

## **MAMPU MEMILIH SOAL BERDASARKAN TINGKAT KESUKARAN**

**Meyke Joice Tresia Rajagukguk**

Pendidikan Agama Kristen, Institut Agama Kristen Negeri Tarutung

[meykerajagukgu@gmail.com](mailto:meykerajagukgu@gmail.com)

**Dorlan Naibaho**

Pendidikan Agama Kristen, Institut Agama Kristen Negeri Tarutung

[dorlannaibaho04@gail.com](mailto:dorlannaibaho04@gail.com)

### **Abstrak**

Kemampuan memilih soal berdasarkan tingkat kesulitannya adalah kemampuan memilih atau membuat soal atau tes dengan tingkat kesulitan yang berbeda-beda untuk disesuaikan dengan kemampuan atau tingkat pengetahuan khalayak sasaran atau siswa. Hal ini memungkinkan pemahaman mereka diukur secara akurat. Belajar tanpa penilaian adalah sebuah kehilangan. Sebab guru tidak mendapat informasi penting tentang tingkat pencapaian tujuan, tingkat penguasaan materi pembelajaran, kelebihan dan kekurangan siswa dalam belajar, serta kelebihan guru terhadap proses pembelajaran yang dapat dikembangkan. Walaupun evaluasi dianggap sebagai pekerjaan penting dan rutin seorang guru, namun dalam kehidupan sehari-hari di lapangan, sistem evaluasi pembelajaran bukannya tanpa permasalahan. Tes obyektif, khususnya tes pilihan ganda di Indonesia, digunakan dalam tes formatif dan sumatif, yang meliputi ujian sekolah, ujian nasional, dan juga digunakan dalam ujian masuk universitas. Keuntungan tes pilihan ganda adalah dapat mencakup berbagai topik dan mudah untuk dinilai. Jenis tes pilihan ganda adalah tes pilihan ganda standar dan tes pilihan ganda terkait. Saat ini tes pilihan ganda asosiasi sudah tidak digunakan lagi dalam ujian sekolah dan ujian negara di tingkat SD, SMP, atau SMA/SMK. Namun ujian masuk perguruan tinggi masih menggunakan tes pilihan ganda dari asosiasi, sehingga biasanya peserta tes SMA/SMK mengalami kesulitan dalam menjawab soal pilihan ganda dari asosiasi karena tidak terbiasa menghadapi soal-soal seperti itu. Soal pilihan ganda asosiatif tidak digunakan di sekolah karena soal pilihan ganda asosiatif dianggap lebih sulit dibandingkan soal pilihan ganda standar baik dari segi pengolahan maupun penanyaannya, namun dapat lebih membedakan siswa berkemampuan tinggi dengan siswa berkemampuan rendah. .

Kata kunci: Pengertian, tujuan, alat ukur.

## **PENDAHULUAN**

Penilaian pembelajaran merupakan suatu sistem yang berarti serangkaian kegiatan yang mencakup beberapa unsur berbeda sebagai satu kesatuan. Setiap elemen mempunyai fungsi dan perannya masing-masing, dan perubahan pada satu elemen akan mempengaruhi elemen lainnya. Penilaian merupakan hal yang tidak terpisahkan dalam dunia pendidikan dan sama pentingnya dalam pembelajaran.

Belajar tanpa penilaian adalah sebuah kehilangan. Sebab guru tidak mendapat informasi penting tentang tingkat pencapaian tujuan, tingkat penguasaan materi pembelajaran, kelebihan dan kekurangan siswa dalam belajar, serta kelebihan guru terhadap proses pembelajaran yang dapat dikembangkan. Walaupun evaluasi dianggap sebagai pekerjaan penting dan rutin seorang guru, namun dalam kehidupan sehari-hari di lapangan, sistem evaluasi pembelajaran bukannya tanpa permasalahan.

Tes obyektif, khususnya tes pilihan ganda di Indonesia, digunakan dalam tes formatif dan sumatif, yang meliputi ujian sekolah, ujian nasional, dan juga digunakan dalam ujian masuk universitas. Keuntungan tes pilihan ganda adalah dapat mencakup berbagai topik dan mudah untuk dinilai. Jenis tes pilihan ganda adalah tes pilihan ganda standar dan tes pilihan ganda terkait. Saat ini tes pilihan ganda asosiasi sudah tidak digunakan lagi dalam ujian sekolah dan ujian negara di tingkat SD, SMP, atau SMA/SMK. Namun ujian masuk perguruan tinggi masih menggunakan tes pilihan ganda dari asosiasi, sehingga biasanya peserta tes SMA/SMK mengalami kesulitan dalam menjawab soal pilihan ganda dari asosiasi karena tidak terbiasa menghadapi soal-soal seperti itu.

Soal pilihan ganda asosiatif tidak digunakan di sekolah karena soal pilihan ganda asosiatif dianggap lebih sulit dibandingkan soal pilihan ganda standar baik dari segi pengolahan maupun penanyaannya, namun dapat lebih membedakan siswa berkemampuan tinggi dengan siswa berkemampuan rendah. .

## **METODE PENELITIAN**

Metode penelitian adalah prosedur atau cara sistematis yang dilakukan oleh peneliti untuk mendapatkan kebenaran dari suatu fenomena melalui pertimbangan logis dan didukung oleh data faktual sebagai bukti konkret (objektif). Penelitian ini menggunakan metode kualitatif dengan meneliti, menganalisa serta menelaah beberapa jumlah dan beberapa

sumber buku yang terkait dengan topik Mampu memilih soal berdasarkan tingkat kesukaran yang bertujuan agar seorang pendidik mampu dalam memilih soal berdasarkan tingkat kesukaran bagi peserta didik dalam pembelajaran.

## **HASIL DAN PEMBAHASAN**

### **Pengertian soal dan tingkat kesukaran?**

Soal adalah suatu pertanyaan atau pernyataan yang dibuat dalam suatu ujian, tes atau latihan yang mengukur pengetahuan, pemahaman atau kemampuan seseorang dalam bidang tertentu. Pada saat yang sama, tingkat keparahan mengacu pada kesulitan atau kompleksitas masalah. Penentuan tingkat kesulitan suatu soal tergantung pada faktor-faktor seperti tingkat pengetahuan yang dibutuhkan, kompleksitas materi dan kemampuan orang dalam menjawab soal. Tingkat kesukaran dibagi menjadi mudah, sedang atau sulit untuk memastikan soal sesuai dengan kemampuan dan tingkat pengetahuan peserta tes atau peserta tes. Mampu memilih soal berdasarkan tingkat kesukaran adalah kemampuan memilih atau mengurutkan soal atau soal yang tingkat kesulitannya sesuai atau sesuai dengan kemampuan atau tingkat pengetahuan individu atau kelompok yang menghadapinya. Hal ini untuk memastikan bahwa soal-soal tersebut tidak terlalu mudah dan tidak terlalu sulit bagi peserta tes.

Tujuan dari analisis kesukaran adalah untuk mengetahui apakah bentuk soal tersebut mudah atau sulit. Kesulitan adalah angka yang menunjukkan betapa sulit atau mudahnya suatu masalah (Arikunto, 1999). Selain itu, tingkat kesulitan suatu soal adalah kemampuan menjawab soal dengan benar pada tingkat keahlian tertentu, biasanya dinyatakan dalam indeks. Soal yang baik adalah soal yang tidak terlalu mudah dan tidak terlalu sulit. Soal yang terlalu mudah tidak dapat mendorong siswa bekerja lebih keras untuk menyelesaikannya. Sebaliknya soal yang terlalu sulit akan membuat siswa putus asa dan tidak mempunyai semangat untuk mencoba karena hal tersebut diluar kemampuannya. Misalnya Guru A memberikan ulangan yang mudah dan sulit, sedangkan Guru B memberikan ulangan yang sulit. Dalam metode ini, siswa akan giat belajar ketika menghadapi ujian dari guru B, dan sebaliknya, ketika mendapat ujian dari guru A, mereka tidak akan mau belajar dengan giat atau bahkan tidak mau belajar.

Angka yang menunjukkan tingkat kesulitan dan kemudahan suatu tugas disebut indeks kesulitan. Indeks keparahan berkisar antara 0,00 hingga 1,00. Indeks kesukaran ini

menunjukkan tingkat kesukaran soal. Soal yang indeks kesukarannya 0,00 menunjukkan soal terlalu sulit, sedangkan indeks kesukaran 1,00 menunjukkan soal terlalu sulit. terlalu mudah 0,00 Berat 0,10 sederhana .Tingkat kesulitan soal bukti biasanya berkaitan dengan tujuan tes. Misalnya untuk ujian semester digunakan soal-soal dengan tingkat kesulitan sedang. Untuk keperluan seleksi digunakan soal-soal tingkat kesulitan tinggi yang sulit, dan untuk tujuan diagnostik umumnya digunakan soal-soal tingkat kesulitan rendah yang mudah (Asrul dkk. 2014).

### **Hakikat tes**

Menurut Anas Sudijono (2006:93)mengatakan bahawa tes adalah alat atau prosedur yang digunakan dalam rangka pengukuran dan penilaian.Sedangkan menurut Ratna sayekti (2008:5),tes adalah seperangkat butir atau diberikan kepada siswa dengan syarat-syarat tertentu atau tes adalah prosedur yang sistematis untuk mengobservasi tingkah laku. Norman E. Gronlund (2002:106) mengemukakan enam prinsip yang perlu diperhatikan ketika merancang tes hasil belajar, yaitu:

- (1) tes hasil belajar harus mengukur hasil belajar yang terdefinisi dengan jelas dan konsisten dengan tujuan pembelajaran
- (2)tes hasil belajar harus mengukur sampel yang representatif dari hasil belajar dan materi yang dibahas dalam penelitian,
- (3) tes hasil belajar harus memuat pertanyaan-pertanyaan yang paling sesuai untuk mengukur hasil belajar yang diinginkan,
- (4) tes hasil belajar harus dirancang sedemikian rupa sedemikian rupa sehingga hasilnya sesuai dengan tujuan tertentu dengan hasil atau sasaran yang diinginkan.
- (5) tes hasil belajar harus diberikan dengan reliabilitas tinggi dan kemudian diinterpretasikan secara cermat,
- (6) tes hasil belajar harus digunakan meningkatkan hasil belajar.

Menurut Slamento (2008:19), syarat tes yang baik adalah:

- (1) valid, yaitu. tes tersebut harus dapat mengukur apa yang seharusnya diukur,
- (2) reliabel, yaitu. jika digunakan dalam kelompok yang sama/serupa. siswa beberapa kali dalam setahun. pada waktu yang berbeda hasilnya sama,
- (3) objektif, artinya tidak dipengaruhi oleh pengaruh subjektif penilai,
- (4) seimbang menurut materi pembelajaran, tingkat kesulitan dan tujuan pengajaran,

- (5) dibedakan. Artinya, hasil tes harus dapat membedakan prestasi belajar setiap siswa yang berhasil, kurang berhasil dan tidak berhasil,
- (6) standar artinya hasil tes mudah diukur menurut standar atau kaidah tertentu,
- (7) adil Artinya pengujian tersebut menimbulkan masalah yang dapat diterima dan tidak ada jebakan.
- (8) praktis, artinya pembiayaan efisien dan pelaksanaannya mudah.

### **Kualiatas tes**

Menurut Wayan Nurkencana (2002:141), tes berkualitas baik dapat dianalisis dari empat kriteria adalah:

- (1) validitas artinya tes tersebut dapat mengukur apa yang hendak diukur,
- (2) reliabilitas artinya tes tersebut dapat menunjukkan hasil yang konsisten (tetap) walaupun sudah berkali-kali tes tersebut dilakukan,
- (3) tingkat daya pembeda artinya tes tersebut dapat membedakan siswa menurut tingkat kemampuannya.
- (4) tingkat kesukaran artinya memiliki keseimbangan antara butir soal sukar, sedang dan mudah.

### **Bentuk tes**

Menurut Ngalim Purwanto(2001:35), tes yang biasa dilakukan guru terdiri dari tes lisan (oral test), tes tertulis (written test) dan tes perbuatan (practice test). Tes tertulis dibagi menjadi tes obyektif dan tes uraian (essay). Tes obyektif terdiri dari tes pilihan ganda, tes benar salah, tes menjodohkan, dan tes isian, sedangkan tes uraian terdiri dari uraian bebas dan uraian terbatas. Bentuk atau teknik tes menurut Depdiknas (2007:4) yaitu unjuk kerja (performance), penugasan (project), hasil kerja (product), tertulis (paper & pen), portofolio (portfolio). Sikap (attitude), penilaian diri (self evaluation). Bentuk tes pilihan ganda terdapat pada teknik tes tertulis. Suharsimi Arikunto (2009:23), secara garis besar alat evaluasi digolongkan menjadi dua.

Menurut Ngalim Purwanto (2001:35) tes yang biasa dilakukan guru terdiri dari tes lisan, tertulis, dan praktik. Tes tertulis dibagi menjadi tes objektif dan esai. Tes obyektif terdiri dari tes pilihan ganda, tes benar-salah, tes korespondensi, dan tes isian, sedangkan tes deskriptif terdiri dari uraian bebas dan uraian terbatas. Tergantung pada bentuk atau

teknik tesnya Departemen Pendidikan (2007:4), yaitu kegiatan, tugas (proyek), hasil kerja (produk), tertulis (kertas dan pena), portofolio (portofolio). Sikap, harga diri. Format ujian pilihan ganda dapat ditemukan pada teknik ujian tertulis. Suharsimi Arikunto (2009:23) secara garis besar membagi alat penilaian menjadi dua jenis, yaitu tes dan non tes.

Tes dapat berupa tes formatif dan sumatif. Tes desain adalah ujian yang diambil pada akhir suatu mata pelajaran atau lebih, seperti ujian harian, UTS (Ujian Tengah Semester) dan latihan-latihan yang bertujuan untuk meningkatkan pembelajaran. Tes formatif termasuk dalam evaluasi pembelajaran. Ujian sumatif adalah ujian akhir semester atau akhir tahun yang berupa UAS (Ujian Akhir Semester), Ujian Nasional (UN) dan Ujian Sekolah (AS) untuk menentukan jenjang/perkembangan kelas dan kelulusan. Tes formulir disertakan evaluasi hasil belajar. Tes juga dapat berupa pre-test dan post-test. Ujian awal merupakan tes yang dilakukan sebelum mata pelajaran diberikan, yang tujuannya adalah untuk mengetahui seberapa banyak informasi yang telah dipelajari siswa tentang mata pelajaran yang diberikan dan untuk mengarahkan perhatian dan konsentrasi siswa pada mata pelajaran yang diberikan. . Post-test adalah tes yang diberikan kepada siswa setelah materi pelajaran disampaikan kepada siswa dan tujuannya untuk mengetahui keberhasilan materi yang diberikan.

### **Analisis tingkat kesukaran soal**

Analisis kesukaran soal artinya melihat soal tes berdasarkan tingkat kesukarannya untuk menentukan mana soal yang mudah, sedang, dan sulit. Untuk mendapatkan soal yang berkualitas, tingkat kesukaran soal diseimbangkan di samping pemenuhan kualifikasi dan kehandalan. Tujuannya adalah untuk menyeimbangkan antara soal yang relatif mudah, sedang, dan sulit. Tingkat kesukaran soal dilihat dari kemampuan siswa dalam menjawab, bukan dari sudut pandang guru sebagai pemberi soal. Sang Pencipta. Dalam menganalisis tingkat kesukaran soal, penentuan proporsi dan kriteria soal mudah, sedang, dan sulit merupakan hal yang penting. Ada beberapa pertimbangan utama dalam menentukan proporsi soal pada kategori mudah, sedang, dan sulit. Aspek pertama adalah keseimbangan, yaitu jumlah pertanyaan yang sama untuk ketiga kategori tersebut.

Cara melakukan analisis untuk mengetahui tingkat kesukaran suatu soal adalah dengan menggunakan rumus:

$I = B \div N =$  indek kesulitan untuk setiap butir soal

B = banyaknya siswa yang menjawab benar setiap butir soal

N = banyaknya siswa yang memberikan jawaban pada soal yang dimaksudkan Kriteria yang digunakan adalah makin kecil indeks yang diperoleh, makin sulit soal tersebut. Sebaliknya, makin besar indeks yang diperoleh, makin mudah soal tersebut. Kriteria indeks kesulitan soal itu adalah sebagai berikut:

0,30-soal kategori sukar

0.31 0.70-soal kategori sedang.

0.71 1.00-soal kategori mudah.

### **Tingkat kesukaran (Difficulty Level)**

Saifudin Azwar (2006:129) mengatakan tingkat kesukaran suatu soal adalah perbandingan jumlah peserta tes yang menjawab soal dengan benar dengan jumlah peserta tes. Artinya, semakin banyak peserta tes yang menjawab soal dengan benar, maka semakin tinggi indeks kesukarannya, yang berarti semakin mudah soal tersebut. Sebaliknya, semakin sedikit peserta tes yang menjawab soal dengan benar, maka semakin sulit soal tersebut. Sedangkan menurut HJX. Dalam Fernandes (2004:60), tingkat kesukaran suatu soal diukur dari persentase siswa yang menjawab soal dengan benar. Jika soalnya mudah maka tingkat kesulitannya lebih tinggi. Soal yang nilai p-nya mendekati 0 merupakan soal yang sangat sulit, sedangkan soal yang nilai p-nya mendekati 1 adalah soal yang sangat mudah. Indeks keparahan 0,3-0,7 sangat bagus. Bistok Sirait (2009:301) berpendapat bahwa tinggi rendahnya tingkat kesulitan pokok soal dapat disebabkan oleh rumitnya topik soal dan kondisi pilihan jawaban yang diberikan. Dalam artian cakupan soal seringkali membingungkan peserta tes, dan alternatif jawaban atau susunan kata pertanyaan yang homogen sulit dipahami atau memiliki arti ganda bagi peserta tes. Menurut Anastas dan Susan Urbina (2007:128), tingkat kesulitan soal berkaitan dengan persentase orang yang menjawab soal dengan benar. Semakin sederhana jiwanya, semakin tinggi persentasenya. Jika tingkat kesulitan 70% ( $p=0.70$ ), maka soal dianggap lebih mudah dibandingkan dengan tingkat kesulitan 15% ( $p=0.15$ ) Yang terbaik adalah soal dengan tingkat kesulitan 0.50.

Wayan Nurkencana (2002:161) menyatakan tentang klasifikasi kesukaran soal yaitu. jika p-value (persentase kesukaran soal) = 0,811,00 maka obyek soal sangat mudah, jika p-value = 0,61 -0,80 maka obyek soal mudah, dengan p-value = 0,41 0,60 rata-rata

soal. jika  $p = 0,21-0,40$  maka soalnya sulit, jika  $p = 0,00-0,20$  maka soalnya sangat sulit. Nana Sujana (2002:135) mengatakan tingkat kesulitan ada yang mudah, sedang, dan sulit. Tingkat kesulitan yang baik adalah  $0,25-0,75$ . Soal yang tingkat kesulitannya di bawah  $0,25$  berarti soal tersebut sulit, sedangkan soal yang tingkat kesulitannya di atas  $0,75$  berarti terlalu mudah.

Berdasarkan teori-teori yang disampaikan, tingkat kesukaran suatu soal adalah seberapa mudah dan sulitnya suatu soal bagi siswa. Tingkat kesukaran dinyatakan dengan persentase siswa yang menjawab soal dengan benar. Semakin tinggi persentase siswa yang menjawab soal dengan benar, maka semakin mudah soal tersebut. Sebaliknya, semakin kecil persentase siswa yang menjawab soal dengan benar, maka soal tersebut semakin sulit. Jadi soal yang dijawab benar oleh 85% siswa memiliki tingkat kesulitan  $0,85$ , soal ini jelas lebih mudah dibandingkan dengan tingkat kesulitan 20% ( $0,20$ ).

### **Tingkat kesukaran soal(Diffyculty index)**

Untuk menghitung tingkat kesukaran soal bentuk objektif dapat digunakan dua cara, yaitu: Cara Pertama, menggunakan rumus tingkat kesukaran (TK):

$$TK = \frac{WL+WH}{nL+nH} \times 100\%$$

Keterangan:

WL Jumlah peserta didik yang menjawab salah dari kelompok bawah

WH: Jumlah peserta didik yang menjawab salah dari kelompok atas

nL jumlah kelompok atas

nH : jumlah kelompok bawah

Sebelum menggunakan rumus diatas, harus melakukan langkah-langkah dibawah ini dahulu:

- 1) Menyusun lembar jawaban peserta didik dari skor tertinggi sampai skor terendah.
- 2) Mengambil 27% lembar jawaban dari atas yang selanjutnya disebut kelompok atas (higher group), dan 27% lembar jawaban dari bawah yang selanjutnya disebut kelompok bawah (lower group). Sisa sebanyak 46% disisihkan.
- 3) Membuat tabel untuk mengetahui jawaban (benar atau salah) dari setiap peserta didik, baik kelompok atas maupun kelompok bawah. Jika jawaban peserta didik benar, diberi tanda + (plus), sebaliknya jika jawaban peserta didik salah, diberi tanda (minus).

### **Daya pembeda soal (Discrimination Power)**

Dali S. Naga (2002:67) mengatakan bahwa daya pembeda soal terletak pada kemampuan soal membedakan peserta tes dari kelompok yang berprestasi tinggi dan rendah. Dengan kata lain, semakin tinggi daya pembeda soal maka semakin banyak peserta ujian pada kelompok tinggi yang dapat menjawab soal dengan benar, dan semakin sedikit peserta ujian pada kelompok rendah yang dapat menjawab soal dengan benar. Nilai D yang dapat diterima (diskriminasi: kemampuan membedakan pertanyaan) adalah 0,30 atau lebih tinggi. Sekaligus bisa dikatakan cukup memuaskan adalah 0,40 ke atas.

Menurut Charles D. Hopkins dan Richard L. Antes (1999: 159), Indeks Diskriminasi Pertanyaan mengukur seberapa baik suatu pertanyaan membedakan tingkat kemampuan siswa. Nilai indeks kekhususan soal adalah -1,00 sampai +1,00. Semakin tinggi nilai D suatu pertanyaan maka semakin baik pertanyaan tersebut membedakan siswa yang berkemampuan tinggi dengan siswa yang berkemampuan rendah. Soal dengan  $D = 0,40$  dinilai sangat efektif dalam membedakan siswa berkemampuan tinggi dengan siswa berkemampuan rendah. dan soal dengan nilai  $D$  0,20 hingga 0,39 dianggap memuaskan. Soal dengan nilai  $D$  kurang dari 0,39 sebaiknya dikoreksi.

Namun Fernandes (2004:27) mengatakan bahwa daya pembeda soal adalah kemampuan memisahkan siswa yang cerdas dengan siswa yang kurang cerdas. Pada saat yang sama, terdapat indeks kekuatan pembeda dari pertanyaan-pertanyaan tersebut persentase perbedaannya sebesar 27% siswa yang meraih nilai tinggi (grup di awal) dan 27% siswa dengan hasil buruk (kelompok bawah). Pertanyaan dengan indeks diskriminasi antara 0,15 dan 0,20 atau lebih tinggi menunjukkan diskriminasi yang baik. Berdasarkan pandangan di atas, daya pembeda soal terletak pada kemampuan soal membedakan antara kelompok tes berkemampuan tinggi dan kelompok tes berkemampuan rendah. Nilai daya diferensial dinyatakan dengan indeks daya diferensial. Semakin tinggi atau besar indeks kekuatan diskriminatif suatu pertanyaan, maka pertanyaan tersebut semakin baik dalam membedakan kelompok tinggi dan rendah. Untuk menghitung daya diferensiasi diambil pertanyaan dari 27% kelompok tinggi dan 27% kelompok rendah. Nilai daya diskriminatif soal berkisar antara -1 hingga +1. Jika siswa pada kelompok tinggi dapat menjawab lebih banyak soal dengan benar dibandingkan siswa pada kelompok rendah, maka indeks diskriminasinya positif. Apabila siswa kelompok tinggi dan rendah dapat menjawab pertanyaan dengan benar, maka pertanyaan tersebut tidak bersifat diskriminatif. Jika siswa

pada kelompok rendah dapat menjawab lebih banyak soal dengan benar dibandingkan siswa pada kelompok tinggi, maka pertanyaan tersebut mempunyai daya pembeda negatif. Soal yang memiliki indeks daya diskriminan negatif dan nol dibuang karena soal tersebut tidak dapat membedakan siswa berkemampuan tinggi dan rendah.

### **Reliabilitas tes**

Nama lain dari keandalan adalah tekad, keyakinan, stabilitas, dapat dipercaya, keteguhan atau stabilitas. Namun yang dimaksud dengan reliabilitas suatu tes adalah sebagai berikut: Menurut Anne Anastase dan Susan Urbina (2007:63), reliabilitas suatu tes adalah kestabilan hasil yang diperoleh orang yang sama ketika diujikan kembali. tes yang sama dalam situasi yang berbeda. Suatu tes dikatakan reliabel apabila hasil tes seseorang sama, meskipun tes tersebut diuji ulang pada waktu yang berbeda.

T. Raka Joni (2006:36) mengatakan tes dikatakan reliabel dan tes ditentukan berdasarkan hasil tes. Uji reliabilitas memberikan setiap hasil tes nilai yang sama untuk orang yang sama.

Sementara itu, Syaifuddin Azwar (2006:180) mengatakan istilah reliabilitas sering disamakan dengan konsistensi, kestabilan. yang pada hakikatnya menunjukkan sejauh mana suatu pengukuran tidak dapat menghasilkan hasil yang relatif berbeda bila pengukuran diulangi pada subjek yang sama. Crocker dan Algina (1996:6) mengemukakan bahwa tinggi rendahnya koefisien reliabilitas dipengaruhi oleh beberapa faktor, antara lain: (1) lama tes. (2) kecepatan pelaksanaan tes (batas waktu), (3) homogenitas kelompok, (4) kesukaran soal tes. Menurut Dali S. Naga (2008:50), metode perhitungan reliabilitas adalah: (1) reliabilitas stabilitas (test-retest), (2) reliabilitas ekivalensi (equivalence test). (3) reliabilitas konsistensi internal (kesetaraan: Spearman Brown, politomi: 1,2,3,4,5.

Berdasarkan teori-teori di atas dapat disimpulkan bahwa reliabilitas tes adalah konsistensi tes. Suatu tes dikatakan reliabel jika memberikan hasil yang sama meskipun tes tersebut diulangi. Reliabilitas dinyatakan dalam angka dan dikenal sebagai koefisien reliabilitas. Semakin tinggi koefisien reliabilitas tes, semakin tinggi pula reliabilitas tes tersebut.

## **KESIMPULAN**

Penilaian pembelajaran merupakan suatu sistem yang berarti serangkaian kegiatan yang mencakup beberapa unsur berbeda sebagai satu kesatuan. Setiap elemen mempunyai fungsi dan perannya masing-masing, dan perubahan pada satu elemen akan mempengaruhi elemen lainnya. Penilaian merupakan hal yang tidak terpisahkan dalam dunia pendidikan dan sama pentingnya dalam pembelajaran.

Belajar tanpa penilaian adalah sebuah kehilangan. Sebab guru tidak mendapat informasi penting tentang tingkat pencapaian tujuan, tingkat penguasaan materi pembelajaran, kelebihan dan kekurangan siswa dalam belajar, serta kelebihan guru terhadap proses pembelajaran yang dapat dikembangkan. Walaupun evaluasi dianggap sebagai pekerjaan penting dan rutin seorang guru, namun dalam kehidupan sehari-hari di lapangan, sistem evaluasi pembelajaran bukannya tanpa permasalahan.

Dalam mengevaluasi alat ukur yang digunakan untuk mengukur prestasi akademik siswa (siswa, siswa, pelajar, dan sebagainya). Alat ukur yang dimaksud adalah tes hasil belajar yang sebagaimana telah kita pahami terdiri dari seperangkat item soal. Dalam penerapannya mempunyai tugas dan peranan yang sangat penting dalam mengetahui tujuan yang ingin dicapai. Penentuan tingkat kesulitan suatu soal tergantung pada faktor-faktor seperti tingkat pengetahuan yang dibutuhkan, kompleksitas materi dan kemampuan orang dalam menjawab soal. Tingkat kesukaran dibagi menjadi mudah, sedang atau sulit untuk memastikan soal sesuai dengan kemampuan dan tingkat pengetahuan peserta tes atau peserta tes. Mampu memilih soal berdasarkan tingkat kesukaran adalah kemampuan memilih atau mengurutkan soal atau soal yang tingkat kesulitannya sesuai atau sesuai dengan kemampuan atau tingkat pengetahuan individu atau kelompok yang menghadapinya. Hal ini untuk memastikan bahwa soal-soal tersebut tidak terlalu mudah dan tidak terlalu sulit bagi peserta tes. Tujuan dari analisis kesukaran adalah untuk mengetahui apakah bentuk soal tersebut mudah atau sulit. Hakikat tes, kualitas tes, bentuk tes, analisis tingkat kesukaran, tingkat kesukaran ( Difficulty Level), Tingkat kesukaran soal( Difficulty index), Daya tarik pembeda soal ( Discrimination Power), Reliabilitas tes. bahwa reliabilitas tes adalah konsistensi tes. Suatu tes dikatakan reliabel jika memberikan hasil yang sama meskipun tes tersebut diulangi. Reliabilitas dinyatakan dalam angka dan dikenal sebagai koefisien reliabilitas. Semakin tinggi koefisien reliabilitas tes, semakin tinggi pula reliabilitas tes tersebut.

## REFERENSI

- Anastasi, Anne dan Susan Urbina. Tes Psikologi. Jakarta: Prenhallindo, 2007.
- Arikunto, Suharsimi. Dasar-dasar Evaluasi Pendidikan. Jakarta: Bumi Aksara, 2009. Azwar, Saifuddin. Tex Prestasi. Yogyakarta: Pustaka Pelajar, 2006..
- Azwar, Saifuddin. Reliabilitas, Validitas, Interpretasi dan Komputasi. Yogyakarta: Liberty, 2006. Depdikbud. Pedoman Penulisan Soal Hasil Belajar IPS. Jakarta: Balitbang, Sisjian, 2007. Depdiknas. Pedoman Penulisan Soal Bentuk Pilihan Ganda IPS SMP. Jakarta: Dinas Pendas, 2004. Blood, Don F. dan Budd William. Educational Measurement and Evaluation. New York: Harper and Row, 2002.
- Chiffs: Prentise-Hall, 2002. Hehahia. A.D. Pengaruh Jumlah Alternatif Tes Obyektif Bentuk Pilihan Ganda Terhadap Reliabilitas. Tingkat Kesukaran dan Daya Pembeda. Tesis. Jakarta: PPSUNJ, 2007.
- Hopkins, Charles D dan Antes, Richard L. Classroom Measurement and Evaluation. Illionis. F.E. Peacock, 1999.
- Joni T. Raka Pengukuran dan Penilaian Pendidikan. Surabaya: Karya Anda, 2006.
- Naga, Dali. S. Pengantar Teori Skor pada Pengukuran Pendidikan. Jakarta: Gunadarma, 2002.
- Nurkencana, Wayan dan Supartana. Evaluasi Hasil Belajar. Surabaya: Usaha Nasional, 2002 Purwanto, Ngalim. Prinsip-Prinsip dan Teknik Evaluasi Pengajaran. Bandung: Remaja Rosdakarya. 2001. Sirait, Bistok.