap dan kemungkinan ada yang turun.

Masalah yang sering timbul dalam melakukan evaluasi terletak pada tujuannya, pendekatan yang digunakan, manfaatnya, dan dampaknya. Selain evaluasi, pendidikan juga harus memberi manfaat kepada peserta didik, lembaga, dan masyarakat. Oleh karena itu, apabila evaluasi pendidikan yang digunakan tidak membantu peningkatan kualitas pendidikan pada suatu sekolah dan tidak memberi manfaat, berarti system evaluasi yang digunakan atau yang dilaksanakan belum berfungsi seperti yang diharapkan (Rasyid dan Mansur, 2007).

Menurut Daryanto, 2007, tidak ada usaha guru yang lebih baik selain usaha untuk selalu meningkatkan mutu tes yang disusunnya. Namun hal ini tidak dilaksanakan karena kecenderungan seseorang untuk beranggapan bahwa yang menjadi hasil karyanya adalah yang terbaik, atau setidak – tidaknya sudah cukup baik.

Dalam membuat butir soal evaluasi, guru harus memperhatikan kurikulum tingkat satuan pendidikan yang digunakan. Menurut Thoha (2003), "Suatu alat evaluasi dikatakan baik yaitu memiliki validitas yang mana alat evaluasi tersebut dapat mengukur apa yang harus diukur dan memiliki realibitas yaitu evaluasi yang memiliki kehandalan dan bersifat konsisten dari suatu pengukuran yang dilakukan.

Menurut Muhammad Nasir dkk (2012) Banyak guru-guru di sekolah tidak melakukan analisis butir soal untuk evaluasi disebabkan karena : 1.Kekurangan waktu untuk melakukan analisis, 2. Analisis butir soal dirasakan masih sukar karena melibatkan rumus-rumus yang komplit dan melelahkan. 3.Tidak adanya/belum adanya alat bantu yang memudahkan untuk melakukan analisis butir soal.

Berdasarkan latar masalah diatas maka pada penelitian ini mencoba untuk membuat solusi yaitu dengan membuat sebuah program (software) untuk menganalisis butir soal. Dengan demikian guru akan lebih mudah untuk menganalisis soal tanpa harus melakukan perhitungan-perhitungan yang kompleks yang kadang membuat kesalahan-kesalahan angka.

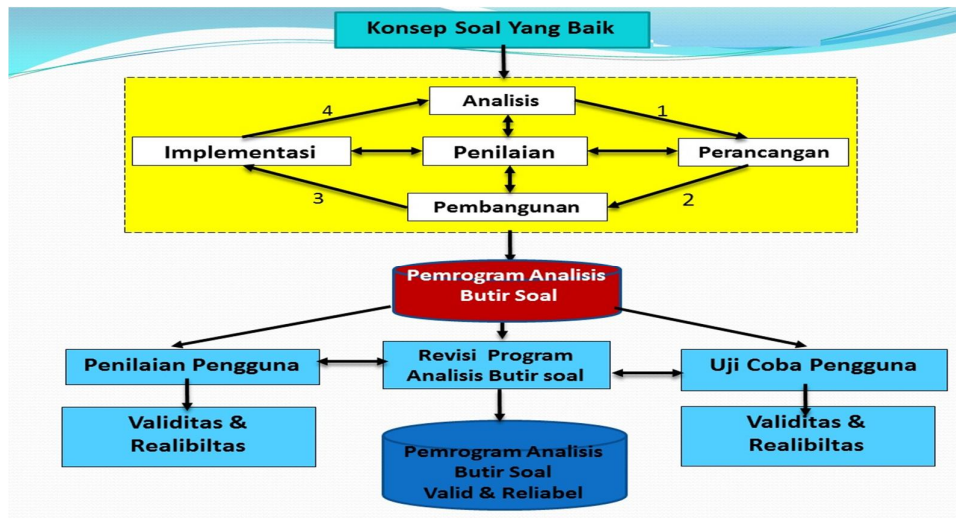
## **2. METODE PENELITIAN**

Penelitian ini menggunakan metode kualitatif melalui pengkajian secara komprehensif terhadap berbagai literatur, menelaah kurikulum, menelaah materi pelajaran Fisika dan teori tentang analisis butir soal, sehingga kemudian dari materi yang telah ditelaah dalam hal ini akan dibahas materi tentang gerak (motion) kemudian di rancang pemrograman analisis butir soal (course ware) yaitu berupa program yang dalam hal ini adalah program yang dirancang dengan ADDIE Model dan dibangun dengan menggunakan Prrogram Microsoft Office excel . Microsoft excell di pilih karena kemudahan mendapatkan dan relatif banyak dikenal oleh kalangan guru sebagai sasaran user dari progem ini. Dengan tahapan kerja seperti berikut :

**1. Tahap Awal**, Tahap ini adalah pengumpulan informasi tentang Konsep analisis butir soal, Rumus matematika yang digunakan dan istilah serta makna dan pengertian yang terkandung dalam istilah istilah dalam analisis butir soal. **2.Tahap kedua**, Setelah didapatkan istilah, teori dan rumusan matematika yang dalam analisis butir soal, kemudian baru di rancang perograman analisis butir soal dengan menggunakan microsoft office excell. Pada tahapan ini dilaksanakan perancangan dengan menggunakan Perancangan *Instructional Design Model ADDIE* (Perancangan Pembelajaran Model ADDIE). **3. Tahap ketiga**, Setelah pemrograman dibuat dilakukan ujicoba disekolah, kemudian dilakukan penilaian oleh pakar pengguna yaitu guru-guru fisika, Pakar pedagogik dan pakar pembelajaran/user dalam hal ini oleh guru senior bidang studi fisika. Pada tahap ini bila masih ada masukan kita di uji coa maka akan dilakukan perbaikan

bilaperbaikan selesai dilakukan maka akan dilanjutkan ketahap keempat yaitu tahapan eksperimen kepada siswa di dalam kelas. **4. Tahap keempat**, Tahap ini merupakan tahapan terakhir yaitu tahapan eksperimen. Eksperimen yang akan dilakukan merupakan eksperimen kuasi, dengan mengambil kelas sebagai kelas eksperimen. Hasil belajar yang akan diamati meliputi Kognitif test tertulis dengan bentuk test objektif pilihan ganda. Hasil test dianalisis dengan menggunakan program analisis butir soal dan hasil belajar yang dianalisis dengan menggunakan analisis secara manual.

Dengan mengamati semua hasil analisis kemudian membandingkannya dan hasil analisis butir secara manual maka didapatkan efektifitas dari program analisis butir soal. Bila terdapat kekurangan pada program akan dilakukan penyempurnaan akan dilakukan perbaikan untuk kemudian akan dijadikan sebuah CD (Compac Disc ) untuk dijadikan sebagai salah satu produk. Untuk lebih jelasnya kerangka konseptual penelitian dapat dilihat seperti gambar 1.



Gambar 1. Kerangka Konseptual Penelitian

Adapun parameter pengamatan pada penelitian ini adalah hasil penilaian pengguna yang merupakan guru-guru SMA Negeri di Pekanbaru sebanyak 28 orang yang dilatihkan menggunakan program ini sebelumnya kemudian diminta mengisi angket penilaian program yang adaptasi dari Muhammad Nasir (2014). Selain mengisi angket tersebut pengguna juga melihat hasil analisis soal yang meliputi :

1. Teknik analisis butir soal secara kuantitatif
  - a. Sebaran Butir Soal IPA Fisika Berdasarkan Ranah Kognitif

Untuk mengetahui sebaran butir soal berdasarkan ranah kognitif dilakukan dengan melihat tingkatan pada tiap soal. Kemudian data yang diperoleh dianalisis dengan menggunakan rumus :

$$P = \frac{F}{N} \times 100 \% \dots\dots\dots(1)$$

Keterangan :

P = Angka persentase yang dicari

F = Frekuensi (sebaran butir soal berdasarkan ranah kognitif)

N = Jumlah soal (Sudijono, 2009)

b. Pengecoh Soal (Distracter Function) dan Omit

Suatu distraktor dapat dikatakan berfungsi baik jika paling sedikit dipilih oleh 5% peserta tes. Sedangkan Omit dikatakan baik jika omitnya tidak lebih dari 10% peserta te (Arikunto, 2012). Untuk melihat baik atau tidaknya suatu distraktor (butir pengecoh) digunakan rumus sebagai berikut :

$$D = \frac{KA+KB}{jumlah\ KA+jumlah\ KB} \times 100 \% \dots\dots\dots(2)$$

Keterangan:

KA = kelompok atas yang memilih pengecoh

KB = kelompok bawah yang memilih pengecoh

D = distraktor

Sedangkan untuk melihat baik atau tidaknya omit digunakan rumus berikut:

$$O = \frac{jumlah\ peserta\ tes\ yang\ tidak\ menjawab}{jumlah\ KA+jumlah\ KB} \times 100\% \dots\dots\dots(3)$$

Keterangan: O= omit

c. Tingkat Kesukaran (*Difficulty Indeks*)

Tingkat kesukaran suatu soal dapat diketahui dengan menganalisis lembar jawaban siswa kemudian data yang diperoleh dianalisis dengan menggunakan rumus:

$$I = \frac{B}{N} \dots\dots\dots(4)$$

Dimana :

I =Indeks kesukaran untk setiap butir soal

B=Banyaknya siswa yang menjawab benar setiap butir soal

N=Jumlah seluruh siswa peserta tes (Sudjana, 2008)

Kriteria indeks tingkat kesukaran diperoleh sebagai berikut:

Tabel 1. Tabel Kriteria indeks tingkat kesukaran soal

No	Indeks Kesukaran (I)	Kategori Soal
1	0,00 – 0,30	Sukar
2	0,31 – 0,70	Sedang
3	0,71 – 1,00	Mudah

(Sudjana, 2008)

d. Daya Pembeda (*Discriminating Power*)

Untuk menghitung daya pembeda soal digunakan rumus sebagai berikut :

$$D = \frac{BA}{JA} - \frac{BB}{JB} = PA - PB \dots\dots\dots(5)$$

Dimana :

- D = Daya pembeda soal
- BP= Jumlah siswa yang menjawab benar dari kelompok atas
- BB= Jumlah siswa yang menjawab benar dari kelompok bawah
- JA= Jumlah peserta dari kelompok atas
- JB= Jumlah peserta dari kelompok bawah
- PA =  $\frac{BA}{JA}$  = Proporsi peserta kelompok atas yang menjawab benar.
- PB =  $\frac{BB}{JB}$  = Proporsi peserta kelompok bawah yang menjawab benar.

Crocker & Algina (1986) menyatakan kriteria koefisien besarnya daya pembeda:

Tabel 2. Kriteria Daya Pembeda

No	Daya Pembeda Soal (D)	Kategori Soal
1	0,40 – 1,00	Baik
2	0,30 – 0,39	Terima
3	0,20 – 0,29	Perlu Revisi
4	-1,00 – 0,19	Tidak Berfungsi

(Rasyid dan Mansur, 2007)

Angka 0,3 merupakan angka kriteria minimal suatu butir soal yang baik (Mardapi dalam Rasyid dan Mansur 2007).

e. Validitas Soal

Sebuah butir soal (item) dikatakan valid apabila mempunyai dukungan yang besar terhadap skor total. Pada butir-butir instrumen dengan bentuk soal objektif (data diskrit murni atau dikotomik), maka digunakan teknik korelasi “point biserial”, yang mempunyai rumus sebagai berikut:

$$Y_{pbis} = \frac{M_p - M_t}{Sd_t} \sqrt{\frac{p}{q}} \dots\dots\dots(6)$$

Keterangan:

- $Y_{pbis}$  = koefisien korelasi biserial
- $M_p$  = rerata skor dari subjek yang menjawab betul bagi item yang dicari validitasnya
- $M_t$  = rerata skor total
- $Sd_t$  = standar deviasi skor total proporsi
- $p$  = proporsi siswa yang menjawab benar

$$p = \frac{\text{banyaknya siswa yang menjawab benar}}{\text{jumlah seluruh siswa}} \dots\dots\dots(7)$$

$q$  = proporsi siswa yang menjawab salah ( $q = 1 - p$ )  
 (Arikunto, 2012).

Patokan yang digunakan untuk menginterpretasikan validitas yaitu:

Tabel 3. Patokan Interpretasi Validitas

No	Validitas Soal	Kategori Soal
1	$Y_{pbis} > r t$	Valid
2	$Y_{pbis} = \text{negatif}$	invalid
3	$Y_{pbis} < r t$	invalid

(Arikunto, 2009)

f. Reliabilitas Tes

Tujuan utama menghitung reliabilitas skor tes adalah mengetahui tingkat ketepatan (*precision*) dan keajegan (*consistency*) skor tes. Untuk mengetahui koefisien reliabilitas tes soal bentuk pilihan ganda dapat digunakan rumus Kuder-Richardson 21 (KR-21) seperti berikut:

$$r_{11} = \frac{n}{n-1} \left[ 1 - \frac{M(n-M)}{nS_t^2} \right] \dots\dots\dots(8)$$

(Arikunto, 2012).

Untuk menentukan tinggi rendahnya koefisien reliabilitas (r) dapat ditentukan berdasarkan kriteria sebagai berikut:

Tabel 4. Klasifikasi Koefisien Reliabilitas Tes

No	Reliabilitas Tes (r)	Kategori instrumen tes
1	$0,00 < r \leq 0,20$	Sangat rendah
2	$0,20 < r \leq 0,40$	Rendah
3	$0,40 < r \leq 0,60$	Sedang
4	$0,60 < r \leq 0,80$	Tinggi
5	$0,80 < r \leq 1,00$	Sangat Tinggi

(Arikunto, 2003)

1) Secara kuantitatif

Sebagai patokan untuk penarikan kesimpulan analisis secara kuantitatif dapat digunakan kriteria pada aspek berikut ini:

a) Tingkat kesukaran soal

- (1) Apabila tingkat kesukaran soal mudah, sedang, dan sukar memenuhi proporsi seimbang yaitu 27%, 46%, 27% maka soal dikategorikan baik.
- (2) Apabila tingkat kesukaran soal mudah, sedang, dan sukar tidak memenuhi proporsi seimbang yaitu 27%, 46%, 27% maka soal dikategorikan tidak baik

b) Daya pembeda soal

- (1) Apabila kriteria daya pembeda  $0,30 - 1,00 \geq 85\%$  maka soal dikategorikan baik.
- (2) Apabila kriteria daya pembeda  $0,30 - 1,00 < 85\%$  maka soal dikategorikan tidak baik.

c) Validitas

- (1) Jika presentase validitas  $\geq 85\%$  maka soal tersebut dikategorikan valid.
- (2) Jika presentase validitas  $< 85\%$  maka soal tersebut dikategorikan tidak valid.

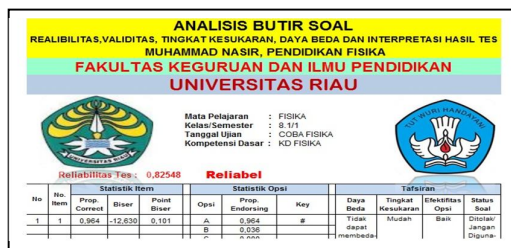
Ketentuan ketentuan tersebut akan menjadi dasar untuk meanalisis, tes hasil ujian siswa dinyatakan baik apabila terdapat dua variable yaitu tingkat kesukaran soal, dan daya pembeda soal dikategorikan baik, dan satu variabel validitas soal dikategorikan valid.

### 3. HASIL DAN PEMBAHASAN

Hasil perancangan program dapat dilihat seperti pada gambar (2),(3) dan (4)



Gambar 2. Tampilan Hasil Program Analisis Butir Soal



Gambar 3. Tampilan Hasil Analisis dari Program Analisis Butir Soal



Gambar 4. Tampilan Hasil Rekap Laporan Program Analisis Butir soal

Hasil penilaian pengguna :

Dari hasil penilaian pengguna di dapati Seperti pada Tabel (1) dan (2) :

Tabel 1. Hasil Penilaian Pengguna

Aspek	Pernyataan	Rata-rata	Standar Deviasi	Varians
Perancangan	1 - 10	4.03	0.52	0.29
Pedagogi	11 - 20	3.97	0.62	0.43
Isi	21 - 30	3.76	0.64	0.42
Teknik	31 - 38	4.02	0.66	0.48
<b>Keselu</b>	<b>1 - 38</b>	<b>3.95</b>	<b>0.40</b>	<b>0.16</b>

Tabel 2. Realibilitas (  $\alpha$  Croncbach)

Cronbach's Alpha	Jumlah Item
.740	38

Berdasarkan pada Tabel (1) dan (2) dapat dikatakan bahwa Program analisis butir soal adalah Valid dan Relabel dengan tingkat releabilitas  $\alpha = 0.74$

#### Analisis Hasil Uji coba secara Kuantitatif

Dari data analisis secara kuantitatif diperoleh hasil analisis butir soal tryout ujian SMA Se-kota Pekanbaru Fisika Tahun 2014 dijelaskan pada aspek berikut:

a. Sebaran Butir Soal berdasarkan ranah kognitif



Tingkat kesukaran berdasarkan taraf penguasaan sebaiknya perbandingan soal kategori pengetahuan: pemahaman + aplikasi: analisis + sintesis + evaluasi (C1 : C2 + C3 : C4 + C5 + C6) adalah 30% soal mudah : 50% soal sedang : 20% soal sukar. Sementara hasil analisis pada taraf penguasaan soal ujian IPA fisika SMP 17 Pekanbaru hanya terdapat perbandingan soal kategori C1: C2+ C3: C4 yaitu: 20% soal mudah: 60% soal sedang: 20% soal sukar. Berdasarkan hal tersebut kualitas soal yang dimiliki belum dikategorikan dengan proporsi soal yang baik, karena terjadi penumpukan pada soal C2 dan C3, sebaiknya proporsi soal dikurangi agar memenuhi proporsi yang seimbang. Sementara soal-soal pada tingkatan C1 dan C4 masih sedikit jumlahnya, sebaiknya soal perlu ditambah agar menjadi proporsi yang seimbang.

a. Keefektifan pengecoh (Distraktor) beserta Omit

Hasil analisis menunjukkan bahwa terdapat 3 dari 20 butir soal (15%) semua distraktor berfungsi dengan baik yaitu pada nomor 13, 14, dan 18. Karena semua distraktor sudah dipilih lebih dari 5% pengikut tes. Butir soal yang distraktornya tidak berfungsi dengan baik misalnya pada nomor satu, *option* D merupakan kunci jawaban dan *option* A, B, dan C merupakan distraktor atau pengecoh soal. Berdasarkan hal tersebut bahwa dari 3 distraktor terdapat 1 distraktor yang tidak berfungsi dengan baik yaitu pada *option* B, karna distraktor dipilih kurang dari 5% pengikut tes. Sedangkan omit sudah dikatakan baik karna tidak ada nilai omit yang lebih dari 10% dari setiap butir soal yang terdapat pada nomor soal 13.

Sebagai tindak lanjut dari hasil penganalisan terhadap fungsi distraktor tersebut maka distraktor yang sudah berfungsi dengan baik dapat digunakan lagi pada ujian semester tahun berikutnya, sedangkan distraktor yang belum berfungsi dengan baik sebaiknya diperbaiki atau diganti dengan distraktor yang lain sehingga dapat digunakan. Menurut Arikunto (2012) sebuah distraktor dapat berfungsi dengan baik apabila distraktor tersebut mempunyai daya tarik yang besar bagi pengikut-pengikut tes yang kurang memahami konsep atau kurang menguasai bahan. Sebaliknya pengecoh yang tidak dipilih sama sekali oleh *testee* berarti pengecoh itu jelek, terlalu menyolok menyesatkan.

b. Tingkat kesukaran soal

Berdasarkan hasil yang diperoleh bahwa tingkat kesukaran soal terdapat 8 butir soal mudah (40%), 11 butir soal sedang (55%) dan 1 butir soal sukar (5%). Jadi, perbandingan proporsi soal ujian semester ganjil IPA fisika tahun 2012 tersebut adalah 40% soal mudah: 55% soal sedang: 5% soal sukar. Menurut Mudjijo (1995), suatu tes hasil belajar yang baik memiliki butir soal yang tingkat kesukarannya seimbang, artinya berdistribusi secara normal masing-masing adalah 27% soal mudah, 46% soal sedang, dan 27% soal sukar. Berdasarkan asumsi tersebut soal tidak berkualitas baik, karena soal

dengan kategori mudah, sedang, dan sukar tidak terdapat secara seimbang yaitu 40% soal mudah: 55% soal sedang: 5% soal sukar. Oleh sebab itu, dalam pembuatan soal proporsi soal yang mudah dan sedang dikurangi sedangkan soal yang sukar harus ditambah. Sehingga didapatkan proporsi soal yang memiliki tingkat kesukaran secara seimbang.

c. Daya Pembeda Soal

Dari hasil analisis terdapat 4 butir soal (20%) yang tidak berfungsi, 4 butir soal (20%) perlu revisi pada nomor 5, 6, 8, dan 11. 7 butir soal (35%) diterima dan tidak perlu revisi, dan 5 butir soal (25%) kategori baik. Menurut Mardapi (dalam Rasyid dan Mansur 2007) angka 0,3 merupakan angka kriteria minimal suatu butir soal yang baik. Berdasarkan hasil data tersebut kualitas soal yang baik terdapat sebanyak 12 butir soal (60%) yaitu pada soal nomor 1, 2, 3, 4, 7, 9, 13, 14, 15, 17, 18, dan 20 karena butir soal tersebut sudah dapat membedakan antara kelompok atas dan kelompok bawah. Tetapi, untuk menentukan kualitas dari instrumen daya pembeda dikategorikan tidak baik, karna kriteria daya pembeda  $0,30 - 1,00 < 85\%$ . Menurut Suryabrata, 1999 (dalam Rasyid dan Mansur, 2007), tujuan pokok mencari daya pembeda ialah untuk menentukan apakah butir soal tersebut memiliki kemampuan membedakan kelompok dalam aspek yang diukur, sesuai dengan perbedaan yang ada pada kelompok tersebut.

d. Validitas item

Menurut Arikunto (2012), suatu item dikatakan valid apabila mempunyai dukungan yang besar terhadap skor total. Artinya, validitas ditentukan apabila  $r$  hitung  $>$   $r$  tabel. Hasil analisis menunjukkan bahwa validitas butir soal 16 soal (80%) dikategorikan valid terdapat pada nomor 1, 2, 3, 4, 5, 6, 7, 8, 9, 11, 13, 14, 15, 17, 18, dan 20. Sedangkan 4 soal (20%) invalid yang terdapat pada nomor 10, 12, 16, dan 19. Sehingga kualitas instrumen tes dikategorikan tidak valid karna nilai persentase validitas  $<$  85%.

Butir yang mempunyai validitas tertinggi adalah butir nomor 14 dengan koefisien korelasi 0,448 dengan kategori sedang. Dan yang paling rendah adalah butir soal nomor 16 dengan koefisien korelasi 0,075 dengan kategori sangat rendah. Pada penelitian ini  $r$  tabel yang diambil berdasarkan sampel yang digunakan sebanyak 151 peserta tes dengan nilai  $r$  tabel sebesar 0,159 dan tingkat kesalahan 5%.

e. Reliabilitas Tes

Hasil analisis reliabilitas tes dihitung dari butir soal yang valid sehingga diperoleh hasil 0,789 reliable dengan kategori instrument reliabel. Menurut Surapranata (2009) besar kecilnya reliabilitas suatu tes ditentukan oleh besar kecilnya nilai korelasi tes yang dinamakan indeks reliabilitas. Indeks reliabilitas ini berkisar antara 0 sampai 1. Semakin

tinggi koefisien reliabilitas suatu tes (mendekati 1), makin tinggi pula keajegan atau ketepatannya.

#### **4. KESIMPULAN DAN SARAN**

##### **1. Kesimpulan**

Berdasarkan data dan hasil pembahasan yang telah dilakukan maka dapatlah disimpulkan bahwa Program Analisis Butir soal telah dirancang dengan menggunakan ADDIE Model dengan tahapan Analisis, Perancangan, Pengembangan, implementasi dimana tiap tahapan perancangan itu dilakukan evaluasi diperoleh hasil bahwa program analisis butir soal ini adalah Valid dan reliabel dengan  $\alpha$  (alpha) Cronchbach = 0,743 dengan kategori reliabel.

##### **2. Saran**

Program ini sudah di ujikan secara empirik, sebaiknya program ini dapat diproduksi secara masal dan dapat di buatkan panduan secara terintegrasi dalam program agar pengguna program ini dapat menggukan secara baik untuk dapat memberikan hasil yang maksimal.

#### **DAFTAR PUSTAKA**

- [1] Anthony J. Nitko 1996, Educational Assessments Students. Prentice Hall PT. New York. 308
- [2] Anne Anastasi , Susana Urbina, 1997. Appropriate for undergraduate/graduate courses in psychological testing, tests and measurements, and test construction. Pearson. 184.
- [3] Linn and Gronlund. 1995. Measurement and assessment in teaching. Englewood Cliffs, N.J Merrill. 315.
- [4] Aiken Lewis R. 1994. Perspective of Individual Difference : Assessment of Intellectual Functioning, Plenium, Plenium Press New York. 63.
- [5] Rasyid, Harun dan Mansur., 2007, Penilaian Hasil Belajar, CV. Wacana Prima, Bandung.
- [6] Daryanto, H., 2007, *Evaluasi Pendidikan* ,Rineka Cipta, Jakarta.
- [7] Thoha, C., 2003, *Teknik Evaluasi Pendidikan*, Raja Grafindo Persada, Jakarta.
- [8] Muhammad Nasir, dkk (2012) Program Perbaikan Mutu Pembelajaran Melalui Analisis Butir Soal Ujian Nasional Di Kabupaten Karimun, Propinsi Kepulauan Riau, Laporan Pengabdian Masyarakat, Sedang Publikasi Jurnal.
- [9] Sudijono., 2009, *Pengantar Statistik Pendidikan*. Raja Grafindo Persada, Jakarta

- [10] Sudjana, Nana., 2008, *Penilaian Hasil Proses Belajar Mengajar*, Pt Remaja Rosdakarya, Bandung.
- [11] Arikunto, S., 2012, *Dasar – Dasar Evaluasi Pendidikan Edisi 2*, PT Bumi Aksara, Jakarta.
- [12] Arikunto, S., 2009, *Dasar – Dasar Evaluasi Pendidikan*, PT Bumi Aksara, Jakarta.
- [13] Arikunto, S., 2003, *Dasar – Dasar Evaluasi Pendidikan*, PT Bumi Aksara, Jakarta.
  
- [14] Yuliana, Nadia., 2009, *Hasil Belajar Keterampilan Kognitif Sains Fisika Siswa Dengan Bantuan Komputer (Computer Assisted Instruction) Menggunakan Media Pembelajaran Animasi Gaya Gravitasi Pada Kelas Xi Sma Muhammadiyah Pekanbaru*, Skripsi, FKIP Universitas Riau, Pekanbaru (tidak diterbitkan).