

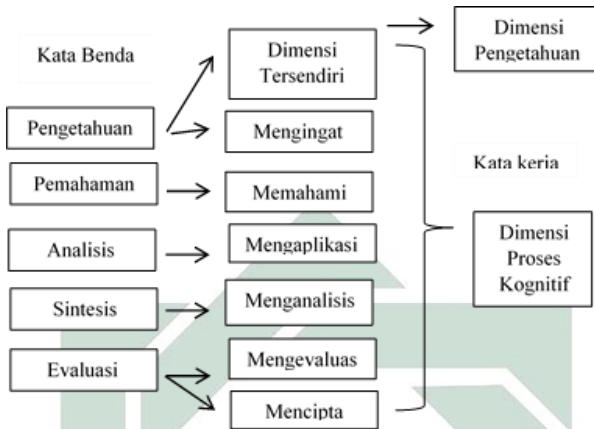
BAB II

KAJIAN PUSTAKA

A. Taksonomi Bloom Revisi

Taksonomi Bloom Revisi merupakan hasil revisi dari Anderson dan Krathwols. Prinsip dasar dari Bloom B.S adalah mengajukan suatu cara untuk mengelompokkan tujuan pendidikan dalam hal yang kompleks secara bertingkat yakni pengetahuan, pemahaman, penerapan, analisis, sintesis dan evaluasi. Sedangkan penyempurnaan taksonomi Bloom dikembangkan oleh Anderson dan Krathwols. Anderson dan Krathwohl mempertahankan ke enam proses kognitif dan melibatkan dua dimensi, dengan enam jenis proses kognitif dan empat jenis pengetahuan.

Domain kognitif ada perubahan kata kunci, dari kata benda menjadi kata kerja. Pada level kesatu yang semula "*knowledge*" berubah menjadi "*remember*" (mengingat). Pada level kedua, yaitu "*comprehension*" dipertegas menjadi "*understand*" (memahami). Level ketiga diubah dari "*application*" menjadi "*apply*" (menerapkan). Level keempat diubah sebutan dari "*analysis*" menjadi "*analyze*" (menganalisis). Perubahan mendasar terletak pada level kelima dan keenam. "*evaluation*" versi lama diubah posisinya dari level keenam menjadi level kelima, dan juga dengan perubahan sebutan dari "*evaluation*" menjadi "*evaluate*" (mengevaluasi). Level kelima lama, yaitu "*synthesis*" hilang, dinaikkan levelnya menjadi level keenam tetapi dengan perubahan mendasar, yaitu dengan nama "*create*" (mencipta). Berikut ini gambar sebagai ilustrasi Revisi Taksonomi Bloom:



Gambar 2.1
Perubahan Kerangka Asli ke Revisi Taksonomi oleh Anderson

Berikut ini merupakan penjelasan dari enam katagori proses kognitif:

1. Mengingat (C1)

Mengingat adalah mendapatkan kembali atau pengambilan pengetahuan relevan yang tersimpan dari memori jangka panjang¹. Dalam katagori mengingat terdapat dua proses. Pertama mengenali atau mengidentifikasi yaitu mengambil pengetahuan yang dibutuhkan dari memori jangka panjang untuk membandingkannya dengan informasi yang baru diterimanya. Dengan mengenali, peserta didik mencari dimemori jangka panjang suatu informasi yang mirip dengan informasi yang baru diterima. Proses kedua adalah mengingat kembali, yaitu mengambil kembali pengetahuan yang dibutuhkan dari memori jangka panjang. Dengan mengingat kembali, peserta didik membawa informasi dari memori jangka panjang dan memprosesnya.

¹ Wowo Sunaryo, *Taksonomi Kognitif*, (Bandung: PT Remaja Rosdakarya, 2012), 115.

2. Memahami (C2)

Memahami adalah mendeskripsikan susunan dalam artian pesan pembelajaran, mencakup oral, tulisan, dan komunikasi grafik.²Memahami juga dapat didefinisikan mengkonstruksi makna dari materi pembelajaran baik secara lisan, tulisan ataupun grafis yang disampaikan melalui pengajaran, buku atau layar komputer. Dalam kategori ini ada tujuh proses kognitif yaitu menafsirkan, mencontohkan, mengklasifikasikan, merangkum, menyimpulkan, membandingkan dan menjelaskan.

Pertama, menafsirkan terjadi ketika peserta didik mengubah informasi dari satu bentuk ke bentuk lainnya. Menafsirkan berupa perubahan kata-kata menjadi kata-kata lain, angka menjadi kata-kata, gambar menjadi angka, dan sebagainya.

Kedua, proses mencontohkan terjadi ketika peserta didik dapat memberikan contoh terhadap suatu konsep. Mencontohkan melibatkan proses identifikasi ciri-ciri pokok dari suatu konsep. Dalam pembelajaran, peserta didik diberi suatu konsep dan peserta didik diharuskan memberi contoh lainnya yang belum pernah dijumpai pada proses pembelajaran.

Ketiga, proses mengklarifikasikan terjadi ketika siswa mengetahui suatu informasi termasuk dalam kategori tertentu. Proses ini juga melibatkan proses identifikasi, mengenali ciri-ciri atau pola-pola terhadap suatu informasi. Mengklasifikasikan melengkapi proses mencontohkan. Mengklasifikasikan dimulai dari peserta didik mencontohkan suatu contoh, kemudian diklasifikasikan sesuai dengan pola-pola atau ciri-ciri suatu konsep.

Keempat, proses kognitif merangkum terjadi ketika peserta didik mengemukakan satu kalimat yang mempresentasikan informasi yang diterima. Pada pembelajaran, peserta didik disajikan suatu informasi kemudian mereka membuat rangkuman dari informasi tersebut.

Kelima, proses kognitif menyimpulkan terjadi ketika peserta didik dapat mengabstraksikan sebuah konsep dengan menerangkan contoh-contohnya dan mencermati ciri-cirinya.

² Ibid, halaman 115.

Proses menyimpulkan melibatkan proses kognitif membandingkan seluruh contohnya

Keenam, proses kognitif membandingkan melibatkan proses mendeteksi persamaan dan perbedaan antara dua atau lebih objek, peristiwa, ide, dan konsep. Membandingkan meliputi pencarian korespondensi satu-satu antara elemen-elemen suatu objek. Tujuan pembelajarannya, peserta didik diberikan informasi baru, mereka akan mendeteksi keteraitan pengetahuan yang sudah familier.

Ketujuh, proses menjelaskan ketika membuat dan menggunakan model sebab-akibat dalam suatu sistem. Nama lain menjelaskan adalah membuat model. Tujuan pembelajarannya, peserta didik diberi gambaran tentang sebuah sistem, peserta dapat menciptakan dan menggunakan model.

3. Mengaplikasi (C3)

Mengaplikasi adalah menggunakan prosedur dalam situasi yang dihadapi.³Dalam kategori ini terdapat dua proses kognitif yaitu mengeksekusi dan mengimplementasi

Pertama, mengeksekusi adalah menerapkan prosedur yang telah familier. Hal tersebut memberikan petunjuk yang cukup untuk memilih prosedur yang tepat dan menggunakannya. Soal yang telah familier adalah soallatihan yang sering dikerjakannya sehingga setelah membaca soal, peserta didik dapat menggunakan prosedur yang benar.

Kedua, mengimplementasikan berlangsung saat peserta didik menggunakan suatu prosedur untuk menyelesaikan tugas yang tidak familier. Karena tidak familier, peserta didik tidak segera mengetahui prosedur yang dilakukan.

4. Menganalisis (C4)

Kemampuan untuk memecah suatu kesatuan menjadi bagian-bagian dan menentukan bagaimana bagian-bagian tersebut dihubungkan antara satu dengan yang lain atau bagian tersebut dengan keseluruhannya. Hal tersebut menekankan pada kemampuan merinci sesuatu unsur pokok menjadi suatu bagian-bagian dan dapat melihat hubungan antar bagian

³ Ibid, halaman 115.

tersebut. Pada tingkat analisis, seseorang akan mampu menganalisa informasi yang masuk, membagi dalam bentuk yang lebih kecil untuk memahami pola atau hubungan serta dapat mengenali dan membedakan faktor-faktor penyebab dan akibatnya. Kategori menganalisa terdiri dari kemampuan membedakan, mengorganisasi, dan memberi simbol.

Pertama, membedakan meliputi proses memilih-milih bagian-bagian yang relevan dari sebuah struktur. Membedakan terjadi pada saat peserta didik mendeskripsikan informasi yang relevan. Membedakan melibatkan proses mengorganisasi secara struktural dan keseluruhannya.

Kedua, mengorganisasi meliputi kemampuan mengidentifikasi unsur-unsur secara bersama-sama menjadi struktur yang saling terkait. Dalam proses mengorganisasi peserta didik dapat membangun hubungan-hubungan dengan sistematis.

Ketiga, mengatribusikan adalah kemampuan siswa menyebutkan tentang sudut pandang, bias, nilai atau maksud dari suatu masalah yang diajukan. Mengatribusikan membutuhkan pengetahuan dasar yang lebih agar dapat menerka maksud dari inti permasalahan dari inti permasalahan yang diajukan.

5. Menilai atau mengevaluasi (C5)

Mengevaluasi didefinisikan sebagai membuat keputusan berdasarkan kriteria dan standar. Kategori dalam evaluasi mencakup *checking* dan *Critiquing*.

Pertama, memeriksa (*checking*) adalah kemampuan untuk mengetes konsistensi internal atau kesalahan pada operasi atau hasil serta mendeteksi keefektifan prosedur yang digunakan. Hal ini terjadi ketika peserta didik menguji apakah kesimpulan sesuai dengan premis-premisnya atau tidak.

Kedua, mengkritik (*critiquing*) adalah kemampuan memutuskan hasil atau operasi berdasarkan kriteria dan standar tertentu, mendeteksi apakah hasil yang diperoleh berdasarkan suatu prosedur menyelesaikan suatu masalah mendekati jawaban yang benar. Dalam mengkritik, peserta didik menilai ciri-ciri positif dan ciri-ciri negatif.

6. Mencipta (*Create*)

Mencipta merupakan menempatkan bagian-bagian secara bersama-sama ke dalam suatu ide, semuanya saling berhubungan untuk membuat hasil yang baik.⁴ Selain itu mencipta didefinisikan menggeneralisasikan ide baru atau cara pandang yang baru, dan produk baru. Siswa dapat dikatakan *create* bila dapat membuat produk baru dengan merombak beberapa bagian kedalam bentuk atau struktur yang belum pernah diterangkan pada Guru sebelumnya. Pada umumnya, proses *create* berhubungan dengan pengalaman belajar siswa sebelumnya. Proses *create* dapat dipecah menjadi tiga fase yaitu merumuskan, merencanakan dan memproduksi.

Pertama, merumuskan melibatkan proses menggambarkan masalah dan membuat pilihan yang memenuhi kriteria tertentu. Dalam pembelajarannya, peserta didik diberi deskripsi tentang suatu masalah dan diharuskan mencari beragam solusinya. Format asesmennya adalah soal yang membutuhkan jawaban singkat yang meminta peserta didik membuat hipotesis.

Kedua, merencanakan adalah mempraktikkan langkah-langkah untuk menciptakan solusi yang nyata bagi suatu masalah.⁵ Merencanakan melibatkan metode penyelesaian masalah yang sesuai dengan kriteria-kriteria masalah. Tujuan pembelajarannya, peserta didik diberikan soal kemudian peserta didik membuat rencana dalam menyelesaikan masalah. Format asesmennya adalah dengan soal yang meminta peserta didik mencari solusi yang reliastis dan mendeskripsikan rencana penyelesaiannya masalah dengan tepat.

Ketiga, memproduksi melibatkan proses melaksanakan rencana untuk menyelesaikan masalah. Nama lain memproduksi adalah mengontruksi. Dalam prosesnya peserta didik diberikan gambaran suatu produk dan harus menciptakan suatu produk sesuai dengan gambaran tersebut. Format asasmennya adalah soal tugas untuk merancang. Secara singkat dapat terurai dari tabel dibawah ini:

⁴Ibid, Halaman 115.

⁵Lorin W Anderson - David R Krathwohl (ed), Kerangka Landasan untuk Pembelajaran, Pengajaran dan Asesmen. (Yogyakarta:Pustaka pelajar, 2010), 131.

Tabel 2.1
Dimensi Proses Kognitif

Kategori dan Proses Kognitif	Nama-Nama Lain	Definisi
1. Mengingat Mengambil pengetahuan dari memori jangka panjang		
1.1 Mengenali	Meng-identifikasi	Menempatkan pengetahuan dalam memori jangka panjang yang sesuai dengan pengetahuan tersebut
1.2 Mengingat kembali	Mengambil	Mengambil pengetahuan yang relevan dari memori jangka panjang
2. Memahami Mengkonstruksi makna dari materi pembelajaran, termasuk apa yang diucapkan, ditulis, dan digambar oleh guru.		
2.1. Menafsirkan	Mengklarifikasi Memparafrasekan Merepresentasi Menerjemahkan	Mengubah satu bentuk gambaran (misalnya angka) jadi bentuk lain (misalnya kata-kata)
2.2. Mencontohkan	Mengilustrasikan Memberi Contoh	Menemukan contoh atau ilustrasi tentang konsep atau prinsip
2.3. Mengklasifikasikan	Mengategorikan Mengelompokkan	Menentukan sesuatu dalam satu kategori
2.4. Merangkum	Mengabstraksi Menggeneralisasi	Mengabstraksikan tema umum atau poin-poin pokok
2.5. Menyimpulkan	Menyarikan Mengekstrapolasi	Membuat kesimpulan yang logis dari informasi yang

	Menginterpolasi Memprediksi	diterima
2.6. Membandingkan	Mengontraskan Memetakan Mencocokkan	Menentukan hubungan antara dua ide, dua objek, dan semacamnya
2.7. Menjelaskan	Membuat model	Membuat model sebab-akibat dalam sebuah sistem
3. Mengaplikasikan Menerapkan atau menggunakan suatu prosedur dalam keadaan tertentu		
3.1. Mengeksekusi	Melaksanakan	Menerapkan suatu prosedur pada tugas yang familier
3.2. Mengimplementasi	Menggunakan	Menerapkan suatu prosedur pada tugas yang tidak familier
4. Menganalisis Memecah-mecah materi jadi bagian-bagian penyusunannya dan menentukan hubungan-hubungan antar bagian itu dan hubungan antar bagian-bagian tersebut dan keseluruhan struktur atau tujuan		
4.1. Membedakan	Menyendirikan Memilah Memfokuskan Memilih	Membedakan bagian materi pelajaran yang relevan dari yang tidak relevan bagian yang penting dari yang tidak penting,
4.2. Mengorganisasi	Menemukan koherensi Memadukan Membuat garis besar Mendeskrripsikan peran Menstrukturkan	Menentukan bagaimana elemen-elemen bekerja atau berfungsi dalam sebuah struktur

4.3. Mengatribusi	Mendekonstruksi	Menentukan sudut pandang, bias, nilai, atau maksud dibalik materi pelajaran
5. Mengevaluasi Mengambil keputusan berdasarkan kriteria atau standar		
5.1. Memeriksa	Mengkoordinasi Mendeteksi Memonitor Menguji	Menemukan inkonsistensi atau kesalahan dalam suatu proses maupun dalam suatu produk; menentukan apakah suatu proses atau produk memiliki konsistensi internal; menemukan efektivitas suatu prosedur yang dipraktikkan
5.2. Mengkritik	Menilai	Menemukan inkonsistensi antar suatu produk dan kriteria eksternal; menentukan apakah suatu proses atau produk memiliki konsistensi internal; menemukan ketepatan suatu prosedur untuk menyelesaikan masalah
6. Mencipta adalah Memadukan bagian-bagian untuk membentuk sesuatu yang baru dan koheren atau untuk membuat suatu produk yang orisinal		
6.1. Merumuskan	Membuat hipotesis	Membuat hipotesis-hipotesis berdasarkan kriteria

6.2. Merencanakan	Mendesain	Merencanakan prosedur untuk menyelesaikan suatu tugas
6.3. Memproduksi	Mengkonstruksi	Menciptakan suatu produk

B. Analisis Butir Soal

Analisis butir soal merupakan kegiatan penting dalam penyusunan soal agar diperoleh soal yang bermutu. Tujuan analisis butir soal menurut Aiken yang dikutip oleh Kusaeri dan Suprananto adalah (1) mengkaji dan menelaah setiap butir soal agar diperoleh soal yang bermutu sebelum digunakan, (2) meningkatkan kualitas butir soal tes melalui revisi atau membuang soal yang tidak efektif, serta (3) mengetahui informasi diagnostik pada siswa apakah mereka telah memahami materi yang telah diajarkan. Menurut Anastasi dan Urbina, analisis butir soal dapat dilakukan secara kualitatif (berkaitan dengan isi dan bentuknya) dan kuantitatif (berkaitan dengan ciri-ciri statistiknya).⁶

Dalam sebuah tes terdapat serangkaian pertanyaan yang harus diberikan tanggapan dengan tujuan untuk mengukur tingkat kemampuan seseorang atau aspek yang diukur. Menurut kamus besar Bahasa Indonesia (KBII) soal adalah apa yang yang menuntut jawaban, hal yang harus dipecahkan dan diselesaikan.⁷ Menurut Thomas Butt yang dikutip Sumardyono menyebutkan klasifikasi soal atau masalah sebagai berikut:⁸

1. Tipe soal ingatan (*recognition*)

Tipe ini biasanya meminta kepada siswa untuk mengenali atau menyebutkan fakta-fakta matematika, definisi, atau pernyataan suatu teorema/dalil. Bentuk soal yang dipakai biasanya bentuk soal benar-salah, pilihan ganda, mengisi

⁶ Kusaeri - Suprananto, *Pengukuran dan Penilaian Pendidikan*, (Yogyakarta:Graha Ilmu, 2012), 163.

⁷Pengertian soal diakses dari: <http://kbbi.web.id/soal> pengertian soal pada tanggal 1 Juni 2015.

⁸ Sumardyono, *Pengertian Dasar Problem Solving*, 2011, (online) (http://erlisilitonga.files.wordpress.com/2011/12/pengertiandasarproblemsolving_smd.pdf, diakses 11 Juni 2015).

yang kosong, atau dengan format menjodohkan. Contohnya meminta menyebutkan teorema Pythagoras.

2. Tipe soal prosedural atau algoritma (*algorithmic*)
Tipe ini menghendaki penyelesaian berupa sebuah prosedur langkah demi langkah, dan seringkali berupa algoritma hitung. Pada soal tipe ini, umumnya siswa hanya memasukkan angka atau bilangan ke dalam rumus, teorema, atau algoritma.
3. Tipe soal terapan (*application*)
Soal aplikasi memuat penggunaan algoritma dalam konteks yang sedikit berbeda. Soal-soal cerita tradisional umumnya termasuk kategori soal aplikasi, dimana penyelesaiannya memuat: (a) merumuskan masalah ke dalam model matematika, dan (b) memanipulasi simbol-simbol berdasarkan satu atau beberapa algoritma. Pada soal tipe ini umumnya siswa mudah mengenal rumus atau teorema yang harus dipergunakan. Satu-satunya keterampilan baru yang harus mereka kuasai adalah bagaimana memahami konteks masalah untuk merumuskannya secara matematis.
4. Tipe soal terbuka (*open search*)
Berbeda dengan tiga tipe soal sebelumnya, maka pada tipe soal terbuka ini strategi pemecahan masalah tidak tampak pada soal. Soal-soal tipe ini umumnya membutuhkan kemampuan melihat pola dan membuat dugaan. Termasuk pada tipe soal ini adalah soal-soal matematika yang berkaitan dengan teka-teki dan permainan.
5. Tipe soal situasi (*situation*)
Salah satu langkah krusial dalam tipe ini adalah mengidentifikasi masalah dalam situasi tersebut sehingga penyelesaian dapat dikembangkan untuk situasi tersebut. Pertanyaan-pertanyaan dalam soal ini antara lain: “Berikan masukan atau pendapat kamu!”, “Bagaimana seharusnya?”, “Apa yang mesti dilakukan?”. Dalam matematika, umumnya soal-soal tipe ini berkenaan dengan kegiatan mandiri atau soal proyek, di mana siswa dituntut untuk melakukan suatu percobaan, penggalan atau pengumpulan data, pemanfaatan sumber belajar baik berupa buku, media, maupun ahli (*expert*).

Bentuk bentuk soal meliputi:

1) Bentuk Soal Pilihan Ganda

Merupakan soal yang jawabannya harus dipilih dari beberapa pilihan jawaban yang disediakan. Pilihan jawaban terdiri dari kunci jawaban dan pengecoh jawaban. Soal pilihan ganda memiliki beberapa kelebihan yakni: (a) mampu mengukur berbagai tingkatan kognitif, (b) penskorannya mudah, cepat, objektif, dan dapat mencakup ruang lingkup bahan atau materi yang luas dalam suatu tes untuk suatu kelas atau jenjang pendidikan, dan (c) lebih tepat untuk ujian yang pesertanya sangat banyak atau massal, tetapi hasilnya harus segera diumumkan. Namun demikian itu, pilihan ganda juga memiliki kekurangan yakni: (a) memerlukan waktu yang relatif lama untuk menulis soalnya, (b) sulit membuat pengecoh yang homogen dan berfungsi dengan baik, dan (c) terdapat peluang untuk menebak jawaban.⁹

2) Bentuk Soal dengan Dua Pilihan Jawaban

Bentuk soal ini menuntut peserta tes untuk memilih dua kemungkinan jawaban. Bentuk kemungkinan jawaban yang digunakan ialah benar dan salah atau ya atau tidak. Keunggulan bentuk tes ini antara lain: (a) dapat mengukur berbagai jenjang kemampuan kognitif, (b) dapat mencakup materi yang luas, (c) dapat diskor dengan mudah, cepat, dan objektif. Sementara itu kelemahan dari bentuk tes ini antara lain: (a) probabilitas menebak dengan benar dan salah yaitu 50%, (b) bentuk soal ini tidak dapat digunakan untuk menanyakan suatu konsep secara utuh karena hanya dituntut menjawab benar atau salah, (c) apabila jumlah butir soal ini sedikit, indeks daya pembeda soal rendah, dan (d) apabila ragu terhadap pernyataan, maka peserta didik cenderung memilih jawaban benar.¹⁰

⁹Kusaeri - Suprananto, Op.Cit., hal 108.

¹⁰Ibid, halaman 123.

- 3) **Bentuk Soal Menjodohkan**
Bentuk soal ini terdiri dari dua kelompok yakni kelompok pertanyaan dan kelompok jawaban. Kelompok pertanyaan biasanya ditulis disebelah kiri sedangkan kelompok jawaban ditulis disebelah kanan. Keunggulan dari bentuk tes ini antara lain: (a) relatif lebih mudah dalam penulisan butir soal, (b) ringkas dan ekonomis dilihat dari segi rumusan butir soal dan dari segi pilihan jawaban, dan (c) dapat dilakukan penskoran dengan mudah, cepat, dan objektif. Kekurangan dari bentuk tes ini antara lain: (a) cenderung mengukur kemampuan mengingat sehingga kurang tepat untuk mengukur kemampuan kognitif yang lebih tinggi dan (b) kemungkinan menebak dengan benar relatif tinggi.¹¹
- 4) **Bentuk Soal Uraian**
Merupakan suatu soal yang menuntut peserta didik untuk mengingat dan mengkoordinasikan gagasan-gagasan atau hal-hal yang telah dipelajari dengan cara mengemukakan gagasan tersebut dalam bentuk tulisan. Kelebihan dari bentuk soal uraian adalah dapat mengukur kemampuan siswa dalam hal menyajikan jawaban terurai secara bebas, mengorganisasikan pikirannya, mengemukakan pendapatnya, dan mengekspresikan gagasan-gagasan dengan menggunakan kalimat peserta didik sendiri. Kelemahan dari bentuk tes ini adalah jumlah materi atau pokok bahasan relatif terbatas, waktu untuk memeriksa jawaban siswa cukup lama penyekorannya relatif subjektif terutama untuk soal uraian non objektif, dan tingkat reliabilitasnya lebih rendah dibandingkan dengan pilihan ganda.¹²
- 5) **Bentuk Isian**
Merupakan soal yang menuntut peserta didik memberikan jawaban singkat, berupa kata, frase, angka atau simbol. Keunggulan yang dimiliki oleh soal isian adalah lingkup materi yang banyak dan dapat diskor

¹¹Ibid, halaman128

¹²Ibid, halaman 137

dengan mudah, cepat dan objektif serta mudah menyusunnya. Kelemahan dari bentuk soal isian ini adalah cenderung mengukur kemampuan mengingat (*simple recall*).¹³

C. Validitas Soal

Validitas adalah suatu konsep yang berkaitan dengan sejauh mana tes telah mengukur apa yang seharusnya diukur. Menurut Anastasi yang dikutip oleh Sumarna validitas adalah suatu tingkatan yang menyatakan bahwa suatu alat ukur telah sesuai dengan alat ukur.¹⁴ Menurut pendapat R.L Thorndike dan H.P. Hagen yang dikutip oleh Zainal Arifin mengatakan bahwa "*Validity is always in relation to a specific decision or use*".¹⁵ Terdapat berbagai macam bentuk pendekatan dalam mengklasifikasikan validitas tes. Pengklasifikasian pertama, validitas dibedakan atas: validitas isi (*content validity*), validitas kriteria (*criteria validity*), dan validitas konstruk (*construct-validity*). Kemudian pengklasifikasian berikutnya menganggap validitas sebagai tiga hal yang berkaitan. Sehingga validitas terbagi atas : validitas terkait isi (*content-related validity*), validitas terkait kriteria (*criterion-related validity*), dan validitas konstruk (*construct-related validity*). Pengklasifikasian yang ketiga, validitas dianggap sebagai konstruk tunggal. Kelompok ini terbagi atas: Bukti validitas yang didasarkan pada isi tes (*validity evidence based on test content*), Bukti validitas yang didasarkan pada keterkaitan variabel lain (*validity evidence based on relations to other variables*), dan Bukti validitas yang didasarkan pada struktur internal (*validity evidence based on internal structure*). Bukti validitas yang didasarkan pada proses menjawab (*validity evidence based on response process*), dan bukti validitas yang didasarkan pada akibat dari pengujian (*validity evidence based on consequences of testing*).¹⁶

¹³Ibid, halaman 145.

¹⁴ Sumarna Suprananta, *Analisis Validasi Reliabilitas dan interpretasi hasil Tes*, (Bandung: PT RemajaRosda Karya, 2012), 50.

¹⁵ Zainal Arifin, *Evaluasi Pembelajaran*, (Bandung: PT Remaja Rosda Karya, 2009), 247.

¹⁶ Kusaeri - Suprananto, Op.Cit., hal 77

Tabel 2.2
Deskripsi dari validitas Tes

Jenis	Maksud	Prosedur
Validitas terkait isi	Bagaimana kemampuan tes merepresantasikan domain yang hendak diukur dengan baik	Membandingkan tes dengan kisi-kisi tes
Validitas terkait kriteria	Bagaimana kemampuan tes memprediksi kemampuan peserta tes dimasa mendatang dengan baik	Membandingkan skor tes dengan skor yang didapatkan pada tes lain dimasa yang akan datang (untuk memprediksi) atau dengan skor pada tes lain sekaligus
Validitas terkait konstruk	Bagaimana kemampuan tes dapat diinterpretasikan sebagai suatu ukuran bermakna dari beberapa karakteristik	Mengkaji teori-teori yang berkaitan dengan konstruk yang relevan diukur oleh tes yang dikembangkan. Berdasarkan telaah teori-teori tersebut diturunkanlah butir-butir tes

1. Validitas Terkait Isi (*Content-Related Validity*)

Validitas isi ialah derajat dimana sebuah tes evaluasi mengukur cakupan substansi yang ingin diukur. Validitas ini mempersoalkan apakah isi butir tes yang diujikan mencerminkan isi kurikulum atau tidak. Cara menguji validitas isi dengan menggunakan pendekatan rasional, yaitu membandingkan antara kisi-kisi soal dimuat data tentang pokok bahasan dan sub pokok bahasan serta aspek kepribadian yang diukur. Upaya mengembangkan kisi-kisi yang cermat sangatlah penting sehingga cakupan isi benar-benar terwujud.

2. Validitas Terkait Kriteria (*Criterion-Related Validity*)

Validitas terkait kriteria adalah pengukuran dengan membandingkan skor tes anak yang dapat memprediksi kemampuan dimasa mendatang dengan hasil dari pengukuran alat ukur lainnya. Pengukuran pada kemampuan kedua (disebut kriteria) mungkin diperoleh dimasa yang akan datang. Contohnya adalah ketika Guru ingin menentukan apakah skor tes potensi akademik (TPA) mampu memprediksi keberhasilan anak pada pelajaran matematika. Tes TPA telah dilaksanakan pada waktu masuk SMP. Untuk kriteria keberhasilan matematika dengan menggunakan tes sumatif yang dilaksanakan pada akhir semester. Dengan demikian hal ini dapat memungkinkan untuk menentukan seberapa baik skor tes potensi akademik memprediksi kemampuan anak di bidang matematika. Caranya adalah membandingkan skor TPA dengan skor Sumatif matematika.

3. Validitas Terkait Konstruk (*Construct-Related Validity*)

Validitas konstruk adalah suatu tes dimana butir soal tersebut membangun setiap aspek berfikir seperti yang disebutkan dalam tujuan instruksional khusus. Konstruksi adalah sesuatu yang berkaitan dengan fenomena dan objek yang abstrak, tetapi gejalanya dapat diamati dan diukur. Validitas Konstruk berarti bahwa suatu alat ukur dikatakan valid apabila telah cocok dengan konstruksi teoritik tes tersebut. Secara umum, proses yang digunakan dalam validitas ini mencakup: (1) mengidentifikasi dan mendeskripsikan (melalui kerangka teoritik) makna konstruk yang diukur, (2) menyusun dugaan (hipotesis) dengan mengacu pada teori yang mendasari konstruk dan (3) menguji kebenaran dugaan secara logis dan empiris.¹⁷ Penjabarannya prosedur menguji validitas terkait konstruks: 1) Mendefinisikan cakupan materi yang diukur. Kisi kisi harus terdefinisi dengan baik sehingga makna konstruk jelas. Sehingga dapat memungkinkan untuk menilai jangkauan tes dan keterwakilan domain yang diukur, 2) Menganalisis proses mental (konstruk) yang mendasari dan diperlukan oleh butir-

¹⁷ Kusaeri - Suprananto, Op.Cit., hal 81.

butir tes, 3) Membandingkannya dengan skor kelompok yang telah diketahui dan 4) Membandingkan skor sebelum dan sesudah diberi beberapa perlakuan. Skor tes terkadang cenderung tetap dan meningkat setelah diberi perlakuan. Berdasarkan teori yang mendasari kemampuan diukur, dapat diprediksi bahwa skor tes tertentu berubah dalam berbagai kondisi.

Menurut Noeng Muhajir yang dikutip oleh Chabib Toha, untuk menguji validitas konstruksi digunakan: (1) pengujian konvergen, (2) pengujian diskriminan, (3) pengujian stabilitas dan keajegan.¹⁸ Uji validitas konvergen adalah cara uji empirik, yaitu mengkorelasikan skor total dengan skor faktor dengan asumsi korelasinya signifikan. Uji validitas diskriminan dengan mengkorelasikan skor faktor yang satu dengan skor faktor yang lain, dengan asumsi setiap faktor tidak berkorelasi secara signifikan. Menguji stabilitas dan keajegan dengan cara tes-retest, uji konsistensi belah dua sedangkan stabilitasnya dilakukan dengan cara membandingkan hasil tes dua kelompok. Sementara itu pendapat menurut Gronlund yang dikutip oleh Zainal Arifin mengemukakan ada tiga faktor yang mempengaruhi validitas hasil tes, yaitu : (1) Faktor instrumen evaluasi, (2) faktor administrasi evaluasi dan penskoran, dan (3) faktor dari jawaban peserta didik.¹⁹

Menurut Kerlinger yang dikutip oleh Zainal Arifin menyatakan bahwa validitas instrumen tidak cukup ditentukan oleh derajat ketepatan instrumen untuk mengukur apa yang seharusnya diukur, tetapi perlu juga dilihat dari tiga kriteria yaitu *appropriatness*, *meaningfulness*, dan *usefulness*.²⁰ *Appropriatness* menunjukkan kelayakan dari tes sebagai suatu instrumen yaitu seberapa besar tes tersebut dapat menjangkau keragaman aspek perilaku peserta didik. *Meaningfulness* menunjukkan kemampuan tes untuk memberikan keseimbangan soal-soal pengukurannya berdasarkan tingkat kepentingan setiap fenomena. Dan *usefulness* menunjukkan sensitif atau tidaknya instrumen dalam menangkap fenomena perilaku dan tingkat ketelitian yang ditunjukkan dalam

¹⁸ Chabib Toha, Teknik Evaluasi pendidikan, (Jakarta: PT Rineka Cipta, 2003), 111.

¹⁹ Zainal Arifin, Evaluasi Pembelajaran, (Bandung: PT Remaja Rosdakarya, 2009), 247.

²⁰ Ibid, Halaman 248.

membuat kesimpulan. Validitas suatu tes dinyatakan dengan angka korelasi koefisien (r). Kriteria korelasi koefisien adalah sebagai berikut:

Tabel 2.3
Kriteria Koefisien Korelasi

Range korelasi validitas	Kategori
$0,00 \leq r \leq 0,20$	Sangat rendah (hampir tidak ada korelasi)
$0,21 \leq r \leq 0,40$	Korelasi rendah
$0,41 \leq r \leq 0,70$	Korelasi cukup
$0,71 \leq r \leq 0,90$	Korelasi tinggi
$0,91 \leq r \leq 1,00$	Korelasi sangat tinggi.

Cara menghitung validitas suatu tes dapat dilakukan dengan koefisien korelasi *point biserial*

$$r_{pbi} = \frac{M_p - Mt}{SD} \sqrt{\frac{p}{q}}$$

Keterangan :

r_{pbi} = koefisien korelasi biserial

M_p = rata-rata skor pada tes dari peserta tes yang memiliki jawaban benar

M_t = rata-rata skor total

SD = standar deviasi skor total

p = proporsi peserta tes yang jawabannya benar pada soal

q = $1 - p$

Lalu interpretasi nilai r_{pbi} maka kita gunakan tabel nilai r Product Moment dengan terlebih dahulu mencari df -nya yakni ($df = N - nr$).

D. Reliabilitas Soal

Reliabilitas menunjukkan derajat konsistensi hasil pengukuran. Reliabilitas bertujuan untuk mengetahui tingkat ketepatan dan keajegan skor tes. Menurut Anastasi yang dikutip oleh Zainal Arifin menyatakan bahwa "*reliability refers to the consistency of scores obtained by the sampersons when reexamined*

*the same test on different occasion ,or with different sets of equivalent items or under other variable examining conditions.*²¹

Menurut Kerlinger yang dikutip oleh Zainal Arifin menyatakan bahwa reliabilitas dapat diukur menjadi tiga kriteria yaitu *stability*, *dependability*, dan *predictability*.²² *Stability* menunjukkan keajegan suatu tes untuk mengukur gejala yang sama pada waktu yang berbeda. *Depandability* menunjukkan kemantapan tes. *Predictability* menunjukkan kemampuan tes yang untuk meramalkan hasil pengukuran gejala selanjutnya. Indeksnya berkisar antara 0 sampai 1. Bila reliabilitasnya mendekati 1 maka semakin tinggi keajekannya.

Beberapa karakteristik reliabilitas adalah: pertama, reliabilitas merujuk pada hasil yang didapat melalui sebuah instrumen tes, bukan merujuk kepada instrumennya sendiri. Sebagaimana yang telah disampaikan oleh Gronlund yaitu *Reliability refers to the result obtained with an evaluation instrument and not to the instrument it self*. Kedua, reliabilitas merupakan syarat perlu, tetapi belum cukup syarat validitas. Ketiga, reliabilitas utamanya berkaitan dengan statistik.

Faktor yang mempengaruhi reliabilitas skor tes antara lain: (a) semakin banyak jumlah butir soal, semakin ajek suatu tes, (b) semakin lama waktu tes, semakin ajek (c) semakin sempit range kesukaran butir soal, semakin besar keajekan, (d) soal-soal yang saling berhubungan akan mengurangi keajekan, dan (e) semakin objektif dalam pemberian skor, semakin besar keajekan.²³

Metode yang digunakan untuk mencari reliabilitas antara lain: (1) Tes-Retes, (2) Bentuk ekuivalen, (3) Tes Retes dengan Bentuk Ekuivalen, (4) Belah Dua (*Split Half*), (5) Koefisien *Alpha* atau *Kuder Richardson*, dan (6) Inter-rater. Adapun penjabarannya sebagai berikut:

1) Motode *Test-Retes Reliability*

Metode ini dalam menguji reliabilitas ini adalah dengan cara memberikan tes kepada sekelompok siswa dalam dua kesempatan yang berlainan (waktu yang berbeda). Skor pada tes pertama dikorelasikan dengan skor tes kedua. Besar

²¹ Ibid.,halaman 258.

²² Zainal Arifin,Op.cit., halaman 258.

²³ Kusaeri - Suprananto,Op,Cit., halaman 177.

kecilnya korelasi menunjukkan reliabilitasnya. Tujuan uji reliabilitas ini untuk mengetahui koefisien stabilitas tes. Adapun langkah-langkahnya dalam menguji reliabilitas ini sebagai berikut: (a) Menyusun sebuah tes yang akan diukur reliabilitasnya, (b) Mengujikan tes yang tersusun, (c) Menghitung skor hasil tes tahap I, (d) Mengujikan ulang tes yang tersusun (tahap II), (e) Menghitung skor tes tahap II, (f) Menghitung reliabilitas tes dengan mengkorelasikan skor tes I dan skor tes ke dengan korelasi Product Moment. Pendekatan Tes-Retes memiliki keterbatasan yakni pengaruh yang dibawa pada tes pertama dan kedua. Anak biasanya mengingat soal yang diujikan pertama kali sehingga dapat meningkatkan skor tes pada saat mengerjakan tes kedua.

2) Metode Ekuivalen

Metode ini menggunakan dua buah tes yang sejenis atau memiliki kesamaan mengenai isi, tingkat kesukaran, tingkat mental yang diukur, jumlah item dan aspek aspek yang lain. Yang berbeda adalah soalnya (soal tidak sama persis). Skor dari tes pertama akan dikorelasikan dengan skor tes yang kedua. Adapun langkah yang ditempuh adalah: (a) Menyusun dua buah tes yang ekuivalen (tes A dan tes B), (b) Menguji kedua tes tersebut (dalam waktu yang bersamaan), (c) Memberikan skor pada tes A dan tes B, dan (d) Mencari korelasi skor tes A dan skor tes B dengan menggunakan korelasi *Product Moment*.

Untuk menyusun dua buah tes yang sama maka yang harus diperhatikan antara lain: 1) materi yang ditanyakan sama, 2) bentuk soal sama, 3) tingkat kesukaran sama, 4) jumlah soal sama, 5) waktu mengerjakan soal sama, dan 6) sistem evaluasinya sama.

3) Tes Retes dengan Bentuk ekuivalen

Metode ini dilakukan dengan cara memberikan dua bentuk tes yang paralel pada sekelompok siswa yang sama dengan jeda waktu yang agak lama antar kedua jenis tes tersebut. Dengan cara ini, hasil koefisien reliabilitas akan memberikan suatu ukuran kemantapan dan ekuivalensi. Metode ini memiliki keunggulan yakni mengurangi pengaruh ingatan

yang dibawa anak pada saat mereka mengerjakan soal yang pertama. Namun ingatan tes pada tes yang kedua tidak dapat dihilangkan. Kelemahan metode ini adalah hanya sedikit tes (baik tes standar ataupun tes yang dibuat guru) yang memiliki ekuivalen. Mengembangkan tes yang ekuivalen memerlukan waktu yang lama.

4) Metode *Split Half Reliability*

Metode ini menggunakan satu tes yang diberikan kepada peserta tes. Lalu tes tersebut dibelah menjadi dua bagian. Teknik belah dua ini dapat dilakukan dengan dua cara. Cara pertama pembelahan soal nomor ganjil dan soal nomor genap (metode ganjil-genap). Sehingga nomor ganjil menjadi satu kelompok dan nomor genap menjadi satu kelompok. Cara kedua adalah pembelahan menurut nomor urut (metode awal-akhir). Maksudnya adalah separuh atau setengah item nomor soal berada dikelompok atas dan separuh item nomor soal berada dikelompok akhir. Adapun langkah-langkah yang ditempuh adalah: (a) Membelah skor tes menjadi dua bagian, (b) Menguji tes, (c) Menghitung skor ganjil dan skor genap, (d) Menghitung koefisien korelasi setengah tes dengan korelasi Product Moment atau mencari deviasi pada belahan ganjil genap, dan (e) Mencari reliabilitas satu tes penuh dengan rumus *Spearman Brown*, Rumus flagan dan Rumus Rulon. Penjabarannya adalah sebagai berikut:

a. Reliabilitas dengan menggunakan rumus *Spearman Brown*

Cara cara yang ditempuh antara lain: 1) Membelah skor tes kedalam skor ganjil dan genap, 2) Skor ganjil menjadi variabel X dan skor Genap menjadi variabel Y, 3) Menghitung koefisien korelasi setengah tes dengan menggunakan korelasi *Product Moment*, 4) Menghitung Korelasi satu tes penuh dengan rumus *Spearman Brown*²⁴, dan 5) Setelah diketahui koefisien korelasi satu tes penuh, dilanjutkan dengan tes signifikasi tabel nilai *r Product Moment*.

b. Mencari Reliabilitas Tes dengan rumus *Flanagan*

²⁴M Ghabib Thoha, *teknik evaluasi pendidikan*(Jakarta: pt raja Grafindo Persada.2003) halaman 125.

Adapun langkah – langkahnya adalah :1) Membelah hasil tes menjadi dua, yaitu belahan ganjil dan belahan genap, 2) Mencari SD^2 pada belahan ganjil dan SD^2 pada belahan genap, dan 3) Menghitung reliabilitas tes dengan rumus *Flanagan*²⁵

c. Mencari Reliabilitas dengan rumus *Rulon*

Dalam rumus *Rulon* menggunakan kuadrat dari deviasi nilai ganjil dan nilai genap dan standar deviasi kuadrat dari skor total. Adapun langkah langkahnya adalah: 1) Membelah tes menjadi dua yaitu belahan genap serta skor total, 2) Mencari deviasi antara skor ganjil dan skor genap, 3) Mencari standar deviasi kuadrat dari deviasi nilai tersebut dan standar deviasi kuadrat dari skor total, 4) Menghitung besarnya menggunakan rumus *Rulon*²⁶

5) Konsistensi Internal

Aspek ini menunjuk pada ketetapan dalam penyelenggaraan suatu tes tertentu menghadapi semua item. Pendekatan ini tidak membelah tes menjadi dua. Hal ini disebabkan oleh dua kemungkinan: 1) jumlah item ganjil, sehingga tidak bisa dibelah dua, 2) komposisi antara item-item ganjil dan genap tidak homogen, sehingga bila dibelah cenderung tidak memiliki korelasi yang positif. Maka itu digunakan rumus:

a. Rumus K-K.R.20

Penemu rumus ini adalah Kuder dan Richardson. Adapun cara menggunakan rumus K-K.R.20 adalah sebagai berikut: 1) Membuat tabel analisis butir tanpa harus dikelompokkan nomor ganjil dan genap, 2) Menghitung proporsi yang menjawab benar dan proporsi yang menjawab salah pada masing masing butir soal, 3) Mengalikan proporsi yang menjawab benar dengan proporsi menjawab salah, 4) Mencari varians dari skor total, dan 5) Menghitung reliabilitas tes dengan menggunakan rumus K-K.R.20.²⁷

²⁵Ibid, halaman 129.

²⁶Ibid, halaman 131.

²⁷Ibid, halaman 134.

b) Rumus Alpha

Menurut Cronbach, rumus *Alpha* dapat digunakan skala *Likert* (skala sikap), tes yang menggunakan bentuk esai sehingga pengukurannya tidak hanya skor benar = 1 dan skor salah = 0. Cara yang ditempuh adalah dengan mencari varians pada tiap butir pada skor totalnya, baru dikalikan dengan proporsi item. Dari segi ketelitian, rumus ini mirip dengan rumus K-R.20. Karena rumus ini banyak menggunakan perhitungan untuk mencari varians masing-masing item.²⁸

6) Metode Inter-Rater

Pada metode ini dilaksanakan satu kali pada sejumlah peserta tes dengan menggunakan dua orang rater. Agar tidak saling mempengaruhi maka masing-masing rater bekerja secara terpisah. Pada penskoran terhadap instrumen non objektif perlu dihitung tingkat persetujuan masing-masing rater. Pada kondisi tertentu, reliabilitas inter-rater tidak terlalu penting pada penskoran soal pilihan ganda atau benar-salah. Disisi lain, hal ini menjadi penting ketika reliabilitas inter-rater rendah maka peserta tes dapat merasakan adanya perbedaan atau kesenjangan yang cukup tinggi dari skor yang diperoleh dari dua penyekor yang berbeda. Secara singkat dapat dilihat pada table 2.4.

²⁸Ibid, halaman 138.

Tabel 2.4
Metode untuk Mengestimasi Reliabilitas Skor Tes

Metode	Tujuan	Banyaknya Format Tes	Prosedur
Tes-Retes	Mengukur Kestabilan	Satu format tes (Sekali diujikan)	Memberikan tes yang sama dua kali pada kelompok yang sama dengan jeda waktu tertentu antar dua tes. Jeda antara kedua penyelenggara tes perlu dipertimbangkan
Bentuk Ekuivalen	Mengukur ekuivalensi	Dua format tes (Sekali diujikan)	Memberikan dua bentuk tes berbeda, tetapi paralel pada sekelompok siswa yang sama, dengan waktu yang sama pula
Tes Retes dengan Bentuk ekuivalen	Mengukur kestabilan dan ekuivalensi	Dua format tes (Dua kali diujikan)	Memberikan dua bentuk tes yang paralel pada sekelompok siswa yang sama dengan jeda waktu yang agak lama antar kedua jenis tes tersebut.
Belah Dua (<i>Split Half</i>)	Mengukur konsistensi internal	Satu format tes (Sekali diujikan)	Memberikan tes sekali. Tes dibagi menjadi dua (misalkan butir tes ganjil dan genap), skor dari masing-masing paruhan dikorekasikan dengan rumus <i>Spearman Brown</i> .
Koefisien Alpha atau <i>Kuder Richardson</i>	Mengukur konsistensi internal	Satu format tes (Sekali diujikan)	Memberikan tes sekali. Skor total tes dikorekasikan dengan rumus <i>Kuder-Richardson</i>
Inter-rater	Mengukur konsistensi penyekoran, bila sebuah tes diskor dua orang atau lebih	Satu format tes (Sekali diujikan)	Menyelenggarakan tes pada sekelompok siswa dalam sekali waktu. Dua atau lebih rater atau penilai menyekor tes secara independen.

E. Daya pembeda

Daya pembeda soal adalah kemampuan suatu butir soal dapat membedakan antara siswa yang telah menguasai materi dan siswa yang belum menguasai materi yang diujikan. Manfaat dari daya pembeda soal yaitu meningkatkan mutu suatu soal dengan data empirik. Sehingga soal tersebut dapat diterima, direvisi dan ditolak. Manfaat lainnya adalah untuk mengetahui sejauh mana setiap butir soal dapat mendeteksi dan membedakan kemampuan siswa. Indeks daya pembeda dinyatakan dalam bentuk proporsi. Semakin tinggi proporsi maka semakin tinggi kemampuan soal tersebut dapat membedakan kemampuan peserta didik. Indeks daya beda berkisar anatar -1 sampai 1. Apabila daya beda negatif maka butir soal sangat jelek. Untuk mencari daya pembeda dalam soal pilihan ganda dapat menggunakan persamaan:

$$D = \frac{BA-BB}{\frac{1}{2}N} \text{ atau } D = \frac{2(BA-BB)}{N}$$

Dengan keterangan:

D = Daya Pembeda

BA = Jumlah jawaban benar pada kelompok atas

BB = Jumlah jawaban benar pada kelompok bawah

N = Jumlah siswa yang mengerjakan tes

Untuk mengetahui daya pembeda soal bentuk uraian maka dapat menggunakan rumus:

$$DP = \frac{(\text{Mean kelompok skor atas} - \text{Mean kelompok bawah})}{(\text{skor maksimum soal})}$$

Rumus tersebut menggambarkan tingkat kemampuan soal dalam membedakan antar peserta tes yang telah memahami materi yang diujikan dan peserta tes yang belum memahami materi yang diujikan. Adapun kriteria yang digunakan menurut Cracker & Algina yang dikutip oleh Kusaeri²⁹

²⁹Kusaeri - Suprananto, Op.Cit., halaman 177.

Tabel 2.5
Kriteria Daya pembeda

Range Daya pembeda	Katagori	Keputusan
$0,40 \leq DP \leq 1,00$	Sangat Memuaskan	Diterima
$0,30 \leq DP \leq 0,39$	Memuaskan	Diterima
$0,20 \leq DP \leq 0,29$	Tidak memuaskan	Ditolak/direvisi
$0,00 \leq DP \leq 0,19$	Sangat Tidak memuaskan	Direvisi Total

F. Tingkat Kesukaran

Tingkat kesukaran soal adalah peluang menjawab benar suatu soal pada tingkat kemampuan tertentu yang biasanya dinyatakan dalam bentuk indeks. Indeks umumnya dinyatakan dalam bentuk proporsi yang besarnya berkisar dari 0 sampai 1. Tingkat kesukaran memiliki dua kegunaan yaitu untuk guru dan untuk pengujian dan pengajaran. Kegunaan bagi guru yaitu: (a) sebagai pengenalan konsep terhadap pembelajaran ulang dan memberi masukan kepada siswa tentang hasil belajar mereka, (b) memperoleh informasi tentang penekanan kurikulum, atau mencurigai terhadap butir soal bias. Kegunaan bagi pengujian dan pengajaran yaitu: (a) pengenalan konsep yang diperlukan untuk diajarkan ulang, (b) tanda-tanda terhadap kelebihan dan kelemahan pada kurikulum sekolah, (c) memberi masukan kepada siswa, (d) tanda-tanda kemungkinan adanya butir soal yang bias dan (e) merakit tes yang memiliki ketepatan data soal.³⁰ Persamaan untuk mengukur tingkat kesukaran dengan proporsi menjawab benar adalah:³¹

$$p = \frac{\sum x}{N}$$

Keterangan :

- P = proporsi menjawab benar atau tingkat kesukaran
- $\sum x$ = Banyaknya peserta tes yang menjawab benar
- N = Jumlah Peserta tes

Untuk soal esai dalam mencari tingkat kesukaran soal adalah sebagai berikut:

³⁰M Ghabib Toha, Op.Cit, halaman 175

³¹M Ghabib Toha, Op.Cit, halaman 12

$$\text{Mean} = \frac{(\text{Jumlah Skor siswa peserta tes pada suatu soal})}{(\text{jumlah peserta didik yang mengikuti tes})}$$

$$\text{Tingkat Kesukaran} = \frac{\text{Mean}}{\text{Skor Maksimum yang ditetapkan}}$$

Tabel 2.6
Kriteria Tingkat Kesukaran Soal

Range Tingkat Kesukaran	Katagori	Keputusan
$0,71 \leq TK \leq 1,00$	Mudah	Ditolak / direvisi
$0,31 \leq TK \leq 0,70$	Sedang	Diterima
$0,0 \leq TK \leq 0,30$	Sulit	Ditolak /direvisi

