

PEMANFAATAN PROGRAM ITEMAN 3.0 UNTUK ANALISIS BUTIR SOAL LOMBA CERDAS CERMAT TEKNOLOGI INFORMASI DAN KOMUNIKASI TINGKAT SMA SEDERAJAT

Ilham Rais Arvianto

Prodi Teknik Informatika STMIK AKAKOM Yogyakarta
ir.arvianto@akakom.ac.id sugeng166oke@gmail.com

Abstrak

Tujuan dari penelitian ini adalah untuk mendeskripsikan kualitas soal penyisihan LCC TIK tingkat SMA sederajat pada kegiatan Akakom IT-Fest 2016. Tergolong jenis penelitian kualitatif dengan menggunakan teknik survey dengan 16 partisipan. Pengumpulan data menggunakan metode dokumentasi. Analisa data menggunakan teknik analisa butir secara kuantitatif menggunakan program ITEMAN 3.0. Kesimpulan dalam penelitian ini adalah soal penyisihan LCC IT Fest Akakom 2016 memiliki kualitas yang tidak baik. Hal ini dapat dilihat berdasarkan kriteria indeks kesukaran, daya beda dan pengecoh diperoleh sebanyak 56% soal yang tidak dipakai, 34% soal harus direvisi terlebih dahulu sebelum dipakai dan hanya 10% soal yang dapat langsung dipakai. Soal tersebut juga tidak reliable, karena memiliki nilai reliabilitas 0,679.

Kata kunci: *Kualitas Soal, Lomba Cerdas Cermat, dan ITEMAN*

Pendahuluan

Dalam kegiatan lomba cerdas cermat (LCC), instrument soal menjadi hal yang vital. Soal digunakan sebagai alat ukur penentu seorang peserta LCC dapat lolos ke babak berikutnya ataupun tidak. Sebagai alat ukur, soal harus benar-benar mengukur kemampuan peserta dengan sebaik-baiknya sehingga dalam penggunaannya akan menghasilkan pengukuran yang objektif.

Dari hasil observasi awal, ditemukan tiga set soal yang digunakan pada kegiatan beregu LCC IT-Fest Akakom 2016 pada tanggal 29 April 2016 tingkat SMA sederajat nasional. Tiga set soal tersebut terdiri dari satu set soal penyisihan berbentuk pilihan ganda (*multiple choice test*) serta satu set soal semifinal dan satu set soal final yang masing-masing berbentuk jawaban singkat. Pada penelitian ini, soal yang

akan dibahas lebih mendalam adalah soal penyisihan. Soal yang digunakan untuk LCC tersebut belum pernah dianalisis secara mendalam dari sisi tingkat kesukaran, daya beda, efektifitas distraktor dan reliabilitas. Berdasarkan skor pada babak penyisihan masih banyak peserta yang memperoleh nilai rendah. Hal ini mengindikasikan soal LCC belum memenuhi standar baku penyusunan tes.

Rustaman (2005) menyebutkan bahwa syarat tes yang baik ada lima, yaitu tes harus valid, reliabel, objektif, bersifat diagnostik serta efisien. Baik dan tidaknya soal dapat diketahui setelah dilakukan analisis terhadap soal yang bersangkutan. Arikunto (2009) menyebutkan kegunaan analisis soal adalah untuk mengidentifikasi butir-butir soal yang baik, cukup baik, atau kurang baik, serta memperoleh informasi yang dapat digunakan untuk

menyempurnakan soal-soal untuk kepentingan lebih lanjut dalam bentuk perbaikan dan memperoleh gambaran tentang karakteristik soal.

Cara menganalisis soal menurut Surapranata (2009) dapat dilakukan melalui dua cara, yaitu analisis kualitatif dan analisis kuantitatif. Analisis kualitatif sebagai validitas logis dalam aspek materi, konstruksi, dan bahasa untuk melihat berfungsi tidaknya sebuah soal serta analisis kuantitatif sebagai validitas empiris untuk melihat lebih berfungsi tidaknya sebuah soal, setelah soal diuji coba pada sampel yang representatif. Analisis kualitatif pada soal LCC telah dilakukan oleh juri lomba LCC. Analisis soal secara kualitatif dilakukan menurut prosedur validasi isi (*content validity*), yaitu melihat apakah soal LCC yang telah dirancang merupakan sampel yang representative dari seluruh materi TIK ataupun belum. Dalam hal ini, soal yang telah dirancang telah dinyatakan oleh juri LCC sudah memenuhi kriteria valid menurut validitas isi. Melihat hal tersebut, maka yang belum dilakukan adalah analisis soal dengan cara kuantitatif, sehingga untuk

penelitian ini akan menggunakan cara kuantitatif untuk menganalisis soal LCC tersebut. Penelitian-penelitian yang sejalan dengan penelitian ini yaitu Sukirno (2006), Masruroh (2012), Otaya (2014), Ruspidu (2014)

Pendekatan analisis dengan cara kuantitatif digunakan untuk melihat apakah suatu soal berfungsi dengan baik atau tidak. Soal yang berfungsi dengan baik harus memenuhi syarat-syarat berikut.

1. Tingkat Kesukaran

Untuk mengetahui tingkat kesukaran, terlebih dahulu harus mencari nilai indeks kesukaran (P). Indeks kesukaran adalah proporsi banyaknya siswa yang menjawab benar pada suatu soal dengan seluruh siswa. Indeks kesukaran setiap butir soal yang diperoleh dengan memakai rumus yang dikemukakan oleh Budiyono (2011).

Cara yang dapat digunakan untuk menafsirkan (menginterpretasikan) angka indeks kesukaran item, menurut Thorndike (Sudijono, 2011) disajikan dalam tabel berikut.

Tabel 1. Interpretasi Indeks Kesukaran

| Nilai Indeks Kesukaran (P) | Tingkat | Interpretasi |
|--------------------------------|---------|--------------|
| $0,00 \leq P < 0,30$ | Sulit | Tidak baik |
| $0,30 \leq P \leq 0,70$ | Sedang | Baik |
| $0,70 < P \leq 1,00$ | Mudah | Tidak baik |

Soal yang baik memiliki interpretasi tingkat indeks kesukaran sedang ($0,30 \leq P \leq 0,70$). Karena dengan tingkat kesukaran sedang dapat mengakibatkan nilai yang berdistribusi normal.

2. Daya Pembeda.

Soal harus memiliki daya beda yang baik, maksudnya adalah soal tersebut

dapat membedakan kategori peserta yang pandai (kelompok atas) dan tidak pandai (kelompok bawah). Salah satu cara untuk menghitung daya pembeda (D) menurut Budiyono (2011) adalah menggunakan rumus korelasi biserial titik (*point biserial correlation*) atau ditulis r_{p-bis} . Interpretasi nilai D tersaji dalam tabel berikut.

Tabel 2. Interpretasi Daya Pembeda

| Nilai Daya Pembeda (D) | Interpretasi |
|----------------------------|--------------|
| $-1,00 \leq P < 0,30$ | Tidak baik |
| $0,30 \leq P \leq 1,00$ | Baik |

Soal dengan daya pembeda bernilai negatif menandakan lebih banyak peserta tidak pandai yang menjawab benar pada soal tersebut. Soal dengan daya pembeda bernilai nol menandakan banyaknya peserta pandai dan tidak pandai yang menjawab benar pada soal tersebut sama banyaknya. Soal yang baik adalah soal yang lebih banyak dijawab benar oleh peserta pandai dibandingkan dengan peserta tidak pandai, yaitu dengan daya pembeda positif. Oleh karena itu, menurut Budiyo (2011) soal yang memiliki daya beda baik adalah soal dengan daya beda positif, atau biasa digunakan nilai ambang $D \geq 0,30$ (daya beda baik).

3. Berfungsinya Pengecoh

Pengecoh (*distractor*) dirancang dengan tujuan untuk mengecoh peserta yang tidak menguasai materi dengan baik. Oleh karena itu, pengecoh yang baik berarti harus dapat mengecoh. Syarat pengecoh yang baik yaitu dipilih minimal 5% peserta (Budiyo, 2011).

4. Reliabilitas

Reliabilitas (*reliability*) biasa ditulis dengan r_{11} menunjukkan

kekonsistensian suatu soal dalam melakukan suatu pengukuran. Soal dikatakan reliabel jika pada waktu kapanpun dilakukan pengukuran, jika objek pengukuran tetap maka hasil pengukurannya juga akan tetap. Terdapat beberapa rumus yang dapat digunakan untuk mencari nilai reliabilitas dari suatu soal, antara lain menggunakan metode satu kali tes, teknik Spearman-Brown, teknik Flanagan, teknik Rulon, teknik Kuder-Richardson, dan teknik Alpha. Nilai reliabilitas yang digunakan untuk menentukan reliabel atau tidaknya suatu soal sebetulnya tidak mempunyai patokan yang baku. Tetapi biasanya diambil nilai reliabilitas $\geq 0,70$. Ini berarti, hasil pengukuran yang mempunyai nilai reliabilitas 0,70 atau lebih dikatakan reliabel (Budiyo, 2003).

Dari kombinasi baik maupun tidak baiknya nilai indeks kesukaran, daya pembeda dan pengecoh dapat memunculkan beberapa kemungkinan dalam pengambilan keputusan. Pedoman pengambilan keputusan yang digunakan disajikan pada tabel berikut.

Tabel 3. Pedoman Pengambilan Keputusan

| Tingkat Kesukaran | Daya Beda | Pengecoh | Keputusan |
|-------------------|------------|------------|---------------|
| Baik | Baik | Baik | Dipakai |
| Baik | Baik | Tidak baik | Dipakai |
| Baik | Tidak baik | Baik | Revisi |
| Tidak baik | Baik | Baik | Revisi |
| Baik | Tidak baik | Tidak baik | Revisi |
| Tidak baik | Baik | Tidak baik | Revisi |
| Tidak baik | Tidak baik | Baik | Tidak Dipakai |
| Tidak baik | Tidak baik | Tidak baik | Tidak Dipakai |

Analisis soal secara kuantitatif dapat dilakukan secara manual maupun dengan menggunakan program komputer. Analisis secara manual dilakukan dengan memanfaatkan rumus yang sudah ada seperti rumus untuk mencari tingkat kesukaran, daya beda, pengecoh maupun reliabilitas. Sedangkan untuk program komputer dilakukan menggunakan program seperti *Anbuso*, *Anatest* dan *ITEMAN*. Tujuan dilakukannya analisis soal menurut Aiken (1994) dalam Masruroh (2012) adalah merevisi atau membuang soal yang tidak efektif dan untuk mengetahui informasi diagnostik pada siswa apakah sudah memahami materi yang diajarkan atau belum. Tes yang baik dan dapat dipergunakan harus reliabel, memiliki daya beda dan tingkat kesukaran yang baik.

Salah satu paket program komputer yang ditawarkan untuk melakukan analisis butir adalah *ITEMAN*, singkatan dari *Item Analysis*. *ITEMAN* dipilih di antara yang lain karena merupakan program komputer yang dipergunakan untuk menganalisis soal secara klasik. Artinya program ini menggunakan rumus-rumus baku yang sudah lazim digunakan dalam melakukan analisis soal. Paket program *ITEMAN* dibuat oleh *Assessment Systems Corporation* di Amerika Serikat. *ITEMAN* Versi 3.0 masih menggunakan sistem operasi DOS.

File input data yang akan dianalisis diketik dalam Notepad dengan ekstension **.dat**, misalkan *DATA.dat*. Berikut ini adalah pedoman penulisan file input data pada *ITEMAN*.

1. Baris pertama berisi kode-kode sebagai berikut.

Tabel 4. Pedoman Penulisan File Input Data *ITEMAN* Baris Pertama

| Kolom | Keterangan | Contoh |
|--------|---|--------|
| 1 – 3 | Banyaknya butir yang dianalisis | 020 |
| 4 | Kosong/spasi | - |
| 5 | Untuk jawaban omit/kosong | O |
| 6 | Kosong/spasi | - |
| 7 | Untuk butir soal yang tidak (belum sempat) dikerjakan | N |
| 8 | Kosong/spasi | - |
| 9 – 10 | Banyaknya kolom yang diperlukan untuk identitas peserta tes | 04 |

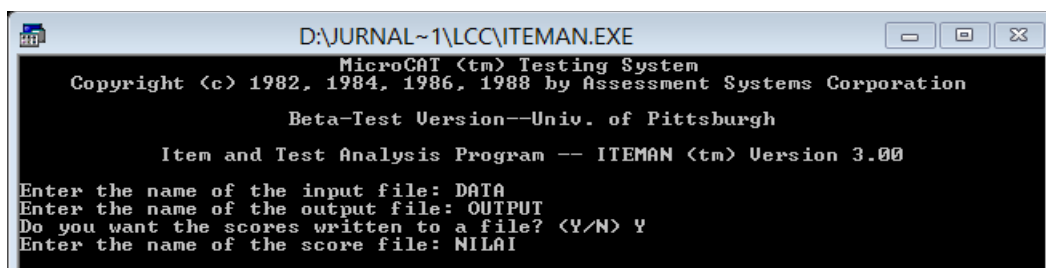
2. Baris kedua berisi kunci jawaban

3. Baris ketiga berisi banyaknya alternatif jawaban

4. Baris keempat: berisi kode: "Y" berarti butir dianalisis, "N" butir tidak dianalisis.

Pada ITEMAN 3.0, file yang harus dieksekusi diberi nama **ITEMAN.EXE**. Program ini bisa diunduh di

<http://www.assess.com/iteman>. Program ITEMAN. EXE ini harus diletakkan dalam satu folder yang sama dengan file input data. Berikut adalah tampilan awal ITEMAN 3.0 ketika dijakankan.



```

D:\JURNAL~1\LCC\ITEMAN.EXE
MicroCAT (tm) Testing System
Copyright (c) 1982, 1984, 1986, 1988 by Assessment Systems Corporation
Beta-Test Version--Univ. of Pittsburgh
Item and Test Analysis Program -- ITEMAN (tm) Version 3.00
Enter the name of the input file: DATA
Enter the name of the output file: OUTPUT
Do you want the scores written to a file? (Y/N) Y
Enter the name of the score file: NILAI

```

Gambar 1 Tampilan Awal ITEMAN 3.0

Petunjuk penggunaan program ITEMAN 3.0 adalah:

- *Enter the name of the input file* (masukkan nama dari file input), isikan dengan nama dari file input, misal DATA. Kemudian tekan tombol Enter.
- *Enter the name of the output file* (masukkan nama file yang akan digunakan sebagai output), misal OUTPUT. Kemudian tekan tombol Enter.
- *Do you want the scores written to a file?* (Y/N) (Apakah Anda ingin menuliskan nilai peserta pada sebuah file), jika Anda menginginkan hal ini isi dengan Y, tetapi jika tidak isi dengan N. Kemudian tekan tombol Enter.
- *Enter the name of the score file* (masukkan nama file yang akan digunakan sebagai nilai), misal NILAI. Kemudian tekan tombol Enter untuk mengakhiri program.

Terkait belum dilakukannya analisis kuantitatif terhadap soal penyisihan LCC tersebut, maka dilakukan penelitian dengan tujuan untuk mendeskripsikan kualitas soal penyisihan LCC TIK tingkat SMA sederajat pada kegiatan Akakom IT-Fest 2016. Adapun materi TIK SMA yang diambil adalah pada kelas X, XI dan XII kurikulum KTSP.

METODE PENELITIAN

Penelitian ini tergolong dalam jenis penelitian deskriptive kualitatif, karena bertujuan mengungkap kualitas dari objek penelitian. Dalam hal ini objek penelitiannya adalah soal penyisihan LCC TIK tingkat SMA sederajat Akakom IT-Fest 2016. Instrumen tes yang dikembangkan mengacu kepada silabi TIK kurikulum KTSP 2006. Soal penyisihan berbentuk pilihan ganda (*multiple choice*) sebanyak 50 soal dengan 5 alternatif jawaban (A, B, C, D dan E) yang dikerjakan dengan durasi 30 menit secara

berkelompok (3 orang tiap kelompok). Proporsi materi TIK yang digunakan adalah 25 soal kelas X, 10 soal kelas XI, dan 15 soal kelas XII.

Penelitian ini dilaksanakan dengan menggunakan teknik survei, yaitu mendapatkan dan mengumpulkan informasi tentang karakteristik dari objek penelitian. Informasi yang diperoleh selanjutnya digunakan untuk mengetahui kualitas, kelayakan dan karakteristik soal LCC TIK tingkat SMA berdasarkan analisis secara kuantitatif.

Penelitian ini menggunakan metode dokumentasi atau telaah dokumen sebagai metode pengumpulan datanya. Data dokumentasi yang diamati adalah data primer karena langsung bersumber dari partisipan, yaitu 16 tim peserta lomba dan panitia

pelaksanaan LCC. Data diperoleh dengan menelusuri arsip-arsip sekolah yang berupa benda mati yaitu berupa soal ujian, kunci jawaban dan hasil jawaban pekerjaan peserta.

Analisa data dalam penelitian ini menggunakan teknik analisis butir secara kuantitatif. Analisis secara kuantitatif dilakukan dengan menggunakan Program ITEMAN (*Item and Test Analysis*) versi 3.00 yang hasilnya meliputi tingkat kesukaran, daya pembeda soal, efektifitas pengecoh serta reliabilitas soal.

Hasil Dan Pembahasan

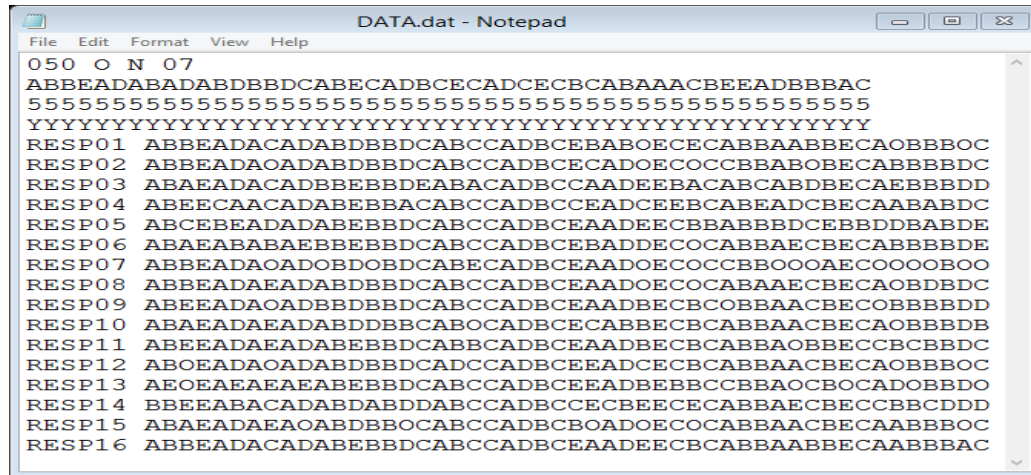
Dari hasil dokumentasi, diperoleh sebaran jawaban dari 16 partisipan yang telah menjawab soal LCC tersebut. dijalankan maka tampilan akan terlihat seperti pada gambar di bawah ini:

Tabel 5. Kunci Jawaban dan Data Sebaran Jawaban dari Partisipan

| KUNCI | A | B | B | E | A | D | A | B | A | D | A | B | D | B | B | D | C | A | B | E | C | A | D | B | C | E | C | A | D | C | E | C | B | C | A | B | A | A | A | C | B | E | E | A | D | B | B | B | A | C |
|--------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| SOAL | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 | 45 | 46 | 47 | 48 | 49 | 50 |
| RESP01 | A | B | B | E | A | D | A | C | A | D | A | B | D | B | B | D | C | A | B | C | C | A | D | B | C | E | B | A | B | O | E | C | E | C | A | B | B | A | A | B | B | E | C | A | O | B | B | O | C | |
| RESP02 | A | B | B | E | A | D | A | O | A | D | A | B | D | B | B | D | C | A | B | C | C | A | D | B | C | E | C | A | D | O | E | C | O | C | C | B | B | A | B | O | B | E | C | A | B | B | B | D | C | |
| RESP03 | A | B | A | E | A | D | A | C | A | D | B | B | E | B | B | D | E | A | B | A | C | A | D | B | C | C | A | A | D | E | E | B | A | C | A | B | C | A | B | D | B | E | C | A | E | B | B | D | D | |
| RESP04 | A | B | E | E | C | A | A | C | A | D | A | B | E | B | B | A | C | A | B | C | C | A | D | B | C | C | E | A | D | C | E | E | B | C | A | B | E | A | D | C | B | E | C | A | A | B | A | B | D | C |
| RESP05 | A | B | C | E | B | E | A | D | A | D | A | B | E | B | B | D | C | A | B | C | C | A | D | B | C | E | A | A | D | E | E | C | B | B | A | B | B | B | D | C | E | B | B | D | D | B | A | B | D | E |
| RESP06 | A | B | A | E | A | B | A | B | A | E | B | B | E | B | B | D | C | A | B | C | C | A | D | B | C | E | B | A | D | D | E | C | O | C | A | B | B | A | E | C | B | E | C | A | B | B | B | D | E | |
| RESP07 | A | B | B | E | A | D | A | O | A | D | O | B | D | O | B | D | C | A | B | E | C | A | D | B | C | E | A | A | D | O | E | C | O | C | C | B | B | O | O | A | E | C | O | O | O | B | O | O | | |
| RESP08 | A | B | B | E | A | D | A | E | A | D | A | B | D | B | B | D | C | A | B | C | C | A | D | B | C | E | A | A | D | O | E | C | O | C | A | B | A | A | E | C | B | E | C | A | O | B | D | B | D | C |
| RESP09 | A | B | E | E | A | D | A | O | A | D | B | B | D | B | B | D | C | A | B | C | C | A | D | B | C | E | A | A | D | B | E | C | B | C | O | B | B | A | A | C | B | E | C | O | B | B | B | D | D | |
| RESP10 | A | B | A | E | A | D | A | E | A | D | A | B | D | D | B | B | C | A | B | O | C | A | D | B | C | E | C | A | B | B | E | C | B | C | A | B | B | A | A | C | B | E | C | A | O | B | B | B | D | B |
| RESP11 | A | B | E | E | A | D | A | E | A | D | A | B | E | B | B | D | C | A | B | B | C | A | D | B | C | E | A | A | D | B | E | C | B | C | A | B | B | A | O | B | B | E | C | C | B | C | B | B | D | C |
| RESP12 | A | B | O | E | A | D | A | O | A | D | A | B | D | B | B | D | C | A | D | C | C | A | D | B | C | E | E | A | D | C | E | C | B | C | A | B | B | A | A | C | B | E | C | A | O | B | B | B | O | C |
| RESP13 | A | E | O | E | A | E | A | E | A | E | A | B | E | B | B | D | C | A | B | C | C | A | D | B | C | E | E | A | D | B | E | B | B | C | C | B | B | A | O | C | B | O | C | A | D | O | B | B | D | O |
| RESP14 | B | B | E | E | A | B | A | C | A | D | A | B | D | A | B | D | A | B | C | C | A | D | B | C | C | E | C | B | E | E | C | E | C | A | B | B | A | E | C | B | E | C | C | B | B | C | D | D | D | |
| RESP15 | A | B | A | E | A | D | A | E | A | O | A | B | D | B | O | C | A | B | C | C | A | D | B | C | B | O | A | D | O | E | C | O | C | A | B | B | A | A | C | B | E | C | A | A | B | B | O | C | | |
| RESP16 | A | B | B | E | A | D | A | C | A | D | A | B | E | B | B | D | C | A | B | C | C | A | D | B | C | E | A | A | D | E | E | C | B | C | A | B | B | A | A | B | B | E | C | A | A | B | B | A | C | |

Data tersebut kemudian ditulis dalam NOTEPAD sesuai dengan pedoman file input data pada ITEMAN. File tersebut kemudian

disimpan dengan nama DATA.dat. Berikut ini adalah tampilan file input datanya.



Gambar 2 Tampilan File Input Data

Selanjutnya, dijalankan program ITEMAN.EXE. Inputkan masukan yang diminta pada program tersebut seperti contoh pada Gambar 1. Setelah selesai, terlihat ada

file baru untuk OUTPUT dan NILAI. Berikut ini adalah potongan tampilan dari file OUTPUT.

| Item Statistics | | | | | Alternative Statistics | | | | |
|-----------------|-------------|---------------|--------|--------------|------------------------|-----------------|--------|--------------|-----|
| Seq. No. | Scale -Item | Prop. Correct | Biser. | Point Biser. | Alt. | Prop. Endorsing | Biser. | Point Biser. | Key |
| 1 | 0-1 | 0.938 | 1.000 | 0.513 | A | 0.938 | 1.000 | 0.513 | * |
| | | | | | B | 0.063 | -1.000 | -0.513 | |
| | | | | | C | 0.000 | -9.000 | -9.000 | |
| | | | | | D | 0.000 | -9.000 | -9.000 | |
| | | | | | E | 0.000 | -9.000 | -9.000 | |
| | | | | | Other | 0.000 | -9.000 | -9.000 | |

Scale Statistics

| | |
|----------------|--------|
| Scale: | 0 |
| N of Items | 50 |
| N of Examinees | 16 |
| Mean | 35.688 |
| Variance | 14.965 |
| Std. Dev. | 3.868 |
| Skew | -0.320 |
| Kurtosis | -1.032 |
| Minimum | 28.000 |
| Maximum | 41.000 |
| Median | 36.000 |
| Alpha | 0.679 |
| SEM | 2.193 |
| Mean P | 0.714 |
| Mean Item-Tot. | 0.281 |
| Mean Biserial | 0.426 |

Output ITEMAN dapat dibedakan ke dalam 2 bagian yaitu: (1) Statistik butir soal dan (2) statistik tes/skala. Berikut menunjukkan hasil analisis statistik butir soal dan hasil analisis statistik tes/skala:

1. Statistik butir soal.

Untuk tes yang terdiri dari butir-butir soal yang bersifat dikotomi misalnya pilihan ganda, statistik berikut adalah output dari setiap butir soal yang dianalisis.

- Seq. No. adalah nomor urut butir soal dalam file data.
- Scala item adalah nomor urut butir soal dalam tes.
- Prop. Correct adalah Proporsi siswa yang menjawab benar butir tes. Nilai ekstrem mendekati nol atau satu menunjukkan bahwa butir soal tersebut terlalu sukar atau terlalu mudah untuk peserta tes. Indeks ini disebut juga indeks tingkat kesukaran soal secara klasikal.
- Biser adalah indeks daya pembeda soal dengan menggunakan koefisien korelasi biserial.
- Point biserial adalah juga indeks daya pembeda soal dan pilihan jawaban (alternatif) dengan menggunakan koefisien korelasi point biserial.

Keterangan di atas merupakan penjelasan untuk bagian *Item Statistics*, Sementara pada bagian *Alternative Statistics* yang perlu diperhatikan adalah Alt. (pilihan jawaban), Prop. Endorsing (proporsi peserta tes yang menjawab alternative jawaban), dan Point Biser.

(daya beda pengecoh). Tanda bintang * yang muncul di sebelah kanan hasil analisis menunjukkan kunci jawaban.

2. Statistik tes.

Disamping menampilkan statistic butir soal, ITEMAN juga menampilkan statistic tes untuk keseluruhan. Statistik tes yang dimaksudkan adalah:

- N of Items: Jumlah butir soal dalam tes yang dianalisis. Untuk tes yang terdiri atas butir-butir soal dikotomi hal ini merupakan jumlah total butir soal dalam tes.
- N of Examinees: Jumlah peserta tes yang digunakan dalam analisis.
- Mean: skor/rata-rata peserta tes
- Variance: Varian dari distribusi skor peserta tes yang memberikan gambaran tentang sebaran skor peserta tes.
- Std. Deviasi: Deviasi standar dari distribusi skor peserta tes. Deviasi standar merupakan akar dari variance.
- Skew. : Kemiringan distribusi skor peserta yang memberikan gambaran tentang bentuk distribusi skor peserta tes. Kemiringan negatif menunjukkan bahwa sebagian besar skor berada pada bagian atas/skor tinggi dari distribusi skor. Sebaliknya kemiringan positif menunjukkan bahwa sebagian besar skor berada pada bagian bawah/skor rendah dari distribusi skor. Kemiringan nol menunjukkan bahwa skor berdistribusi secara simetris di sekitar rata-rata/mean.

- Kurtosis: Puncak distribusi skor yang menggambarkan kelandaian distribusi skor dibanding dengan distribusi normal. Nilai positif menunjukkan distribusi yang lebih lancip/memuncak dan nilai negatif menunjukkan distribusi yang lebih landai/merata. Kurtosis untuk distribusi normal adalah nol.
 - Minimum: Skor terendah peserta dalam tes.
 - Maximum: Skor tertinggi peserta dalam tes.
 - Median: Skor tengah dimana 50% berada pada/lebih rendah dari skor tersebut.
 - Alpha: Koefisien reliabilitas alpha untuk tes yang merupakan indeks homogenitas tes. Koefisien alpha bergerak dari 0.0 sampai 1.0. Koefisien alpha hanya cocok digunakan pada tes yang bukan mengukur kecepatan dan hanya mengukur satu dimensi (single trait). Reliabilitas adalah kestabilan skor yang diperoleh orang yang sama ketika diuji ulang dengan tes yang sama pada situasi yang berbeda. Jadi reliabilitas dapat dikatakan sebagai tingkat konsistensi hasil dua pengukuran terhadap hal yang sama. Dengan perangkat soal yang reliable apabila tes diberikan dua kali pada orang yang sama tetapi dalam selang waktu yang berbeda sepanjang tidak ada perubahan kemampuan maka skor yang diperoleh akan konstan.
 - SEM: kesalahan pengukuran standar untuk setiap tes. SEM merupakan estimasi dari deviasi standar kesalahan pengukuran dalam skor tes. Kesalahan pengukuran membantu pemakai tes dalam memahami kesalahan yang bersifat random yang mempengaruhi skor peserta dalam pelaksanaan tes.
 - Mean p: rata-rata tingkat kesukaran semua butir soal dalam tes secara klasikal dihitung dengan cara mencari rata-rata proporsi peserta tes yang menjawab benar untuk semua butir soal dalam tes.
 - Mean item tot: Nilai rata-rata indeks daya pembeda dari semua soal dalam tes yang diperoleh dengan menghitung nilai rata-rata point biserial dari semua soal dalam tes.
 - Mean biserial: Nilai rata-rata indeks daya pembeda yang diperoleh dengan menghitung nilai rata-rata korelasi biserial dari semua butir soal dalam tes.
 - Scale intercorelation: Indeks korelasi antara skor-skor peserta tes yang diperoleh dari setiap subtes.
- Hasil analisis secara kuantitatif dari 50 soal pilihan ganda menggunakan program ITEMAN 3.0 menurut indeks kesukaran, daya beda dan pengecoh tersaji pada tabel berikut.

Tabel 6. Hasil Analisis dengan ITEMAN 3.0

| No. Soal | Tingkat Kesukaran | | Daya Pembeda | | Pengecoh (%) | | | | | | Keputusan | |
|--------------|-------------------|------------|--------------|------------|--------------|------|------|------|------|------------|---------------|--|
| | Angka | Kategori | Angka | Kategori | A | B | C | D | E | Kategori | | |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | |
| 1 | 0.938 | Tidak baik | 0.513 | Baik | * | 6.3 | 0.0 | 0.0 | 0.0 | Tidak baik | Revisi | |
| 2 | 0.938 | Tidak baik | 0.246 | Tidak baik | 0.0 | * | 0.0 | 0.0 | 0.0 | Tidak baik | Tidak Dipakai | |
| 3 | 0.313 | Baik | 0.403 | Baik | 25.0 | * | 6.3 | 0.0 | 25.0 | Tidak baik | Dipakai | |
| 4 | 1.000 | Tidak baik | -9.000 | Tidak baik | 0.0 | 0.0 | 0.0 | 0.0 | * | Tidak baik | Tidak Dipakai | |
| 5 | 0.875 | Tidak baik | 0.311 | Baik | * | 6.3 | 6.3 | 0.0 | 0.0 | Tidak baik | Revisi | |
| 6 | 0.688 | Baik | 0.643 | Baik | 6.3 | 12.5 | 0.0 | * | 12.5 | Baik | Dipakai | |
| 7 | 1.000 | Tidak baik | -9.000 | Tidak baik | * | 0.0 | 0.0 | 0.0 | 0.0 | Tidak baik | Tidak Dipakai | |
| 8 | 0.063 | Tidak baik | -0.046 | Tidak baik | 0.0 | * | 31.3 | 6.3 | 31.3 | Tidak baik | Tidak Dipakai | |
| 9 | 1.000 | Tidak baik | -9.000 | Tidak baik | * | 0.0 | 0.0 | 0.0 | 0.0 | Tidak baik | Tidak Dipakai | |
| 10 | 0.813 | Tidak baik | 0.127 | Tidak baik | 0.0 | 0.0 | 0.0 | * | 12.5 | Tidak baik | Tidak Dipakai | |
| 11 | 0.750 | Tidak baik | 0.289 | Tidak baik | * | 18.8 | 0.0 | 0.0 | 0.0 | Tidak baik | Tidak Dipakai | |
| 12 | 1.000 | Tidak baik | -9.000 | Tidak baik | 0.0 | * | 0.0 | 0.0 | 0.0 | Tidak baik | Tidak Dipakai | |
| 13 | 0.563 | Baik | 0.287 | Tidak baik | 0.0 | 0.0 | 0.0 | * | 43.8 | Tidak baik | Revisi | |
| 14 | 0.813 | Tidak baik | 0.417 | Baik | 6.3 | * | 0.0 | 6.3 | 0.0 | Tidak baik | Revisi | |
| 15 | 1.000 | Tidak baik | -9.000 | Tidak baik | 0.0 | * | 0.0 | 0.0 | 0.0 | Tidak baik | Tidak Dipakai | |
| 16 | 0.813 | Tidak baik | -0.080 | Tidak baik | 6.3 | 6.3 | 0.0 | * | 0.0 | Tidak baik | Tidak Dipakai | |
| 17 | 0.875 | Tidak baik | 0.556 | Baik | 0.0 | 0.0 | * | 6.3 | 6.3 | Tidak baik | Revisi | |
| 18 | 1.000 | Tidak baik | -9.000 | Tidak baik | * | 0.0 | 0.0 | 0.0 | 0.0 | Tidak baik | Tidak Dipakai | |
| 19 | 0.938 | Tidak baik | -0.355 | Tidak baik | 0.0 | * | 0.0 | 6.3 | 0.0 | Tidak baik | Tidak Dipakai | |
| 20 | 0.063 | Tidak baik | -0.313 | Tidak baik | 6.3 | 6.3 | 75.0 | 0.0 | * | Tidak baik | Tidak Dipakai | |
| 21 | 1.000 | Tidak baik | -9.000 | Tidak baik | 0.0 | 0.0 | * | 0.0 | 0.0 | Tidak baik | Tidak Dipakai | |
| 22 | 1.000 | Tidak baik | -9.000 | Tidak baik | * | 0.0 | 0.0 | 0.0 | 0.0 | Tidak baik | Tidak Dipakai | |
| 23 | 1.000 | Tidak baik | -9.000 | Tidak baik | 0.0 | 0.0 | 0.0 | * | 0.0 | Tidak baik | Tidak Dipakai | |
| 24 | 1.000 | Tidak baik | -9.000 | Tidak baik | 0.0 | * | 0.0 | 0.0 | 0.0 | Tidak baik | Tidak Dipakai | |
| 25 | 1.000 | Tidak baik | -9.000 | Tidak baik | 0.0 | 0.0 | * | 0.0 | 0.0 | Tidak baik | Tidak Dipakai | |
| 26 | 0.750 | Tidak baik | 0.438 | Baik | 0.0 | 6.3 | 18.8 | 0.0 | * | Tidak baik | Revisi | |
| 27 | 0.125 | Tidak baik | 0.275 | Tidak baik | 43.8 | 12.5 | * | 0.0 | 25.0 | Tidak baik | Tidak Dipakai | |
| 28 | 0.938 | Tidak baik | 0.513 | Baik | * | 0.0 | 6.3 | 0.0 | 0.0 | Tidak baik | Revisi | |
| 29 | 0.813 | Tidak baik | 0.085 | Tidak baik | 0.0 | 18.8 | 0.0 | * | 0.0 | Tidak baik | Tidak Dipakai | |
| 30 | 0.125 | Tidak baik | 0.177 | Tidak baik | 0.0 | 25.0 | * | 6.3 | 25.0 | Tidak baik | Tidak Dipakai | |
| 31 | 1.000 | Tidak baik | -9.000 | Tidak baik | 0.0 | 0.0 | 0.0 | 0.0 | * | Tidak baik | Tidak Dipakai | |
| 32 | 0.813 | Tidak baik | 0.375 | Baik | 0.0 | 12.5 | * | 0.0 | 6.3 | Tidak baik | Revisi | |
| 33 | 0.500 | Baik | 0.145 | Tidak baik | 6.3 | * | 0.0 | 0.0 | 12.5 | Tidak baik | Revisi | |
| 34 | 0.938 | Tidak baik | 0.313 | Baik | 0.0 | 6.3 | * | 0.0 | 0.0 | Tidak baik | Revisi | |
| 35 | 0.750 | Tidak baik | 0.140 | Tidak baik | * | 0.0 | 18.8 | 0.0 | 0.0 | Tidak baik | Tidak Dipakai | |
| 36 | 1.000 | Tidak baik | -9.000 | Tidak baik | 0.0 | * | 0.0 | 0.0 | 0.0 | Tidak baik | Tidak Dipakai | |
| 37 | 0.063 | Tidak baik | 0.288 | Tidak baik | * | 81.3 | 6.3 | 0.0 | 6.3 | Tidak baik | Tidak Dipakai | |
| 38 | 0.875 | Tidak baik | 0.458 | Baik | * | 6.3 | 0.0 | 0.0 | 0.0 | Tidak baik | Tidak Dipakai | |
| 39 | 0.375 | Baik | 0.630 | Baik | * | 12.5 | 0.0 | 12.5 | 18.8 | Tidak baik | Dipakai | |
| 40 | 0.625 | Baik | -0.129 | Tidak baik | 0.0 | 18.8 | * | 6.3 | 0.0 | Tidak baik | Revisi | |
| 41 | 0.875 | Tidak baik | 0.458 | Baik | 6.3 | * | 0.0 | 0.0 | 6.3 | Tidak baik | Revisi | |
| 42 | 0.875 | Tidak baik | 0.409 | Baik | 0.0 | 6.3 | 0.0 | 0.0 | * | Tidak baik | Revisi | |
| 43 | 0.000 | Tidak baik | -9.000 | Tidak baik | 0.0 | 6.3 | 93.8 | 0.0 | * | Tidak baik | Tidak Dipakai | |
| 44 | 0.688 | Baik | 0.538 | Baik | * | 0.0 | 12.5 | 6.3 | 0.0 | Tidak baik | Dipakai | |
| 45 | 0.125 | Tidak baik | -0.409 | Tidak baik | 18.8 | 31.3 | 0.0 | * | 6.3 | Tidak baik | Tidak Dipakai | |
| 46 | 0.813 | Tidak baik | 0.334 | Baik | 0.0 | * | 6.3 | 0.0 | 0.0 | Tidak baik | Revisi | |
| 47 | 0.688 | Baik | 0.503 | Baik | 12.5 | * | 6.3 | 6.3 | 0.0 | Tidak baik | Dipakai | |
| 48 | 0.938 | Tidak baik | 0.513 | Baik | 0.0 | * | 0.0 | 6.3 | 0.0 | Tidak baik | Revisi | |
| 49 | 0.063 | Tidak baik | 0.355 | Baik | * | 0.0 | 0.0 | 68.8 | 0.0 | Tidak baik | Revisi | |
| 50 | 0.500 | Baik | 0.695 | Baik | 0.0 | 6.3 | * | 18.8 | 12.5 | Tidak baik | Dipakai | |
| Σ Baik | | 9 | | 20 | | | | | | | 1 | |
| Σ Tidak Baik | | 41 | | 30 | | | | | | | 49 | |

Dari kategori indeks kesukaran, soal yang baik adalah soal yang memiliki tingkat kesukaran sedang nilai proporsi $0,30 \leq P \leq 0,70$. Soal LCC IT ini termasuk soal dengan kualitas tidak baik dari segi indeks kesukaran karena dari 50 soal, terdapat 9 soal (18%) yang memenuhi tingkat kesukaran sedang dan 41 soal tidak memenuhi. Menurut Arikunto (2009) soal yang baik adalah soal yang tidak terlalu mudah atau terlalu sukar karena soal yang terlalu mudah tidak akan mendorong siswa untuk memecahkannya serta soal yang terlalu sulit akan menyebabkan siswa putus asa dan tidak mau mencobanya sebab itu diluar batas kemampuannya. Jika dianalisis dari soal yang tidak memenuhi kriteria, sebanyak 31 soal (62%) memiliki kategori mudah ($0,70 < P \leq 1,00$) dan 10 soal (20%) memiliki kategori sulit. Ini menandakan, kualitas sebagian besar soal penyisihan LCC IT terlalu mudah. Hal ini kemungkinan disebabkan peserta LCC berasal dari seleksi yang ketat dari tiap sekolahnya. Oleh karena itu tingkat kemampuannya tergolong tinggi. Sebuah soal jika diberikan kepada siswa yang berkemampuan tinggi tingkat kesukarannya akan menjadi kecil (terasa mudah) sebaliknya jika diberikan kepada siswa yang berkemampuan rendah tingkat kesukarannya akan menjadi tinggi (terasa sulit). Sehingga soal yang disusun perlu ditingkatkan tingkat kesukarannya. Salah satu caranya adalah dengan memberikan soal yang sifatnya aplikatif dan pengembangan.

Daya beda suatu butir tes merupakan kemampuan tes tersebut dalam membedakan

peserta pandai (kelompok atas) dengan peserta tidak pandai (kelompok bawah). Dari data di atas diperoleh 20 soal (40%) memiliki daya beda baik, sedangkan 30 soal (60%) memiliki daya beda tidak baik. Dengan demikian dapat dikatakan bahwa soal LCC TI tersebut memiliki kualitas daya beda yang tidak baik. Pada set soal terdapat 20 soal yang memiliki indeks daya beda negatif, yaitu nomor 4, 7, 8, 9, 12, 15, 16, 18, 19, 20, 21, 22, 23, 24, 25, 31, 36, 40, 43 dan 45. Butir soal seperti ini menunjukkan bahwa peserta tes yang menjawab benar soal tersebut memiliki skor yang relatif rendah sehingga butir soal tersebut tidak bisa membedakan siswa berkemampuan tinggi dengan rendah. Beberapa faktor yang menyebabkannya adalah soal yang kurang tegas dan jelas perumusannya, soal yang bersifat mendua dapat menyebabkan pengertian yang berbeda di siswa.

Suatu distraktor dikatakan baik apabila dipilih oleh minimal 5% peserta tes (Budiyono, 2011). Pada soal LCC TI ini hanya 1 soal (2%) saja yang memiliki pengecoh yang baik, yaitu pada soal nomor 6. Selain itu, sebanyak 49 soal (98%) memiliki pengecoh yang kurang/tidak berfungsi dengan baik. Dengan demikian dapat dikatakan bahwa soal LCC TI tersebut memiliki kualitas distractor/pengecoh yang tidak baik. Distraktor yang kurang berfungsi dengan baik sebanyak 21 soal (42%), artinya dari 4 distraktor yang direncanakan hanya 2 atau 3 distraktor saja yang berfungsi. Ini menandakan distractor yang sebagian tidak berfungsi perlu dilakukan revisi. Berfungsi

tidaknya suatu distraktor banyak ditentukan oleh cara penyusunan suatu tes. Distraktor yang disusun tanpa memperhatikan homogenitas dari alternatif jawaban akan menyebabkan tidak dipilihnya option tersebut, akibatnya tidak dapat berfungsi sebagaimana yang diharapkan. Alternatif

jawaban yang terlalu panjang juga menjadi salah satu penyebab kurang berfungsinya kualitas distraktor.

Dari Tabel 6, selanjutnya jika dilakukan rekap menurut soal yang dipakai, revisi dan tidak dipakai. Hasilnya tersaji pada tabel berikut.

Tabel 7. Rekap Hasil Analisis dengan ITEMAN 3.0

| Keputusan | No. Soal | Banyak Soal | Persentase Soal |
|---------------|---|-------------|-----------------|
| Dipakai | 3, 6, 39, 47, 50 | 5 | 10% |
| Revisi | 1, 5, 13, 14, 17, 26, 28, 32, 33, 34, 40, 41, 42, 44, 46, 48, 49 | 17 | 34% |
| Tidak dipakai | 2, 4, 7, 8, 9, 10, 11, 12, 15, 16, 18, 19, 20, 21, 22, 23, 24, 25, 27, 29, 30, 31, 35, 36, 37, 38, 43, 45 | 28 | 56% |

Berdasarkan Tabel 6, terlihat bahwa 56% soal tidak dipakai, 34% soal harus direvisi terlebih dahulu sebelum dipakai dan 10% soal yang dapat langsung dipakai. Melihat persentase yang begitu besar dari keputusan soal yang tidak dipakai, maka secara umum dapat dikatakan soal LCC IT ini masih banyak kekurangan (kurang memenuhi kualitas).

Dari output pada bagian Scale Statistics dapat dilihat nilai Alpha 0.679 yang merepresentasikan nilai reliabilitas soal. Reliabilitas soal terkait dengan tingkat keajegan soal dalam menilai kemampuan siswa. Hasil analisis menunjukkan soal LCC IT termasuk dalam kategori kualitas reliabilitas rendah atau tidak reliable, karena $r_{11} < 0,7$. Faktor yang mempengaruhi tinggi rendahnya reliabilitas soal antara lain banyaknya butir soal dimana semakin banyak butir soal yang digunakan, maka reliabilitasnya semakin tinggi. Luas tidaknya sampling yang digunakan, semakin luas akan

semakin tinggi reliabilitasnya. Perbedaan bakat dan kemampuan murid yang dites, jika tes diberikan kepada siswa yang kemampuannya lebih heterogen maka reliabilitas soalnya akan semakin tinggi (Purwanto, 2009).

SIMPULAN DAN SARAN

Berdasarkan hasil penelitian dan pembahasan dapat disimpulkan bahwa soal penyisihan LCC IT Fest Akakom 2016 memiliki kualitas yang tidak baik. Hal ini dapat dilihat berdasarkan kriteria indeks kesukaran, daya beda dan pengecoh diperoleh sebanyak 56% soal yang tidak dipakai, 34% soal harus direvisi terlebih dahulu sebelum dipakai dan hanya 10% soal yang dapat langsung dipakai. Soal tersebut juga tidak reliable, karena memiliki nilai reliabilitas dibawah standar.

Dari kesimpulan tersebut, dapat disarankan bahwa dalam membuat soal LCC IT

berikutnya sebaiknya memperhitungkan analisis secara kuantitatif dan kualitatif.

DAFTAR PUSTAKA

- Arikunto, S. 2009. *Dasar-dasar Evaluasi Pendidikan (Edisi Revisi)*. Jakarta: Bumi Aksara.
- Budiyono. 2011. *Penilaian Hasil Belajar*. Surakarta: Program Pascasarjana Universitas Sebelas Maret Surakarta.
- Masruroh, dkk. 2012. Analisis Soal Ulangan Semester Gasal Biologi Kelas X di Kecamatan Petanahan Kebumen. *Unnes Journal of Biology Education*, Vol. 1(2), 116-121.
- Otaya, Lian G. 2014. Analisis Kualitas Butir Soal Pilihan Ganda Menurut Teori Tes Klasik dengan Menggunakan Program ITEMAN. *TADBIR Jurnal Manajemen Pendidikan*, Vol. 2(2), 231-251.
- Purwanto, N. 2009. *Prinsip-prinsip dan Teknik Evaluasi Pengajaran*. Bandung: PT Remaja Rosdakarya.
- Ruspidu, MT. 2014. Analisis Butir Soal Ujian Akhir Mata Pelajaran Ekonomi Akuntansi Menggunakan *Iteman Microcat*. *Kajian Pendidikan Akuntansi Indonesia*, 55-66
- Rustaman, N. 2009. *Strategi Belajar Mengajar Biologi*. Malang: Penerbit Universitas Negeri Malang.
- Sudijono. 2011. *Pengantar Evaluasi Pendidikan*. Jakarta: PT Raja Grafindo Persada.
- Sukirno. 2006. Pengembangan Tes Formatif untuk Mata Kuliah Akuntansi Manajemen Jurusan Pendidikan Akuntansi FIS UNY. *Jurnal Pendidikan Akuntansi Indonesia*, Vol. 5(1), 1-18.
- Supranata, S. 2009. *Analisis, Validitas, Reliabilitas, dan Interpretasi Hasil Tes*. Bandung: PT Remaja Rosdakarya