

PERSEPSI KONSEPTUALITAS PERAKITAN BUTIR SOAL**OLEH :****MUSA MARENGKE**

IAIN TERNATE

(musamarengke.mm@gmail.com)

ABSTRACT:

Development test basically have to make two decisions, namely what to measure and how to measure it. Assembly of items relating to the second decision. If a teacher works based on the specification item, then by itself the structure and format of grain should have been set. But if the teacher is not based on the specification items, it is necessary to establish the format if items that must be followed before the grains begin writing stage. This is where the expertise and precision required of teachers. Teachers need to know the procedure of assembling items, both conceptual and technical-operational. Conceptually, the stages of assembly problems have offered ideas on how to measure the level of difficulty, different power levels, about the validity and reliability. These procedures need to be analyzed in this essay by analyzing the concepts on theories of items. It can be analyzed with glasses-descriptive qualitative and quantitative-mathematical. Thought is intended to analyze the problem and establish the position of matter at each of the criteria. Problem qualified one of which can be measured from the element representation of aall competence prikomotoriknya. In addition, the question must be valid and reliable, which has a certain degree of validity and reliability. This paper aims to check your assembly procedure in a matter of perception konseptualitas. Then this article will examine the items in the perception of the use of items, quality problems and the steps of assembly and analysis.

Keywords: Conceptuality and Problem Items

ABSTRAK:

Tes pengembangan pada dasarnya harus membuat dua keputusan, yaitu apa yang diukur dan bagaimana mengukurnya. Perakitan item yang berkaitan dengan keputusan kedua. Jika seorang guru bekerja berdasarkan item spesifikasi, maka dengan sendirinya struktur dan format gandum harus ditetapkan. Tetapi jika guru tidak berdasarkan spesifikasi item, perlu untuk menetapkan format jika item yang harus diikuti sebelum butir mulai menulis tahap. Di sinilah keahlian dan ketelitian yang dibutuhkan guru. Guru perlu mengetahui prosedur mengumpulkan barang, baik secara konseptual maupun teknis-operasional. Secara konseptual, tahapan masalah perakitan telah menawarkan ide tentang bagaimana mengukur tingkat kesulitan, tingkat daya yang berbeda, tentang validitas dan reliabilitas. Prosedur ini perlu dianalisis dalam esai ini dengan menganalisis konsep-konsep tentang teori item. Hal ini dapat dianalisis dengan kaca-mata-deskriptif kualitatif dan kuantitatif-matematis. Pikiran bermaksud menganalisis masalah dan menetapkan posisi materi pada setiap kriteria. Masalah yang memenuhi syarat salah satunya dapat diukur dari elemen representasi dari semua kompetensi prikomotoriknya. Selain itu, pertanyaan harus valid dan dapat dipercaya, yang memiliki tingkat validitas dan reliabilitas tertentu. Makalah ini bertujuan untuk memeriksa prosedur perakitan Anda dalam masalah persepsi konseptualitas. Maka artikel ini akan memeriksa item dalam persepsi penggunaan item, masalah kualitas dan langkah-langkah perakitan dan analisis.

Kata Kunci : Konseptualitas dan Butir Soal

A. Latar Belakang Masalah

Penilaian merupakan komponen penting dalam penyelenggaraan pendidikan. Upaya meningkatkan kualitas pendidikan dapat ditempuh melalui peningkatan kualitas pembelajaran dan kualitas sistem penilaiannya. Keduanya saling terkait, sistem pembelajaran yang baik akan menghasilkan kualitas belajar yang baik. Selanjutnya sistem penilaian yang baik akan mendorong guru untuk menentukan strategi mengajar yang baik dan dapat memotivasi peserta didik untuk belajar yang lebih baik. Penilaian selanjutnya mencakup semua cara, terfokus pada peserta didik dan prosesnya meliputi pengumpulan bukti-bukti dari progress kemampuan peserta didik. Bukti atau dokumen penilaian peserta didik tidak selalu diperoleh melalui tes, tetapi juga dapat dilakukan melalui pengamatan dan laporan diri.

Penilaian itu berkaitan dengan proses pemberian atau penentuan nilai peserta didik berdasarkan kriteria dan syarat tertentu, sementara butir soal berkaitan dengan penyusunan pertanyaan atau pernyataan berdasarkan kriteria dan syarat tertentu pula. Penilaian dibutuhkan proses interpretasi yang diakhiri dengan *judgement* yang tema penilaiannya mengimplikasikan adanya suatu perbandingan antara kriteria dan kenyataan konteks. Disinilah letak perlunya ketelitian dan kecermatan dalam keterampilan perakitan butir soal. Ketelitian dan kecermatan butir soal adalah langkah utama ketika dunia pendidikan ingin mengukur tingkat perkembangan dan kemajuannya. Kualitas pendidikan dan pembelajaran dapat ditimbang ke dalam berbagai ukuran dan variabel, diantaranya adalah perakitan soal itu sendiri.

Penyusunan butir soal yang baik akan menghasilkan skor yang seimbang dan baik. Karena itu, perlunya seorang guru mengetahui kaidah-kaidah penyusunan butir soal. Karena penyusunan soal dibutuhkan pengetahuan dan keterampilan guru yang tinggi. Apabila guru mengetahui dan memahami prosedur perakitan butir soal akan dengan mudah ia memberikan penilaian yang obyektif. Akan tetapi fenomena dan realitas menunjukkan bahwa terkadang guru tidak mampu menyusun atau merakit butir soal sehingga penilaian hasil belajar pun tidak menyentuh aspek yang diukur, penyusunan soal pun asal-asal tanpa melewati sebuah proses analisis butir yang baik dan benar sehingga tidak heran jika hasil penilaian peserta didik pun tidak memiliki bobot skor yang sebenarnya.

Menurut data dari BERMUTU bahwa kemampuan guru dalam menyusun soal baru mencapai 25 persen dari total materi yang telah diberikan. Guru hanya pandai menyusun soal atau pertanyaan berdasarkan buku paket, tidak ditelaah atau diuji butir soal baik Tingkat Kesukaran (TK), Dapa Pembeda (DP) maupun Mengecoh Soal (PS) atau dari sisi konten maupun konstruksinya. Menyusun soal memerlukan pemahaman dan keterampilan guru yang baik dan benar. Pemahamannya berkaitan dengan teori dan konsep tentang tingkat kesukaran dan tingkat daya beda soal, sementara keterampilan berkaitan dengan kemampuan guru dalam menyusun butir-butir soal, baik melalui proses manual maupun proses komputerisasi. Kesalahan dalam menyusun soal berakibat fatal pada hasil belajar siswa yang tidak valid. Dengan demikian pertanyaan yang perlu diajukan adalah bagaimana persepsi konseptualitas perakitan butir soal yang harus diketahui dan dilaksanakan oleh guru.

B. Manfaat Soal yang telah ditelaah

Kegiatan menganalisis butir soal merupakan kegiatan yang harus dilakukan guru untuk meningkatkan mutu soal yang telah ditulis. Kegiatan ini menyangkut proses pengumpulan,

peringkasan dan penggunaan informasi dari jawaban peserta didik untuk membuat keputusan tentang setiap penilaian. Tujuan dari penelaahan adalah mengkaji setiap butir soal agar diperoleh soal yang bermutu sebelum soal digunakan, juga untuk membantu meningkatkan tes melalui revisi soal-soal yang dipandang eror atau tidakberfungsi serta memberikan informasi diagnostik pada peserta didik apakah sudah memahami atau belum materi yang diberikan guru. Tujuan tersebut maka kegiatan analisis butir soal memiliki beberapa manfaat, diantaranya : (1) dapat membantu para pengguna tes, (2) sangat relevan bagi penyusunan tes informal dan tes yang disiapkan guru untuk peserta didik di kelas, (3) secara materi dapat memperbaiki tes di kelas, (4) meningkatkan validitas dan realibilitas soal. Sehingga dalam pelaksanaan kegiatan analisis butir soal perlu seorang guru mendesain untuk menjawab beberapa pertanyaan butir yaitu : apakah fungsi soal sudah tepat?apakah soal ini telah memiliki tingkat kesukaran yang tepat?apakah soal bebas dari hal-hal yang tidak relevan?dan apakah pemilihan jawabannya efektif?

Pertanyaan-pertanyaan tersebut berkaitan dengan tingkat kemampuan guru dalam menganalisis butir soal sehingga yang diharapkan adalah butir-butir soal yang dapat bermanfaat bagi guru dan siswa misalnya hasil perakitan butir dapat bermanfaat bagi kerja-kerja remedial, desain pembelajaran dan bahan masukan untuk evaluasi progress guru itu sendiri.

C. Soal Yang Bermutu

Dalam dunia pengukuran dan penilaian, masalah penyusunan soal atau test berkaitan dengan aspek kemampuan kognitif, afektif dan psikomotorik peserta didik. Ketiga aspek ini adalah dasar dan patokan dalam proses penyusunan soal dan proses penilaian. Menurut Nana Sudjana, berdasarkan teori Benyamin Bloom bahwa *aspek kognitif* berkaitan dengan pengetahuan atau ingatan, pemahaman, aplikasi, analisis, sintesis dan evaluasi, sedangkan *aspek afektif* berkaitan dengan sikap yakni menyangkut penerimaan, jawaban atau reaksi, penilaian, organisasi dan internalisasi dan aspek psikomotorik berkaitan dengan hasil belajar keterampilan yang terdiri dari penilaiangerakan reflex, keterampilan gerakan dasar, kemampuan perseptual, keharmonisan atau ketepatan, gerakan keterampilan kompleks dan gerakan ekspresif dan interpretatif.

Soal yang bermutu salah satunya dapat diukur dari unsur keterwakilan seluruh kompetensi peserta didik, baik kompetensi kognitif, afektif maupun kompetensi psikomotoriknya. Adalah soal yang dapat memberikan informasi setepat-tepatnya sesuai dengan tujuannya diantaranya dapat menentukan peserta didik mana yang sudah atau belum menguasai materi yang diajarkan guru. Disini dibutuhkan kualitas guru dalam menyusun soal yang bermutu. Soal yang bermutu dapat membantu guru meningkatkan pembelajaran dan memberikan informasi dengan tepat tentang peserta didik mana yang belum atau sudah mencapai kompetensi. Salah satu ciri soal yang bermutu adalah dapat membedakan setiap kemampuan (kompetensi) peserta didik. Semakin tinggi kemampuan peserta didik dalam memahami materi pelajaran, semakin tinggi pula peluang menjawab benar soal atau mencapai kompetensi yang ditetapkan, demikian sebaliknya, semakin rendah kemampuan peserta didik dalam memahami materi pelajaran, makin kecil pula peluang menjawab benar soal untuk mengukur pencapaian kompetensi yang ditetapkan.

Soal yang bermutu adalah soal yang harus sah dan handal, yakni memiliki tingkat validitas dan realibilitas tertentu. Validitas didefinisikan sebagai ukurna seberapa cermat suatu soal test melakukan fungsi ukurnya. Suatu soal disebut validitas apabila tes tersebut dapat mengukur apa yang hendak diukur. Misalnya mistar hanya mengukur panjang, timbangan hanya mengukur beratnya dan lain sebagainya. Lindgen, H.C mengatakan ...validity the wxtent to which they measure what they are expected to measure. Valisitas itu merupakan tingkat ketepatan soal dalam mengukur materi dan

perilaku yang harus diukur. Validitas tersebut berkaitan dengan mengukur validitas isi (*content validity*), validitas kriteria (*criterion validity*) dan validitas kontruk (*construct validity*).

Menurut Mudjijo, validitas isi berkaitan dengan soal yang ditetapkan menurut analisis rasional yakni menyusun butir-butir soal didasarkan pada materi pelajaran (subject matter) dan perubahan tingkah laku (behavioral changes). Sementara validitas kriteria diuji secara empiris dengan menggunakan teknik statistic yakni teknik korelasi antara skor test (X) dengan kroteri (Y). Dan validitas berkaitan dengan suatu alat ukur mengukur kontruk teori yang menjadi dasar penyusunan butir soal yang didasarkan pada aspek psikologis dimana suatu butir soal mampu atau dapat mengukur kualitas psikologis dimana suatu butir soal mampu atau dapat mengukur kualitas psikologis yang mencakup atau terdapat dalam suatu aspek perilaku peserta didik yang hendak dan harus diukur oleh test yang bersangkutan. Suatu alat ukur yang tinggi validitasnya akan memiliki skor pengukuran skor pengukuran yang kecil, artinya skor setiap subyek yang diperoleh oleh alat ukur tersebut tidak jauh berbeda dari skor yang sesungguhnya dengan demikian secara keseluruhan soal tes yang bersangkutan akan menghasilkan varians eror yang kecil pula.

Syarat soal yang bermutu tersebut yang harus mempunyai standar reliabilitas tertentu. Reliabilitas berhubungan dnegan konsistensi dan keajegan. Kalau kita mengukur panjang sebuah meja dengan menggunakan sebuah meteran berulang-ulang, baik dalam tenggang waktu yang singkat maupun tenggang waktu lama. Maka hasil ukur dapat dipastikan selalu menunjukkan angka yang sama selama panjang meja tersebut belum berubah. Maka diaktakan bahwa meteran tersebut reliable atau konsistensi. Untuk itu, untuk menghasilkan soal yang valid (sahih) dan reliable (handal), maka guru dapat merumuskan kisi-kisi dan menulis soal berdasarkan kaidah penulisan soal yakni pertimbangan factor materi, factor psikologis, factor lingkungan dan situasi test. Dimana factor yang mempengaruhi reliabilitas yang berhubungan dnegan tes adalah (1) banyaknya butir, (2) homogenitas materi tes, (3) homogenitas karakteristik butir, dan (4) variabilitas skor, sedangkan factor yang berhubungan dengan peserta didik adalah dipengaruhi oleh beberapa factor diantara : (1) homogenitas kelompok, (2) pengalaman peserta didik mengikuti tes dan motivasi peserta didik (3). Faktor yang memperngaruhi reliabilitas yang berhubungan dengan administrasi adalah batas waktu dan kesempatan peserta didik. Reliabilitas skor tes dalam teori respon butir adalah penggunaan fungsi informasi test. Menurut Hambleton dan Swaminathan, pengukuran fungsi informasi tes lebih akurat bila dibandingkan dengan penggunaan reliabilitas karena bentuknya tergantung hanya pafa butir-butir soal test dan mmepunyai estimasi kesalahan pengukuran pada setiap level abilitas.

Secara empiric, tinggi rendahnya reliabilitas ditunjukkan oleh suatu angka yang disebut koefisien reliabilitas. Menurut Saifuddin Azwar, koefisien reliabilitas berkisar dari 0,00 sampai 1,00 akan tetapi pada kenyataannya koefisien sebensar 1,0 dan sekecil 0,00 tidak pernah dijumpai... Koefisien reliabilitas $r_{xx} = 1,0$ berarti adanya konsistensi yang sempurna pada hasil ukur yang bersangkutan. Kesempurnaan itu tidak dapat terjadi dalam pengukuran aspek-aspek psikologis dan soal yang menggunakan manusia (peserta didik) dikarenakan terdapat berbagai sumber eror dalam pengukuran tes. Dengan demikian, untuk meningkatkan validitas dan reliabilitas soal test yang dikatakan bermutu perlu dilakukan analisis butir soal.

D. AnalisisButir Soal

Telaah butir soal dilakukan terhadap aspek materi, aspek kontruksi dan aspek bahasa. Aspek materi berkait dengan substansi keimuan yang ditanayakan serta tingkat berpikir yang terlihat. Aspek konstruksi berkaitan dengan teknik penulisan soal, baik bentuk obyektif, maupun yang non-obyektif. Untuk mendapatkan soal yang baik perlu dilakukan pemetaan soal. Ada dua perspektif

dalam menganalisis soal yakni analisis kualitatif dan analisis kuantitatif. Analisis butir secara kualitatif dilaksanakan berdasarkan kaidah penulisan soal (tes tertulis, tes perbuatan dan tes sikap). Penelaahan biasanya dilakukan sebelum soal digunakan atau diujikan. Analisis ini melihat setiap butir soal dari sisi materi, konstruksim bahasa-buadya dan kunci jawaban atau pedoman penskorannya, dimana dalam prosedur analisisnya menggunakan format penelaahan butir soal serta bahan-bahan yang dibutuhkan adalah kisi-kisi soal, kurikulum yang digunakan, buku sumber dan kamus bahasa Indonesia. Sedangkan analisis butir secara kuantitatif adalah penelaahan butir soal didasarkan pada data empiric dari butir soal setelah soal diujikan. Model analisisnya menggunakan teori Klasik dan Teori Modern. Diama butir-butir soal dianalisis mulai dari Tingkat Kesukaan (TK), Daya Pembeda (DP), distribusi jawaban, reliabilitas skor tes, reliabilitas tes (soal bentuk pilihan ganda) maupun di analisis butir soal dari sisi Item Respon Theory (IRT) dan Kalibrasi butir soal.

Analisis operasional yang sering digunakan pula adalah menggunakan pendekatan penilaian acuan kriteria, maka butir soal yang digunakan harus memenuhi standar butir sola *Acuan Kriteria*(Criterion referended test), demikian pula jika menggunakan pendekatan penilaian *Acuan Norma* , maka butir soal *Acuan Norma* (norm referenced test). Model Acuan Kriteria dan Acuan Nirma diperuntukkan untuk melihat lebih dalam mengenai status soal dan tingkat validitas dan reliabilitas soal. Soal yang baik adalah soal yang diawali dengan pemetaan konsep dan selanjutnya proses analisis soal. Untuk meningkatkan butir soal perlu dilakukan mellalui tiga komponen analisis butir yaitu tingkat kesukaran, daya pembeda, disamping pengecoh soal dan validitas dan reliabilitas. Menganalisis tingkat kesukaran soal adalah mengkaji soal-soal dari segi kesulitannya sehingga dapat diperoleh soal-soal yang berkategori soal mudah ,sedang dan sukar, sedangkan menganalisis daya pembeda artinya mengkaji soal-soal dari segi kesanggupan tes dalam membedakan peserta didik yang termasuk ke dlaam ketegori lemah atau rendah dan kategori kuat atau tinggi prestasinya.

1. Analisis Tingkat Kesukaran Butir Soal

Persepsi Tingkat kesukaran (TK) butir soal adalah peluang untuk menjawab benar suatu soal pada tingkat kemampuan tertentu. Jika indeks berkisar 0,00 berarti peserta didik tidak ada yang menjawab benar soal tersebut dan jika indeks berkisar 1,00 berarti peserta didik menjawab benar soal tersebut. Semakin besar indeks tingkat kesukaran soal yang diperoleh maka berarti soal tersebut semakin mudah. Persepsi ini mempunyai asumsi yang digunakan untuk memperoleh kualitas soal yang baik, disamping memenuhi validitas dan reliabilitas, adalah adanya keseimbangan dari tingkat kesulitan butir soal. Keseimbangan yang dimaksud adalah adanya soal-soal yang termasuk berkategori *soal mudah, soal sedang dan soal sukar*, secara proporsional. Dalam taraf kesukaran ini soal-soal terakhir hanya peserta didik yang pandai-pandai saja yang mengerjakannya. Pendapat lain mengatakan bahwa soal-soal tes diatur menurut aturan denan kategori ringan, sedang dan berat, sedang dan ringan. Dasar pemilihan ini didsarkan pada aspek psikologis yakni peserta didik ke luar dari ruangan tes dengan pikiran yang sudah ringan.

Terlepas dari asumsi dan beda pendapat naun yang jelas analisis butir soal untuk diletakkan soal pada ketegori tertentu memerlukan beberapa pertimbangan, pertama adanya keseimbangan dimana jumlah soal sama untuk ketegori mudah, sedang dan sukar, misalnya pelajaran matematika dibuat 60 soal pilihan ganda maka soal kategori mudah sebanyak 20, soal yang sedang sebanyak 20 dan soal kategori sukar 20. Pertimbangan ini dipakai karena dalam analisis butirnya tidak didasarkan pada analisis kurva normal. Sedangkan pertimbangan kedua adalah proporsi jumlah soal untuk ketiga kategori tersebut didasarkan atas kurva normal artinya sebagian besar soal berada pada ketegori sedang dan sebagian lagi berkategori mudah dan sukar dengan proporsi yang seimbang, biasanya dibuat konsep 3-4-3. Artinya 30% soal ketegori mudah, 40% kategori soal sedang dan 30% kategori

soal sukar atau bisa juga dipakai konsep 3-5-2. Konsep 3-4-3 dan 3-5-2 dalam menjadi paradigme yang menjadi acuan ketika seornag guru ingin menggunakan analisis kurva normal. Pertimbangan pemilihan soal berkategori bisa juga dilakukan melalui *judgement*. Cara menganalisisnya dengan menentukan tingkat kesukaran soal yakni $I = B/N$ yakni semakin kecil indeks yang diperoleh, makin sulit soal tersebut, sebaliknya makin besar indeks yang diperoleh, makin mudah pula soal tersebut. Kriteria indeks tingkat kesukaran soal adalah sebagai berikut :

0,00 – 0,30 = soal kategori sukar

0,31 – 0,70 = soal kategori sedang

0,71 – 1,00 = soal kategori mudah

Penghitungan indeks Tingkat Kesukaran dilakukan untuk setiap nomor soal. Pada prinsipnya, skor rata-rata yang diperoleh peserta didik pada butir soal yang bersangkutan. Rumus yang digunakannya adalah untuk soal-soal yang berdimensi soal obyektif yaitu :

$$TK = n/N$$

n = jumlah siswa yang menjawab benar

N = jumlah sisea yang mengikuti tes

Fungsi tingkat kesukaran butir soal biasanya dikaitkan dengan tujuan tes. Misalnya untuk keperluan ujian semester digunakan butir soal yang tngkat kesukarannya sedang untuk keperluan seleksi digunakan tingkat kesukaran nya tinggi/sukar dna memiliki tingkat kesukaran rendah/mudah dipergunakan untuk keperluan diasnogtik. Analisis tingkat kesukaranjuga bisa dilakukan pada butir-butir soal uraian dengan analiais pemikiran kuantitatif.

Mean = jumlah skor tes siswa pada suatu soal / jumlah yang mengikuti tes

Tingkat kesukaran = Mean/ skor maksimum yang ditetapkan

Pendapat lain mengatakan bahwa kriteria indeks tingkat kesukaran soal bisa menggunakan rumus SR+ST (SR adalah peserta didik yang menjawab salah dari kelompok rendah dan ST adalah peserta didik yang menjawab salah dari kelompok tinggi, dilakukan dengan menggunakan Tabel Rose dan Stanleys, sebagai berikut :

Persentase	Option				Kategori
	2	3	4	5	
16	0,16 n	0,213 n	0,24 n	0,256 n	Mudah
50	0,50 n	0,667 n	0,75 n	0,80 n	Sedang
84	0,84 n	0,20 n	1,26 n	1,344 n	Sukar

Keterangan :

Option 2 adalah bentuk soal benar salah

Option 3,4 dan 5 adalah bentuk soal pilihan ganda

N adalah 27% dari banyaknya peserta didik yang mengikuti tes

2. Analisis daya pembeda

Dalam pemikiran kuantitatif, analisis butir soal dapat diukur dengan menggunakan kesukaran rata-rata skor dan varias kemudia indeks daya pembeda seperti korelas, indeks reliabilitas butir, point besimal dan lain sebagainya. Menurut Muhammad Nur, analisis butir soal merupakan istilah yang secara luas digunakan untuk mendefinisikan penghitungan dan pengkajian setiap sifat statistic jawaban peserta didik terhadap suatu butir tes tertentu dengan tiga parameter yaitu (1) parameter yang menunjukkan sebaran jawaban untuk setiap butir (parameter itu adalah rata-rata dan varians jawaban butir), (2) parameter yang menunjukkan tingkat hubungan antara jawaban butir dan kriteria yang dipilih, (3) parameter yang merupakan fungsi dari varians butir dan hubungan dengan suatu kriteria.

Analisis daya pembeda adalah mengkaji butir secara deskriptif dengan tujuan untuk mengetahui kesanggupan soal dalam membedakan peserta didik. Daya Pembeda (DP) adalah selisish proporsi jawaban benar pada siswa berkemampuan tinggi atau kelompok atas (KA) dengan siswa berkemampuan rendah atau kelompok bawah (KB). Soal yang baik adalah soal yang dapat membedakan kelompok peserta didik yang berkemampuan tinggi dan peserta yang berkemampuan rendah. Artinya apabila soal tersebut diberikan kepada siswa yang mampu, hasilnya menunjukkan prestasi yang tinggi dan apabila diberikan kepada siswa yang lemah, hasilnya rendah. DP soal berkisar antara -1 sampai dengan +1. Tanda negative berarti siswa KB yang menjawab benar soal tertentu lebih banyak dari siswa KA. Jika $DP = 0$, artinya soal tersebut tidak dapat memebedakan siswa KA dengan siswa KB. Dengan demikian, soal yang baik adalah soal dengan DP soal bertanda positif (+) sedangkan DP pengecoh diharapkan negative (-). Ketentuan tersebut didasarkan pada analisis $DP = 2 (KA-KB) / n$ atau biasa melalui $Dp = KA/nA! - KB / nB'$, dengan kriteria DP sebagai berikut :

Kriteria Daya Pembeda

Kriteria DP	Keterangan
$DP > 0,25$	Soal Diterima
$0 < DP < 0,25$	Soal Diperbaiki
$DP \leq 0$	Soal tiak dipakai / dibuang / ditolak

Cara lain yang dilakukan dalam analisis day apembeda adalah denan menggunakan rumus tingkat kesukaran soal yakni $SR - ST$ dengan kriteria yang digunakan Tabel Ross dan Stanley, sebagai berikut :

Jumlah Test (N)	n / 27% N	Option			
28 – 31		2	3	4	5
32 – 35	8	4	5	5	5
36 – 38	9	5	5	5	5
Dst....	10	5	5	5	5

Bila $SR-ST$ sama atau lebih besar dari nilai table, artinya butir soal itu mempunyai daya pembeda. Langkah penyelesaian daya pembeda dilakukan beberap tahap antara lain (1) Memeriksa jawaban soal semua peserta tes, (2) Membuat daftar peringkat hasil tes berdasarkan skor yang dicapai (3) menentukan jumlah sampel sebanyak 27% dari jumlah peserta tes untuk siswa pandai atau (KA) dan 27% untuk kelompok siswa kurang, atau (KB). (4) Menghitung jumlah siswa yang menjawab salah dari semua nomor soal, baik (KA) maupun (KB). (5) Menghitung jumlah selisih siswa yang

salah menjawab pada kelompok kurang dengan kelompok pandai. (SR-ST) (6) Membandingkan nilai selisih yang diperoleh dengan nilai Tabel Ross dan Stanley, (7) Menentukan ada tidaknya daya pembeda pada setiap nomor soal dengan kriteria “memiliki daya pembeda”. Bila nilai selisih jumlah siswa yang menjawab benar – salah antara kelompok kurnag dengan kelompok panai (SR-ST) sama atau lebih besar dari table/

Dengan demikian, tes dikatakan tidak memiliki daya pembeda apabila tes tersebut jika diujikan kepada siswa berprestasi tinggi hasilnya rendah, tetapi bila diberikan kepada siswa yang lemah hasilnya lebih tinggi. Atau bila diberikan kepada kedua kategori kelompok siswa tersebut, hasilnya sama saja. Jadi butir soal yang tidak memiliki daya pembeda diduga terlalu mudah atau terlalu sukar sehingga perlu diperbaiki atau diganti. Untuk mengetahui seberapa jauh setiap butir soal dapat mendekteksi atau membedakan kemampuan siswa yaitu siswa telah memahami atau belum memahami materi yang diajarkan guru. Apabila suatu butir soal tidak dapat membedakan kedua kemampuan siswa itu, maka butir soal itu dapat dicurigai “kemungkinannya”, disebabkan antara lain :(1) Kunci jawaban butir soal itu tidak tepat, (2) butir soal itu memiliki 2 atau lebih kunci jawaban yang benar. (3) kompetensi yang diukur tidak jelas, (4) pengecoh soal tidak berfungsi (5) materi yang ditanyakan terllau sult sehingga banyak siswa yang menebak (6)sebagian besar siswa yang memahamimateri yang ditanyakan berpikir ada yang salah informasi dalam butir soalnya.

E. Analisis Prosedur Perakitan Soal

Merakit soal adalah persoalan penyusunan soal yang memiliki prosedur dan tujuan yang sesuai dnegan apa yang diukur. Pengukuran soal dapat dibaca dari tingkat kesukaran, tingkat daya beda, tingkat pengecoh dan validitas soal seta reliabilitasnya yang tentunya masing-masing mempunyai langkah-langkah penyelesaian berdasarkan kriteria indeks masing-masing pula. Kaidah-kaidah seperti ini, didasarkan pada pemikiran kualitatif-deskriptif dan pemikiran kuantitaif – matematis dimana masing-masing soal, baik soal pilihan ganda, obyektif, uraian dan lain sebagainya diperuntukkan kepada keperluan tes itu sendiri. Masing-masing tes, baik tes maupun non tes mempunyai prosedur perakitan soal yang berbeda-beda. Ada yang dianalisa butir soal dengan teknik moderator dan teknik panel., juga menggunakan teknik IRT dan Kalibrasi.

Untuk itu, guru dapat bisa memahami dan mampu merakit soal pada suatu tes yang tepat sesuai dnegan langkah-langkahnya antara lain : (1) mengelompokkan soal-soal yang mengukur kompetensi dan materi yang sama, kemudian soal-soal itu ditempatkan dalam urutan yang sama., (2) memberi nomor urut soal didasarkan nomor urut soal dalam kisi-kisi soal, (3) mengecek setiap soal dalam satu paket tes apakah soal-soalnya sudah bebas dari kaiadh (setiap soal tidak boleh memberi petunjuk jawaban terhadap soal yang lain. (4) membuat petunjuk umum dan khusus untuk mengerjakan soal (5) membuat format lembar jawaban (6)membuat lembar kunci jawaban dan petunjuk penilaiannya. (7) menentukan dan menghitung penyebaran kunci jawaban (untuk pilihan ganda) dengan kaidah rumus penyelesaiannya (8) menentukan soal inti sebanyak 10% dari jumlah soal dalam satu paket. (9) menentukan besarnya bobot setiap soal (untuk soal bentuk uraian).

Perlu diketahui bahwa bobot soal adalah besarnya angka yang ditetapkan oleh guru untuk suatu butir soal dalam perbandingan rasio dengan butir soal lainnya dalam satu perangkat tes. Penentuan besar kecilnya bobot soal didasarkan atas tingkat ke dalam dan keluasan materi yang disoalkan atau kompleksitas jawaban yang dituntut oleh suatu soal. Analisis prosedur berikutnya adalah menyusun table konversi skor, terutama dalam suatu tes terdiri dari dua bentuk soal, misalnya bentuk pilihan ganda dan bentuk uraian atau tes tertulis dan tes praktik. Hal ini penting, karena selama ini guru dengan mudah menggabungkan antara masing-masing tes dan masing-masing soal.

Padahal tidak diperkenankan dalam beberapa pertimbangan. Skor dari soal bentuk pilihan ganda tidak dapat langsung digabung dengan skor bentuk uraian, sebab tingkat keluasan dan kedalaman materi yang ditanyakan tidak sama. Nilai keduanya bida digabung setelah keduanya ditetapkan bobot skornya. Ketentuan seperti ini dapat mempermudah menyusun table konversi skor.

Dalam ilmu pengukuran, konversi dapat disusun melalui konversi biasa (Teori Klasik) dan konversi yang terkalibrasi dengan model respon butir (Teori Modern). Menurut ilmu evaluasi bahwa konversi nilai yang terkalibrasi (ditempatkan pada hasil perhitungan) adalah konversi nilai yang disusun berdasarkan kemampuan peserta didik dari tingkat kesukaran butir soal yang terkalibrasi dengan model Rasch. Untuk memahami model ini diperlukan pengertian dari guru untuk memahaminya. Setiap jumlah jawaban benar soal misalnya 1 sampai 50, masing-masing butir memiliki tingkat kemampuan (untuk teori klasik tidak ada tetapi dalam teori modern-respon butir ada) dengan memakai model Rasch. Model ini menekankan nilai abilitas (misal -3,00 sampai +3,00) ditransformasikan ke dalam skala 0-10, 0-100 atau 0-4. Contoh untuk dapat ditransformasi ke dalam skala 0-100 diperlukan rata-rata 50 dan standar deviasi 5, sehingga untuk membuat table konversi menggunakan rumus : $Y=50+rY$. Y = nilai peserta didik dan X = adalah nilai abilitas. Sehingga dalam analisis prosedur ini konversi terkalibrasi dapat disusun. Sedangkan dalam analisis prosedur teori klasik, model konversi tidak sama dengan model teori modern. Konversi ala prosedur pengukuran klasik tidak susah dan biasa digunakan guru di sekolah, yaitu untuk memperoleh nilai murni peserta didik. Bila guru menghendaki skor maksimum 10 digunakan prosedur klasik (skor perolehan : skor maksimum) x 10 dan bila menggunakan skor maksimum 100 digunakan nilai konversi dengan prosedur analisisnya (skor perolehan : skor maksimum x 100 atau apabila menggunakan skor maksimum 4 digunakan nilai konversi dengan ketentuan (skor perolehan : skor maksimum) x 4.

Meskipun demikian, sebagian kalangan menilai bahwa konversi dengan prosedur analitik klasik dipandang memiliki dua kelemahan, *pertama*, setiap butir soal dihitung memiliki tingkat kesukaran yang sama artinya peserta didik manapun yang menjawab benar 40 dari 50 butir soal dalam satu tes misalnya yang bersangkutan akan memperoleh nilai 8 (untuk konversi skor maksimum 10), 80 (untuk konversi skor maksimum 100) dan 0,2 (untuk konversi skor maksimum 4). *Kedua*, tingkat kesukaran butir soal tidak ditempatkan /dikalibrasi pada skala yang sama artinya bahwa butir-butir soal tidak disusun berdasarkan tingkat kesukarannya dan kemampuan peserta didik sehingga model konversi ini belum bisa menentukan nilai murni peserta didik yang sebenarnya. Kelemahan ini tidak berlaku pada analisis prosedur model Rasch atau Teori Modern. Kelebihannya adalah tidak didasarkan pada grup dependen, skor peserta didik dideskripsikan bukan tes dependen dan model teori modern ini menekankan pada tingkat butir soal bukan pada tes yang tidak memerlukan parallel tes dalam menentukan reliabilitas tes, dimana model ini juga memberikan suatu pengukuran ketepatan untuk setiap skor tingkat kemampuan peserta didik.

Dengan demikian, prosedur-prosedur perakitan soal tersebut menekankan adanya ketelitian yang hati-hati dimana masing-masing butir soal memiliki eksistensi skor yang berbeda-beda yang tentunya model analisis juga berbeda-beda. Perbedaan ini dipicu atas pandangan bahwa dalam penyusunan soal akan melibatkan pengaruh psikologis dan social yang dengan itu bisa membuat pemetaan-pemetaan berdasarkan rumus kualitatif dan rumus kuantitatif terhadap butir-butir soal baik materi, konstruksi, bahasa maupun tingkat kesukaran dan daya pembeda.

DAFTAR PUSTAKA

- Azwar Saifuddin, *Reliabilitas dan Validitas*, cet ke-6, Yogyakarta : Pustaka Pelajar, 2006
- Direktorat Pembinaan Pendidikan dan Pelatihan , direktorat Jenderal PMPTK bekerjasama dengan THE WORL BANK, *Panduan BERMUTU (Better Education Through Reformed Management and Universal Techer Upgrading)* 2008.
- Departemen Pendidikan Nasional; Direktorat Jenderal Manajemen Pendidikan Dasar dan Menengah, *Pedoman Analisis Butir Soal*, Jakarta : direktorat Pembinaan Sekolah Menengah Atas, 2008
- Djemari Mardapi, *Teknik Penyusunan Instrumen Tes dan Nontes*, cet. Ket-1. Yogyakarta : Mitra Cendekia, 2008.
- Lendgren, H.G, *Educational Pskychology in the Classroom*, New York. 1967. H.445
- Mudjijo, *Hasil Belajar*, cet I, Jakarta : Bumi Aksara, 1995
- Nur Muhammad, *Pengantar Teori Tes; Naskah bahan Pengajaran/buku teks program refresher*, Jakarta : Direktorat Jenderal Pendidikan Tinggi, Departemen Pendiikan and Kebudayaan
- Nurkencana, Wayan an PPN Sumartana, *Evaluasi Pendidikan*, cet ke-1 Suarabaya: Usaha Nasional.
- Sudjana, Nana. *Penilaian Hasil Proses Belajar Mengajar*, Cet ke-10, Jakarta : Remaja Rosdakarya, 2005