ISBN: 978-602-98559-2-0

JILI

PROSIDING SEMINA

Bidang Biol

SEMINAR DAN RAPAT TAHUNA

BIDANG ILMU MIPA 20

BKS PTN BAR

Universitas Lampung, 10-12 Mei 20

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN
UNIVERSITAS LAMPUNG

Universitas Lampung, 10-12 Mei 2013

Didukung oleh:

FEI

ambiValue
If ambition & value goes together

PHENOMWORLD

PAN

Hubungan Antara Validitas Butir, Reliabilitas, Tingkat Kesukaran Dan Daya Pembeda Soal Ujian Semester Genap Bidang Studi Biologi Kelas Xi Sma/Ma Negeri Di Kota Padang Tahun Pelajaran 2010/2011*

Anizam Zein**, Muhyiatul Fadillah**, Rahma Novianti***

Abstract. Test is set question that requires an answer is right or wrong. A test is said to be good as a measurement when it meets the test requirements, such as having validity, reliability, objectivity, and economic practicalities. Thus the test as a measure of learning outcomes should be in accordance with these standards. This study aims to determine the relationship between the validity, reliability, level of difficulty, and the distinguishing power of semester exam study field biology class XI SMA / MA State in Padang school year 2010/2011. This research is a descriptive study. The study population is a matter of the answer sheet along with semester exams of class XI field of biology which amounts to about 1 set 1 set about SMA and MA State Padang academic year 2010/2011 which consists of 40 items and 35 items on the MA about SMA. In this study, all members of the population sampled (total sampling). The data were processed with items useful analysis using SPSS 16.From the analysis of the data obtained that 1) difficulty level correlated with the validity of (66.67%), 2) difficulty level correlated with the validity of (66.67%). 3) Power difference correlated with the validity of (100%), 4) reliability value is different between the value of all items with a value of r valid point only.furthermore, it can be concluded that 1) There was no relationship between the level of difficulty by distinguishing the extent siginifikansi 1%, 2) there is a direct relationship and level of difficulty distinguishing the validity of the significance level of 1%, 3) The relationship between reliability with validity described by reliability tests that have valid points higher than the test that all grain o valid.

Keywords: test, reliability, validity, distinguishing

PENDAHULUAN

Perubahan kurikulum yang terlalu cepat membuat guru dan siswa menjadi kebingungan. Belum lagi jumlah mata pelajaran yang terlalu banyak dan tidak berfokus pada sasaran yang tepat. Selain itu, kurangnya tindakan evaluasi membuat dunia pendidikan tidak mengalami perbaikan.

Perbaikan pada mutu penddikan harus dimulai dari hal yang paling penting yaitu lingkungan sekolah. Pada lingkungan sekolah inilah terjadi proses`pembelajaran Dalam kegiatan proses pembelajaran, terdapat berbagai unsur yang menjadi bagian pembelajaran tersebut, mulai dari tujuan, pelaksanaan hingga pencapaian hasil belajar. Untuk mengetahui pencapaian hasil belajar perlu diadakan sebuah evaluasi.

Evaluasi merupakan cara yang dapat dilakukan untuk mengetahui hasil belajar dan efektifitas pelaksanaan proses pmbelajaran. Melalui kegiatan evaluasi ini dapat ditentukan tindak lanjut dari hasil belajar siswa dan perbaikan pelaksanaan pengajaran oleh guru. Joesmani (1999) mengemukakan bahwa evaluasi adalah proses menentukan sampai berapa jauh kemampuan yang dapat dicapai siswa dalam proses pembelajaran.

Evaluasi tes bertujuan untuk mengetahui kualitas butir tes, agar berkualitias baik, sebelum digunakan butir-butir tes dianalisis terlebih dahulu. (Depdiknas, 2002). Evaluai terdiri dari pengukuran dan penilaian Kurniawan (2009) menyatakan bahwa pengukuran (measurement) merupakan

proses pemberian angka atau usaha memperoleh deskrispi numerik sejauh mana peserta didik telah mencapai suatu tingkatan. Penilaian (assessment) merupakan dasar untuk memperoleh balikan dalam memperbaiki proses`pembelajaran dan sistem pembelajaran secara keseluruhan yang dilakukan secara berkesinambungan, baik terhadap proses itu sendiri maupun terhadap hasil yang dicapai. Penilaian proses dapat dilakukan melalui pengamatan terhadap kegiatan siswa dalam belajar dengan menggunakan panduan pengamatan, sedangkan penilaian hasil belajar dapat dilakukan dengan teknik-teknik tertentu, baik teknik tes maupun non tes.

Evaluasi memiliki tujuan, seperti yang dkemukakan Tayibnapis (2008) menyatakan bahwa evaluasi memiliki dua fungsi yaitu fungsi formatif dan fungsi sumatif. Fungsi formatif mencakup perbaikan dan pengembangan kegiatan yang sedang berjalan, sedangkan fungsi sumatif mencakup pertanggungjawaban, keterangan, seleksi atau lanjutan. Jadi evaluasi hendaknya dapat membantu pengembangan, implementasi, kebutuhan atau program, perbaikan program, pertanggungjawaban, seleksi, motivasi, menambah pengetahuan dan dukungan dari mereka yang terlibat.

Salah satu cara melakukan evaluasi adalah dengan soal ujian atau tes. baik itu soal objektif, maupun essay. Tujuan tes antara lain untuk : 1) mengetahui pengetahuan awal siswa, 2) Tingkat pencapaian standar kompetensi, 3) mengetahui perkembangan kemampuan siswa, 4) mendiagnosis kesulitan belajar siswa, 5) mengetahui hasil suatu poses pembelajaran, 6) memotivasi siswa belajar, dan 7) memberi umpan balik kerja guru untuk mengetahui program pembelajarannya (Depdiknas, 2002).Soal yang dominan digunakan di sekolah adalah soal objektif, terutama pilihan ganda (multiple choisce). Soal pilihan ganda dapat mewakili lebih banyak materi, ekonomis dan efisien, serta memudahkan penilaian oleh guru.

Soal yang digunakan untuk mengevaluasi hasil belajar siswa ini hendaknya sesuai dengan standar yang telah ditentukan. Menurut Arikunto (2008), sebuah tes yang dapat dikatakan baik sebagai alat pengukur harus memenuhi persyaratan tes, yaitu memiliki validitas, reliabilitas, okjektifitas, praktikabilitas, dan ekonomis. Soal sebagai alat ukut hasil belajar siswa hendaknya sesuai dengan standar tersebut. Begitu juga halnya dengan soal-soal yang digunakan guru se Kota Padang.

Berdasarkan pengamatan awal dan wawancara dengan beberapa orang guru SMA/MA Negeri di Kota Padang pada bulan Juli tahun 2011 diketahui ternyata sedikit dari guru tersebut yang melakukan analisis awal untuk mengetahui kualitas soal ujian. Seyogyanya soal ujian merupakan instrumen bagi seorang guru untuk mengetahui hasil belajar siswanya. Jika hasil belajar siswa rendah berarti seorang guru harus mampu menganalisis bagian mana dari rencana pembelajarannya yang masih memerlukan perbaikan, termasuk instrumen penilaian, Menurut Arifin (2009) rencana pelaksanaan pembelajaran merupakan bagian integral dari tugas guru sebagai pelaksana dan pengelola pembelajaran.

Menurut Gronlund (1985) dalam Arifin (2009), ada beberapa langkah yang harus ditempuh dalam perencanaan penilaian hasil belajar yaitu merumuskan tujuan penilaian, mengidentifikasi kompetensi dan hasil belajar, menyusun kisi-kisi atau blueprint, mengembangkan draft instrumen, uji coba dan analisis instrumen, revisi, dan merakit instrument baru. Depdiknas (2002) menambahkan langkah-langkah penyusunan tes antara lain, menetapkan spesifikasi tes, menulis butir soal tes, menelaah soal tes, meakukan uji

coba tes, menganalisis butir soal, memperbaiki soal tes, merakittes, melaksanakan tes, dan menganalisis hasil tes Akan tetapi seringkali seorang guru tidak melakukan langkah uji coba, analisis instrumen, revisi, dan merakit instrumen baru. Hal ini terjadi karena kurangnya panduan dan keinginan guru untuk melaksanakan analisis soal.

Analisis soal dilakukan untuk mengetahui berfungsi atau tidaknya suatu soal. Menurut Surapranata (2005) Analisis soal dilakukan melalui dua cara, yaitu analisis kualitatif (validitas logis) dan analisis kuantitatif (validitas emperis). Analisis kualitatif dapat dilakukan sebelum soal diujikan sedangkan analisis kuantitatif dilakukan sesudah soal diujikan pada sampel yang representative..Masrun (1982) menambahkan tujuan analisis item adalah untuk menentukan apakah item-item itu merupakan item yang baik atau tem yang jelek. Lebih lanjut Surapranata (2005) menambahkan bahwa analisis kualitatif disebut juga validitas logis yang dimaksudkan untuk menganalisis soal dari segi teknis, isi, dan editorial. Analisisi teknis berkaitan dengan prinsip pengukuran dan teknik penulisan soal. Analisis ini berkaitan dengan kalayakan pengetahuan yang ditanyakan. Analisis editorial berkaitan dengan penggunaan bahasa Indonesia yang baik dan benar menutut EYD.

Analsis kuantitatif adalah penelaahan butir soal berdasarkan pada karakteristik internal tes melalui data yang diperoleh secara emperis. Karakteristik internal yang dimaksud meliputi parameter soal tingkat lesukaran, daya pembeda, validitas butir dan reliabilitas. Analisis kuantitatif ini dilakukan setelah soal diujikan (Kurniawan, 2009). Dua jenis analisis yang telah dijabarkan di atas, analisis kuantitatif menjadi fokus dalam penelitian ini. Analisis kuantitatif dilakukan dengan menelaah data yang dihasilkan langsung oleh testee setelah instrumen tersebut dipakaikan.

Namun analisis kuantitatif ini jarang dilakukan oleh guru. Fakta dilapangan menunjukkan terdapat beberapa hal yang menyebabkan guru tidak melalukan analisis soal, antar lain : 1) guru merasa terbebani oleh proses analisis soal sehingga enggan melalukan dan cendrung mengabaikan, 2) guru tidak menyadari bias ketidak vakid-an hasil evaluasi hasil belajar yang ditimbulkan oleh soal tes yang tidak baik, 3) guru merasa yakin dengan kualitas soal tes yang dibuat sehingga merasa tidak perlu melakukan penelaahan lebih lanjut. Padahal soal tes yang tidak baik dapat mencerminkan gambaran hasil belajar yang keliru.

Selain beberapa masalah yang telah disebutkan mengenai analisis soal, masih terdapat permasalahan lain yang ditemukan oleh peneliti, yaitu penentuan tindakan lajutan terhadap soal-soal yang telah diujikan. Secara umum, tindakan yang bisa diambil ada tiga, yaitu memakai butir soal, memperbaiki.. atau tidak, memakai butir soal tersebut. Namun, disinilah letak kesulitannya. Sampai saat ini peneliti belum menemukan sebuah acuan yang pasti dalam menentukan syarat sebuah soal dapat dipakai, diperbaiki, atau tidak dipakai dengan mempertimbangkan faktor validitas, reliabilitas, tingkat kesukaran, dan daya pembeda soal secara bersamaan

Untuk itu peneliti merasa perlu menemukan hubungan dan mendeskripsikan bentuk hubunan yang terdapat antara keempat faktor dalam fungsinya sebagai penentu kualitas soal yang biasa dipakai. Dengan diperolehnya bentuk hubungan tersebut, maka pada tindakan lanjutnya bentuk hubungan tersebut dapat dkaai ebagai riteria nentuan keterkapaian butir soal. Penelitian in memfokuskan pada penentuan hubungan dan bentuk hubungan antara faktor analisis.

Penelitian ini bertujuan untuk mengetahui hubungan yang terdapat antara validitas butir, reliabilitas, tingkat kesukaran, dan daya pembeda soal ujian semester genap bidang Studi Biologi kelas XI SMA/MA Negeri Kota Padang tahun pelajaran 2010/2011.

METODE PENELITIAN

Penelitian ini adalah penelitian deskriptif. Populasi dari penelitian adalah seluruh set soal beserta lembar jawaban ujian semester genap bidang studi biologi kelas XI yang berjumah 2 set soal SMA (paket A dan Pekat B) dan 1 set soal MA Negeri di Kota Padang tahun pelajaran 2120/2011. Bentuk tes yang dianalisis adalah tes pilihan ganda atau multiple choice. Sedangkan sampel dalam penelitian ini adalah keseluruhan anggota populasi atau total sampling, dengan rincian SMAN 15 dua set soal dan jumlah lembar jawaban adalah 99, MAN I satu set soal dengan jumlah lembar jawaban adalah 62. Variabal dalam penelitian adalah validitas butir soal, reliabilitas, tingkat kesukaran, dan daya pembeda/Analisis data dilakukan dengan menggunakan program SPSS 16.. Ada tidaknya hubungan antara faktor analisis menggunakan analisis korelasi bivariate

HASIL DAN PEMBAHASAN

Soal ujian semester genap bidang studi biologi kelas XI SMA.MA I Padang tahun pelajaran 2010/2011 jumlah soal yang dianalisis adalah SMA Paket A 31 butir, SMA Paket B 31 butir dan MA 40 butir. Pada data yang diolah dengan program SPSS 16, nomor soal yang tidak lengkap tidak ada, namun nomor soal tersebut diganti dengan nomor soal yang berada dinomor berikutnya yaitu nomor 6 dengan soal nomor 7 dan seterusnya sehingga soal objekif yang dianalisis berjumlah 31 butir.

Hubungan antara korelasi antara tingkat kesukaran dengan daya pembeda diperoleh dua dari tiga kelompok data (66,67%) menunjukkan tingkat kesukaran tidak berkorelasi dengan daya pembeda pada taraf sigifikansi 1%. Data yang tidak berkorelasi terdapat pada set soal SMA paket A dan SMA`paket B, sedangkan set *soal MA data berkorelasi sebesar 0,431.*

Hasil analisis hubugan antara tingkat kesukaran dengan validitas butir diperoleh dari tiga kelompok (88,67%) data menunjukkan tingkat kesukaran berkorelasi dengan validitas butir pada taraf signifikansi 1%. Data yang berkorelasi terdapat pada set soal MA sebesar 0,508 dan SMA paket A sebesar 0,526, sedangkan soal SMA pada paket B data tidak berkorelasi.

Hasil uji korelasi antara daya pembeda dengan validitas butir menunjukkan bahwa terdapat korelasi antara daya pembeda degan validitas butir pada taraf signifikansi1% disemua kelompok data (100%). Pada paket MA terdapat korelasi sebesar 0,831, SMA paket A berkorelasi sebesar 0,796, dan SMA paket B sebesar 0,895.

Nilai koefisien realibilitas tes

Hasil uji reliabilitas menunjukkan nilai koefisien reliabilitas (r) berbeda antara semua butiur dengan butir yang valid saja. Perbedaan nilai koefisien reliabilitias dapat dilihat pada tabel berikut

Tabel :1. Nilai koefisien reliablitias soal

Set soal	Nilai r semua	Nilai r butir yang valid
MAN	0.390	0,673
SMA paket A	0.635	0,647
SMA paket B	0,438	0,647

Hasil uji reliabilitas menunjukkan nilai r untuk semua butir lebih rendah dari pada nilai r butir yan valid. Akan tetapi, koefisien reliabilitas masih lebih kecil dari pada 0.700 yang berarti soal belum memiliki reliabilitas yang tinggi.

Hubungan antara tingkat Kesukaran dengan daya pembeda

Tingkat kesukaran menunjukkan derajat kesulitan item tes. Tingkat keukaran diwakili oleh satu indeks. Indeks setiap item diperoleh dari jumlah skor siswa terhadap item tersebut dibandigkan dengan jumlah siswa yang menjawab item tersebut.

Indeks tingkat kesukaran pada penelitian ini diperoleh dari jumlah skor benar dibandingkan dengan jumlah tester. Semakin sedikit jumlah skor benar menunjukkan semakin kecil tester yang mampu menjawab dengan benar item tersebut. Inilah yang memberi makna bahwa item tersebut sukar. Semakin kecil jumah skor benar, semakin kecil indes kesukaran yang diperoleh.

Daya pembeda menunjukkan kemampuan suatu item membedakan kemampuan tester. Indeks daya pembeda diperoleh dari selisih skor benar tester kelompok atas dan tester kelompok bawah. Semakin besar skor benar tester kelompok atas, semakin besar indeks daya pembeda.

Berdasarkan hasi analisis data diketahui bahwa satu indeks kesukaran pada satu butir soal berpeluang memiliki beberapa indeks daya pembeda. Suatu soal yang memiliki tingkat kesukaran dengan kategori sukar, belum dapat dipastikan memiliki suatu item dengan tingkat kesukaran tertentu dapat diamati pada tabel berikut :

Tabel 2. Indek tingkat kesukaran dan indeks daya pembeda

No	Nomor Butir Soal	Indeks tingkat Kesukaran	Makna	Indeks daya pembeda	Makna
1	2	3	4	5	6
1.	17	0,918	mudah	-0,003	tidak baik
2.	31	0,918	mudah	-0,003	tidak baik
3.	6	0,837	mudah	0,075	jelek
4.	23	0,816	mudah	0,115	jelek
5.	29	0,776	mudah	0,358	cukup
6.	5	0,735	mudah	0,030	jelek
7.	9	0,735	mudah	-0,052	tidak baik
8.	22	0,735	sedang	0,357	cukup
9.	19	0,673	sedang	0,558	baik
10.	10	0,633	sedang	0,312	cukup
11.	16	0,512	sedang	0,197	jelek
12.	12	0,592	sedang	0,065	jelek
13.	27	0,571	sedang	0,178	jelek
14.	8	0,490	sedang	0,020	jelek
15.	13	0,490	sedang	0,183	jelek
16.	3	0,469	sedang	0,142	jelek
17.	28	0,388	sedang	0,138	jelek
18.	30	0,388	sedang	0,138	jelek
19.	25	0,367	sedang	-0,148	tidak baik
20.	24	0,347	sedang	0,218	cukup
21.	26	0,327	sedang	0,340	cukup
22.	18	0,306	sedang	0,380	cukup
23.	7	0,286	sukar	0,257	cukup
24.	15	0,286	sukar	0,093	jelek
25.	14	0,245	sukar	0,173	jelek
26.	21	0,224	sukar	0,295	cukup
27.	20	0,184	sukar	0,212	cukup
28.	1	0,163	sukar	0,088	jelek
29.	2	0,163	sukar	0,170	jelek
30.	4	0,122	sukar	0,168	jelek
31.	11	0,102	sukar	0,208	cukup

Item-item dengan kategori sukar bisa saja memiliki daya pembeda jelek atau cukup. Item-item yang memiliki katogori sedang pun bisa memiliki daya pembeda jelek, cukup, atau baik. Jadi dapat disimpulkan bahwa tidak terdapat hubungan antara tingkat kesukaran dan daya pembeda pada taraf signifikansi 1%.

Hubungan antara Tingkat Kesukaran dengan Validitas Butir

Hasil analisis korelasi menginformasikan bahwa 66,67% sampai menunjukkan adanya korelasi antara tingkat kesukaran dengan validitas butir pada taraf signifikansi 1 %. Sisanya sebesar 33,33% sampai tidak menunjukkan adanya korelasi pada taraf signifikansi 1%.

Tingkat kesukaran menunjukkan derajat kesulitan item tes yang diwakili suatu jndeks. Indeks setiap item diperoleh dari

jumlah skor siswa terhadap item tersebut dibandingkan dengan jumlah siswa yang menjawab item tersebut. Ariyono (2012) menyatakan bahwa suatu item yang memiliki tiga kesukaran dan kualitas soal yang baik dapat ditunjukkan dari satu koefisien yang dikenal dengan koefisien korelasi poin biserial. Koefisien poin biserial pada dasarnya digunakan untuk melihat hubungan antar skor atau hasil jawaban pada masing-masing item. Nilai koefisien yang besar menunjukkan testee dapat menjawab dengan baik satu item, sebaliknya nilai koefisien yang kecil menujukkan bahwa item pertanyaan tidak dapat dijawab dengan baik oleh testee. Nilai koefisen poin biserrial yang semakin besar menunjukkan bahwa suatu item memiliki tingkat kesukaran dan kualitas soal yang semakin baik.

Koefisien validitas butir menunjukkan seberapa kuat skor total satu item mendukung terhadap skor total seluruh item. Besarnya koefisien validitas butir ini juga dapat ditentukan dengan menggunakan teknik korelasi biserial dan pint biserial. Korelsi biserial maupun korelasi poin biserial adalah korelasi product moment yang diterapkan pada data, dimana variable-variabel yang dikorelasikan sifatnya masing-masing berbeda satu sama lain. Variabel butir soal bersifat dikotom, sedangkan variable skor total berifat kontinum (Surapranata, 2005).

Berdasarkan hasil analisis diperoleh nilai poin biseral yang cukup tinggi, yaitu 0,0508 pada soal MA da ,526 pada soal SMA paket A. Dengan demikian, diketahui bahwa terdapat korelasi antara tingkat kesukaran dengan validitas butir pada taraf signifikansi 1%. Bentuk korelasi yang terlibat adalah korelasi searah, dimana semakin besar indeks tingkat kesukaran, maka semakin besar pada koefisien validitas butir. Namun bila indeks tingkat kesukaran dan koefisien validitas butit dimasukkan dalam tabel, terlihat seperti berikut :

Tabel 3. Distribusi indeks tingkat kesukaran dan koefisien validitas butir pada soal MA

No	Nomor Butir Soal	Indeks Tingkat Kesukaran	Koefisien Validitas butir
1	2	3	4
1.	29	0,968	0,300
2.	20	0,887	0,370
3.	7	0,839	0,340
4..	5	0,806	0,477
5..	31	0,758	0,261
6.	10	0,726	0,423
7.	15	0,694	0,319
8.	8	0,677	0,377
9.	28	0,661	0,203
10.	37	0,661	0,323
11.	24	0,645	0,113
12.	23	0,597	-0,021
13.	16	0,565	0,327
14.	26	0,565	0,277
1	2	3	4
15.	21	0,516	0,278
16.	33	0,500	0,118
17.	2	0,468	0,528
18.	34	0,468	-0,031
19.	40	0,435	0,257
20.	9	0,419	0,508
21.	12	0,419	0,183
22.	36	0,403	0,252
23.	38	0,387	-0,114
24.	39	0,339	0,046
25.	4	0,274	0,181
26.	3	0,258	0,348
27.	6	0,242	0,478
28.	11	0,242	0,125
29.	27	0,242	-0,030
30.	14	0,226	0,265
31.	25	0,210	0,018
32.	17	0,149	-0,184
33.	32	0,194	0,103
34.	1	0,194	-0,040
35.	35	0,145	0,107
36.	30	0,129	0,199
37.	13	0,065	0,184
38.	18	0,048	0,-0,118
39.	22	0,032	0,101
40.	19	0,016	0,033

Pada tabel di atas, nilai indeks tingkat kesukaran yang benar tidak menunjukkan keofisien validitas butir yang semakin besar pula. Hal ini sesuai dengan penelitian Hal ini sesuai dengan penelitian Muhyiatul Fadillah (2011), menyatakan bahwa tidak terdapat korelasi antara tingkat kesukaran dengan validitas butir pada taraf signifikansi 1 % dengan jumlah sampel yang lebih besar dan luas ditemukan 92,8% sampel yang tidak berkorelasi. Ditambahkan juga bahwa validitas butir tidak tergantung dari banyaknya jumlah jawaban benar. Validitas butir berasal dari koefisien yang menunjukkan seberapa kuat skor total satu butir mendukung terhadap skor total seluruh butir. Suatu butir yang dijawab benar atau salah oleh semua testeee belum tentu menghasilkan nilai koefisien yang semakin besar atau kecil. Sabaliknya pada tingkat kesukaran jika semakin banyak jumlah jawaban benar, maka butir tersebut akan semakin mudah.

Hubungan antara daya Pembeda dengan Validitas Butir

Indeks daya pembeda diperoleh dari selisih akar benar testee kelompok atas dan testee kelompok bawah dibandingkan dengan jumlah testee kelompok atas dan jumlah testee kelompok bawah. Daya pembeda juga dapat dihitung secara statistik dengan menghitung nilai korelasi poin biserial. Korelasi poin biserial dapat digunakan untuk melihat hubungan antara skor atau hasil jawaban pada masing-masing item dan pola jawaban testee seperti yang telah dijabarkan sebelumnya. Pola jawaban testee ini dapat mengidenifikasikan tingkat kesukaran dan daya pembeda dari satu item (Ariyoso, 2012).

Penentuan nilai daya pembeda dan validitas butir memiliki proedur yang sama, yaitu menggunakan prinsip korelasi. Hal ini yang menjelaskan terdapatnya hubungan antara daya pembeda dengan validias butir pada tingkat signifikansi 1%. Bentuk korelasi yang ditunjukkan adalah korelasi searah, dimana semakin besar indeks daya pembeda, maka semakin besar pula koefisien validitas butir. Bentuk korelasi searah ini data dilihat pada tabel berikut.

Tabel 4. Distribusi indeks daya pembeda dan validitas butir pada soal MA

No	Nomor Butir Soal	Indeks Daya Pembeda	Koefisien Validitas Butir
1	2	3	4
1.	9	0,452	0,508
2.	2	0,355	0,528
3.	3	0,323	0,384
4.	21	0,323	0,278
5.	6	0,290	0,478
6.	10	0,290	0,423
7.	5	0,258	0,447
8.	8	0,258	0,377
9.	15	0,226	0,319
10.	16	0,226	0,372
11.	26	0,226	0,277
12.	28	0,226	0,203
13.	33	0,226	0,118
14.	36	0,226	0,252
15.	37	0,226	0,323
16.	4	0,161	0,181
17.	20	0,161	0,370
18.	40	0,161	0,257
19.	12	0,129	0,183
20.	24	0,129	0,113
21.	11	0,097	0,125
22.	23	0,097	-0,021
23.	31	0,097	0,261
24.	34	0,097	-0,031
25.	7	0,065	0,394
26.	13	0,065	0,184
27.	14	0,065	0,265
28.	29	0,065	0,300
29.	30	0,065	0,199
30.	19	0,032	0,033
31.	25	0,032	0,018
32.	27	0,032	-0,030
33.	39	0,032	0,046
34.	22	0,000	0,101
1	2	3	4
35.	32	0,000	0,103
36.	18	-0,032	-0,118
37.	35	-0,032	0,107
38.	1	-0,129	-0,040
39.	17	-0,129	-0,184
40.	18	-0,129	-0,114

Dari tabel diatas terlihat bahwa semakin besar ideks daya pembeda, maka semakin besar pula koefisien validitas butir. Dengan demikian dapat disimpulkan bahwa antara daya pembeda dengan validitas butir terdapat korelasi searah pada taraf signifikansi 1 %.

Hubungan antara Validitas Butir dengan Reabilitas

Hubungan antara validitas butir dengan reabilitas dianalisis dengan cara yang berbeda dari analisis korelasi sebelumnya. Analisis korelasi antara tingkat kesukaran-daya pembeda, tingkat kesukaran validitas butir, dan daya pembeda-validitas butir dilakukan dengan memasangkan data antara masing-masing faktor. Hasil perhitungan reabilitas hanya memberikan satu nilai untuk semua butir dalam satu tes. Untuk mengetahui ada atau tidaknya hubungan antara validitas butir dengan reabilitas dilakukan degan membandingan koefisien reabilitas antara semua item dengan item yang valid saja. Menurut Scarvin Anderson dalam Arikunto (2008) persyaratan bagi tes yaitu validitas dan reliabilitas ini penting. Dalam hal ini validitas lebih penting dan reliabilitas ini perlu karena menyokong terbentuknya dengan butir yang valid validitas. Sebuah tes mungkin reliabel tetapi tidak valid. Sebaliknya sebuah tes yang valid bisanya reliabel. Berdasarkan uji reliabilitas semua butir dengan reliabilitias butir yang valid saja, diperoleh bahwa nilai r yang valid lebih besar dibanding dengan nilai r semua butir. Hasil uji validitas butir dengan reliabilitas dapat dilihat pada tabel berikut : Dari tabel 5 dapat diketahui bahwa koefisien reliabilitas (nilai r) yang disumbangkan oleh butir-butir yang valid lebih besar dari pada koefisien reliabilitas semua butir. Dengan demikian dapat disimpulkan bahwa validitas butir memiliki hubungan dengan reliabilitas, dimana semakin banyak jumlah butir yang valid, semakin besar pula koefisien reliabilitas`suatu tes.

Tabel 5. Perbandingan nilai r semua butir dengan butir yang valid

Sampel	Nilai r semua butir	Nilai r butir yang valid
MAN	0,390	0,673
SMA paket A	0,635	0,547
SMA paket B	0,439	0,642

Penentuan indeks tingkat kesukaran dan daya pembeda sangat dipengaruhi oleh kondisi siswa saat melakukan tes. Jika soal diujikan pada dua kelomok siswa yang memiliki kemampuan yang berbeda, maka dapat menunjukkan indeks yang berbeda pula. Akan teapi, tidak demikian halnya pada validitas butir dan reliabilitas. Kondisi siswa tidak berpengaruh terhadap nilai koefisien yang dihasilkan pada suatu tes.

Secara umum validitas butir memiliki hubungan timbal balik dengan tiga faktor lainnya yaitu tingkat kesukaran, daya pembeda, dan reliabilitas. Dua faktor yang tidak memiliki hubungan adalah tingkat kesukaran dengan daya pembeda. Dengan demikian kriteria validitas pada suatu butir diprediksi sebagai kriteria penentu suatu item yang baik.hanya saja besarnya kontribusi masing-masing tidak dikaji secara mendalam pada penelitian ini,. Perlu penelitian lebih lanjut mengenai hubungan antara tingkat kesukaran dengan validitas butir karena diperoleh temuan yang berbeda dengan literatur dari penelitian sebelumnya.

SIMPULAN DAN SARAN

Berdasarkan hasil penelitian dapat disimpulkan bahwa tingkat kesukaran sebagian besar tidak berkorelasi signifikan dengan daya pembeda pada taraf signifikansi 1%. Tingkat kesukaran sebagian besar berkorelasi signifikan dengan validitas butir pada taraf signifikansi 1%. Daya pembeda berkorelasi signifikan dengan validitas butir pada taraf signifikansi 1%. Reliabilitas suatu soal

lebih tinggi jika soal tersebut memiliki item-item yang lebih valid.Disarankan penelitian lanjutan untuk hubungan (multi variabel) faktor analisis soal secara kuanititafif dan fakor kualitas option sebagai salah satu aspek dalam analisis kuantitatif dan semua faktor secara komprehensif.

DAFTAR PUSTAKA

Arifin, Zainal. 2009. Evaluasi Pembelajaran. Bandung. PT. Remaja Rosdakarya.

Arikunto, Suharsimi. 2008. Dasar dasar Evaluasi Pendidikan. Jakarta. Bumi Aksara.

Arikunto, Suharsimi. 2006. Prosedur Penelitian Suatu Pendekatan Praktik. Jakarta. PT Asdi Mahasatya.

Ariyoso. 2011. Korelasi Biserial. http://aroyoso.wprdpress.com/.2011/07/07/krelasi-biserial/ diaksses. Tanggal 14 Maret 2012.

Darsono. 2002. *Belajar dan Pembelajaran*. Semarang. IKIP Semarang Press.

Daryanto. 2005. *Evalasi Penddikan*. Jakarta. PT Rineka Cipta.

Depdiknas. 2002. *Pedoman Khusus* Pola Induk Pengembangan Silabus Berbasis Keterampilan Dasar Sekolah Menengah Umum (SMU). Jakarta. Dirjen Dikdasmen.

Joesmani. 1998. *Pengukuran dan Evaluasi dalam Pengajaran*. Jakarta. Rajawali Press.

Kuniawan, Endang dan Endah Mutagimah. 2009. *Penilaian*, Jakarta Depdiknas.

Nurkancana, Wayan. 1986. *Evaluasi Pendidikan*. Surabaya. Usaha National.

Masrun. 1982. *Analisis Item*. Yogyakarta. Fakultas Psikologi.

Purwanto, Ngalim. 1994.*Prinsip-Prinsip dan Teknik Evauasi Pengajaran*. Bandung. Remajs Rodakarya.

Sudjionon, Anas. 2009. *Pengantar Evaluasi Pendidikan*. Jakarta. Raja Grafindo Persada.

Sudjana, nana. 2004. *Penilaian hasil dan Proses Belajara Mengajar*. Bandung. Remaja Risdakarya.

Surappranta, Sumarna. 2005. *Analisis, Validitas, Reliabilitas dan Interprestasi Hasil tes*. Bandung. Pt Remaja Rsda Karya

Tayibnapis, Farda Yusuf. 2008. *Evaluasi Program dan Instrumen Evaluasi untuk Program Pendidikan dan Penelitian. Jakarta. Rineka Cipta.*

Thoha, Chabib. 2003. Teknik Evaluasi Pendidikan. Jakarta. PT Raja Grafindo Persada.

Sertifikat

BADAN KERJASAMA
PERGURUAN TINGGI NEGERI WILAYAH BARAT (BKS-B)
BIDANG ILMU MIPA
diberikan kepada:

BKS PTN Barat
Bidang Ilmu MIPA

Drs. Anizam Zein, M.Si

sebagai: **Pemakalah**

Pada kegiatan:
SEMINAR NASIONAL DAN RAPAT TAHUNAN BIDANG ILMU MIPA
Tema: **"Peran Ilmu MIPA dalam Pemanfaatan Sumber Daya Alam untuk Menunjang Percepatan Pembangunan Ekonomi Indonesia".**

Di Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Lampung, 10-12 Mei 2013

BKS PTN Barat
Koordinator Bidang MIPA,

Dr. Sutarman, M.Sc
NIP.196310261991031001

Ketua Panitia

Prof. Sutopo Hadi,M.Sc.,Ph.D
NIP. 197104151995121001

FEI PANalytical PHENOMWORLD ambiValue

Semairata