



Dr. Rahmi, M.A.
Prof. Dr. Martin Kustati, M.Pd.
Dr. Hadel, M.A., M.Pd.

EVALUASI

PENDIDIKAN

PERSPEKTIF ISLAM

**EVALUASI PENDIDIKAN
PERSPEKTIF ISLAM**

UU No 28 tahun 2014 tentang Hak Cipta

Fungsi dan sifat hak cipta Pasal 4

Hak Cipta sebagaimana dimaksud dalam Pasal 3 huruf a merupakan hak eksklusif yang terdiri atas hak moral dan hak ekonomi.

Pembatasan Pelindungan Pasal 26

Ketentuan sebagaimana dimaksud dalam Pasal 23, Pasal 24, dan Pasal 25 tidak berlaku terhadap:

- i. Penggunaan kutipan singkat Ciptaan dan/atau produk Hak Terkait untuk pelaporan peristiwa aktual yang ditujukan hanya untuk keperluan penyediaan informasi aktual;
- ii. Penggandaan Ciptaan dan/atau produk Hak Terkait hanya untuk kepentingan penelitian ilmu pengetahuan;
- iii. Penggandaan Ciptaan dan/atau produk Hak Terkait hanya untuk keperluan pengajaran, kecuali pertunjukan dan Fonogram yang telah dilakukan Pengumuman sebagai bahan ajar; dan
- iv. Penggunaan untuk kepentingan pendidikan dan pengembangan ilmu pengetahuan yang memungkinkan suatu Ciptaan dan/atau produk Hak Terkait dapat digunakan tanpa izin Pelaku Pertunjukan, Produser Fonogram, atau Lembaga Penyiaran.

Sanksi Pelanggaran Pasal 113

1. Setiap Orang yang dengan tanpa hak melakukan pelanggaran hak ekonomi sebagaimana dimaksud dalam Pasal 9 ayat (1) huruf i untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama 1 (satu) tahun dan/atau pidana denda paling banyak Rp100.000.000 (seratus juta rupiah).
2. Setiap Orang yang dengan tanpa hak dan/atau tanpa izin Pencipta atau pemegang Hak Cipta melakukan pelanggaran hak ekonomi Pencipta sebagaimana dimaksud dalam Pasal 9 ayat (1) huruf c, huruf d, huruf f, dan/atau huruf h untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama 3 (tiga) tahun dan/atau pidana denda paling banyak Rp500.000.000,00 (lima ratus juta rupiah).

EVALUASI PENDIDIKAN PERSPEKTIF ISLAM

Dr. Rahmi, M.A.

Prof. Dr. Martin Kustati, M.Pd.

Dr. Hadeli, M.A., M.Pd.



EVALUASI PENDIDIKAN PERSPEKTIF ISLAM

**Rahmi
Martin Kustati
Hadeli**

Desain Cover :
Ali Hasan Zein

Sumber :
www.shutterstock.com

Tata Letak :
Amira Dzatin Nabila

Proofreader :
Annisa Nurisnaini K.P.

Ukuran :
viii, 146 hlm, Uk: 15.5x23 cm

ISBN :
978-623-02-4381-3

Cetakan Pertama :
Maret 2022

Hak Cipta 2022, Pada Penulis

Isi diluar tanggung jawab percetakan

Copyright © 2022 by Deepublish Publisher
All Right Reserved

Hak cipta dilindungi undang-undang
Dilarang keras menerjemahkan, memfotokopi, atau
memperbanyak sebagian atau seluruh isi buku ini
tanpa izin tertulis dari Penerbit.

PENERBIT DEEPUBLISH
(Grup Penerbitan CV BUDI UTAMA)

Anggota IKAPI (076/DIY/2012)

Jl.Rajawali, G. Elang 6, No 3, Drono, Sardonoharjo, Ngaglik, Sleman

Jl.Kaliurang Km.9,3 – Yogyakarta 55581

Telp/Faks: (0274) 4533427

Website: www.deepublish.co.id

www.penerbitdeepublish.com

E-mail: cs@deepublish.co.id

KATA PENGANTAR PENERBIT

Membaca adalah sarana ekspresi diri dalam berkomunitas serta untuk terus maju menuju pencerdasan dan pencerahan. Ini menjadi sebuah motivasi dan dorongan bagi kami di Penerbit Deepublish untuk ikut berikhtiar dalam mencerdaskan dan memuliakan umat manusia, serta memanfaatkan ilmu pengetahuan dan teknologi untuk menciptakan industri *processing* berbasis sumber daya alam (SDA) Indonesia. Berdasarkan pandangan, sikap dasar, tujuan itu, maka buku yang berjudul Evaluasi Pendidikan Perspektif Islam ini diterbitkan.

Buku yang berjudul Evaluasi Pendidikan Perspektif Islam, berisi/membahas evaluasi yang merupakan bagian yang sangat penting dalam komponen pendidikan, karena untuk melihat mutu pendidikan suatu bangsa dapat dilihat dari hasil evaluasinya. Evaluasi dan komponen pendidikan lainnya sangat berpengaruh dalam kondisi politik suatu negara dan juga berpengaruh terhadap tujuan suatu negara.

Jika pembelajaran merupakan proses transfer ilmu dari pendidik kepada peserta didik yang bertujuan adanya perubahan dan pendewasaan dari diri peserta didik, maka evaluasi merupakan komponen yang akan menjadi penambah dan pengubah perilaku peserta didik.

Buku ini terdiri atas bab: Konsep Evaluasi Pendidikan Perspektif Islam, Teknik Penyusunan dan Pelaksanaan Hasil Belajar, Teknik Penganalisisan Item Tes Hasil Belajar, Teknik Tes dan Teknik Nontes sebagai Alat Evaluasi Hasil Belajar, Teknik Pengujian Validitas Tes dan Validitas Item Hasil Belajar, dan Contoh-Contoh Pengolahan Evaluasi Tes dan Nontes.

Kami sadar masih terdapat berbagai kekurangan dalam buku ini. Namun, kami mencoba untuk terus mengembangkan diri, dan mencoba memperkecil kesalahan-kesalahan.

Kami mengucapkan terima kasih kepada penulis buku Rahmi, Martin Kustati & Hadeli yang telah memberikan perhatian, kepercayaan, dan kontribusi demi kesempurnaan buku ini. Dan kepada pihak-pihak lainnya yang terus menjadi inspirasi dan memberikan semangat dalam menerbitkan buku yang berkualitas dan bermanfaat.

Dengan dukungan dari pembaca, kami dapat terus memberikan kontribusi bagi upaya mencerdaskan dan memuliakan umat manusia, serta memanfaatkan ilmu pengetahuan dan teknologi. Semoga buku ini dapat memperkaya khazanah dan memberi manfaat bagi para pembaca.

Hormat Kami,

Penerbit Deepublish

DAFTAR ISI

KATA PENGANTAR PENERBIT	v
DAFTAR ISI.....	vii
BAB I KONSEP EVALUASI PENDIDIKAN	
PERSPEKTIF ISLAM.....	1
A. Pengertian Evaluasi.....	1
B. Term-Term atau Istilah Evaluasi di dalam Al- Qur'an.....	5
C. Evaluasi dalam Pendidikan	8
D. Tujuan Evaluasi Pendidikan.....	14
E. Sasaran Evaluasi	18
F. Prinsip-Prinsip Evaluasi	20
G. Unsur-Unsur Evaluasi Pendidikan	22
H. Evaluasi yang Dilakukan Rasulullah saw.....	23
I. Syarat-Syarat Evaluasi Pendidikan Islam.....	33
J. Kedudukan Evaluasi Pendidikan Islam	34
BAB II TEKNIK PENYUSUNAN DAN PELAKSANAAN	
HASIL BELAJAR.....	40
A. Pengertian Tes Hasil Belajar	40
B. Ciri-Ciri Tes Hasil yang Baik.....	44
C. Prinsip-Prinsip Dasar dalam Penyusunan Tes Hasil Belajar.....	46
D. Bentuk-Bentuk Tes Hasil Belajar dan Teknik Penyusunannya	47
E. Teknik Pelaksanaan Tes Belajar.....	54
BAB III TEKNIK PENGANALISISAN ITEM TES HASIL	
BELAJAR	59
A. Pengantar	59
B. Teknik Menganalisis <i>Item</i> Tes Hasil Belajar.....	60
C. Soal untuk Melatih dan Penugasannya Struktur di Luar Kelas.....	71

BAB IV	TEKNIK TES DAN TEKNIK NONTES SEBAGAI ALAT EVALUASI HASIL BELAJAR.....	77
	A. Teknis Tes	77
	B. Teknik Nontes	81
BAB V	TEKNIK PENGUJIAN VALIDITAS TES DAN VALIDITAS ITEM TES HASIL BELAJAR	92
	A. Pengertian Validitas.....	92
	B. Teknik-Teknik Pengujian Validitas Tes Hasil Belajar	95
	C. Teknik Pengujian Validitas <i>Item</i> Tes Hasil Belajar	99
	D. Faktor-Faktor yang Memengaruhi Validitas Tes	100
BAB VI	CONTOH-CONTOH PENGOLAHAN EVALUASI TES DAN NONTES	102
	A. Evaluasi	102
	B. Tes	109
	C. Nontes.....	123
	DAFTAR PUSTAKA	142
	TENTANG PENULIS	144

BAB I

KONSEP EVALUASI PENDIDIKAN PERSPEKTIF ISLAM

PENDAHULUAN

Evaluasi merupakan bagian yang sangat penting dalam komponen pendidikan, karena untuk melihat mutu pendidikan suatu bangsa dapat dilihat dari hasil evaluasinya. Evaluasi dan komponen pendidikan lainnya sangat berpengaruh dalam kondisi politik suatu negara dan juga berpengaruh terhadap tujuan suatu negara.

Jika pembelajaran merupakan proses transfer ilmu dari pendidik kepada peserta didik yang bertujuan adanya perubahan dan pendewasaan dari diri peserta didik, maka evaluasi merupakan komponen yang akan menjadi penambah dan pengubah perilaku peserta didik.

A. Pengertian Evaluasi

1. Evaluasi menurut etimologi

Evaluasi berasal dari kata “*to evaluate*” yang berarti “menilai”. Istilah nilai (*value/al qimat*) pada mulanya dipopulerkan oleh Filosof dan Platola yang mula-mula mengemukakannya.¹

Evaluasi secara bahasa berasal dari bahasa Inggris *evaluation* yang berarti penilaian.² Sedangkan *evaluation* (penilaian) berasal dari kata dasar

¹ Ensiclopedia Americana (New York, 1997, vol. 27: 86)

value (nilai) yang berarti nilai atau harga.³ Penilaian berarti sebuah proses yang dilakukan untuk melihat sejauh mana keberhasilan telah dicapai, sehingga dapat ditentukan nilai pencapaian dari proses tersebut.

Nilai dalam bahasa Arab lebih dikenal dengan *Al-Qiimah* (الْقِيَمَةُ) dan *at-Taqdiir* (التَّقْدِيرُ) yang berarti penilaian, secara harfiah evaluasi pendidikan dapat dikatakan *At-Taqdiir At-tarbiyah* (التقدير التربوية) yang berarti penilaian dalam bidang pendidikan atau penilaian terhadap hal-hal yang berkaitan dengan pendidikan.⁴ Dengan demikian, penilaian dapat terjadi dalam dunia pendidikan dan non pendidikan, bahkan dapat dilakukan di lembaga-lembaga tertentu yang tidak berkaitan dengan pendidikan. Pada dasarnya evaluasi merupakan proses untuk mengetahui nilai yang telah dicapai dan kekurangan yang perlu untuk diperbaharui agar lebih baik dari sebelumnya.

2. Evaluasi Menurut Terminologi

Evaluasi pendidikan merupakan proses kegiatan untuk penentuan majunya suatu pendidikan dibandingkan dengan tujuan yang sudah ditentukan dan juga usaha yang diperoleh dari informasi berupa ucapan baik bagi kesempurnaan pendidikan.⁵

Evaluasi secara terminologi atau secara istilah dapat dilihat dari beberapa pengertian, yaitu:

a) Evaluasi dalam arti pengukuran

Pengukuran adalah proses menentukan suatu gejala menurut aturan tertentu.⁶ Pengukuran dapat menggunakan tes atau nontes, bisa bersifat kuantitatif atau kualitatif. Kuantitatif hasilnya berupa angka, sedangkan kualitatif hasilnya bukan angka (berupa predikat,

² Anas Sudiono, *Pengantar Evaluasi Pendidikan*, (Jakarta: RajaGrafindo Persada, 2006), hlm. 1

³ Ramayulis, *Metodologi Pendidikan Agama Islam*, (Jakarta: Kalam Mulia, 2013), hlm. 398

⁴ *Ibid*, hlm. 398

⁵ Anas Sudijono, *Pengantar Evaluasi Pendidikan*, (Jakarta: PT RajaGrafindo Persada, 2008) hlm. 2

⁶ Guilford, *Evaluasi Hasil Belajar*, (Jakarta: Depdiknas, Dirjen PD dan PM, Direktorat Pembinaan SMA, 2008), hlm. 4

pernyataan kualitatif, misalnya sangat baik, baik, cukup, kurang, sangat kurang), disertai deskripsi penjelasan prestasi peserta didik.⁷ Pengukuran ini merupakan proses memperoleh gambaran berupa angka dan ciri tindakan yang dimiliki individu, kegiatan ini adalah membandingkan sesuatu dengan sesuatu yang lain.⁸ Dengan demikian pengukuran lebih dipahami dengan penilaian yang bersifat kuantitatif.

b) Evaluasi dalam arti penilaian

Penilaian adalah proses mengumpulkan, menganalisis, dan menginterpretasikan informasi guna menetapkan keluasan pencapaian tujuan oleh individu.⁹ Menurut Mardapi penilaian adalah kegiatan menafsirkan atau mendeskripsikan hasil pengukuran.¹⁰ Hasil pengukuran yang bersifat kuantitatif dari pengukuran kemudian ditafsirkan dalam bentuk nilai.

Evaluasi merupakan suatu kegiatan yang harus dilakukan oleh seorang guru dalam kegiatan pembelajaran. Sebab dengan begitu guru akan mengetahui perkembangan proses dan hasil belajar, bakat khusus, minat, hubungan sosial, sikap dan kepribadian peserta didik.

Jika kata evaluasi dikaitkan dengan kata pendidikan, maka dapat diartikan sebagai proses membandingkan situasi yang ada dengan kriteria tertentu terhadap masalah-masalah yang berkaitan dengan pendidikan. Untuk evaluasi tidak hanya menilai tentang hasil belajar para siswa dalam suatu jenjang pendidikan tertentu. Melainkan berkenaan dengan berbagai aspek yang memengaruhi proses belajar siswa tersebut, seperti evaluasi terhadap guru, kurikulum, metode, sarana dan prasarana, lingkungan dan sebagainya.¹¹

Di samping itu evaluasi juga diartikan sebagai pengumpulan kenyataan secara sistematis untuk menetapkan apakah dalam

⁷ Ramayulis, *Op. cit.*, hlm. 399

⁸ Suharsimi Arikunto, *Dasar-dasar Evaluasi Pendidikan*, (Jakarta: Bumi Aksara, 2008), hlm. 2

⁹ Ramayulis, *Metodologi Pendidikan Agama Islam*, (Jakarta: Kalam Mulia, 2005), hlm. 332

¹⁰ Djemari Mardapi, *Teknik Penyusunan Instrumen Test dan Non-test*, (Yogyakarta: Mitra Cendekia Press, 2008), hlm. 6

¹¹ Abuddin Nata, *Filsafat Pendidikan Islam*, (Jakarta: Logos, 1997), hlm. 131

kenyataannya terjadi perubahan dalam diri siswa dan menetapkan sejauh mana tingkat perubahan dalam diri siswa.¹² Slameto juga mendefinisikan evaluasi sebagai suatu proses usaha yang dilakukan individu untuk memperoleh suatu perubahan tingkah laku yang baru secara keseluruhan, sebagai hasil pengalaman individu dalam interaksinya dengan lingkungan.¹³

Sedangkan dalam “Standar Penilaian Pendidikan”, kata-kata “evaluasi” disebut dengan “penilaian” yang merupakan salah satu standar pendidikan nasional. Standar penilaian pendidikan adalah “standar nasional pendidikan yang berkaitan dengan mekanisme, prosedur, dan instrumen penilaian hasil belajar peserta didik.” Sedangkan penilaian pendidikan adalah proses pengumpulan dan pengolahan informasi untuk menentukan pencapaian hasil belajar peserta didik.¹⁴

Berdasarkan uraian di atas evaluasi pembelajaran atau pendidikan adalah suatu kegiatan yang berisi mengadakan pengukuran dan penilaian terhadap keberhasilan pendidikan dari berbagai aspek yang berkaitan dengan kegiatan pendidikan.¹⁵ Dimaksudkan untuk melihat perubahan-perubahan yang terjadi pada diri anak didik, baik berupa penambahan ilmu pengetahuan atau informasi (kognitif) maupun berupa peningkatan aspek tingkah laku atau sikap (afektif) dan keterampilan (psikomotorik) mengerjakan sesuatu sebagai konsekuensi dari aktivitas belajar yang telah dilakukannya. Sehingga perubahan-perubahan tersebut pada gilirannya memberikan dampak positif terhadap pribadi siswa itu sendiri bagi kebahagiaan dan kesejahteraan hidupnya sekarang dan masa yang akan datang.

Pelaksanaan evaluasi pendidikan Islam ada dua cara yang dapat ditempuh di antaranya: *pertama*, evaluasi kuantitatif, yaitu cara untuk mengetahui sebuah hasil pendidikan dengan cara memberikan penilaian dalam bentuk angka. *Kedua*, evaluasi kualitatif yaitu suatu

¹² Daryanto, *Evaluasi Pendidikan*, (Jakarta: Rineka Cipta, 1999), hlm. 1

¹³ Slameto, *Belajar dan Faktor-faktor yang Memengaruhinya*, (Jakarta: Rineka Cipta, 1988), hlm. 2

¹⁴ Permendiknas RI No. 20 Tahun 2007 tentang Standar Penilaian Pendidikan

¹⁵ Abuddin Nata, *Op. cit*, hlm. 133-134

cara untuk mengetahui hasil pendidikan yang diberikan dengan cara memberikan pernyataan verbal dan sejenisnya.

B. Term-Term atau Istilah Evaluasi di dalam Al-Qur'an

1. Term Soal

وَقِفُّهُمْ أَيْهُمْ مَسْئُولُونَ ﴿٣٧﴾

“dan tahanlah mereka (di tempat perhentian) karena Sesungguhnya mereka akan ditanya...” (QS. Ash-Shaffat [37]: 24)¹⁶

Pada dua ayat sebelumnya Allah menggambarkan kondisi orang-orang kafir di hari pengadilan nanti. Pada saat itu mereka akan dikumpulkan bersama dengan teman sejawatnya, kemudian mereka akan ditunjukkan jalan menuju neraka. Setelah itu Allah akan mengajukan pertanyaan berkaitan dengan kondisi mereka yang tidak lagi melakukan tolong-menolong dengan sesama kawannya.

Dari sini terlihat bahwa relasi antara ayat ini dengan ayat sebelumnya dan sesudahnya adalah bahwa walaupun mereka di akhirat kelak berada dalam satu kelompok yang sama—biasanya di kehidupan dunia mereka senantiasa saling menolong antara sesamanya—maka pada saat itu mereka tidak lagi menghiraukan teman sejawatnya. Bahkan mereka akan sibuk dengan pertanyaan-pertanyaan yang diajukan pada dirinya.

Dengan demikian, evaluasi akhir yang diajukan Allah kepada manusia harus dijawab sendiri, karena memang pada saat itu tidak seseorang pun mungkin meminta bantuan kepada orang lain. Kenyataan ini sangat bertolak belakang dengan kehidupan dunia yang terkadang pada saat-saat yang sangat menentukan ini masih saja banyak orang yang memberikan bantuan untuk menjawab soal-soal yang diajukan, sehingga dapat dikatakan bahwa kualitas nilai yang dihasilkannya pun masih dipertanyakan.¹⁷

¹⁶ QS. Ash-Shaffat [37]: 24

¹⁷ Nurwadjah Ahmad, *Tafsir Ayat-ayat Pendidikan*, (Bandung: Marja, 2007) hlm. 97-98

2. Term Al Hisab

Memiliki makna mengira, menafsirkan, menghitung dan menganggap.

لَلَّهِ مَا فِي السَّمٰوٰتِ وَمَا فِي الْاَرْضِ ۗ وَاِنْ تُبَدُّواْ مَا فِىْ اَنْفُسِكُمْ اَوْ تَخَفُوْهُ يُحَاسِبْكُمْ
بِهٖ ۗ اَللّٰهُ ۙ فَيَغْفِرُ لِمَنْ يَّشَآءُ وَيُعَذِّبُ مَنْ يَّشَآءُ ۗ وَاللّٰهُ عَلٰى كُلِّ شَيْءٍ قَدِيْرٌ ﴿٢٨٤﴾

“Kepunyaan Allah-lah segala apa yang ada di langit dan apa yang ada di bumi. Dan jika kamu melahirkan apa yang ada di dalam hatimu atau kamu menyembunyikan, niscaya Allah akan membuat perhitungan dengan kamu tentang perbuatanmu itu. Maka Allah mengampuni siapa yang dikehendaki-Nya dan menyiksa siapa yang dikehendaki-Nya; dan Allah Maha Kuasa atas segala sesuatu” (QS. Al-Baqarah [2]: 284).

3. Term Al Bala'u, memiliki makna coba ujian, sebagaimana terdapat dalam surah Al-Mulk [67]: 2.

الَّذِىْ خَلَقَ الْمَوْتَ وَالْحَيٰوةَ لِيَبْلُوَكُمْ اَيْكُمْ اَحْسَنُ عَمَلًا ۗ وَهُوَ الْعَزِيْزُ الْغَفُوْرُ ﴿٢﴾

“Yang menjadikan mati dan hidup, supaya Dia menguji kamu, siapa di antara kamu yang lebih baik amalnya. dan Dia Maha Perkasa lagi Maha Pengampun” (QS. Al-Mulk [67]: 2).

4. Term Al Hukm, memiliki makna putusan atau vonis, sebagaimana terdapat dalam surah An-Naml [27]: 78.

اِنَّ رَبَّكَ يَقْضِىْ بَيْنَهُمْ بِحُكْمِهٖ ۗ وَهُوَ الْعَزِيْزُ الْعَلِيْمُ ﴿٧٨﴾

“Sesungguhnya Tuhanmu akan menyelesaikan perkara antara mereka dengan keputusan-Nya, dan Dia Maha Perkasa lagi Maha mengetahui.” (QS. An-Naml [27]: 78).

5. *Term Al Qadha*, memiliki arti putusan sebagaimana terdapat dalam surah Thaha ayat [20]: 72.

فَالْوَالُونَ أَنْ يُكْفَرُوا أَوْ يَكْفُرُوا بِالَّذِينَ كَفَرُوا عَلَيْهِمْ قَوْلًا مَّا أُنزِلَتْ فِي ذَٰلِكَ مِنْ آيَاتِنَا ۗ إِنَّمَا تَقْضَىٰ هَذِهِ الْحَيَاةُ الدُّنْيَا ﴿٧٢﴾

“Mereka berkata: “Kami sekali-kali tidak akan mengutamakan kamu daripada bukti-bukti yang nyata (mukjizat), yang telah datang kepada Kami dan daripada Tuhan yang telah menciptakan kami; Maka putuskanlah apa yang hendak kamu putuskan. Sesungguhnya kamu hanya akan dapat memutuskan pada kehidupan di dunia ini saja” (QS. Thaha [20]:72).

6. *Term Al Nazhir*, memiliki makna melihat, sebagaimana terdapat dalam surah An-Naml [27]: 27.

﴿٢٧﴾ قَالَ سَنَنْظُرُ أَصَدَقْتَ أَمْ كُنْتَ مِنَ الْكَاذِبِينَ

“Berkata Sulaiman: “Akan Kami lihat, apa kamu benar, atukah kamu Termasuk orang-orang yang berdusta” (QS. An-Naml [27]: 27).

7. *At-Taqdiir*, memiliki makna menetapkan, mengukur, mengira terdapat dalam surah Yasin ayat 38.

﴿٣٨﴾ وَالشَّمْسُ تَجْرِي لِمُسْتَقَرٍّ لَهَا ۚ ذَٰلِكَ تَقْدِيرُ الْعَزِيزِ الْعَلِيمِ

“Dan matahari berjalan ditempat peredarannya. Demikianlah ketetapan yang Maha Perkasa lagi Maha Mengetahui” (QS. Yasin: 38).

Beberapa ayat ini pada dasarnya tidak memiliki padanan yang pasti terkait makna evaluasi, namun demikian makna dari kalimat-kalimat tertentu yang terdapat dalam ayat tersebut mengarah pada makna evaluasi. Beberapa makna kalimat dalam ayat tersebut dapat terlihat kesepadannya dengan beberapa pengertian evaluasi secara terminologi berikut ini.

C. Evaluasi dalam Pendidikan

Benyamin S. Bloom dan kawan-kawannya berpendapat bahwa taksonomi (pengelompokan) tujuan pendidikan harus senantiasa mengacu kepada tiga jenis domain (daerah binaan atau ranah) yang melekat pada diri peserta didik, yaitu (1) ranah proses berpikir (*cognitive domain*), (2) ranah nilai atau sikap (*affective domain*), dan (3) ranah keterampilan (*psychomotor domain*). Dalam konteks evaluasi hasil belajar, ketiga domain itu harus dijadikan sasaran dalam kegiatan evaluasi hasil belajar. Adapun sasaran tersebut, yaitu sebagai berikut.

1. Apakah peserta didik sudah dapat memahami semua materi pelajaran yang telah diberikan kepada mereka?
2. Apakah peserta didik sudah dapat menghayatinya?
3. Apakah materi pelajaran yang telah diberikan sudah dapat diamalkan secara konkret dalam kehidupan sehari-hari?¹⁸

1. Evaluasi Ranah Kognitif (*An-Nahiyah Al-Fikriyah*)

Ranah kognitif adalah ranah yang mencakup kegiatan mental (otak). Menurut Bloom, segala sesuatu yang menyangkut aktivitas otak termasuk ke dalam ranah ini. Dalam ranah kognitif terdapat enam jenjang proses berpikir, mulai dari jenjang terendah sampai jenjang yang tertinggi. Keenam jenjang dimaksud adalah (1) pengetahuan, hafalan, ingatan (*knowledge*); (2) pemahaman (*comprehension*); (3) penerapan (*application*), (4) analisis (*analysis*), (5) sintesis (*synthesis*); dan (6) penilaian (*evaluation*). Sehubungan dengan evaluasi ranah kognitif ini, ditemukan hadis berikut.

عَنْ مُعَاذِ بْنِ جَبَلٍ أَنَّ رَسُولَ اللَّهِ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ لَمَّا آرَادَ أَنْ يَبْعَثَ مُعَاذًا إِلَى
الْيَمَنِ قَالَ كَيْفَ تَقْضِي إِذَا عَرَدَ لَكَ قَضَاءٌ قَالَ أَقْضِي بِكِتَابِ اللَّهِ. قَالَ فَإِنْ لَمْ تَجِدْ فِي
كِتَابِ اللَّهِ قَالَ فَبِسُنَّةِ رَسُولِ اللَّهِ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ قَالَ فَإِنْ لَمْ تَجِدْ فِي سُنَّةِ رَسُولِ اللَّهِ

¹⁸ Bukhari Umar, Hadis Tarbawi, (Jakarta: Sinar Grafika Offset, 2012) hlm. 192-208

صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ وَلَا فِي كِتَابِ اللَّهِ قَالَ أَجْتَهُدُ رَأْيِي وَلَا آلُو. فَضَرَبَ رَسُولُ اللَّهِ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ صَدْرَهُ وَقَالَ الْهَمْدُ لِلَّهِ الَّذِي وَفَّقَ رَسُولَ رَسُولِ اللَّهِ لِمَا يُرْضَى رَسُولًا لِلَّهِ

Mu'adz bin Jabal meriwayatkan bahwa Rasulullah ketika akan mengutus Mu'adz ke Yaman, beliau bertanya kepadanya, "Bagaimana kamu mengadili perkara, jika dihadapkan kepadamu suatu perkara pengadilan?" Mu'adz menjawab, "Saya mengadili (perkara itu) dengan Kitab Allah Swt. (Al-Qur'an)." Rasulullah bertanya lagi, "Bagaimana jika kamu tidak menjumpai (petunjuk) dalam kitab Allah?" Mu'adz menjawab, "Saya mengadili dengan sunah Rasulullah saw." Rasulullah bertanya lagi, "Bagaimana jika kamu tidak menjumpai petunjuk dalam sunah Rasulullah dan tidak menjumpainya dalam kitab Allah?" Mu'adz menjawab, "Saya berijtihad sekuat akal pikiran saya." Rasulullah menepuk dada Mu'adz sambil bersabda, "Segala puji milik Allah yang telah memberi petunjuk kepada utusan Rasulullah terhadap apa yang Rasulullah berkenan terhadapnya (HR. Abu Dawud, At-Tarmidzi, Ahmad, dan Ad-Darimi).¹⁹

Di antara informasi yang terkandung dalam hadis di atas adalah (1) Rasulullah saw., bermaksud mengutus Mu'adz ke Yaman (untuk memimpin umat); (2) beliau bertanya kepada Mu'adz tentang dasar yang digunakan dalam memutuskan perkara peradilan; (3) Mu'adz menjawab dengan urutan; *pertama* dengan kitab Allah, *kedua* dengan sunah Rasulullah, dan *ketiga* dengan ijtihad; serta (4) setelah jawaban Mu'adz selesai, beliau menepuk dada Mu'adz karena senang lalu memuji Allah Swt.

Dalam hadis di atas terlihat bahwa beliau menguji kemampuan dan pengetahuan seorang sahabat sebelum memberikan tugas kepadanya. Setelah ia berhasil menjawab secara benar sesuai dengan keinginan, beliau memperlihatkan rasa senangnya dengan memberikan ganjaran yang menyenangkan dan memuji Allah Swt. Pujian kepada-Nya di sini dapat diartikan sebagai rasa syukur atas keberhasilan dalam mendidik sahabat.

Ujian yang diberikan oleh Rasulullah dalam hadis di atas berkaitan dengan tugas yang akan diemban oleh Mu'adz. Beliau baru akan

¹⁹ Bukhari Umar, Hadis Tarbawi (Jakarta: Sinar Grafika Offset, 2012) 192-208

menyerahkan suatu tugas kepada sahabat apabila ia menguasai (memiliki ilmu) tentang persoalan tugas yang akan diembannya.

2. Evaluasi Ranah Afektif (*An-Nahiyah Al-Mauqifiyah*)

Ranah afektif adalah ranah yang berkaitan dengan sikap dan nilai. Beberapa pakar mengatakan bahwa sikap seseorang dapat diramalkan perubahannya apabila seseorang tersebut telah memiliki penguasaan kognitif tingkat tinggi. Ciri-ciri hasil belajar afektif akan tampak pada peserta didik dalam berbagai tingkah laku, seperti perhatiannya terhadap pelajaran agama Islam, kedisiplinannya dalam mengikuti pelajaran agama di sekolah, motivasinya yang tinggi untuk tahu lebih banyak mengenai pelajaran agama yang diterimanya, dan penghargaan atau rasa hormatnya terhadap guru agama. Sehubungan dengan ranah ini, ditemukan hadis sebagaimana yang tertera sebagai berikut.

عَنْ جُبَيْرٍ قَالَ قَالَ رَسُولُ اللَّهِ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ إِنَّ اللَّهَ عَزَّ وَجَلَّ يَبْتَلِي عَبْدَهُ بِالسَّقَمِ هَتَّى يُكْفَرَ عَنْهُ كُلَّ ذَنْبِهِ.

Jubair berkata, “Rasulullah saw. bersabda, “Sesungguhnya Allah ‘azza wa jalla menguji seorang hamba-Nya dengan suatu penyakit hingga Dia mengampuni semua dosanya,” (HR. Ath-Thabrani).²⁰

أَبِي هُرَيْرَةَ عَنِ النَّبِيِّ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ قَالَ مَا يُصِيبُ الْمُسْلِمَ مِنْ نَصَبٍ وَلَا وَصَبٍ وَلَا هَمٍّ وَلَا هُزْنٍ وَلَا أَدَى وَلَا غَمٍّ هَتَّى الشُّوْكَةِ يُشَاكُهَا إِلَّا كَفَّرَ اللَّهُ بِهَا مِنْ خَطَايَاهُ

Dari Abu Hurairah dari Nabi Muhammad saw., beliau bersabda, “Setiap musibah yang menimpa seorang muslim yang berupa kelelahan, penyakit kronis, kegalauan pikiran, kegelisahan hati, sampai kena duri; akan dihapus Allah kesalahannya” (HR. Al-Bukhari).²¹

Semua materi ujian dalam hadis ini berada di wilayah domain afektif, yaitu kesabaran. Apabila seorang muslim mampu menerima ujian tersebut dengan penuh kesabaran, maka Allah Swt. akan menghapus

²⁰ *Ibid*

²¹ *Ibid*

kesalahan-kesalahan yang telah dilakukannya. Ini merupakan hadiah dari Allah untuk hamba-Nya yang lulus. Dalam hadis ini disebutkan bahwa manusia akan diuji oleh Allah dengan penyakit. Sasarannya adalah kesabaran yang termasuk domain afektif. Selain itu, dalam hadis ini disebutkan ganjaran yang akan diberikan oleh Allah kepada manusia yang lulus dalam ujian kesabaran menghadapi penyakit yang dideritanya.

3. Evaluasi Ranah Psikomotor (*An-Nahiyah Al-Harakah*)

Ranah psikomotor adalah ranah yang berkaitan dengan keterampilan (*skill*) atau kemampuan bertindak setelah seseorang menerima pengalaman belajar tertentu. Hasil belajar ranah psikomotor dikemukakan oleh Simpson (1956) yang menyatakan bahwa hasil belajar psikomotor ini tampak dalam bentuk keterampilan dan kemampuan bertindak individu. Hasil belajar psikomotor ini sebenarnya merupakan kelanjutan dan hasil belajar kognitif (memahami sesuatu) dan hasil belajar afektif (yang baru tampak dalam bentuk kecenderungan-kecenderungan untuk berperilaku). Sehubungan dengan ini ditemukan hadis berikut:

عَنْ أَبِي هُرَيْرَةَ أَنَّ رَسُولَ اللَّهِ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ دَخَلَ الْمَسْجِدِ فَدَجَلَ رَجُلٌ فَصَلَّى
فَسَلَّمَ عَلَى النَّبِيِّ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ فَرَدَّ وَقَالَ ارْجِعْ فَصَلِّ فَإِنَّكَ لَمْ تُصَلِّ فَرَجَعَ يُصَلِّي
كَمَا صَلَّى ثُمَّ جَاءَ فَسَلَّمَ عَلَى النَّبِيِّ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ فَقَالَ ارْجِعْ فَصَلِّ فَإِنَّكَ لَمْ تُصَلِّ
ثَلَاثًا فَقَالَ الَّذِي بَعَثَكَ بِحَقِّ مَا أَحْسِنُ غَيْرُهُ فَعَلَّمَنِي فَقَالَ إِذَا قُمْتَ إِلَى الصَّلَاةِ فَكَبِّرْ ثُمَّ
اقْرَأْ مَا تيسَّرُ مَعَكَ مِنَ الْقُرْآنِ ثُمَّ ارْكَعْ حَتَّى تَظْمِنَ رَاكِعًا ثُمَّ ارْفَعْ حَتَّى تَعْدِلَ قَائِمًا ثُمَّ
اسْجُدْ حَتَّى تَظْمِنَ سَاجِدًا ثُمَّ ارْفَعْ حَتَّى تَظْمِنَ جَالِسًا وَافْعَلْ ذَلِكَ فِي صَلَاتِكَ كُلِّهَا.

Dari Abu Hurairah, bahwasanya Rasulullah saw. masuk masjid lalu masuk pula seorang laki-laki yang kemudian salat dan memberikan salam kepada nabi saw. Beliau menjawab salam dan berkata, "Ulangi salatmu karena sesungguhnya kamu belum salat." Laki-laki itu mengulangi salatunya seperti salatunya tadi. Kemudian ia datang dan mengucapkan salam kepada nabi saw. Beliau berkata lagi, "Ulangi salatmu karena sesungguhnya kamu belum salat." Laki-laki itu kembali salat seperti

salatnya tadi. Setelah itu, ia kembali dan mengucapkan salam kepada nabi. Kemudian nabi berkata lagi, "Ulangi salatmu karena sesungguhnya kamu belum salat." Begitulah sampai tiga kali. Lalu laki-laki itu berkata, "Demi Zat yang telah mengutusmu dengan benar, sungguh aku tidak dapat berbuat yang lebih baik lagi daripada itu. Oleh karena itu, ajarilah aku." Lalu nabi bersabda, "Apabila kamu berdiri untuk salat, maka takbirlah. Lalu bacalah ayat yang mudah bagimu, kemudian rukuklah hingga tumakninah. Kemudian bangkitlah sehingga iktidal dalam keadaan berdiri, kemudian sujudlah sehingga tumakninah dalam keadaan sujud. Kemudian bangkitlah hingga tumakninah dalam keadaan sujud. Kemudian bangkitlah hingga tumakninah dalam keadaan duduk, kemudian sujudlah sehingga tumakninah dalam keadaan sujud. Kemudian berbuatlah yang demikian itu dalam semua salatmu" (HR. Al-Bukhari).²²

Dalam hadis ini, Rasulullah saw. menguji sahabat dalam mendirikan salat. Ini berada di wilayah psikomotor. Teknik yang digunakan observasi. Beliau mengamati salat yang dilakukan oleh sahabat. Setelah melihat adanya kekeliruan, beliau langsung menyuruhnya untuk mengulanginya. Jadi, segera ada perbaikan setelah terjadinya kesalahan. Menurut Anas Sudijono, secara umum, pengertian observasi adalah cara menghimpun bahan-bahan keterangan (data) yang dilakukan dengan mengadakan pengamatan dan pencatatan secara sistematis terhadap fenomena-fenomena yang sedang dijadikan sasaran pengamatan. Observasi sebagai alat evaluasi banyak digunakan untuk menilai tingkah laku individu atau proses terjadinya suatu kegiatan yang dapat diamati, baik dalam situasi yang sebenarnya maupun dalam situasi bantuan. Observasi dapat mengukur atau menilai hasil dan proses belajar, misalnya tingkah laku peserta didik ketika guru agama menyampaikan pelajaran di kelas; tingkah laku peserta didik ketika istirahat atau pada saat terjadinya kekosongan pelajaran; dan perilaku peserta didik pada saat salat jemaah di musala sekolah, ceramah-ceramah keagamaan, upacara bendera, serta ibadah salat tarawih.

²² *Ibid*

4. Kualitas Ujian Sesuai dengan Tingkat Keberagamaan
Berkaitan dengan hal ini ditemukan hadis sebagai berikut.

عَنْ سَعْدِ قَالَ قُلْتُ يَا رَسُولَ اللَّهِ أَيُّ النَّاسِ أَشَدُّ بَلَاءً قَالَ الْأَنْبِيَاءُ ثُمَّ الْأَمْثَلُ فَالْأَمْثَلُ
فَيُبْتَلَى الرَّجُلُ عَلَى حَسَبِ دِينِهِ فَإِنْ كَانَ دِينُهُ صُلْبًا اشْتَدَّ بَلَاؤُهُ وَإِنْ كَانَ فِي دِينِهِ رِقَّةٌ ابْتُلِيَ
عَلَى حَسَبِ دِينِهِ فَمَا يَبْرُحُ الْبَلَاءُ بِالْعَبْدِ حَتَّى يَتْرُكَهُ يَمْشِي عَلَى الْأَرْضِ مَا عَلَيْهِ خَطِيئَةٌ .

Sa'ad meriwayatkan, "Saya bertanya kepada Rasulullah, "Siapa manusia yang mendapat ujian yang paling kuat?" Beliau menjawab. 'Para nabi, kemudian yang sepertinya dan kemudian yang sepertinya. Jika agamanya kuat, maka ujiannya akan lemah pula. Ujian itu senantiasa diberikan kepada manusia sampai ia tidak berbuat kesalahan lagi'" (HR. At-Tarmidzi).²³

Dalam Tuhfah Al-Ahwazi dijelaskan bahwa manusia yang paling banyak dan sulit ujian dan cobaannya adalah para nabi. Mereka banyak diuji karena mereka senang dengan ujian itu sebagaimana orang lain senang dengan nikmat. Apabila tidak diuji, mereka meragukan kecintaan Tuhan dan kesabarannya lemah dalam menghadapi umat. Semakin kuat ujiannya, mereka semakin tawaduk dan berharap kepada Allah Swt.

Al-amstal adalah orang-orang paling utama, paling tinggi kedudukan dan posisinya. Mereka paling dekat dengan Allah dan diberi ujian yang berat supaya mendapat pahala yang banyak.

Seseorang diuji sesuai dengan tingkat (ukuran) agamanya. Artinya, sesuai dengan kelemahan, kekuatan, kekurangan, dan kesempurnaan agamanya. Jika ia kuat dalam beragama, maka ujiannya kuat pula. Sebaliknya, jika agamanya lemah, maka ujiannya akan lemah pula. Ujian itu senantiasa diberikan kepada manusia sampai ia tidak berbuat kesalahan lagi. Menurut Arifin, sistem evaluasi dalam sunah nabi yang bersifat makro adalah untuk mengetahui kemajuan belajar manusia, termasuk nabi sendiri. Hal ini sebagaimana kisah kedatangan Malaikat Jibril yang menguji beliau dengan pertanyaan-pertanyaan tentang rukun Islam dan setiap jawaban beliau selalu dibenarkan oleh Malaikat Jibril. Peristiwa

²³ *Ibid*

lainnya, yaitu Malaikat Jibril yang mendatangi nabi untuk menguji sejauh mana hafalan ayat-ayat Al-Qur'an, juga konsentrasi, dan validitas ingatan beliau. Dalam melaksanakan kegiatan dakwah dan pengajaran, nabi juga sering sekali mengadakan evaluasi terhadap hasil belajar para sahabatnya dengan sistem tanya-jawab dan musyawarah. Tujuannya adalah mengetahui mana di antara para sahabat yang cerdas, patuh, dan saleh atau mana yang kreatif dan aktif-responsif terhadap pemecahan problem-problem yang dihadapi bersama nabi dalam keadaan mendesak.

Dari hadis dan penjelasan di atas dapat dipahami bahwa Rasulullah telah mengevaluasi proses dan hasil pembelajaran para sahabat. Evaluasi yang beliau lakukan mencakup ranah kognitif, afektif dan psikomotor walaupun dalam bentuk pelaksanaan yang masih sederhana sesuai dengan kebutuhan pada waktu itu. Dengan demikian, seyogianya pendidik dalam lembaga pendidikan Islam saat ini mengembangkan sistem evaluasi yang komprehensif dan menyeluruh. Jangan berpuas diri dengan hasil evaluasi ranah kognitif saja. Ranah yang lainnya juga harus diperhatikan dan diutamakan sesuai dengan tujuan pendidikan Islam.

D. Tujuan Evaluasi Pendidikan

Menurut Abuddin Nata tujuan evaluasi adalah untuk mengetahui kadar pemahaman peserta didik terhadap materi pelajaran, melatih keberanian dan mengajak peserta didik untuk mengingatkan kembali materi yang telah diberikan dan mengetahui pula tingkat perubahan perilakunya. Selain itu, evaluasi bertujuan untuk mengetahui siapa di antara peserta didik yang cerdas dan dilakukan tindakan yang tepat bagi mereka. Lebih dari itu, sasaran evaluasi tidak hanya mengevaluasi pendidik, materi pendidikan, proses penyampaian materi pelajaran dan berbagai aspek lainnya yang berkaitan dengan materi pendidikan. Hal ini perlu dilakukan karena satu komponen dengan komponen lainnya saling berkaitan. Satu komponen pendidikan yang lemah akan berpengaruh terhadap komponen lainnya, begitu juga sebaliknya jika satu komponen kuat maka akan berpengaruh terhadap hasil pendidikan.²⁴

²⁴ Abuddin Nata, *Ilmu Pendidikan Islam*, (Jakarta: Kencana, 2010), hlm. 308

Dengan demikian, tujuan utama evaluasi dalam pembelajaran adalah untuk mendapatkan informasi yang akurat mengenai tingkat pencapaian tujuan pembelajaran oleh peserta didik sehingga dapat diupayakan tindak lanjutnya. Tindak lanjut dari kegiatan evaluasi sebagai aktivitas untuk memperoleh informasi yang akurat mengenai tingkat pencapaian tujuan pembelajaran oleh siswa merupakan fungsi evaluasi yang masing-masing dapat dilakukan melalui tes.

Tujuan evaluasi menurut ajaran Islam, berdasarkan pemahaman terhadap ayat-ayat Al-Qur'an antara lain adalah:

1. Untuk menguji daya kemampuan manusia beriman terhadap berbagai macam problema kehidupan yang dialaminya.
2. Untuk mengetahui sampai sejauh mana hasil pendidikan wahyu yang telah ditetapkan Rasulullah saw. terhadap umatnya.
3. Untuk menentukan klasifikasi atau tingkatan atau keimanan manusia, sehingga diketahui manusia yang paling mulia di sisi Allah, yaitu manusia yang paling bertakwa kepadanya, manusia yang sedang dalam iman dan ketakwaannya, dan manusia yang ingkar kepada ajaran Islam.²⁵

Sedangkan menurut Mukhtar Bukhari dalam M. Chatib Thaha bahwa ada dua tujuan evaluasi, yaitu:

1. Untuk mengetahui kemajuan belajar peserta didik setelah menyadari pendidikan selama jangka waktu tertentu.
2. Untuk mengetahui tingkat efisien metode pendidikan yang dipergunakan dalam jangka waktu tertentu.²⁶

Uraian di atas dapat dipahami bahwa tujuan dari evaluasi setidaknya untuk mengetahui kemampuan, kemajuan, hasil belajar yang telah dilaksanakan serta efisiensi metode dalam proses pembelajaran tersebut. Dengan mengetahui hal-hal tersebut maka akan ada tindak lanjut untuk meningkatkan menjadi lebih baik dalam segala aspek tersebut.

Menurut Athiyah al-Abrasy bahwa tujuan dari evaluasi adalah untuk mengetahui kadar pemahaman anak didik untuk meningkatkan kembali

²⁵ *Ibid*, hlm. 136-137

²⁶ M. Chatib Thaha, *Teknik Evaluasi Pendidikan*, (Jakarta: PT. RajaGrafindo Persada, 1966), hlm. 6

materi yang telah diberikan, selain itu program evaluasi bertujuan untuk mengetahui siapa di antara anak didik yang lemah, sehingga yang lemah diberi perhatian khusus agar ia dapat mengejar kekurangannya, sehingga naik tingkat kelas maupun tamat sekolah. Pada akhirnya sasaran evaluasi tidak bertujuan mengevaluasi anak didik saja, tetapi bertujuan untuk mengevaluasi pendidik yaitu sejauh mana ia sungguh-sungguh dalam menjalankan tugasnya untuk mencapai tujuan pendidikan.²⁷

Dengan demikian, evaluasi tidak hanya bertujuan mengevaluasi peserta didik melainkan secara menyeluruh tentang segala hal yang berkaitan dengan proses belajar mengajar menjadi target utama dari sebuah tujuan evaluasi, selanjutnya bagaimana kemudian secara menyeluruh dipecahkan segala macam kelemahan-kelemahan yang terjadi dan kemudian memberikan prioritas tertentu terhadap elemen yang memungkinkan menjadi penyebab dari tidak keserasian tujuan pendidikan itu sendiri.

Secara khusus tujuan evaluasi pendidikan memberikan standar keberhasilan baik dari sisi aspek kognitif, aspek afektif, dan aspek psikomotorik kemudian berimplikasi pada penanganan siswa oleh pendidiknya untuk lebih bisa memberikan perhatian lebih baik secara fisik maupun secara psikis tergantung dari gejala yang dialami peserta didik agar tidak salah dalam tindak lanjut. Dalam pendidikan Islam faktor afektif dan psikomotorik agar lebih diperhatikan dibandingkan dengan faktor kognitifnya, karena penekanan ini bertujuan untuk mengetahui kemampuan peserta didik yang secara garis besarnya meliputi empat hal, di antaranya:

- a. Sikap dan pengalaman terhadap hubungan pribadinya dengan Tuhannya.
- b. Sikap dan pengalaman terhadap arti hubungannya dengan masyarakat.
- c. Sikap dan pengalamannya terhadap arti hubungan kehidupannya dengan alam sekitarnya.

²⁷ Muhaimin-Abdul Mujib, *Pemikiran Pendidikan Islam*, (Bandung: Trigenda Karya, 1993), hlm. 277

- d. Sikap dan pandangannya terhadap diri sendiri selaku hamba Allah, anggota masyarakat, serta khalifatullah.²⁸

Selanjutnya empat di atas dapat diklasifikasikan menjadi beberapa kemampuan teknis, antara lain:

- a. Sejauh mana loyalitas pengabdianya kepada Allah dengan indikasi lahiriah berupa tingkah laku yang mencerminkan keimanan dan ketakwaan kepada Allah Swt.
- b. Sejauh mana peserta didik dapat menerapkan nilai-nilai agamanya dan kegiatan hidup masyarakat.
- c. Bagaimana peserta didik memelihara dan mengelola serta menyesuaikan diri dengan alam sekitarnya.
- d. Sejauh mana dan bagaimana memandang diri sendiri sebagai hamba Allah dalam menghadapi kenyataan masyarakat yang beraneka ragam budaya, suku bahkan agamanya.²⁹

Selain tujuan evaluasi di atas, evaluasi juga memiliki beberapa fungsi, di antaranya:

1. Evaluasi berfungsi sebagai selektif

Dengan mengadakan evaluasi guru mempunyai cara untuk mengadakan seleksi atau penilaian terhadap siswanya, evaluasi tersebut mempunyai berbagai tujuan, antara lain:

- a. Untuk memilih siswa yang dapat diterima di sekolah tertentu
- b. Untuk memilih siswa yang dapat naik kelas atau tingkat berikutnya
- c. Untuk memilih siswa yang seharusnya mendapat beasiswa
- d. Untuk memilih siswa yang sudah berhak meninggalkan sekolah dan sebagainya

2. Evaluasi berfungsi sebagai diagnostik

Apabila alat yang digunakan dalam evaluasi cukup memenuhi persyaratan, maka dengan melihat hasilnya, guru akan mengetahui kelemahan siswa, sebab kelemahan siswa. Dengan demikian mengadakan evaluasi, sebenarnya guru mengadakan diagnosis kepada siswa tentang kebaikan dan kelemahannya. Dengan diketahui sebab dan musabab

²⁸ Samsul Nizar, *Filsafat Pendidikan Islam*, (Jakarta: Ciputat Press, 2002), hlm. 80

²⁹ Arifin, *Ilmu Pendidikan Islam*, (Jakarta: Bumi Aksara, 1991), hlm. 239-240

kelemahan tersebut maka akan lebih mudah dicari cara untuk mengatasinya.

3. Evaluasi berfungsi sebagai penempatan

Evaluasi ini adalah untuk menentukan dengan pasti di kelompok mana siswa harus ditempatkan. Sekelompok siswa yang mempunyai hasil penilaian yang sama akan berada dalam kelompok yang sama dalam belajar.

4. Evaluasi berfungsi sebagai pengukur keberhasilan

Fungsi ini bertujuan untuk mengetahui sejauh mana program ditetapkan.³⁰

Beberapa uraian di atas dipahami bahwa kegiatan evaluasi bertujuan dan berfungsi untuk mengetahui tepat atau tidaknya guru dalam memilih bahan, metode, dan berbagai penyesuaian di dalam kelas. Sehingga dapat diketahui pula tingkat efisiensi metode-metode yang dipergunakan pendidikan selama jangka waktu tertentu dalam mendidik kita menginginkan tercapai hasil yang sebesar-besarnya. Tidak mustahil jika suatu metode yang dipilih tidak menghasilkan yang diharapkan, maka ketepatan dalam memilih dan menggunakan metode dapat diketahui dengan diadakannya evaluasi pada periode-periode tertentu.

E. Sasaran Evaluasi

Dalam standar penilaian pendidikan ada beberapa macam penilaian yang dilakukan pada satuan pendidikan, di antaranya:

- a. Ulangan adalah proses untuk mengukur pencapaian kompetensi peserta didik secara berkelanjutan dalam pembelajaran, memantau kemajuan, melakukan perbaikan pembelajaran, dan menentukan keberhasilan belajar peserta didik.
- b. Ulangan harian adalah untuk mengukur pencapaian kompetensi peserta setelah menyelesaikan satu kompetensi dasar atau lebih.
- c. Ulangan tengah semester adalah untuk mengukur pencapaian kompetensi peserta didik setelah melaksanakan 8-9 minggu kegiatan

³⁰ Abuddin Nata, *Op. cit.*, hlm. 136-137

- pembelajaran. Cakupan ulangan meliputi seluruh indikator yang mempresentasikan seluruh kemampuan dasar pada periode tersebut.
- d. Ulangan akhir semester adalah kegiatan yang dilakukan oleh pendidik untuk mengukur pencapaian kompetensi peserta didik di akhir semester.
 - e. Ulangan kenaikan kelas adalah kegiatan untuk mengukur pencapaian kompetensi peserta didik di akhir semester genap pada satuan pendidikan yang menggunakan sistem paket.
 - f. Ujian sekolah/madrasah adalah kegiatan pengukuran pencapaian kompetensi peserta didik untuk memperoleh pengakuan prestasi belajar dan merupakan persyaratan kelulusan dari satuan pendidikan.
 - g. Ujian nasional atau UN adalah kegiatan mengukur pencapaian kompetensi peserta didik pada beberapa mata pelajaran tertentu dalam rangka mencapai penilaian Standar Nasional Pendidikan.³¹

Berdasarkan PP di atas pelaksanaan evaluasi pada satuan pendidikan, dilakukan oleh pendidik, satuan pendidikan dan pemerintah dalam kerangka untuk menjamin mutu pendidikan secara nasional.

Langkah yang harus ditempuh seorang pendidik dalam mengevaluasi adalah menetapkan apa yang menjadi sasaran evaluasi tersebut. Sasaran evaluasi sangat penting untuk diketahui agar memudahkan pendidik dalam menyusun alat-alat evaluasinya. Pada umumnya ada tiga sasaran pokok evaluasi, yaitu:

1. Segi tingkah laku, artinya yang menyangkut sikap, minat, perhatian, keterampilan siswa sebagai akibat dari proses belajar mengajar.
2. Segi pendidikan, artinya penguasaan pelajaran yang diberikan oleh guru dalam proses belajar mengajar.
3. Segi yang menyangkut proses belajar mengajar yaitu bahwa proses belajar mengajar perlu diberi penilaian secara objektif dari guru. Sebab baik tidaknya proses belajar mengajar akan menentukan baik tidaknya hasil dicapai oleh murid.³²

³¹ Permendiknas RI No. 20 Tahun 2007 tentang Standar Penilaian Pendidikan

³² Abuddin Nata, *Filsafat Pendidikan Islam, Op.cit*, hlm. 143

Dengan menetapkan sasaran di atas, maka pendidik lebih mudah mengetahui alat-alat evaluasi yang dipakai baik dengan tes maupun dengan nontes. Di antaranya:

- a. Kedudukan akademis setiap murid, baik dibandingkan dengan teman-teman sekelasnya, sekolahnya, maupun sekolah-sekolah lain.
- b. Kemajuan belajar dalam satu pelajaran tertentu, misalnya fikih, tauhid, tarikh dan lainnya.
- c. Kelemahan dan kelebihan murid.³³

F. Prinsip-Prinsip Evaluasi

Menurut Abuddin Nata penilaian akan berhasil jika dilakukan dengan sesuai prinsip-prinsip, yaitu:

1. Prinsip berkesinambungan, penilaian hendaknya dilakukan secara berkesinambungan. Sebagaimana dijelaskan dalam surah Al-Fushilat ayat 30.

إِنَّ الَّذِينَ قَالُوا رَبُّنَا اللَّهُ ثُمَّ اسْتَقَمُوا تَتَنَزَّلُ عَلَيْهِمُ الْمَلَائِكَةُ أَلَّا تَخَافُوا وَلَا تَحْزَنُوا وَأَبْشِرُوا بِالْجَنَّةِ الَّتِي كُنتُمْ تُوعَدُونَ ﴿٣٠﴾

Artinya: “Sesungguhnya orang-orang yang mengatakan: “Tuhan Kami ialah Allah” kemudian mereka meneguhkan pendirian mereka, Maka Malaikat akan turun kepada mereka dengan mengatakan: “Janganlah kamu takut dan janganlah merasa sedih; dan gembirakanlah mereka dengan jannah yang telah dijanjikan Allah kepadamu” (QS. Al-Fushilat: 30).

2. Prinsip menyeluruh, maksudnya penilaian harus mengumpulkan data mengenai seluruh aspek kepribadian sebagaimana firman Allah Swt. dalam surah Al-Zalzalah ayat 7-8.

فَمَنْ يَعْمَلْ مِثْقَالَ ذَرَّةٍ خَيْرًا يَرَهُ ﴿٧﴾ وَمَنْ يَعْمَلْ مِثْقَالَ ذَرَّةٍ شَرًّا يَرَهُ ﴿٨﴾

³³ Ibid

Artinya: “(7) Barang siapa yang mengerjakan kebaikan seberat zarah pun, niscaya Dia akan melihat (balasan)nya. (8) Dan barang siapa yang mengerjakan kejahatan sebesar zarah pun, niscaya Dia akan melihat (balasan)nya pula” (QS. AL-Zalzalah: 7-8).

3. Prinsip objektif, penilaian harus diusahakan seobjektif mungkin sebagaimana firman Allah dalam surah Al-Maidah ayat 8.

يَتَّيِبُهَا لِلَّذِينَ ءَامَنُوا كُونُوا قَوَّامِينَ لِلَّهِ شُهَدَاءَ بِالْقِسْطِ ۗ وَلَا يَجْرِمَنَّكُمْ شَنَاٰنُ
قَوْمٍ عَلٰٓى اٰلٍ اَلَّا تَعْدِلُوْا ۗ اَعْدِلُوْا هُوَ اَقْرَبُ لِلتَّقْوٰى ۗ وَاتَّقُوا اللّٰهَ ۗ اِنَّ اللّٰهَ خَبِيْرٌۢ بِمَا
تَعْمَلُوْنَ ﴿٨﴾

Artinya: “Hai orang-orang yang beriman hendaklah kamu jadi orang-orang yang selalu menegakkan (kebenaran) karena Allah, menjadi saksi dengan adil. Dan janganlah sekali-kali kebencianmu terhadap sesuatu kaum, mendorong kamu untuk berlaku tidak adil. Berlaku adillah, karena adil itu lebih dekat kepada takwa. Dan bertakwalah kepada Allah, Sesungguhnya Allah Maha Mengetahui apa yang kamu kerjakan” (QS. Al-Maidah: 8).

4. Prinsip sistematis, penilaian harus dilakukan secara sistematis dan teratur.³⁴

Sedangkan dalam Standar Penilaian Pendidikan ada beberapa prinsip penilaian hasil belajar peserta didik pada jenjang pendidikan dasar dan menengah didasarkan prinsip-prinsip sebagai berikut:

1. Prinsip sah
2. Prinsip objektif
3. Prinsip adil
4. Prinsip terpadu
5. Prinsip terbuka
6. Prinsip menyeluruh dan berkesinambungan
7. Prinsip sistematis

³⁴ Ibid, hlm. 140

8. Prinsip beracuan kriteria
9. Prinsip akuntabel

Prinsip ini lebih banyak dipraktikkan pada pendidikan umum, namun demikian, dalam penyelenggaraan evaluasi pendidikan Islam tidak salah apabila menggunakan prinsip-prinsip tersebut. Sebab Islam merupakan agama yang fleksibel dan dapat menerima nilai-nilai luar yang tidak bertentangan dengan ajaran Islam.

G. Unsur-Unsur Evaluasi Pendidikan

Unsur-unsur evaluasi sangat diperhatikan dalam Islam, hal tersebut terlihat di dalam surah Al-Baqarah ayat 31-32.

وَعَلَّمَ آدَمَ الْأَسْمَاءَ كُلَّهَا ثُمَّ عَرَضَهُمْ عَلَى الْمَلَائِكَةِ فَقَالَ أَنْبِئُونِي بِأَسْمَاءِ هَؤُلَاءِ
 إِنْ كُنْتُمْ صَادِقِينَ قَالُوا لَا نَعْلَمُ لَنَا إِلَّا مَا عَلَّمْتَنَا إِنَّكَ أَنْتَ الْعَلِيمُ
 الْحَكِيمُ

Artinya: Dan dia mengajarkan kepada Adam nama-nama (benda-benda) seluruhnya, kemudian mengemukakannya kepada para malaikat lalu berfirman: "Sebutkanlah kepada-Ku nama benda-benda itu jika kamu memang benar orang-orang yang benar!" Mereka menjawab: "Mahasuci Engkau, tidak ada yang kami ketahui selain dari apa yang telah Engkau ajarkan kepada Kami. Sesungguhnya Engkau adalah Yang Maha Mengetahui lagi Maha Bijaksana." Sebenarnya terjemahan Hakim dengan Maha Bijaksana kurang tepat, karena arti hakim ialah yang mempunyai hikmah. Hikmah ialah penciptaan dan penggunaan sesuatu sesuai dengan sifat, guna dan faedahnya. Di sini diartikan dengan Maha Bijaksana karena dianggap arti tersebut hampir mendekati arti hakim (QS. Al-Baqarah: 31-32).

Jika dihubungkan dengan evaluasi, maka dapat dikemukakan beberapa unsur evaluasi sebagai berikut:

1. Unsur evaluator dan pengembang, dalam hal ini Allah Swt. sebagai guru yang mendidik Nabi Adam a.s.

2. Unsur materi, dalam hal ini segala sesuatu yang diajarkan kepada Nabi Adam a.s.
3. Unsur kesahihan hasil evaluasi, dalam hal ini pengakuan dan penilaian yang jujur dari para malaikat yang mengakui kemampuan Nabi Adam a.s. sebagai hasil didikan Allah Swt.
4. Unsur pengakuan terhadap hasil evaluasi dan pengembangan, dalam hal ini para malaikat menyatakan hormat dan apresiasi yang tinggi terhadap Nabi Adam a.s.³⁵

Uraian di atas dapat dipahami bahwa unsur evaluasi, setidaknya harus ada orang yang mengevaluasi (*evaluator*) dalam hal ini pendidik, guru atau dosen dan lainnya, orang yang dievaluasi dalam hal ini peserta didik, materi yang akan dievaluasi yaitu materi yang telah diberikan kepada peserta didik. Dengan demikian, akan didapatkan hasil evaluasi yang sesuai dengan ketentuan yang telah ditetapkan.

H. Evaluasi yang Dilakukan Rasulullah saw.

Pada masa Rasulullah saw. ada beberapa bentuk praktik evaluasi, hanya saja evaluasi pada masa Rasulullah saw. belum terstruktur sebagaimana evaluasi pada masa modern saat ini. Adapun evaluasi tersebut di antaranya sebagai berikut:

1. Mengevaluasi hafalan ayat-ayat Al-Qur'an yang telah disampaikan kepada sahabat. Rasulullah mengevaluasi hafalan para sahabat dengan cara menyuruh para sahabat membacakan ayat-ayat Al-Qur'an kemudian dengan membetulkan bacaan mereka yang keliru, sebagaimana hadis Rasulullah saw.:

حَدَّثَنَا عُمَرُ بْنُ حَفْصِ بْنِ غِيَاثٍ حَدَّثَنَا أَبِي حَدَّثَنَا الْأَعْمَشُ قَالَ حَدَّثَنِي إِبْرَاهِيمُ عَنِ الْأَسْوَدِ عَنْ عَبْدِ اللَّهِ رَضِيَ اللَّهُ عَنْهُ قَالَ بَيْنَمَا نَحْنُ مَعَ النَّبِيِّ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ فِي غَارٍ يَمْنَى إِذْ نَزَلَ عَلَيْهِ وَالْمُرْسَلَاتِ وَأَنَّهُ لَيَتْلُوهَا وَإِنِّي لَأَتَلَّقَاهَا مِنْ فِيهِ وَإِنَّ فَاهُ لَرَطْبٌ بِهَا

³⁵ Abuddin Nata, *Ilmu Pendidikan Islam*, Op.cit, hlm. 309-310

إِذْ وَثَبَتْ عَلَيْنَا حَيَّةٌ فَقَالَ النَّبِيُّ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ أَقْتُلُوهَا فَاثْبَتْرَنَاهَا فَذَهَبَتْ فَقَالَ النَّبِيُّ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ وَثَبَتْ شَرُّكُمْ كَمَا وَثَبَتْ شَرُّهَا (رواه البخاري)

Artinya: Telah menceritakan kepada kami Umar bin Hafsh bin Ghiyats telah menceritakan kepada kami bapakku telah menceritakan kepada kami A'masy berkata, telah menceritakan kepada saya Ibrahim dari al-Aswad dari Abdullah Ra berkata: ketika kami sedang bersama Rasulullah saw. di dalam gua di mina, wahyu turun kepada beliau yaitu surah Al-Mursalat. Saat itu beliau membacakannya dan gerak bibir mulut beliau, dan aku mengambil bacaan surat tersebut langsung dari gerak bibir beliau. Ketika bibir beliau masih basah membacakan surat tersebut tiba-tiba ada seekor ular melompat kepada kami, maka Rasulullah saw. Berkata "Bunuhlah ular itu". Maka kami mengejar ular itu namun ular itu sudah pergi, maka Rasulullah saw. berkata ular itu telah lolos dari kejahatan kalian sebagaimana kalian telah lolos dari kejahatannya (HR. Bukhari).³⁶

Berdasarkan hadis di atas dapat dipahami Rasulullah saw. mengevaluasi hafalannya dengan memperdengarkan (menyimakkan) bacaan Al-Qur'an di hadapan para sahabat, para sahabat menyaksikan langsung bacaan ayat Al-Qur'an yang dilafazkan oleh Rasulullah saw.

2. Evaluasi terkait pemahaman ajaran Islam yang telah disampaikan Rasulullah kepada sahabat.

Evaluasi ini dilakukan dengan cara bertanya tentang suatu masalah hukum secara langsung kepada Rasulullah saw. lalu Rasulullah menjawabnya, sebagaimana riwayat berikut ini:

حَدَّثَنَا قُتَيْبَةُ، حَدَّثَنَا إِسْمَاعِيلُ بْنُ جَعْفَرٍ، عَنْ عَبْدِ اللَّهِ دِينَارٍ، عَنْ أَبِي عُمَرَ قَالَ، قَالَ رَسُولُ اللَّهِ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ، "إِنَّ مِنْ شَجَرٍ شَجْرَةً لَا يَسْقُطُ وَرَقُهَا، وَأَنَّهَا مِثْلُ الْمُسْلِمِ فَحَدِّثُونِي مَا هِيَ؟ فَوَقَعَ النَّاسُ فِي شَجَرَةِ الْيُودِيِّ، قَالَ عَبْدُ اللَّهِ، وَوَقَعَ فِي نَفْسِي أَنَّهَا النَّخْلَةُ، هِيَ فَاسْتَحْيَيْتُ. ثُمَّ قَالُوا، حَدَّثَنَا مَا هِيَ يَا رَسُولَ اللَّهِ، قَالَ النَّخْلَةُ (رواه البخاري)

³⁶ Al-Imam Al-Bukhari, *Terjemahan Shahih Bukhari*, hadis no. 1699

Artinya: menceritakan kepada kami Qutaibah, menceritakan kepada kami Ismail ibn ja'far dari Abdullah ibn dinar dari ibn Umar ia berkata Rasulullah saw. bersabda, sesungguhnya di antara pepohonan ada satu pohon yang daunnya tidak jatuh ke tanah secara berguguran. Pohon itu bagaikan seorang muslim. Jelaskanlah kepadaku pohon apa itu? Orang-orang mengatakan pohon itu terdapat di pedalaman, Abdullah berkata, dalam benakku terbetik pikiran bahwa yang dimaksud adalah pohon kurma. Akan tetapi aku malu menjawabnya” orang-orang berkata, beritahukanlah kami pohon apakah itu ya Rasulullah? Beliau menjawab pohon kurma (HR. Bukhari).

Hadis ini menunjukkan bahwa evaluasi yang dilakukan oleh Rasulullah saw. dengan memberikan respons atau memancing pertanyaan sehingga menimbulkan pertanyaan, yang pada akhirnya Rasulullah saw. memberikan jawaban kepada yang bertanya-tanya akan maksud Rasulullah saw. Hadis lain yang berbicara tentang evaluasi pemahaman (*kognitif*) para sahabat tentang ajaran Islam. Seperti halnya dialog Rasulullah saw. dengan Muadz bin Jabal ketika akan diutus sebagai Qadhi ke Negeri Yaman. Sebelum Rasulullah saw. menguji pengetahuan terkait dengan dasar dan dijadikan rujukan bila ditemukan persoalan di tengah masyarakat yang membutuhkan penyelesaian, Muadz menjawabnya dengan tiga rujukan, sebagaimana hadis berikut ini:

عَنْ مُعَاذِ بْنِ جَبَلٍ أَنَّ رَسُولَ اللَّهِ -صلى الله عليه وسلم- لَمَّا أَرَادَ أَنْ يَبْعَثَ مُعَاذًا إِلَى الْيَمَنِ قَالَ « كَيْفَ تَقْضِي إِذَا عَرَضَ لَكَ قَضَاءٌ ». قَالَ أَقْضِي بِكِتَابِ اللَّهِ. قَالَ « فَإِنْ لَمْ تَجِدْ فِي كِتَابِ اللَّهِ ». قَالَ فَيَسْتَنْتِ رَسُولَ اللَّهِ -صلى الله عليه وسلم-. قَالَ « فَإِنْ لَمْ تَجِدْ فِي سُنَّةِ رَسُولِ اللَّهِ -صلى الله عليه وسلم- وَلَا فِي كِتَابِ اللَّهِ ». قَالَ أَجْتَهُدُ رَأْيِي وَلَا أَلُو. فَضَرَبَ رَسُولُ اللَّهِ -صلى الله عليه وسلم- صَدْرَهُ وَقَالَ « الْحَمْدُ لِلَّهِ الَّذِي وَفَّقَ رَسُولَ رَسُولِ اللَّهِ لِمَا يُرْضَى رَسُولَ اللَّهِ. (رواه أبو داود والترمذی وأحمد والدارمی)

Artinya: Mu'az (bin Jabal) meriwayatkan bahwa Rasulullah saw. ketika akan mengutus Mu'az ke Yaman, beliau bertanya (kepada Mu'az), “Bagaimana engkau mengadili perkara, jika dihadapkan kepadamu suatu perkara pengadilan?” Mu'az menjawab: “Saya mengadili (perkara itu)

dengan kitab Allah (Al-Qur'an)". (Rasulullah bertanya lagi, "Maka bagaimana jika kamu tidak menjumpai (petunjuk) dalam kitab Allah (Al-Qur'an)? Mu'az menjawab, "Maka (saya mengadili) dengan sunah Rasulullah saw." Rasulullah bertanya lagi, (bagaimana bila) kamu tidak menjumpai (petunjuk dalam sunah Rasulullah saw. dan (tidak menjumpainya) dalam kitab Allah (Al-Qur'an)?" Mu'az menjawab, "Saya berjihad sekuat akal pikiran saya". Maka Rasulullah saw. menepuk dadaku sambil bersabda: Segala puji milik Allah yang telah memberi petunjuk kepada utusan Rasulullah terhadap apa yang Rasulullah berkenan terhadapnya (HR. Abu Daud At-Turmudzi dan Ahmad dan Ad-Darimiy).³⁷

Di antara informasi yang terkandung dalam hadis di atas adalah: (1). Rasulullah saw. bermaksud akan mengutus Mu'az ke Yaman (untuk memimpin umat), (2). Rasulullah saw. bertanya kepada Mu'az tentang dasar yang digunakan dalam memutuskan perkara peradilan, (3). Mu'az menjawab dengan urutan: *pertama* dengan kitab Allah, *kedua* dengan sunah Rasulullah, *ketiga* dengan ijtihad, (4). Setelah selesai jawaban Mu'az, Rasulullah saw. menepuk dada Mu'az karena senang, dan memuji Allah.

Dalam hadis di atas terlihat bahwa Rasulullah saw. menguji kemampuan dan pengetahuan seorang sahabat sebelum memberikan tugas kepadanya. Setelah sahabat berhasil menjawab secara benar sesuai dengan keinginan Rasulullah saw., beliau memperlihatkan rasa senangnya dengan memberikan ganjaran yang menyenangkan sahabat dan beliau memuji Allah. Pujian kepada Allah di sini dapat diartikan sebagai rasa syukur atas keberhasilan Rasulullah saw. mendidik sahabatnya.

Ujian yang diberikan oleh Rasulullah saw. dalam hadis di atas berkaitan dengan tugas yang akan diemban oleh Mu'az bin Jabal. Rasulullah saw. baru akan menyerahkan suatu tugas kepada sahabat bila sahabat tersebut menguasai (memiliki ilmu) tentang persoalan tugas yang akan diembannya.

³⁷ Sunan Abu Dawud., *Kitab As-Sunan Abi Dawud.*, Juz 3, hlm. 303

3. Evaluasi yang dilakukan dengan melalui perantara malaikat Jibril a.s. mengenai kemampuan hafalan ayat-ayat Al-Qur'an beliau.

Setiap bulan Ramadan Rasulullah kedatangan malaikat Jibril a.s. yang kedatangannya untuk menguji sekaligus untuk mengadakan pengulangan hafalan Rasulullah saw. dengan memperdengarkan Al-Qur'an yang telah diturunkan tersebut. Hal ini sesuai dengan sebuah hadis yang diriwayatkan oleh Imam Bukhari, Muslim, dan Nasa'i, yang artinya:

كَانَ النَّبِيُّ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ أَحْوَدَ النَّاسِ وَأَجْوَدُ مَا يَكُونُ فِي رَمَضَانَ حِينَ يَلْقَاهُ وَكَانَ جِبْرِيلُ عَلَيْهِ السَّلَامُ وَيَلْقَاهُ فِي كُلِّ لَيْلَةٍ مِنْ رَمَضَانَ فَيُدَارِسُهُ الْقُرْآنَ فَرَسُولُ اللَّهِ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ أَحْوَدُ بِالْخَيْرِ مِنَ الرِّيحِ الْمُرْسَلَةِ (رواه البخاري)

Artinya: "Nabi Muhammad saw. adalah manusia yang paling dermawan, dan kedermawanannya semakin menjadi-jadi saat Ramadan apalagi jika Jibril menemuinya dan Jibril menemuinya setiap malam bulan Ramadan dan bertadarus Al-Qur'an bersamanya, maka Rasulullah benar-benar sangat dermawan dengan kebaikan melebihi angin yang berembus (HR. Bukhari).³⁸

Hadis di atas menggambarkan secara jelas bahwa Rasulullah saw. selalu melakukan evaluasi terhadap hafalan Al-Qur'an yang telah disampaikan oleh malaikat Jibril a.s., evaluasi tersebut bertujuan untuk menguatkan hafalan dan menjaga kemurnian ayat-ayat Al-Qur'an yang telah dihafalnya dari malaikat Jibril a.s. Dari riwayat ini terlihat bahwa sebagai pengujinya Malaikat Jibril materi ujiannya *tahfizh Al-Qur'an* peserta ujiannya adalah Rasulullah saw. Hal ini mengisyaratkan bahwa dalam evaluasi harus ada evaluator, materi evaluasi, dan orang yang akan dievaluasi.

4. Evaluasi yang dilakukan Rasulullah terkait dengan perang.

Rasulullah saw. menguji kemampuan sahabat berkaitan dengan keterampilan berperang.

³⁸ Al-Imam Al-bukhari, *Terjemahan Shahih Bukhari*, hadis no. 3220

Rasulullah melakukan hal ini saat akan berangkat perang, sebagaimana riwayat berikut ini, yang artinya:

حَدَّثَنَا مُحَمَّدُ بْنُ عَبْدِ اللَّهِ بْنِ نُمَيْرٍ، حَدَّثَنَا أَبِي، حَدَّثَنَا عَبْدُ اللَّهِ، عَنْ نَافِعٍ، عَنِ بْنِ عُمَرَ
قَالَ، عَرَضَنِي رَسُولُ اللَّهِ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ يَوْمَ أُحُدٍ فِي الْقِتَالِ، وَأَنَا بِنُ أَرْبَعِ عَشْرَةَ، فَلَمْ يَجْزِنِي،
وَعَرَضَنِي يَوْمَ الْخَنْدَقِ وَأَنَا بِنُ خَمْسِ عَشْرَةَ سَنَةً فَأَجْزَنِي. (رواه المسلم)

Artinya: “Menceritakan kepada kami Muhammad bin Abdullah bin Numair menceritakan kepada kami ayahku, menceritakan kepada kami Abdullah, dari nafi’ dari Ibnu Umar berkata, “Rasulullah saw. menguji kemampuku berperang pada perang uhud, ketika aku berusia empat belas tahun lalu beliau tidak mengizinkanku, dan beliau mengujiku kembali pada hari perang khandak ketika aku telah berusia lima belas tahun, lalu beliau mengizinkanku” (HR. Muslim).

Hadis lainnya yang berkaitan dengan perang mengisyaratkan tentang ranah afektif, yaitu ketika Rasulullah melarang untuk bertemu dengan musuh, hal ini sesuai dengan berikut ini:

عَنْ أَبِي هُرَيْرَةَ أَنَّ النَّبِيَّ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ قَالَ: لَا تَمَنَّوْا لِقَاءَ الْعَدُوِّ فَإِذَا لَقِيْتُمُوهُمْ فَاصْبِرُوا.
(رواه المسلم)

Artinya: dari Abu Hurairah r.a, katanya, Nabi saw. bersabda: janganlah kamu bertemu dengan musuh, jika kamu bertemu dengan musuh maka teguhkanlah hatimu (sabarlah) (HR. Bukhari).³⁹

Hadis ini dapat dipahami bahwa sabar merupakan salah satu cara untuk menghadapi segala ujian. Mengutip pernyataan dari Prof. Dr. H. Quraisy Shihab bahwa sabar “usaha maksimal untuk meningkatkan prestasi dan kualitas diri serta kebaikan.”⁴⁰

³⁹ Al-Imam Abi Husain Muslim Ibn Al-Hujjaj Al-Qusyairiyu An-Naisaburiyu, *Terjemahan Shahih Muslim*, hadis no. 1723

⁴⁰ Quraisy Shihab, *Tausiyah Ramadan*, di Metro TV

Berdasarkan pernyataan ini dapat disimpulkan bahwa selain sabar merupakan sifat yang termasuk pada bagian dari ranah afektif, sabar juga sebuah proses dan usaha evaluasi untuk meningkat atau menjadi lebih baik.

5. Evaluasi mengenai metode “hendaknya menggunakan metode lemah lembut dalam mendidik”

Evaluasi terkait metode ini sesuai dengan riwayat hadis berikut ini:

حَدَّثَنَا أَبُو جَعْفَرٍ مُحَمَّدُ بْنُ الصَّبَّاحِ وَأَبُو بَكْرِ بْنِ أَبِي شَيْبَةَ وَتُقَارِبًا فِي لَفْظِ الْحَدِيثِ قَالَ حَدَّثَنَا إِسْمَاعِيلُ بْنُ إِبْرَاهِيمَ عَنْ حَجَّاجِ الصَّوَّافِ عَنْ يَحْيَى بْنِ أَبِي كَثِيرٍ عَنْ هِلَالِ بْنِ أَبِي مَيْمُونَةَ عَنْ عَطَاءِ بْنِ يَسَارٍ عَنْ مُعَاوِيَةَ بْنِ الْحَكَمِ السَّلْمِيِّ قَالَ بَيْنَا أَنَا أُصَلِّي مَعَ رَسُولِ اللَّهِ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ إِذْ عَطَشَ رَجُلٌ مِنَ الْقَوْمِ فَقُلْتُ “يَرْحَمُكَ اللَّهُ” فَرَمَى الْقَوْمُ بِأَبْصَارِهِمْ فَقُلْتُ وَاتَّكَلْتُ أُمِّيَاءَهُ مَا شَأْنُكُمْ تَنْظُرُونَ إِلَيَّ فَجَعَلُوا يَضْرِبُونَ بِأَيْدِيهِمْ عَلَى أَفْخَادِهِمْ فَلَمَّا رَأَيْتُهُمْ فَصُمْتُوَنِي لِكَيْ سَكَتٌ فَلَمَّا صَلَّعَ فَأَبَى هُوَ وَأُمِّي مَا رَأَيْتُ مُعَلِّمًا قَبْلَهُ وَبَعْدَهُ أَحْسَنَ تَعْلِيمًا مِنْهُ فَوَاللَّهِ مَا كَهَرَنِي وَلَا صَرَبَنِي وَلَا شَتَمَنِي قَالَ إِنَّ هَذِهِ الصَّلَاةَ لَا يَصْلُحُ فِيهَا شَيْءٌ مِنْ كَلَامِ النَّاسِ إِنَّمَا هُوَ التَّسْبِيحُ وَالتَّكْبِيرُ وَقِرَاءَةُ الْقُرْآنِ. (رواه المسلم)

Artinya: “dari Abu Ja’far Muhammad Ibn Shahabah dan Abu Bakar Ibn Aby Syaibah, dari Ismail ibn Ibrahim dari Hajjaj as-Shawwaf dari Yahya ibn Abi Katsir dari Hilal ibn Abi Maimunah dari Atha’ ibn Yasar dari Mu’awiyah ibn Hakam, berkata: ketika saya salat bersama Rasulullah saw., seorang dari jemaah bersin dan saya membaca yarhamukallah. Orang-orang mencelaku dengan pandangan mereka, saya berkata: Celaka kenapa kalian memandangiku? Mereka memukul paha dengan tangan mereka ketika saya memandang mereka, mereka menyuruh saya diam dan saya diam. Setelah Rasulullah saw. selesai salat (aku bersumpah) demi ayah dan ibuku, saya tidak pernah melihat guru sebelum dan sesudahnya yang lebih baik pengajarannya daripada beliau, demi Allah, beliau tidak membentak, memukul dan mencela saya, Rasulullah saw. hanya bersabda

“*sesungguhnya salat ini tidak boleh di dalamnya sesuatu kecuali dari pembicaraan manusia, ia hanya tasbih, takbir, dan membaca Al-Qur’an*” (HR. Muslim).

Riwayat ini tentu mengajarkan kepada kita akan metode yang digunakan sebaiknya metode lemah lembut, sebab metode ini sangat penting dalam mendidik anak, karena dengan lemah lembut pendidikan akan diterima dan menyenangkan bagi siswa. Dengan lemah lembut ini juga menghindarkan peserta didik dari rasa ketakutan terhadap pendidiknya. Sikap lemah lembut juga merupakan kunci kesuksesan Rasulullah saw. dalam mendakwahkan ajaran Islam kepada orang-orang yang menentanginya.

Selain hadis di atas riwayat lain yang berkaitan dengan sikap Rasulullah saw. dalam menghadapi ujian yaitu dengan kesabaran terhadap ujian yang diberikan, sesuai dengan hadis berikut ini:

عن جبير قال: قَالَ رَسُولُ اللَّهِ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ: «إِنَّ اللَّهَ عَزَّ وَجَلَّ يَبْتَلِي عَبْدَهُ بِالسَّقَمِ حَتَّى يُكَفِّرَ عَنْهُ كُلَّ ذَنْبِهِ». (رواه الطبراني)

Artinya: *Sesungguhnya Allah Swt. menguji seorang hamba-Nya dengan suatu penyakit sehingga Ia mengampuni semua dosanya* (HR. At-Thabary).⁴¹

أَبِي هُرَيْرَةَ عَنِ النَّبِيِّ - صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ - قَالَ « مَا يُصِيبُ الْمُسْلِمَ مِنْ نَصَبٍ وَلَا وَصَبٍ وَلَا هَمٍّ وَلَا حُزْنٍ وَلَا أَذًى وَلَا غَمٍّ حَتَّى الشَّوْكَةِ يُشَاكُهَا ، إِلَّا كَفَّرَ اللَّهُ بِهَا مِنْ خَطَايَاهُ. (رواه البخاري)

Artinya: *Dari Abu Hurairah, dari nabi saw., beliau bersabda, setiap musibah yang menimpa seorang muslim yang berupa penyakit, penyakit kronis, kegalauan pikiran, kegelisahan hati, sampai kena duri, akan dihapus Allah kesalahannya* (HR. Bukhari).⁴²

⁴¹ Al-Thabrâniy, *Al-Mu'jam al-Kabîr li al-Thabrâniy*, Juz 2, hlm. 169 dalam *Al-Maktabah al-Syâmilah*

⁴² Al-Bukhariy, *Op.cit.*, Juz 4, hlm. 2316

Berdasarkan hadis tersebut terlihat bahwa yang menjadi sasaran evaluasinya adalah ranah afektif, yaitu sikap kesabaran. Bila seorang muslim mampu menerima ujian tersebut dengan penuh kesabaran, maka Allah akan menghapus kesalahan-kesalahan yang dilakukannya. Ini merupakan hadiah dari Allah untuk hamba-Nya yang lulus.

Hadis ini juga mengisyaratkan bahwa manusia akan diuji oleh Allah dengan penyakit. Sasarannya adalah kesabaran yang termasuk wilayah domain afektif. Selain itu, dalam hadis ini disebut ganjaran yang akan diberikan oleh Allah kepada manusia yang lulus dalam ujian tersebut. Kesabaran adalah cara utama menghadapi penyakit yang dideritanya.

Jika hasil belajar kognitif dan hasil belajar afektif dengan materi tentang kedisiplinan menurut ajaran Islam sebagaimana telah dikemukakan pada pembicaraan terdahulu, maka wujud nyata dan hasil belajar psikomotor.⁴³

عَنْ أَبِي هُرَيْرَةَ أَنَّ رَسُولَ اللَّهِ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ دَخَلَ الْمَسْجِدَ فَدَخَلَ رَجُلٌ فَصَلَّى فَسَلَّمَ عَلَى النَّبِيِّ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ فَرَدَّ وَقَالَ ارْجِعْ فَصَلِّ فَإِنَّكَ لَمْ تُصَلِّ فَارْجِعْ يُصَلِّي كَمَا صَلَّى ثُمَّ جَاءَ فَسَلَّمَ عَلَى النَّبِيِّ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ فَقَالَ ارْجِعْ فَصَلِّ فَإِنَّكَ لَمْ تُصَلِّ ثَلَاثًا فَقَالَ وَالَّذِي بَعَثَكَ بِالْحَقِّ مَا أَحْسِنُ غَيْرَهُ فَعَلَّمَنِي فَقَالَ إِذَا قُمْتَ إِلَى الصَّلَاةِ فَكَبِّرْ ثُمَّ اقْرَأْ مَا تيسَّرَ مَعَكَ مِنَ الْقُرْآنِ ثُمَّ ارْكَعْ حَتَّى تَظْمِنَ رَاكِعًا ثُمَّ ارْفَعْ حَتَّى تَعْدِلَ فَإِذَا قُمْتَ اسْجُدْ حَتَّى تَظْمِنَ سَاجِدًا ثُمَّ ارْفَعْ حَتَّى تَظْمِنَ جَالِسًا وَافْعَلْ ذَلِكَ فِي صَلَاتِكَ كُلِّهَا. (رواه البخارى)

Artinya: Dari Abu Hurairah, sesungguhnya Nabi saw. pernah masuk Masjid, lalu ada seorang laki-laki masuk pula dan salat. Kemudian, ia datang kepada Nabi lalu mengucapkan salam. Kemudian Nabi berkata: "Ulangi salatmu lagi karena sesungguhnya kamu belum salat". Laki-laki itu mengulangi salatnya seperti salatnya tadi. Kemudian, ia datang dan mengucapkan salam kepada Nabi. Nabi berkata lagi: "Ulangi salatmu karena kamu belum salat". Maka laki-laki itu kembali salat seperti

⁴³ *Ibid.*, hlm. 57

salatnya tadi. Setelah itu, ia kembali dan mengucapkan salam kepada Nabi. Kemudian, Nabi berkata lagi: "Ulangi salatmu karena sesungguhnya kamu belum salat". Begitulah sampai tiga kali, lalu laki-laki tersebut berkata: Demi Zat yang telah mengutusmu dengan benar, sungguh aku tidak dapat berbuat yang lebih baik lagi daripada itu. Oleh karena itu, ajarilah aku! Maka Nabi bersabda: "Apabila kamu berdiri untuk salat, maka takbirlah, lalu bacalah ayat yang mudah bagimu, kemudian rukuklah sehingga tumakninah, kemudian bangkitlah sehingga iktidal dalam keadaan berdiri, kemudian sujudlah sehingga tumakninah dalam keadaan sujud, kemudian bangkitlah sehingga tumakninah dalam keadaan duduk, kemudian sujudlah sehingga tumakninah dalam keadaan sujud, kemudian berbuatlah yang demikian itu dalam semua salatmu". (HR. Bukhari, Muslim dan Ahmad, tetapi dalam Muslim tidak terdapat sebutan sujud kedua).⁴⁴

Selain Rasulullah saw. menguji ranah kognitif dan afektif Rasulullah saw. menguji psikomotorik para sahabat, dalam hal ini Rasulullah saw. menguji melalui mendirikan salat. Teknik yang digunakan observasi (nontes). Rasulullah saw. mengamati pekerjaan-pekerjaan yang dilakukan oleh sahabat dalam salatnya. Setelah melihat kekeliruan sahabat, beliau langsung menyuruhnya untuk mengulangi lagi. Jadi, ada perbaikan segera setelah terjadinya kesalahan.

Berdasarkan hadis di atas juga dapat diambil pelajaran bahwa Rasulullah saw. dalam bentuk yang sederhana telah menggunakan observasi sebagai teknik tes kemampuan ranah psikomotor, kendatipun belum menggunakan perencanaan tertulis dan pencatatan lapangan. Pada zaman modern ini, observasi digunakan sebagai instrumen pengukuran kemampuan kerja seseorang dan dilengkapi dengan catatan-catatan yang diperlukan.

Menurut Anas Sudijono, secara umum, pengertian observasi adalah cara menghimpun bahan-bahan keterangan (data) yang dilakukan dengan mengadakan pengamatan dan pencatatan secara sistematis terhadap fenomena-fenomena yang sedang dijadikan sasaran pengamatan. Observasi sebagai alat evaluasi banyak digunakan untuk menilai tingkah

⁴⁴ Al-Bukhariy, *Op.cit.*, Juz 1, hlm. 297

laku individu atau proses terjadinya suatu kegiatan yang dapat diamati, baik dalam situasi yang sebenarnya maupun dalam situasi buatan.⁴⁵

Observasi dapat mengukur atau menilai hasil dan proses belajar; misalnya tingkah laku peserta didik pada waktu guru pendidikan agama menyampaikan pelajaran di kelas, tingkah laku peserta didik pada jam-jam istirahat atau pada saat terjadinya kekosongan pelajaran, perilaku peserta didik pada saat salat jemaah di musala atau masjid sekolah, ceramah-ceramah keagamaan, upacara bendera, ibadah salat tarawih dan sebagainya.

I. Syarat-Syarat Evaluasi Pendidikan Islam

Evaluasi harus memiliki syarat-syarat agar menghasilkan evaluasi yang diakui kebenarannya dan mencerminkan keadaan yang sebenarnya, serta dapat memberi kepuasan kepada peserta didik yang menjadi sasaran dari evaluasi tersebut. Syarat-syarat evaluasi di antaranya *validity*, *reliable*, efisiensi, *ta'abbudiyah* dan ikhlas.

1. Validity

Yaitu pelaksanaan tes harus berdasarkan pada yang seharusnya dievaluasi, yang meliputi seluruh bidang tertentu yang diinginkan dan diselidiki sehingga tidak hanya mencakup satu bidang saja. Soal-soal tes harus memberi gambaran keseluruhan (*representative*) dari kesanggupan peserta didik mengenai bidang itu.

2. Reliable

Yaitu dapat dipercayai yakni dengan memberikan ketelitian dan keterangan mengenai kesanggupan peserta didik. Soal yang ditampilkan tidak membawa tafsiran yang bermacam-macam, sehingga mudah dimengerti oleh peserta didik.

3. Efisiensi

Yaitu tes yang dilakukan merupakan tes yang mudah administrasinya, penilaiannya, dan interpretasinya (penafsirannya). Selain

⁴⁵ Anas Sudijono, *Op.cit.*, hlm. 76

itu, evaluasi dilaksanakan secara cermat dan tepat pada sasarannya. Sesuai dengan surah Al-Insyiqaq ayat 8.

فَسَوْفَ تُحَاسَبُ حِسَابًا يَسِيرًا ﴿٨﴾

Artinya: “Maka Dia akan diperiksa dengan pemeriksaan yang mudah” (QS. Al-Insyiqaq: 8).

4. Ta’abbudiyah dan Ikhlas

Yaitu evaluasi yang dilakukan dengan penuh ketulusan dan pengabdian kepada Allah Swt. Apabila prinsip ini dilakukan maka upaya evaluasi akan membuahkan kesan prasangka baik, terjadi perbaikan tingkah laku secara positif dan menutupi rahasia buruk pada diri seseorang.

J. Kedudukan Evaluasi Pendidikan Islam

Evaluasi pendidikan memiliki kedudukan atau posisi yang strategis, karena hasil dari kegiatan evaluasi bisa dipergunakan sebagai input untuk melakukan perbaikan kegiatan dan proses pembelajaran.⁴⁶

Di dalam Al-Qur’an juga dijelaskan bahwasanya evaluasi sangatlah penting. Sebagaimana dijelaskan dalam surah Al-Baqarah ayat 31-32.

وَعَلَّمَ آدَمَ الْأَسْمَاءَ كُلَّهَا ثُمَّ عَرَضَهُمْ عَلَى الْمَلَائِكَةِ فَقَالَ أَنْبِئُونِي بِأَسْمَاءِ هَٰؤُلَاءِ
إِنْ كُنْتُمْ صَادِقِينَ قَالُوا ﴿٣١﴾ سُبْحٰنَكَ لَا عِلْمَ لَنَا إِلَّا مَا عَلَّمْتَنَا ۗ إِنَّكَ أَنْتَ الْعَلِيمُ

الْحَكِيمُ ﴿٣٢﴾

Artinya: “(31) dan Dia mengajarkan kepada Adam nama-nama (benda-benda) seluruhnya, kemudian mengemukakannya kepada para malaikat lalu berfirman: “Sebutkanlah kepada-Ku nama benda-benda itu jika kamu memang benar orang-orang yang benar! (32) mereka menjawab: “Mahasuci Engkau, tidak ada yang kami ketahui selain dari apa yang

⁴⁶ Abudin Nata, *Filsafat Pendidikan Islam* (Logos Wacana Ilmu: Jakarta, 1992) hlm. 134

telah Engkau ajarkan kepada kami. Sesungguhnya Engkaulah Yang Maha Mengetahui lagi Maha Bijaksana” (QS. Al-Baqarah:31-32).

Dari ayat di atas dijelaskan bahwasanya Allah sebagai pendidik dan Adam sebagai peserta didik. Allah menyuruh Adam mendemonstrasikan ajaran yang diterimanya di hadapan para malaikat. Ayat ini juga mengisyaratkan bahwa materi evaluasi atau materi yang disajikan haruslah materi yang pernah diajarkan.

Selanjutnya, di dalam Al-Qur’an juga dijelaskan bagaimana Nabi Sulaiman mengevaluasi seekor burung hud-hud yang memberitahukan tentang adanya kerajaan yang diperintah oleh seorang wanita cantik yang dikisahkan dalam surah An-Naml ayat 27.

﴿ قَالَ سَنَنْظُرُ أَصَدَقْتَ أَمْ كُنتَ مِنَ الْكٰذِبِينَ ﴾

Artinya: “Berkata Sulaiman: “Akan Kami lihat, apa kamu benar, ataukah kamu termasuk orang-orang yang berdusta” (QS. An-Naml: 27).

Jenis-jenis Evaluasi Pendidikan Islam

Evaluasi (penilaian) pendidikan Islam bisa dilihat dari segi fungsi dan tujuannya evaluasi dan pengembangan pendidikan Islam terdiri dari berbagai macam sebagai berikut:

1. Evaluasi Formatif

Evaluasi ini bertujuan untuk mengetahui hasil kegiatan belajar mengajar yang telah dilakukan oleh pendidikan dan dicapai oleh peserta didik. Evaluasi ini dilakukan karena manusia memilih banyak kelemahan, sebagaimana dijelaskan dalam surah An-Nisa ayat 28 dan berawal dari ketidaktahuan surah An-Nahl ayat 78.

﴿ يُرِيدُ اللّٰهُ اَنْ يُخَفِّفَ عَنْكُمْ وَّخُلِقَ الْاِنْسٰنُ ضَعِيْفًا ﴾

Artinya: “Allah hendak memberikan keringanan kepadamu [286], dan manusia dijadikan bersifat lemah” (QS. An-Nisa: 28).

وَاللَّهُ أَخْرَجَكُمْ مِنْ بُطُونِ أُمَّهَاتِكُمْ لَا تَعْلَمُونَ شَيْئًا وَجَعَلَ لَكُمُ السَّمْعَ وَالْأَبْصَرَ

وَالْأَفْئِدَةَ لَعَلَّكُمْ تَشْكُرُونَ ﴿٧٨﴾

Artinya: “dan Allah mengeluarkan kamu dari perut ibumu dalam keadaan tidak mengetahui sesuatu pun, dan Dia memberi kamu pendengaran, penglihatan dan hati, agar kamu bersyukur” (QS. An-Nahl: 78).

Jadi dapat disimpulkan, penilaian formatif yaitu penilaian untuk mengetahui hasil belajar yang dicapai oleh anak didik setelah menyelesaikan program dalam setelah menyelesaikan program dalam satuan pelajaran pada suatu mata pelajaran tertentu: asumsi yang mendasari evaluasi ini adalah manusia (peserta didik) diciptakan dengan beberapa kelemahan, dan semula tidak mengetahui apa-apa sehingga pengetahuan, sikap dan keterampilan tidak akan abadi kecuali dibiasakan dan diulang-ulang. Sedangkan evaluasi pendidikan merupakan bagian dari pembiasaan dan pengulangan itu. Evaluasi formatif mempunyai:

a. Fungsi

Untuk memperbaiki proses belajar mengarahkan ke arah yang lebih baik dan efisien atau memperbaiki proses pembelajaran tersebut.

b. Tujuan

Untuk mengetahui hingga di mana penguasaan peserta didik tentang penguasaan bahwa pembelajaran yang diajarkan dalam satu program pembelajaran serta sesuai tidak dengan tujuan.

c. Aspek-aspek yang dinilai

Aspek-aspek yang dinilai pada penilaian formatif ialah hasil kemajuan belajar peserta didik yang meliputi, pengetahuan (kognitif, keterampilan (psikomotor), sikap (afektif) terhadap pelajaran yang disajikan.⁴⁷

2. Evaluasi Sumatif

Yaitu evaluasi yang dilakukan untuk mengetahui hasil belajar yang dicapai peserta didik setelah melakukan pembelajaran dalam satu

⁴⁷ Ramayulis, *Ilmu Pendidikan Islam*, (Kalam Mulia: Jakarta, 2001) hlm. 204

caturwulan, satu semester atau akhir tahapan dalam rangka menentukan jenjang berikutnya. Evaluasi ini sejalan dengan semangat Al-Qur'an dalam surah Al-Insyiqaq ayat 19⁴⁸ dan surah Al-Qamar ayat 49.

لَتَرْكَبُنَّ طَبَقًا عَنْ طَبَقٍ ﴿١٩﴾

Artinya: "Sesungguhnya kamu melalui tingkat demi tingkat (dalam kehidupan)" (QS. Al-Insyiqaq:19).

إِنَّا كُلَّ شَيْءٍ خَلَقْنَاهُ بِقَدَرٍ ﴿٤٩﴾

Artinya: "Sesungguhnya Kami menciptakan segala sesuatu menurut ukuran" (QS. Al-Qamar: 49).

Penilaian sumatif mempunyai:

a. Fungsi

Untuk menentukan angka atau nilai peserta didik setelah mengikuti program bahan pelajaran dalam satu semester.

b. Tujuan

Untuk mengetahui taraf hasil belajar yang dicapai oleh peserta didik setelah menyelesaikan proses belajar mengajar dalam semester, akhir tahun.

c. Aspek-aspek yang dinilai

Aspek-aspek yang dinilai ialah kemajuan hasil belajar meliputi, kognitif, psikomotor, dan afektif.

d. Waktu pelaksanaan

Penilaian sebaiknya dilakukan sebelum anak didik mengikuti proses belajar mengajar permulaan atau anak tersebut baru akan mengikuti pendidikan di suatu tingkat tertentu.

3. Evaluasi Diagnostik

Yaitu penilaian yang dilakukan terhadap hasil penganalisan tentang keadaan belajar anak didik baik yang merupakan kesulitan-kesulitan atau hambatan yang ditemui dalam situasi belajar mengajar.

⁴⁸ Abudin Nata, *Ilmu Pendidikan Islam*, (Kencana: Jakarta, 2010) hlm. 310

Asumsi yang mendasar evaluasi ini adalah bahwa pengalaman pahit masa lalu dapat dijadikan “guru” untuk memperbaiki masa depan. Setiap kegiatan dalam proses pembelajaran tidak terlepas dari kesulitan dan hambatan. Apabila seseorang peserta didik dapat menyelesaikan dan memecahkan hambatan dan kesulitan itu, maka ia akan memperoleh kemudahan dalam kegiatan berikutnya. Banyak firman Allah Swt. yang mengisyaratkan asumsi ini, seperti peringatan-Nya dalam cerita-cerita kaum terdahulu yang mana kaum itu hancur dikarenakan membuat kesulitan sendiri,

Salah satu firman Allah Swt.:

يٰۤاَيُّهَا الَّذِيْنَ ءَامَنُوْا اتَّقُوا اللّٰهَ وَتَنْظُرُوْا نَفْسَكُمْ مَّا قَدَّمْتُمْ لِغَدٍ وَاتَّقُوا اللّٰهَ ۚ اِنَّ اللّٰهَ

حَبِيْرٌۢ بِمَا تَعْمَلُوْنَ ﴿١٨﴾

Artinya: “Hai orang-orang yang beriman, bertakwalah kepada Allah dan hendaklah setiap diri memperhatikan apa yang telah diperbuatnya untuk hari esok (akhirat); dan bertakwalah kepada Allah, Sesungguhnya Allah Maha Mengetahui apa yang kamu kerjakan” (QS. Al-Hasyr:18).

Seperti juga jenis evaluasi yang lain evaluasi diagnostik ini mempunyai:

a. Fungsi

Untuk mengetahui masalah-masalah yang diderita atau mengganggu peserta didik, sehingga ia mengalami kesulitan, hambatan atau gangguan ketika mengikuti program bahan pengajaran dalam satu bidang studi tertentu atau keseluruhan bidang studi. Kesulitan anak tersebut harus diusahakan pemecahannya.

b. Tujuan

Untuk membantu kesulitan atau mengatasi hambatan yang dialami anak didik waktu mengikuti kegiatan belajar pada suatu mata pelajaran atau keseluruhan program pembelajaran.

c. Aspek-aspek yang dinilai

Aspek-aspek yang dinilai termasuk hasil belajar yang diperoleh murid, latar belakang kehidupannya, serta semua aspek yang menyangkut kegiatan belajar mengajar.

d. Waktu pelaksanaan

Pelaksanaan evaluasi diagnostik ini sesuai dengan keperluan pembinaan dari suatu lembaga pendidikan, dalam rangka meningkatkan mutu pengetahuan murid-muridnya.

Menurut Muzayyin, meskipun dalam sumber ilmu pendidikan Islam, klasifikasi jenis penilaian di atas tidak kita temukan secara eksplisit, namun dalam praktik dapat diketahui bahwa pada prinsipnya jenis penilaian seperti ini seringkali ditemukan. Di samping itu di dalam pendidikan Islam kita bisa saja mengadopsi hal-hal yang positif yang datang dari luar untuk diterapkan pula dalam pendidikan Islam selama yang diadopsi itu tidak bertentangan dengan prinsip kependidikan dalam Islam.

BAB II

TEKNIK PENYUSUNAN DAN PELAKSANAAN HASIL BELAJAR

A. Pengertian Tes Hasil Belajar

Kata tes berasal dari bahasa Prancis Kuno yang berarti piring untuk menyisihkan logam-logam mulia yang dimaksud di sini adalah dengan menggunakan alat berupa piring akan dapat diperoleh jenis-jenis logam mulia yang bernilai tinggi. Dalam perkembangannya dan seiring kemajuan zaman, tes berarti ujian atau percobaan. Ada beberapa istilah yang memerlukan penjelasan sehubungan dengan uraian di atas yaitu test, testing, tester dan *testee*, yang masing-masing mempunyai pengertian berbeda namun erat kaitannya dengan tes.

1. Testing berarti saat dilaksanakannya pengukuran dan penilaian atau saat pengambilan tes.
2. Tester artinya orang yang melaksanakan tes atau orang yang disertai untuk melaksanakan pengambilan tes terhadap para responden.
3. *Testee* adalah pihak yang sedang dikenai tes.

Ada beberapa pendapat dari beberapa ahli tentang pengertian tes, menurut Anne Anastasi dalam karya tulisnya yang berjudul *Psychological Testing*, yang dimaksud dengan tes adalah alat pengukur yang mempunyai standar yang objektif sehingga dapat digunakan secara meluas, serta dapat digunakan sebagai cara untuk mengukur dan membandingkan keadaan

psikis atau tingkah laku individu.⁴⁹ Menurut Douglas Brown, dalam kitab *Usus al Talimu*, tes adalah cara mengukur pengetahuan individu (*testee*) dalam bidang tertentu⁵⁰

Menurut Lee J. Cronbach dalam bukunya berjudul *Essential of Psychological Testing*, tes merupakan suatu prosedur yang sistematis untuk membandingkan tingkah laku dua orang atau lebih. Sedangkan menurut Goodenough, tes adalah suatu tugas atau serangkaian tugas yang diberikan kepada individu atau kelompok individu, yang dimaksud untuk membandingkan kecakapan satu sama lain.

Dari pengertian dari para ahli tersebut dalam dunia pendidikan dapat disimpulkan bahwa tes adalah cara yang digunakan atau prosedur yang ditempuh dalam rangka pengukuran dan penilaian di bidang pendidikan, yang memberikan tugas dan serangkaian tugas yang diberikan oleh guru sehingga dapat dihasilkan nilai yang melambangkan tingkat laku atau prestasi peserta didik

Tes sebagai salah satu teknik pengukuran dapat didefinisikan *A test will be defined as a systematic procedure for measuring a sample of an individual's behaviour*. Definisi tersebut mengandung dua hal pokok yang perlu diperhatikan dalam memahami makna tes, yaitu pertama adalah kata *systematic procedure* yang artinya bahwa suatu tes harus disusun, dilaksanakan (diadministrasikan) dan diolah berdasarkan aturan-aturan tertentu yang telah ditetapkan. Sistematis di sini meliputi tiga langkah, yaitu:

1. sistematis dalam isi, artinya butir-butir soal (*item*) suatu tes hendaknya disusun dan dipilih berdasarkan kawasan dan ruang lingkup tingkah laku yang akan dan harus diukur atau dites, sehingga tes tersebut benar-benar tingkat validitasnya dapat dipertanggungjawabkan.
2. sistematis dalam pelaksanaan (administrasi) artinya tes itu hendaknya dilaksanakan dengan mengikuti prosedur dan kondisi yang telah ditentukan.

⁴⁹ Abdul Kadir, "Menyusun dan menganalisis tes hasil belajar", Jurnal Al-Ta'dib, vol.8, no.2.2015.

⁵⁰ Chabib Thoha, "Teknik Evaluasi Pendidikan", (Jakarta: Rajawali, 1991), hlm.43

3. sistematis di dalam pengolahannya, artinya data yang dihasilkan dari suatu tes diolah dan ditafsirkan berdasarkan aturan-aturan dan tolak ukur (norma) tertentu.

Kedua adalah *measuring of an individual's is behaviour* yang artinya bahwa tes itu hanya mengukur suatu sampel dari suatu tingkah laku individu yang dites. Tes tidak dapat mengukur seluruh (populasi) tingkah laku, melainkan terbatas pada isi (butir soal) tes yang bersangkutan.

Suatu tes akan berisikan pertanyaan-pertanyaan dan atau soal-soal yang harus dijawab dan atau dipecahkan oleh individu yang dites (*testee*), maka disebut tes hasil belajar (*achievement test*). Hal ini sependapat dengan seorang ahli yang menyatakan bahwa *The type of ability test that describes what a person has learned to do is called an achievement test*. Berdasarkan pendapat itu, tes hasil belajar biasanya terdiri dari sejumlah butir soal yang memiliki tingkat kesukaran tertentu (ada yang mudah, sedang, dan sukar). Tes tersebut harus dapat dikerjakan oleh siswa dalam waktu yang sudah ditentukan. Oleh karena itu, tes hasil belajar merupakan *power test*. Maksudnya adalah mengukur kemampuan siswa dalam menjawab pertanyaan atau permasalahan.⁵¹

Tes dapat dibedakan sebagai berikut:

1. Berdasarkan objek pengukurannya
 - Yaitu terdiri atas kepribadian (*personality test*) dan hasil belajar (*achievement test*)
 - a. Tes kepribadian adalah tes yang ditujukan untuk mengukur salah satu lebih dekat aspek-aspek non intelektual dari susunan mental dan psikologis individu. Yang termasuk dalam jenis tes ini dan sering dilakukan dalam pendidikan adalah pengukuran sikap, pengukuran minat, pengukuran bakat dan tes inteligensi.
 - b. Tes hasil belajar adalah tes yang biasa digunakan untuk mengungkap tingkat pencapaian atau prestasi belajar. Tes hasil belajar dilakukan untuk mengukur sejauh mana hasil belajar yang telah dicapai oleh peserta didik didorong untuk menunjukkan penampilan maksimalnya. Dari penampilan maksimal yang ditujukan dalam jawaban tes hasil belajar, maka dapat diketahui

⁵¹ Abdul Kadir, *Opcit.*,

penguasaan peserta didik terhadap materi yang diajarkan dan dipelajari.⁵²

2. Berdasarkan fungsinya

Berdasarkan fungsinya tes dibedakan menjadi empat jenis antara lain:

a. Tes penempatan

Tes penempatan adalah tes untuk mengukur kemampuan dasar yang dimiliki oleh anak didik kemampuan tersebut dapat dipakai meramalkan kemampuan peserta didik pada masa mendatang, sehingga kepadanya dapat dibimbing, diarahkan atau ditempatkan pada jurusan yang sesuai dengan kemampuan dasarnya.

b. Tes formatif

Tes formatif adalah tes yang dilaksanakan untuk memantau kemajuan belajar peserta didik selama proses belajar berlangsung. Untuk memberikan balikan bagi penyempurnaan program pembelajaran, serta untuk mengetahui kelemahan-kelemahan yang memerlukan perbaikan, sehingga hasil belajar peserta didik dan proses pembelajaran guru menjadi lebih baik.

Jadi sebenarnya penilaian formatif tidak hanya berbentuk tes tertulis dan dilakukan tiap akhir pelajaran, tetapi dapat tertulis dan dilakukan tiap akhir pelajaran, tetapi ada juga dalam bentuk pertanyaan lisan atau tugas-tugas yang diberikan selama pelajaran berlangsung ataupun sesudah pelajaran selesai.

c. Tes *diagnostic*

Tes diagnostik yaitu tes yang digunakan untuk mengetahui kelemahan-kelemahan peserta didik sehingga dapat dilakukan pemberian perlakuan yang tepat.

d. Tes sumatif

Tes sumatif adalah tes yang dilakukan untuk memperoleh data atau informasi sampai di mana penguasaan atau pencapaian belajar peserta didik terhadap bahan pelajaran yang telah dipelajarinya selama jangka waktu tertentu. Tes ini biasanya dilaksanakan pada akhir unit program yaitu akhir semester atau akhir tahun. Hasil tes

⁵² Anas Sudijono. "Pengantar Evaluasi Pendidikan". (Jakarta: Rajawali Pres. 2009) hlm. 73.

sumatif ini dipakai untuk membuat keputusan penting bagi peserta didik, misalnya penentuan kenaikan kelas, kelulusan sekolah, dan membuat keputusan lainnya yang terkait dengan kepentingan peserta didik.⁵³

3. Berdasarkan tingkatannya

a. Tes standar

Tes standar adalah tes yang telah mengalami proses standarisasi, yakni proses validasi dan andal untuk satu tujuan dan bagi suatu kelompok tertentu. Jadi tes standar merupakan tes yang telah memiliki validitas reliabilitas yang cukup tinggi berdasarkan atas percobaan yang telah dilakukan terhadap sampel yang cukup luas dan representatif.

b. Tes nonstandar

Tes nonstandar adalah tes yang disusun oleh seorang pendidik yang belum memiliki keahlian profesional dalam menyusun tes secara baik, mengujicobakan, melakukan analisis sehingga validitas dan reliabilitasnya belum dapat dipertanggungjawabkan.

B. Ciri-Ciri Tes Hasil yang Baik

Setidak-setidaknya ada empat ciri yang harus dimiliki oleh tes hasil belajar, sehingga tes tersebut dapat dinyatakan sebagai tes yang baik yaitu:

1. Validitas

Kata valid sering diartikan dengan tepat, benar, sah, absah. Jadi kata validitas dapat diartikan dengan ketepatan, kebenaran, kesahihan, dan keabsahan. Untuk menetapkan apakah sebuah tes hasil belajar dapat dinyatakan memiliki validitas atau daya ketepatan mengukur ataukah belum dapat dilakukan dengan penganalisan secara rasional (logika) dan dapat pula dilakukan dengan penganalisan secara empiris.

2. Reliabilitas

Kata ini sering diterjemahkan dengan keajekan (*stability*) atau kemantapan (konsistensi). Tes hasil belajar dapat dinyatakan *reliable*,

⁵³ Daryanto, "Evaluasi Pendidikan", (Jakarta: PT Rineka cipta. 2008) hlm.37.

apabila hasil pengukuran yang dilakukan dengan menggunakan tes tersebut secara berulang kali terhadap subjek yang sama, senantiasa menunjukkan hasil yang tetap sama atau yang sifatnya stabil. Ada 3 jenis pendekatan yang digunakan untuk mengetahui apakah sebuah tes hasil belajar itu memiliki reliabilitas yang tinggi atau rendah yaitu:

1. Pendekatan *single test (single trial)*
2. Pendekatan tes *retest*
3. Pendekatan *alternate forms*.

3. Objektif

Objektif berarti tidak adanya unsur pribadi yang memengaruhinya, bukan subjektif. Sebuah tes dapat dikatakan memiliki objektivitas apabila dalam melaksanakan tes tersebut tidak ada faktor subjektif yang memengaruhi terutama dalam sistem skornya. Apabila dikaitkan dengan reliabilitas maka objektivitas menekankan pada ketetapan dalam sistem *skorsing*, sedangkan reliabilitas menekankan pada ketetapan dalam hasil tes.

4. *Practicability* (praktis) dan ekonomis

Bersifat praktis mengandung pengertian bahwa tes hasil tes belajar tersebut dapat dilaksanakan dengan mudah karena tes itu:

- a. Bersifat sederhana, dalam arti tidak memerlukan peralatan yang banyak atau peralatan yang sulit pengadaannya.
- b. Lengkap, dalam arti bahwa tes tersebut telah dilengkapi dengan petunjuk mengenai cara mengerjakannya, kunci jawabannya, dan pedoman *skorsing* serta penentuan nilainya.
- c. Bersifat ekonomis, artinya bahwa tes hasil belajar tersebut tidak memakan waktu yang panjang dan tidak memerlukan tenaga serta biaya yang banyak.⁵⁴

⁵⁴ Suharsimi arikunto, “*Dasar-dasar evaluasi pendidikan*”, (Jakarta: Bumi Aksara, 2012), hlm. 235-238.

Zainal Arifin menyatakan bahwa tes hasil belajar yang baik yaitu:

1. Validitas

Ada 2 unsur penting dalam validitas yaitu:

- a. Validitas menunjukkan suatu derajat kesempurnaan tes sebagai ukuran.
- b. Validitas selalu dihubungkan dengan suatu putusan yang spesifik.

2. Reliabilitas

Adalah derajat konsistensi dari suatu tes. Menurut Kerlinger (1886) mengemukakan reliabilitas dapat diukur dari tiga kriteria yaitu:

- a. *Stability* menunjukkan kestabilan suatu tes dalam mengukur gejala yang sama pada waktu berbeda.
- b. *Dependability*, menunjukkan kestabilan suatu tes seberapa jauh tes dapat diandalkan.
- c. *Predictability* menunjukkan kemampuan suatu tes.

3. Kepraktisan

Dimiyanti dan Mujiono 1994 mengemukakan faktor-faktor yang memengaruhi kepraktisan tes meliputi: kemudahan mengadministrasi, kemudahan menskor, kemudahan interpretasi dan aplikasi, tersedianya bentuk tes yang ekuivalen dan sebanding.⁵⁵

C. Prinsip-Prinsip Dasar dalam Penyusunan Tes Hasil Belajar

Ada beberapa prinsip dasar yang perlu dicermati di dalam menyusun tes hasil belajar agar tes tersebut dapat mengukur tujuan instruksional khusus untuk mata pelajaran yang telah diajarkan, atau mengukur kemampuan dan keterampilan peserta didik yang diharapkan setelah mereka menyelesaikan suatu unit pengajaran tertentu.

Pertama, tes hasil belajar harus dapat mengukur secara jelas hasil belajar (*learning out comes*) yang telah ditetapkan sesuai dengan tujuan instruksional. Kejelasan mengenai pengukuran hasil belajar yang dikehendaki akan memudahkan bagi guru dalam menyusun butir-butir soal tes hasil belajar.

⁵⁵ Zainal Arifin, "Evaluasi pembelajaran", (Bandung: PT Remaja Rosda Karya, 2011), hlm.214-215.

Kedua, butir-butir soal tes hasil belajar harus merupakan sampel yang representatif dari populasi bahan pelajaran yang telah diajarkan, sehingga dapat dianggap mewakili seluruh performa yang telah diperoleh selama peserta didik mengikuti suatu unit pengajaran.

Ketiga, bentuk soal yang dikeluarkan dalam tes hasil belajar harus dibuat bervariasi, sehingga betul-betul cocok untuk mengukur hasil belajar yang berupa keterampilan misalnya, tidak tepat jika hanya menggunakan soal-soal *essay*. Demikian pula untuk mengukur kemampuan menganalisis suatu prinsip, tidak cocok jika digunakan butir-butir soal yang berbentuk *objective test* yang pada dasarnya hanya mengungkap daya ingat peserta didik.

Keempat, tes hasil belajar harus didesain sesuai dengan kegunaannya untuk memperoleh hasil yang diinginkan. Pernyataan tersebut mengandung makna, bahwa desain tes hasil belajar harus disusun relevan dengan kegunaan yang dimiliki oleh masing-masing jenis tes.

Kelima, tes hasil belajar harus memiliki reliabilitas yang dapat diandalkan. Artinya, setelah tes hasil belajar itu dilaksanakan berkali-kali terhadap subjek yang sama, hasilnya selalu sama atau relatif sama. Dengan demikian tes hasil belajar itu hendaknya memiliki keajekan hasil pengukuran yang tidak diragukan lagi.

Keenam, tes hasil belajar di samping harus dapat dijadikan alat pengukur keberhasilan belajar siswa, juga harus dapat dijadikan alat untuk mencari informasi yang berguna untuk memperbaiki cara belajar siswa dan cara mengajar guru itu sendiri.⁵⁶

D. Bentuk-Bentuk Tes Hasil Belajar dan Teknik Penyusunannya

1. Tes Hasil Belajar dalam Bentuk Uraian

Tes uraian sering dikenal dengan istilah tes subjektif, yaitu salah satu jenis tes hasil belajar yang memiliki karakteristik sebagai berikut:

1. Tes berbentuk pertanyaan atau perintah yang menghendaki jawaban berupa uraian atau paparan kalimat yang pada umumnya cukup panjang.

⁵⁶ Surakhmad, “*Pengantar Filsafat Pendidikan dan Cara Belajar*”, (Jakarta: Bina Aksara, 2000), hlm 98.

2. Bentuk pertanyaan itu menuntut kepada *taste* untuk memberikan penjelasan, komentar, penafsiran, membandingkan, membedakan dan sebagainya.
3. Jumlah butir soal umumnya terbatas, yaitu berkisar antara lima sampai dengan sepuluh butir.
4. Butir-butir soal umumnya diawali dengan kata-kata “jelaskan, terangkan, uraikan, mengapa, bagaimana, dll.”

Adapun kelebihan dan kelemahan yang dimiliki oleh tes uraian:

1. Kelebihan tes uraian
 - a. Tes uraian adalah jenis tes hasil belajar yang pembuatannya dapat dilakukan dengan mudah dan cepat. Hal ini disebabkan karena kalimat-kalimat soal pada tes uraian itu adalah cukup pendek, sehingga dalam penyusunannya tidak terlalu sulit dan tidak terlalu banyak memakan waktu, tenaga, pikiran, peralatan, dan biaya.
 - b. Tes uraian dapat memberikan jawaban yang tepat dan benar.
 - c. Penyusunan soal akan dapat mengetahui seberapa jauh tingkat kedalaman dan tingkat penguasaan *testee* dalam memahami materi yang ditanyakan dalam tes tersebut.
 - d. Tes uraian dapat mendorong siswa untuk berani dan terbiasa mengemukakan pendapat dengan menggunakan susunan kalimat dan gaya bahasa yang merupakan hasil olahannya sendiri.
2. Kelemahan tes uraian
 - a. Tes uraian pada umumnya kurang dapat menampung atau mencakup dan mewakili isi dan luasnya materi atau bahan pelajaran yang telah diberikan kepada *testee*, yang seharusnya diujikan dalam tes hasil belajar.
 - b. Cara mengoreksi jawaban soal tes uraian cukup sulit. Karena sekalipun butir soalnya sangat terbatas, namun jawabannya bisa panjang lebar dan sangat bervariasi, sehingga pekerjaan koreksi akan banyak menyita waktu, tenaga, dan pikiran.
 - c. Hasil tes uraian sulit untuk diserahkan kepada orang lain karena pada tes uraian orang yang paling tahu mengenai jawaban yang sempurna adalah penyusun tes itu sendiri.

- d. Daya ketetapan mengukur (validasi) dan daya keajekan mengukur (reliabilitas) yang dimiliki oleh tes uraian pada umumnya rendah sehingga kurang dapat diandalkan sebagai alat pengukur hasil belajar yang baik.⁵⁷

2. Tes Hasil Belajar Bentuk Objektif (*Objective Test*)

Tes objektif (*objective test*) yang juga dikenal dengan istilah tes jawaban pendek (*short answer test*), tes “ya-tidak” (*yes-no test*) dan test model baru (*new type test*), adalah salah satu jenis tes hasil belajar yang terdiri dari butir-butir soal (*items*) yang dapat dijawab oleh *testee* dengan jalan memilih salah satu (atau lebih) di antara beberapa kemungkinan jawaban yang telah dipasangkan pada masing-masing *items*; atau dengan jalan menuliskan (mengisikan) jawabannya berupa kata-kata atau simbol-simbol tertentu pada tempat atau ruang yang telah disediakan untuk masing-masing butir *item* yang bersangkutan.⁵⁸

Sebagai salah satu jenis tes hasil belajar, tes objektif dapat dibedakan menjadi lima golongan, yaitu:

1) Tes Objektif Bentuk Benar-Salah (*True-False Test*)

Tes objektif bentuk *true-false* adalah salah satu bentuk tes objektif di mana butir-butir soal yang diajukan dalam tes hasil belajar itu berupa pernyataan (*statement*), pernyataan mana ada yang benar dan ada yang salah.

Jadi, tes objektif itu bentuknya adalah kalimat atau pernyataan yang mengandung dua kemungkinan jawab: benar atau salah, dan *testee* diminta menentukan pendapatnya mengenai pernyataan-pernyataan tersebut dengan cara seperti yang ditentukan dalam petunjuk cara mengerjakan soal. Maka ada beberapa petunjuk yang perlu diperhatikan dalam menyusun butir-butir soal tes objektif bentuk *true-false*, yaitu:

- a) Seyogianya tuliskan huruf B - S di depan masing-masing pernyataan dan jangan di belakangnya.

⁵⁷ Anas Sudijono, “*Pengantar Evaluasi Pendidikan*”, (Jakarta: PT RajaGrafindo Persada, 2012), hlm. 101

⁵⁸ *Ibid.*, hlm. 106.

- b) Jumlah butir soal hendaknya berkisar antara 10 sampai dengan 20 butir.
- c) Jumlah butir soal yang jawabannya Betul (B) sebaiknya sama atau seimbang dengan jumlah butir soal yang jawabannya Salah (S).
- d) Urutan soal-soal yang jawabannya Betul (B) dan yang jawabannya Salah (S) hendaknya jangan dibuat ajek; buatlah berselang-seling sehingga dapat mencegah timbulnya permainan spekulasi di kalangan *testee*.

Contoh:

B - S - S - B - B - B - S - S - B - S - B - S

- e) Butir-butir soal yang jawabannya Betul (B) sebaiknya tidak mempunyai corak berbeda dari soal-soal yang jawabannya Salah (S).
 - f) Hindarilah pernyataan-pernyataan yang susunan kalimatnya persis seperti yang dimuat dalam buku (bahan tes).
 - g) Dalam menyusun butir-butir soal tes objektif bentuk *true-false* hendaknya dapat dihindari sejauh mungkin, agar tidak ada butir-butir soal yang jawabannya bersifat relatif.⁵⁹
- 2) Tes Objektif Bentuk *Matching*

Tes objektif bentuk *matching* sering dikenal dengan istilah tes menjodohkan, tes mencari pasangan, tes menyesuaikan, tes mencocokkan dan tes mempertandingkan. Tes objektif bentuk *matching* merupakan salah satu bentuk tes objektif dengan ciri-ciri sebagai berikut:

- a) Tes terdiri dari satu seri pertanyaan dan satu seri jawaban.
- b) Tugas *testee* adalah mencari dan menempatkan jawaban-jawaban yang telah tersedia, sehingga sesuai atau cocok atau merupakan pasangan, atau “jodoh” dari pertanyaan.

Contoh:

	Daftar I	Daftar II
1. B	Salat sunah yang dilaksanakan pada tiap malam bulan Ramadan.	A. <i>Istiqa'</i> B. Tarawih

⁵⁹ *Ibid.*, hlm. 107-110.

- | | | |
|---------|---|--------------------------------------|
| 2. | Salat sunah yang dilakukan sewaktu memasuki masjid. | C. Rawatib
D. Mutlak |
| 3. | Salat sunah yang tidak ditentukan Waktunya dan tidak pula ditentukan Bilangan rakaatnya | E. Khauf
F. Istikharah
G. Duha |
| 4. |dan seterusnya..... ⁶⁰ | |

3) Tes Objektif Bentuk *Fill In*

Tes objektif bentuk *fill in* (=bentuk isian) ini biasanya berbentuk cerita atau karangan. Kata-kata penting dalam cerita atau karangan itu beberapa di antaranya dikosongkan (tidak dinyatakan), sedangkan tugas *testee* adalah mengisi bagian-bagian yang telah dikosongkan itu. Dalam penyusunan butir-butir *item* tes objektif bentuk *fill in* perlu diperhatikan pedoman berikut ini:

- a. Jawaban yang harus diisikan ditulis pada lembar jawaban atau pada tempat yang terpisah.
- b. Ungkapan cerita yang dijadikan bahan tes hendaknya disusun secara ringkas dan padat, demi menghemat tempat atau kertas serta waktu penyusunannya.
- c. Sejauh mungkin supaya diusahakan agar butir-butir *item* yang diajukan dalam tes objektif bentuk *fill in* ini adalah butir-butir *item* yang selain mengungkap pengetahuan atau pengenalan juga dapat mengungkap taraf kompetensi lain yang sifatnya lebih mendalam.
- d. Apabila jenis mata pelajaran yang akan diteskan itu memungkinkan, penyajian soal juga dapat dituangkan dalam bentuk gambar, peta dan sebagainya, sehingga kalimat cerita dapat dipersingkat.

Contoh:

Isilah titik-titik berikut ini dengan jawaban yang tepat!

Pengertian pendidikan Islam menurut Syekh Anwar Jundi ialah (1) konsep pendidikan Islam tersebut di atas mengandung pengertian bahwa pendidikan Islam itu berlangsung (2) Syekh Anwar Jundi selanjutnya merumuskan tujuan

⁶⁰ *Ibid.*, hlm. 111-112.

pendidikan Islam, yaitu (3) sedangkan menurut Syekh Dr. 'Athiyah al-Abrasyi, tujuan pendidikan Islam itu ialah (4).⁶¹

4) Tes Objektif Bentuk *Completion*

Tes objektif bentuk *completion* sering dikenal dengan istilah tes melengkapi atau menyempurnakan, yaitu salah satu jenis tes objektif yang memiliki ciri-ciri sebagai berikut.

- Tes tersebut terdiri atas susunan kalimat yang bagian-bagiannya sudah dihilangkan.
- Bagian-bagian yang dihilangkan itu diganti dengan titik-titik (.....).
- Titik-titik itu harus diisi atau dilengkapi atau disempurnakan oleh *testee*, dengan jawaban –(yang oleh tester)– telah dihilangkan.

Contoh:

Isilah titik-titik berikut ini dengan jawaban yang benar dan tepat.

1. Aliran Jabariah terkenal dengan pahamnya.....
 2. Lembaga keilmuan terkenal di masa kejayaan Khalifah Al-Ma'mun bernama.....
 3. Buku Al-Syifa' merupakan ensiklopedi tentang.....⁶²
- #### 5) Tes Objektif Bentuk *Multiple Choice Item*

Tes objektif bentuk *multiple choice item* sering dikenal dengan istilah tes objektif bentuk pilihan ganda, yaitu salah satu bentuk tes objektif yang terdiri atas pertanyaan atau pernyataan yang sifatnya belum selesai, dan untuk menyelesaikan harus dipilih salah satu (atau lebih) dari beberapa kemungkinan jawab yang telah disediakan pada tiap-tiap butir soal yang bersangkutan.

Contoh 1: Bentuk pertanyaan

Pilih satu jawaban yang tepat!

1. Pada saat yang manakah bacaan Talbiyah dibaca oleh para jemaah haji?
 - A. Sai
 - B. Wukuf
 - C. Tawaf

⁶¹ *Ibid.*, hlm.114-115.

⁶² *Ibid.*, hlm. 116-117.

- D. Berangkat ke Arafah
- E. Irham

Contoh 2: Berbentuk pertanyaan (*statement*)

Pilihlah satu jawaban yang tepat!

1. Orang yang menggantikan puasa Ramadan dengan memberi makan kepada fakir miskin, disebut membayar:
 - A. Jariah
 - B. Fidiah
 - C. Sedekah
 - D. Hibah
 - E. Wakaf

Seperti dapat diperiksa pada dua contoh di atas, maka tes objektif bentuk *multiple choice item* terdiri atas dua bagian, yaitu:

- a) *Item* atau soal, yang dapat berbentuk pertanyaan dan dapat pula berbentuk pernyataan.
- b) *Option* atau alternatif, yaitu kemungkinan-kemungkinan jawab yang dapat dipilih oleh *testee*.

Dalam perkembangannya, sampai saat ini tes objektif bentuk *multiple choice item* dapat dibedakan menjadi sembilan model, yaitu:

1. Model melengkapi lima pilihan
2. Model asosiasi dengan lima atau empat pilihan
3. Model melengkapi berganda
4. Model analisis hubungan antarhal
5. Model analisis khusus
6. Model hal kecuali
7. Model hubungan dinamik
8. Model pemakaian diagram, grafik, peta atau gambar.⁶³

⁶³ *Ibid.*, hlm. 118-120.

E. Teknik Pelaksanaan Tes Belajar

1. Teknik Pelaksanaan Tes Tertulis

a. Merujuk pada Silabus/SAP

Biasanya suatu sekolah/lembaga pendidikan telah mempunyai silabus dan SAP untuk setiap mata pelajaran. Silabus berisikan pokok-pokok bahasan yang akan diajarkan dalam satu semester. Silabus diperlukan pada waktu membuat kisi-kisi soal agar soal yang dibuat mewakili semua pokok bahasan yang ada sehingga akhirnya dapat dilihat apakah tujuan pembelajaran tercapai atau tidak.

b. Menyusun Kisi-Kisi Soal

Menyusun kisi-kisi merupakan langkah awal yang harus dilakukan setiap kali menyusun tes dan menulis soal. Dengan adanya kisi-kisi, penyusunan soal dapat menghasilkan tes yang relatif sama. Kisi-kisi tes adalah suatu format atau matriks yang memuat kriteria butir soal yang diperlukan dalam menyusun tes. Oleh karena itu, kisi-kisi yang baik harus memenuhi beberapa kriteria, yaitu:

1. Dapat menggambarkan keterwakilan isi kurikulum;
2. Komponen yang membentuk kisi-kisi harus jelas, rinci, dan mudah dipahami;
3. Setiap indikator dapat dituliskan butir soalnya.

c. Menyusun Soal

Soal dapat disusun dalam bentuk tes objektif maupun tes esai. Sebagai bahasan dalam tulisan ini penulis memilih bentuk tes objektif dengan bentuk soal tes pilihan ganda. Jumlah soal yang disusun harus melebihi jumlah yang dibutuhkan dan disusun sesuai kisi-kisi. Sukar atau mudahnya suatu soal bukan semata-mata ditentukan oleh materi soal, akan tetapi ditentukan juga oleh teknik penyusunannya. Beberapa butir pernyataan yang merupakan bagian pokok dalam pedoman umum penulisan butir soal tes pilihan ganda adalah sebagai berikut:

- a) Butir soal harus sesuai dengan indikator.

- b) Pokok soal dan pilihan jawaban harus dirumuskan secara jelas, singkat, padat, dan tegas, sehingga perumusan tersebut hanya mencakup pernyataan yang diperlukan saja.
 - c) Pokok soal jangan memberi petunjuk ke arah jawaban yang benar.
 - d) Pokok soal dan pilihan jawaban tidak mengandung pernyataan yang bersifat negatif ganda.
 - e) Pilihan jawaban yang merupakan kunci jawaban harus menunjukkan kebenaran mutlak dan terbaik.
 - f) Pilihan jawaban harus homogen dan logis secara materi dan bahasa.
 - g) Panjang rumusan pilihan jawaban harus relatif sama.
 - h) Pilihan jawaban sebaiknya jangan memakai bunyi “semua pilihan jawaban di atas salah” atau “semua pilihan jawaban di atas benar”.
 - i) Pilihan jawaban berbentuk angka harus disusun berdasarkan urutan kecil ke besar atau sebaliknya.
- d. Melaksanakan Uji Coba Tes

Agar memperoleh soal/tes yang baik maka soal/tes tersebut harus diuji coba terlebih dahulu dan hasilnya dianalisis sehingga memenuhi syarat-syarat tes yang baik. Peserta uji coba misalnya adalah siswa, maka siswa tersebut harus mempunyai status sama dengan peserta tes yang sebenarnya.

e. Membuat Skor

Setelah soal diuji coba maka selanjutnya dibuat skor masing-masing siswa (peserta yang diuji coba). Bila siswa menjawab benar diberi skor 1, dan bila siswa menjawab salah atau tidak menjawab diberi skor 0. Semua skor yang diperoleh untuk setiap siswa dibuat dalam bentuk tabel.⁶⁴

2. Teknik Pelaksanaan Tes Lisan

Tes lisan merupakan tes yang berupa pertanyaan-pertanyaan yang diberikan guru secara lisan dan peserta didik merespons pertanyaan

⁶⁴ Abdul Kadir, *Opcit.*, hlm. 72-73.

tersebut secara lisan juga, sehingga menumbuhkan sikap berani berpendapat. Jawaban dari pertanyaan tersebut dapat berupa fakta, frase, kalimat, maupun paragraf.⁶⁵

Ada beberapa petunjuk yang dapat digunakan dalam pelaksanaan tes lisan, di antaranya yaitu:

- Sebelum tes lisan dilaksanakan, seyogianya tester sudah melakukan inventarisasi sebagai jenis soal yang akan diajukan kepada *testee* dalam tes lisan tersebut, sehingga tes lisan dapat diharapkan memiliki validitas yang tinggi, baik dari segi isi maupun konstruksinya.
- Setiap butir soal yang telah ditetapkan untuk diajukan dalam tes lisan itu, juga harus disiapkan sekaligus pedoman atau ancar-ancar jawaban betulnya. Hal ini dimaksudkan agar tester di samping mempunyai kriteria yang pasti dalam memberikan skor atau nilai kepada *testee* atas jawaban yang mereka berikan dalam tes lisan tersebut, juga tidak akan terpukau atau terkecoh dengan jawaban panjang lebar yang diberikan oleh *testee*, yang menurut anggapan *testee* merupakan jawaban betul dan tepat, padahal menurut kriteria yang ditentukan sesungguhnya sudah menyimpang atau tidak ada hubungannya dengan soal yang diajukan kepada *testee*.
- Jangan sekali-sekali menentukan skor atau nilai hasil tes lisan setelah seluruh *testee* menjalani tes lisan. Skor atau nilai hasil tes lisan harus sudah dapat ditentukan di saat masing-masing *testee* selesai dites. Hal ini dimaksudkan agar pemberian skor atau nilai hasil tes lisan yang diberikan kepada *testee* itu tidak dipengaruhi oleh jawaban yang diberikan oleh *testee* yang lain.
- Tes hasil belajar yang dilaksanakan secara lisan hendaknya jangan sampai menyimpang atau berubah arah dari evaluasi menjadi diskusi. Tester harus senantiasa menyadari bahwa *testee* yang ada di hadapannya adalah *testee* yang sedang diukur dan dinilai prestasi belajarnya setelah mereka menempuh proses pelajaran dalam jangka waktu tertentu.

⁶⁵ Umi Salamah, “Penjaminan Mutu Penilaian Pendidikan”, Jurnal Evaluasi Vol. 2, No. 1, 2018, hlm. 284.

- Dalam rangka menegakkan prinsip objektivitas dan prinsip keadilan, dalam tes yang dilaksanakan secara lisan itu, tester hendaknya jangan sekali-kali memberikan angin segar atau memancing-mancing dengan kata-kata, kalimat-kalimat, atau kode-kode tertentu yang sifatnya menolong *testee* tertentu dengan alasan kasihan atau karena tester menaruh rasa simpati kepada *testee* yang ada di hadapannya itu.
- Tes lisan harus berlangsung secara wajar. Pernyataan tersebut mengandung makna bahwa tes lisan itu jangan sampai menimbulkan rasa takut, gugup atau panik dikalangan *testee*.
- Sekalipun sulit untuk diwujudkan, namun sebaiknya tester mempunyai pedoman yang pasti, berapa waktu yang disediakan bagi setiap peserta tes dalam menjawab soal-soal atau pertanyaan-pertanyaan pada tes lisan tersebut.
- Pertanyaan-pertanyaan yang diajukan dalam tes lisan hendaknya dibuat bervariasi, dalam arti bahwa sekalipun inti persoalan yang ditanyakan itu sama, namun cara pengajuan pertanyaannya dibuat beragam.

Sejauh mungkin dapat diusahakan agar tes lisan itu berlangsung secara individual agar tidak memengaruhi mental *testee* yang lain.⁶⁶

3. Prosedur/Teknik Pelaksanaan Tes Perbuatan

Tes perbuatan pada umumnya digunakan untuk mengukur taraf kompetensi yang bersifat keterampilan (psikomotorik), di mana penilaiannya dilakukan terhadap proses penyelesaian tugas dan hasil akhir yang dicapai oleh tester setelah melaksanakan tugas tersebut.

Karena tes ini bertujuan ingin mengukur keterampilan, maka sebaiknya tes perbuatan ini dilaksanakan secara individual. Hal ini dimaksudkan agar masing-masing individu yang dites akan dapat diamati dan dinilai secara pasti, sejauh mana kemampuan atau keterampilannya

⁶⁶ Anas Sudijono, “*Pengantar Evaluasi Pendidikan*” (Jakarta: RajaGrafindo Persada, 2001), hlm. 154-156.

dalam melaksanakan tugas yang diperintahkan kepada masing-masing individual tersebut.⁶⁷

Dalam melaksanakan tes perbuatan itu, ada beberapa hal yang perlu diperhatikan oleh tester:

- a. Pertama, tester harus mengamati dengan secara teliti, cara yang ditempuh oleh *testee* dalam menyelesaikan tugas yang telah ditentukan.
- b. Kedua, agar dapat dicapai kadar objektivitas setinggi mungkin, hendaknya tester jangan berbicara atau berbuat sesuatu yang dapat memengaruhi *testee* yang sedang mengerjakan tugas tersebut.
- c. Ketiga, dalam mengamati *testee* yang sedang melaksanakan tugas itu, hendaknya tester telah menyiapkan instrumen berupa lembar penilaian yang di dalamnya telah ditentukan hal-hal apa sajakah yang harus diamati dan diberikan penilaian.

Tes tindakan adalah tes yang persoalan atau pertanyaan disampaikan dalam bentuk tugas yang harus dikerjakan oleh peserta didik. Alat yang dapat digunakan tes ini adalah berupa observasi atau pengamatan terhadap tingkah laku tersebut, yang hasilnya kemudian diserahkan pada guru.⁶⁸

⁶⁷ M.Purwanto Ngalim, “*Prinsip-Prinsip dan Teknik Pengajaran*”, (Bandung: PT Remaja Roskarya, 2006), hlm. 156-157.

⁶⁸ Khaeruddin, “*Administrasi, Analisis Butir, dan Kaidah Penulisan Tes*”, Jurnal Madaniyah, Vol. 1 Ed XII, 2017, hlm. 104.

BAB III

TEKNIK PENGANALISISAN ITEM TES HASIL BELAJAR

A. Pengantar

Adapun yang dimaksud dengan *item* atau soal-soal yang baik adalah yang tingkat kesukarannya dapat diketahui tidak terlalu sukar dan tidak terlalu mudah. Sebab tingkat kesukaran *item* memiliki korelasi dengan daya pembeda. Bilamana *item* memiliki tingkat kesukaran maksimal, maka daya pembedanya akan rendah, demikian pula bila *item* itu terlalu mudah juga tidak akan memiliki daya pembeda. Karenanya, sebaiknya tingkat kesukaran soal itu dipertahankan dalam batas yang mampu memberikan daya pembeda.⁶⁹ Analisis butir soal atau *item-item* dapat digunakan secara kualitatif dan kuantitatif yang masing-masing mempunyai keunggulan dan kelemahan. Oleh karenanya, teknik terbaik adalah menggabungkan keduanya:

1. Teknik Analisis secara Kualitatif

Ada beberapa teknik yang digunakan untuk menganalisis secara kualitatif, yaitu dengan teknik moderator dan teknik panel. Teknik moderator yaitu menganalisis dengan cara berdiskusi yang mana ada satu orang sebagai pencegah. Dengan teknik ini, setiap butir soal didiskusikan secara bersama-sama dengan beberapa ahli seperti guru yang mengajarkan

⁶⁹ M. Chabib Thoha, MA, *Teknik Evaluasi Pendidikan*, (Jakarta: CV Rajawali, 1991), hlm. 145.

materi, ahli materi, penyusun atau pengembang kurikulum. Teknik panel yaitu satu teknik menelaah butir soal berdasarkan kaidah penulisan butir soal. Kaidah itu di antaranya materi, konstruksi, bahasa atau budaya, kebenaran kunci jawaban atau pedoman penskoran. Dalam penganalisan butir soal dengan teknik kualitatif, penggunaan format penelaahan soal akan sangat membantu dan mempermudah prosedur pelaksanaannya.

2. Teknik Analisis secara Kuantitatif

Analisis soal dengan kuantitatif yaitu penelaahan butir soal didasarkan pada data empiris. Di mana data empiris ini didapatkan dari soal yang sudah diujikan. Ada 2 pendekatan dalam analisis secara kuantitatif, yaitu pendekatan secara klasik dan modern. Analisis soal secara klasik adalah proses penelaahan butir soal melalui informasi dari jawaban peserta didik tes guna meningkatkan mutu butir soal yang bersangkutan dengan menggunakan teori tes klasik. Aspek yang sangat penting untuk ditelaah secara klasik ini terlihat dari segi, tingkat kesukaran, daya pembeda butir, fungsi pengecoh pada setiap pilihan jawaban. Reliabilitas dan validitas soal untuk mengetahui lebih dalam apakah soal-soal (*item*) yang diidentifikasi menjalankan fungsinya sebagai alat pengukur hasil belajar yang memadai, penulis akan membahas lebih dalam tentang menganalisis soal secara klasik yaitu, mengetahui taraf kesukaran soal, daya pembeda serta fungsi distraktor.

B. Teknik Menganalisis *Item* Tes Hasil Belajar

a. Analisis Kesukaran Soal

Tingkat kesulitan *item* atau disebut juga indeks kesulitan *item* menurut Sukardi (2011: 136) adalah angka yang menunjukkan proporsi siswa yang menjawab betul dalam satu soal yang dilakukan dengan menggunakan tes objektif.⁷⁰ Asumsi yang digunakan untuk memperoleh kualitas soal yang baik, di samping memenuhi validitas dan reliabilitas, adalah adanya keseimbangan dari tingkat kesulitan soal tersebut. Keseimbangan yang dimaksudkan adalah adanya soal-soal yang termasuk

⁷⁰ Jurnal analisis tingkat kesukaran, daya pembeda dan fungsi distractor soal ujian semester ganjil mata pelajaran produktif di SMK negeri 1 Indralaya Utara tahun pelajaran 2012/2013

mudah, sedang dan sukar secara proporsional. Tingkat kesukaran soal dipandang dari kesanggupan atau kemampuan siswa dalam menjawabnya, bukan dilihat dari sudut guru sebagai pembuat soal.⁷¹ Soal yang baik adalah soal yang tidak terlalu mudah atau tidak terlalu sukar. Soal yang terlalu mudah tidak merangsang siswa untuk mempertinggi usaha memecahkannya. Sebaliknya soal yang terlalu sukar akan menyebabkan siswa menjadi putus asa dan tidak mempunyai semangat untuk mencoba lagi karena di luar kemampuannya.⁷²

Fungsi tingkat kesukaran butir soal biasanya dikaitkan dengan tujuan tes. Misalnya untuk keperluan ujian semester digunakan butir soal yang memiliki tingkat kesukaran sedang, untuk keperluan seleksi digunakan butir soal yang memiliki tingkat kesukaran tinggi/sukar, dan untuk keperluan diagnostik biasanya digunakan butir soal yang memiliki tingkat kesukaran rendah/mudah. Menganalisis tingkat kesukaran soal adalah untuk menentukan kualitas soal yang baik, mengetahui klasifikasi soal mudah, sedang dan sukar.⁷³ Apabila seluruh *testee* tidak dapat menjawab soal dengan betul (karena terlalu sukar) tidak dapat disebut sebagai *item* yang baik. Atau apabila seluruh *testee* dapat menjawab dengan betul (karena soal terlalu mudah) juga tidak dapat dimasukkan dalam kategori yang baik.⁷⁴

1. Analisis Kesukaran Soal Pilihan Ganda

Seperti yang diketahui, bahwa jenis soal itu bermacam-macam di antaranya soal pilihan ganda dan esai. Di sini akan dijelaskan analisis kesukaran soal pada soal pilihan ganda. Bilangan yang menunjukkan sukar dan mudahnya suatu soal disebut indeks kesukaran (*difficulty indeks*) Besarnya indeks kesukaran antara 0,00–1,00. Indeks kesukaran ini menunjukkan taraf kesukaran soal, sehingga soal dengan indeks 0,00 menunjukkan bahwa soal itu terlalu sukar, sebaliknya indeks 1,0 menunjukkan bahwa soalnya

⁷¹ Sulistyorini, *Evaluasi Pendidikan Dalam Meningkatkan Mutu Pendidikan*, (Yogyakarta: Teras, 2009), hlm.174

⁷² Suharsimi Arikunto, *Dasar-Dasar Evaluasi Pendidikan*, (Yogyakarta: PT. BINA AKSARA, 1984), hlm. 159

⁷³ Jurnal Menyusun Dan Menganalisis Tes Hasil Belajar Vol.8.No.2, Juli-Desember,2015

⁷⁴ Purwanto, *Evaluasi Hasil Belajar*, (Yogyakarta: Pustaka Pelajar,2009), hlm. 97.

terlalu mudah. Dalam evaluasi, indeks kesukaran ini diberi simbol P , singkatan dari “*proporsia*”. Angka indeks kesukaran *item* dapat diperoleh dengan rumus yang dikemukakan oleh Du Bois, yaitu ⁷⁵:

$$P = \frac{Np}{N}$$

di mana:

P : Proporsi atau proporsia atau angka indeks kesukaran *item*.

Np : Banyaknya *testee* yang dapat menjawab dengan betul terhadap butir *item*.

N : Jumlah *testee* yang mengikuti tes hasil belajar.

Rumus lainnya adalah:

$$P = \frac{B}{JS}$$

dengan

P : Indeks kesukaran.

B : Banyaknya siswa yang menjawab soal itu dengan benar.

JS : Jumlah seluruh siswa peserta tes.

Sedangkan kriteria yang digunakan adalah makin kecil indeks yang diperoleh, makin sulit soal tersebut. Sebaliknya, makin besar indeks yang diperoleh, makin mudah soal tersebut.

Perlu diketahui bahwa soal-soal yang terlalu mudah atau terlalu sukar, lalu tidak berarti tidak boleh digunakan. Hal ini tergantung dari penggunaannya. Jika dari pengikut yang banyak, kita menghendaki yang lulus hanya sedikit, kita ambil siswa yang paling top. Untuk itu lebih baik diambilkan soal-soal tes yang sukar. Sebaliknya, jika kekurangan pengikut ujian, kita pilihkan soal-soal yang mudah. Selain itu, soal yang sukar akan menambah semangat belajar bagi siswa yang pandai, sedangkan soal-soal yang terlalu mudah, akan membangkitkan semangat siswa yang tidak pandai.

Maka dari itu, dalam menyusun suatu naskah ujian sebaiknya digunakan butir soal yang mempunyai tingkat kesukaran berimbang,

⁷⁵ Anas Sudjiono, *Pengantar Evaluasi Pendidikan*, (Jakarta: Rajawali Press, 2016), hlm. 371

yaitu soal berkategori sukar sebanyak 25%, kategori sedang 50%, dan kategori mudah 25%.

Analisis Kesukaran Soal Tes Uraian atau *Essay*.⁷⁶ Tidak hanya pada soal pilihan ganda saja kita dapat menganalisis tingkat kesukaran soal, begitupun pada soal uraian atau esai yang mana indeks tingkat kesukaran ini umumnya dinyatakan dalam bentuk proporsi yang besarnya kisaran 0,00-1,00. Yang mana jika semakin besar indeks tingkat kesukaran yang diperoleh, maka semakin mudah soal itu. Karena fungsi kesukaran soal biasanya dikaitkan dengan tujuan tes. Misalnya untuk keperluan ujian semester digunakan butir soal yang memiliki tingkat kesukaran sedang, untuk keperluan seleksi digunakan butir soal yang memiliki tingkat kesukaran tinggi, dan untuk keperluan diagnostik biasanya digunakan butir soal yang memiliki tingkat kesukaran rendah. Dan untuk mengetahui tingkat kesukaran soal bentuk uraian digunakan rumus berikut ini:

$$\text{Mean} = \frac{\text{Jumlah skor siswa peserta tes pada butir soal tertentu}}{\text{Banyak siswa yang mengikuti tes}}$$

Hasil perhitungan dengan menggunakan rumus di atas menggambarkan tingkat kesukaran soal itu. Klasifikasi tingkat kesukaran soal dapat dicontohkan seperti berikut:

- 0,00 – 0,30 : Soal tergolong sukar
- 0,31 – 0,70 : Soal tergolong sedang
- 0,71 – 1,00 : Soal tergolong mu

Tingkat kesukaran butir soal memiliki 2 kegunaan, yaitu bagi siswa sebagai pengujian dan pengajaran. Bagi guru, sebagai pengenalan konsep terhadap pembelajaran ulang dan memberi masukan kepada siswa tentang hasil belajar mereka, memperoleh informasi tentang penekanan kurikulum atau mencurigai terhadap butir soal yang biasa. Adapun kegunaannya bagi pengujian dan pengajaran yaitu, sebagai pengenalan konsep yang diperlukan untuk diajarkan ulang, tanda-tanda terhadap kelebihan dan kelemahan kurikulum sekolah, memberi masukan kepada siswa, tanda-tanda

⁷⁶ Daryanto, *Evaluasi Pendidikan*, (Jakarta: Rineka Cipta, 2012), hlm.180

kemungkinan ada soal yang biasa dan merakit tes yang memiliki ketepatan data soal.

b. Analisis Daya Pembeda

Dalam mengetahui intensitas sebuah soal dalam hal kesukaran dibutuhkan sebuah daya pembeda, yaitu kemampuan antara butir soal dapat membedakan antara peserta didik yang menguasai materi yang diujikan dan peserta didik yang belum menguasai materi yang diujikan. Menurut Zainul, daya beda butir soal ialah indeks yang menunjukkan tingkat kemampuan butir soal membedakan kelompok yang berprestasi tinggi dari kelompok yang berprestasi rendah di antara para peserta tes. Angka yang menunjukkan besarnya daya pembeda disebut indeks diskriminasi. Yang berkisar antara 0,00 sampai 1,00. Pada indeks ini kemungkinan adanya tanda negatif manakala suatu tes terbalik menunjukkan kualitas tes yaitu anak pandai disebut tidak pandai dan sebaliknya.

Bagi suatu soal yang dapat dijawab benar oleh siswa pandai maupun tidak pandai, maka soal itu tidak baik karena tidak mempunyai daya pembeda. Demikian sebaliknya, jika siswa yang pandai maupun tidak pandai tidak bisa menjawab benar, soal tersebut juga tidak baik karena tidak mempunyai daya pembeda.

Tes yang baik adalah tes yang dapat dijawab dengan benar oleh siswa yang pandai saja. Contohnya, jika suatu kelompok anak yang berprestasi tinggi dapat menjawab dengan benar suatu tes dan seluruh atau hampir suatu kelompok yang berprestasi rendah menjawab salah, dikatakan bahwa soal itu memiliki indeks diskriminasi (D) terbesar. Sebaliknya jika kelompok yang berprestasi rendah seluruhnya menjawab soal dengan benar sedangkan kelompok berprestasi tingginya menjawab dengan salah, maka indeks diskriminasi (D) soal tersebut -1,00. Sedangkan jika antara kedua kelompok sama-sama menjawab dengan benar, berarti indeks diskriminasi (D) soal tersebut 0,00 atau tidak memiliki daya

pembeda.⁷⁷ Untuk mengetahui besar kecilnya angka indeks diskriminasi *item* dapat digunakan 2 macam rumus berikut ini.⁷⁸

a. Rumus Pertama

$$D = PA - PB \text{ atau } D = PH - PL$$

Di mana:

D : *Discriminatory power* (angka indeks diskriminasi *item*)

PA atau PH : Proporsi *testee* kelompok atas yang dapat menjawab dengan betul butir *item* yang bersangkutan.

Dapat diperoleh dengan rumus:

$$PA = PH = \frac{BA}{JA}$$

dengan:

BA : Banyaknya *testee* kelompok atas yang menjawab dengan betul butir *item* bersangkutan.

JA : Jumlah *testee* yang termasuk dalam kelompok atas

PB atau PL : Proporsi *testee* kelompok bawah yang dapat menjawab dengan betul butir *item* yang bersangkutan.

Dapat diperoleh dengan rumus:

$$PB = PL = \frac{BB}{JB}$$

dengan:

BB : Banyaknya *testee* kelompok bawah yang dapat menjawab dengan benar butir *item* yang bersangkutan

JB : Jumlah *testee* yang termasuk dalam kelompok bawah.

⁷⁷ Syamsudin, *Pengukuran Daya Pembeda, Taraf Kesukaran, dan Pol Jawaban Tes (Analisis Butir Soal)*, Jurnal Ilmu Tarbiyah "At-Tajdid", Vol. 1, No.2, Juli 2012, hlm. 189.

⁷⁸ Anas Sudjiono, *Pengantar Evaluasi Pendidikan*, (Jakarta: Raja Wali Pers, 2009) hlm. 390-391

- b. Rumus Kedua, angka indeks diskriminasi *item* diperoleh dengan menggunakan teknik korelasi Phi (ϕ) dengan rumus,

$$\phi = PH - PL/2 \cdot (P)(q)$$

Di mana:

ϕ : Angka indeks korelasi Phi, yang dalam hal ini dianggap sebagai angka indeks diskriminasi *item*.

PH : Proportion of the higher group

PL : Proportion of the lower group

2 : Bilangan konstan

P : Proporsi seluruh *testee* yang jawabannya betul

Q : Proporsi seluruh *testee* yang jawabannya salah, di mana $q = (1 - p)$

Cara Menentukan Daya Pembeda⁷⁹ membedakan menjadi kelompok kecil (kurang dari 100) dan kelompok besar (100 ke atas):

- i. Kelompok kecil (kurang 100), seluruh kelompok tes terbagi 2 sama besar, separuh kelompok atas dan separuh kelompok bawah
- ii. Kelompok besar (100 ke atas), untuk memudahkan analisis cukup diambil kedua kutub atas dan bawahnya saja, masing-masing 27% sebagai JA dan JB nya. Dari tabel kelompok atas dan bawah itu dicari menggunakan rumus:

$$D = \left(\frac{Ba}{JA} \right) - \left(\frac{Bb}{JB} \right) = Pa - Pb$$

dengan

D : Daya Pembeda

J : Jumlah Peserta

JA : Jumlah Peserta Atas

JB : Jumlah Peserta Bawah

Bb : Jumlah Peserta Kelompok bawah menjawab benar

Ba : Jumlah peserta kelompok atas menjawab benar

⁷⁹ Syamsudin, *Pengukuran Daya Pembeda, Taraf Kesukaran, dan Pola Jawaban Tes (Analisis Butir Soal)*, hlm. 190

$PB - \frac{BB}{JB}$: Proporsi peserta kelompok bawah yang menjawab benar

$PA - \frac{BA}{JA}$: Proporsi peserta kelompok atas yang menjawab benar.

Sebagai tindak lanjut atas hasil analisis daya pembeda adalah⁸⁰:

- a. Butir-butir *item* yang sudah memiliki daya pembeda baik (*satisfactory, good, dan excellent*) hendaknya dimasukkan dalam bank soal. Dan dapat dikeluarkan lagi pada tes berikutnya karena kualitasnya sudah cukup memadai.
- b. Butir-butir *item* yang daya pembedanya masih rendah (*poor*), ada 2 kemungkinan:
 1. Ditelusuri untuk kemudian diperbaiki, dan setelah diperbaiki dapat diajukan lagi dalam tes hasil belajar, yang kemudian dianalisis lagi apakah meningkat atau tidak.
 2. Dibuang atau didrop dan tidak dikeluarkan lagi untuk tes hasil belajar.
- c. Khusus butir-butir *item* yang angka indeks diskriminasi *item*-nya bertanda negatif, sebaiknya pada tes hasil belajar tidak usah dikeluarkan lagi, sebab butir yang demikian kualitasnya sangat jelek (*testee* yang pandai lebih banyak menjawab salah ketimbang *testee* yang tidak pandai, justru hanya sedikit yang menjawab salah).

Daya pembeda *item* adalah kemampuan suatu butir *item* tes hasil belajar untuk dapat membedakan (mendiskriminasi) antara *testee* yang berkemampuan tinggi (pandai), dengan *testee* yang berkemampuan rendah (bodoh). Mengetahui daya pembeda *item* itu penting sekali, sebab salah satu dasar yang dipegang untuk menyusun butir-butir *item* tes hasil belajar adalah adanya anggapan, bahwa kemampuan antara *testee* yang satu dengan *testee* yang lain itu berbeda-beda, dan bahwa butir-butir *item* tes hasil belajar itu haruslah mampu memberikan hasil tes yang mencerminkan adanya perbedaan-perbedaan kemampuan yang terdapat di kalangan *testee* tersebut. Daya pembeda *item* dapat diketahui dengan cara

⁸⁰ Anas Sudjiono, *Pengantar Evaluasi Pendidikan*, (Jakarta: Raja wali Pers,2009) hlm. 409.

melihat besar kecilnya angka indeks diskriminasi *item*, yaitu angka yang menunjukkan besar kecilnya daya pembeda (diskriminasi *power*) butir *item*. Untuk mengetahui angka indeks ini dapat menggunakan rumus:

$$D = PA - PB \text{ atau } D = PH - PL$$

Di mana:

D = discriminatory power

PA atau PH = proporsi *testee* kelompok atas yang dapat menjawab betul butir *item* yang bersangkutan.

PA atau PH ini dapat diperoleh dengan rumus:

QUOTE

Di mana:

BA = banyaknya *testee* kelompok atas yang dapat menjawab dengan betul butir *item* yang bersangkutan.

JA = jumlah *testee* yang termasuk dalam kelompok atas.

PH atau PL = proporsi tes kelompok bawah yang dapat menjawab dengan betul butir *item* yang bersangkutan.

PH atau PL diperoleh dengan cara:

QUOTE

Di mana:

BB = banyaknya *testee* kelompok bawah yang dapat menjawab dengan betul butir *item* yang bersangkutan.

JA = jumlah *testee* yang termasuk dalam kelompok bawah.

Rumus kedua untuk mengetahui angka indeks daya pembeda adalah dengan menggunakan teknik korelasi Phi (\emptyset).

QUOTE

Di mana:

\emptyset = angka indeks korelasi Phi;

p = proporsi seluruh *testee* yang jawabannya betul;

PH = proportion of the higher group;

PL = proportion of the lower group; 2 = bilangan konstan;

q = proporsi seluruh *testee* yang jawabannya salah, di mana

q = (1-p) (jurnal dr cerlilia)

Dalam kegiatan analisis kualitas tes dan butir soal terdapat manfaat daya pembeda butir soal sebagaimana kami kutip berdasarkan pendapat Karjono Natar berikut ini.

- 1) Untuk meningkatkan mutu setiap butir soal melalui data empirisnya. Berdasarkan indeks daya pembeda, setiap butir soal dapat diketahui apakah butir soal itu baik, direvisi, atau ditolak.
- 2) Untuk mengetahui seberapa jauh setiap butir soal dapat mendeteksi/membedakan kemampuan siswa, yaitu siswa yang 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” seperti berikut ini:
 - (a) Kunci jawaban butir soal itu tidak tepat
 - (b) Butir soal itu memiliki 2 (dua) atau lebih kunci jawaban yang benar
 - (c) Kompetensi yang diukur tidak jelas
 - (d) Pengecoh tidak berfungsi
 - (e) Materi yang ditanyakan terlalu sulit, sehingga banyak siswa yang menebak
 - (f) Sebagian besar siswa yang memahami materi yang ditanyakan berpikir ada yang salah informasi dalam butir soalnya.⁸¹

c. Analisis Fungsi Distraktor

Dalam setiap tes objektif selalu digunakan alternatif jawaban yang mengandung dua unsur sekaligus, yaitu jawaban tepat dan jawaban yang salah sebagai penyesat (distraktor).⁸² Pengertian distraktor yaitu, “*Distractor are classified as the incorrect answer in a multiple-choice question.*”⁸³ Dalam setiap tes objektif selalu digunakan alternatif jawaban

⁸¹ Karjono Natar, *Panduan Analisis Butir Soal* (Lampung: UNILA Press, 2011), hlm 12

⁸² Thoha, M. Chabib. *Teknik Evaluasi Pendidikan*. Jakarta: RajaGrafindo Persada, 1996. hlm. 149

⁸³ Shafizan Sabri, *Item Analysis of student Comprehensive Test for Research Teaching Begiliner String Ensemble Using Model Based Teaching Among Music Students In Public Universities*”, *International Journal of Rducation and Research*, 1 (12), hlm. 1.

yang mengandung 2 unsur sekaligus, yaitu jawaban tepat dan jawaban yang salah sebagai penyesat (*distraktor*).

Tujuan pemakaian distraktor menurut Thoha (2001: 149) adalah mengecohkan mereka yang kurang mampu (tidak tahu) untuk dapat dibedakan dengan yang mampu. Arikunto (2008:220) menjelaskan bahwa sebuah distraktor (pengecoh) dapat dikatakan berfungsi dengan baik apabila distraktor tersebut mempunyai daya tarik yang besar bagi pengikut-pengikut tes yang kurang memahami konsep atau kurang menguasai bahan. Tujuan pemakaian distraktor ini adalah mengecohkan mereka yang kurang mampu atau tidak tahu untuk dapat dibedakan dengan yang mampu. Dari beberapa pendapat tersebut dapat disimpulkan bahwa fungsi distraktor adalah kemampuan suatu opsi pengecoh dalam mengecohkan *testee* agar *testee* tertarik untuk memilihnya.⁸⁴ Oleh karena itu, distraktor yang baik adalah yang dapat dihindari oleh anak-anak yang pandai dan terpilih oleh anak-anak yang kurang pandai.⁸⁵ Dan apabila terpilih minimal 5% dari jumlah peserta.

Option atau alternatif yaitu beberapa kemungkinan jawaban berjumlah kisaran antara 3 sampai dengan 5 buah, dan dari kemungkinan-kemungkinan jawaban yang terpasang pada setiap soal, salah satu di antaranya adalah merupakan jawaban betul, sedangkan sisanya jawaban salah.⁸⁶ Jawaban di setiap soal mempunyai pola jawaban. Yang dimaksud pola jawaban di sini adalah distribusi *testee* dalam hal menentukan pilihan jawaban pada soal bentuk pilihan ganda. Dan untuk memperolehnya dengan menghitung banyaknya *testee* yang memilih pilihan jawaban *a*, *b*, *c*, atau *d* yang tidak memilih pilihan mana pun. Dari pola jawaban soal dapat ditentukan apakah pengecoh (*distractor*) berfungsi sebagai pengecoh dengan baik atau tidak. Pengecoh yang tidak dipilih sama sekali oleh *testee* berarti bahwa pengecoh itu jelek, terlalu mencolok menyesatkan. Sebaliknya sebuah distraktor dikatakan berfungsi dengan baik apabila

⁸⁴ Jurnal analisis tingkat kesukaran, daya pembeda dan fungsi distractor soal ujian semester ganjil mata pelajaran produktif di SMK Negeri 1 Indralaya Utara tahun pelajaran 2012/2013

⁸⁵ M. Chabib Thoha, M.A, *Teknik Evaluasi Pendidikan...*, hlm. 149.

⁸⁶ Meita Fitriawanawati, *Peran Analisis Butir Soal Guna Meningkatkan Kualitas Butir Soal, Kompetensi Guru dan Hasil Belajar Peserta Didik, Tulisan Artikel di Seminar Nasional Pendidikan PGSD UMS & HDPGSDI Wilayah Jawa*, hlm. 287.

distraktor tersebut mempunyai daya tarik yang besar bagi pengikut tes yang kurang memahami konsep atau kurang menguasai materi.⁸⁷ Suatu distraktor dapat diperlakukan dengan 3 cara:

- a. Diterima, karena sudah baik.
- b. Ditolak, karena tidak baik.
- c. Ditulis kembali, karena kurang baik.

Analisis distraktor diperlukan hanya untuk pembuat soal. Selain menghitung indeks kesukaran dan daya beda dalam analisis butir juga perlu diketahui apakah distraktor atau pengecoh yang disediakan tepat atau tidak benar. Apakah semua pilihan yang disediakan dipilih semua karena dianggap betul, jawaban terkumpul pada pilihan tertentu atau pilihan yang sama sekali tidak ada pemilihnya. Dengan memeriksa pola pilihan jawaban, dapat ditentukan hal-hal, yaitu;

- 1) Berapa jumlah subjek yang menjawab betul
- 2) Distraktor mana yang terlalu jelas atau menyolok sehingga sangat sedikit yang terkecoh untuk memilihnya
- 3) Distraktor mana yang justru menyesatkan subjek yang termasuk kelompok tinggi yang seharusnya tidak terkecoh, dan
- 4) Distraktor mana yang dapat menarik bagi subjek kelompok rendah, tetapi tidak cukup menarik bagi subjek dari kelompok tinggi. Pengecoh dikatakan efektif bila minimal dijawab oleh 5% peserta dan lebih banyak dipilih oleh kelompok siswa yang belum paham materi, berarti $5\% \times (A + B)$.

C. Soal untuk Melatih dan Penugasannya Struktur di Luar Kelas

Penelitian ini dilaksanakan dalam dua siklus. Pembelajaran pada siklus I berlangsung dalam empat kali pertemuan, tiga pertemuan proses pembelajaran meliputi tiga kegiatan yaitu kegiatan pendahuluan, kegiatan inti, dan kegiatan penutup, satu pertemuan ulangan harian. Pada kegiatan inti dijelaskan materi pembelajaran dan memperjelas umpan balik yang telah dituliskan di kertas tugas masing-masing siswa. Dalam setiap pertemuan baik siklus I maupun siklus II dengan memberikan tugas

⁸⁷ Suharsimi Arikunto, *Dasar-Dasar Evaluasi Pendidikan*, (Jakarta: Bumi Aksara, 2003) hlm. 173

terstruktur di setiap akhir pertemuan, dan telah diperiksa/dikoreksi serta dikembalikan sebelum pertemuan berikutnya dimulai. Data kualitatif berupa deskripsi kegiatan siswa dan guru selama tiga kali pertemuan yang diperoleh dari hasil catatan observasi dan diperkuat dengan hasil dokumentasi foto pembelajaran. Berdasarkan hasil observasi pembelajaran yang dilakukan pada pelaksanaan tindakan, pertama guru membuka pembelajaran dengan apersepsi dan membangkitkan motivasi belajar siswa untuk mengikuti pelajaran serta mengemukakan tujuan dan tema pembelajaran. Selanjutnya guru memberi pemahaman awal kepada siswa dan menjelaskan materi sesuai indikator yang akan dibahas. Pembelajaran dengan pemberian tugas terstruktur disertai umpan balik terdiri atas aktivitas siswa yang diamati meliputi:

- (1) Merespons pembahasan tugas.
- (2) Merespons motivasi dan apersepsi.
- (3) Menyimak tujuan pembelajaran yang disampaikan oleh guru.
- (4) Memperhatikan penjelasan guru.
- (5) Teliti dalam mengerjakan soal-soal latihan/LKS.
- (6) Berani mengemukakan pendapat/ide/menjawab (pertanyaan yang dilontarkan guru.
- (7) Berani maju dalam menjawab soal di papan tulis.
- (8) Memberikan umpan balik terhadap tugas (soal latihan) yang diberikan.
- (9) Mencatat tugas yang diberikan oleh guru.
- (10) Disiplin dalam mengumpulkan tugas-tugas yang diberikan.

Pada tahap ini guru, siswa dan *observer* merefleksi semua kegiatan yang terjadi pada siklus pertama dalam hubungannya dengan hasil belajar yang diperoleh terhadap hasil observasi guru, siswa, motivasi, karakter dan keterampilan sosial. Apabila dilihat dari ketuntasan belajar siswa hanya ada 14 orang siswa (45,16%) yang tuntas, dan 17 orang siswa (54,84%) yang tidak tuntas, hal ini menandakan bahwa belum mencapai target yang telah ditetapkan yakni minimal ketuntasan klasikal 80%. Ketidaktuntasan hasil belajar siswa dapat dipengaruhi oleh dua aspek, yakni aspek dari guru dan aspek dari siswa. Dari aspek guru yang perlu ditingkatkan adalah kegiatan apersepsi, artinya guru masih perlu memperkaya pengetahuan baik materi yang telah diajarkan maupun materi yang belum diajarkan,

membahas tugas dan menyampaikannya di hadapan kelas, tugas yang telah dikembalikan ke siswa agar dibahas di kelas sebelum memulai materi selanjutnya dan menyampaikan hasilnya, mendemonstrasikan pengetahuan dan keterampilan, pada tahap ini guru perlu memaksimalkan segala kemampuan agar penjelasan/demonstrasi yang dilakukan dapat dipahami siswa, merencanakan dan memberi bimbingan latihan awal, bimbingan yang diberikan kepada siswa harus efisien dan tepat sasaran, jangan melakukan bimbingan ke satu orang siswa dengan menggunakan waktu yang lama, karena yang lain akan dibimbing, mengecek pemahaman dan memberikan umpan balik, soal yang diberikan harus yang betul yang dapat mengukur kemampuan siswa, melakukan refleksi atau membuat rangkuman dengan melibatkan siswa, memberi kesempatan untuk pelatihan lanjutan dan penerapan, di sini dituntut agar guru dapat melihat penyebaran kemampuan siswa. Aspek siswa yang perlu diperbaiki adalah adanya siswa yang memiliki aktivitas sangat kurang dan kurang, yang menyebabkan hal itu terjadi karena masih ada beberapa siswa yang tidak disiplin mengikuti kegiatan pembelajaran, tidak memperhatikan tujuan, motivasi, serta penjelasan guru.

Aspek yang lain adalah tidak teliti dalam mengerjakan latihan dan LKS, belum berani mengemukakan pendapat dan tampil di depan kelas, tidak memperhatikan latihan dan tugas-tugasnya yang telah diperiksa (yang telah diberikan catatan-catatan). Masih ada juga siswa yang malas mencatat, dan terlambat mengumpulkan tugas. Perbaikan yang dilakukan untuk aspek siswa adalah

- 1) Menyampaikan semua hasil yang diperoleh kepada siswa, mulai dari nilai tugas, nilai ulangan harian, nilai proses, nilai karakter, dan nilai keterampilan proses.
- 2) Memberikan arahan kepada siswa yang masih memiliki kedisiplinan rendah.
- 3) Memotivasi siswa yang belum berani mengemukakan pendapat, dan tampil di depan kelas.
- 4) Menilai dan mengembalikan latihan/tugas yang telah siswa kerjakan.

- 5) Memberikan arahan agar dapat memperhatikan penjelasan guru, serta mencatat hal yang penting penjelasan guru, serta berusaha menemukan sendiri jawaban terhadap tugas yang diberikan.
- 6) Memberikan nilai plus bagi siswa yang mengumpulkan tugasnya dengan tepat waktu.

Pada pembelajaran siklus kedua ini tampak sebagian besar siswa telah siap untuk mengikuti pembelajaran, apabila dilihat dari ketuntasan belajar siswa ada 25 orang (80,64%) yang tuntas, dan 6 orang siswa (19,35%) yang tidak tuntas, hal ini menandakan bahwa telah terjadi peningkatan dari siklus pertama ke siklus kedua dan telah mencapai target yang telah ditetapkan yakni ketuntasan klasikal 80%. Ketuntasan hasil belajar siswa dapat dipengaruhi oleh dua aspek, yakni aspek dari guru dan aspek dari siswa.

Aspek dari guru berhubungan dengan proses pembelajaran yang dilakukan oleh guru, berdasar pada hasil observasi dapat diketahui bahwa guru telah melakukan perbaikan dalam proses pembelajaran, dari aspek guru adalah kegiatan apersepsi, guru telah mampu menghubungkan dengan baik materi yang telah diajarkan maupun materi yang belum diajarkan, membahas tugas dan menyampaikannya di hadapan kelas, tugas yang telah dikembalikan ke siswa dibahas di kelas sebelum memulai materi selanjutnya dan menyampaikan hasilnya, sehingga siswa yang masih ingin memperjelas pemahaman mereka tentang tugasnya masih punya kesempatan. Perbaikan selanjutnya adalah mendemonstrasikan pengetahuan dan keterampilan. Pada tahap ini guru telah memaksimalkan segala kemampuan agar penjelasan/demonstrasi yang dilakukan dapat dipahami siswa, sehingga jika ada soal yang diberikan dapat dijawab oleh siswa dengan baik, merencanakan dan memberi bimbingan latihan awal, membimbing siswa saat memberikan soal latihan secara efisien dan tepat sasaran, mengecek pemahaman dan memberikan umpan balik, diberikan beberapa soal dan ketika ada siswa yang dapat menjawab dengan benar, maka guru memberikan pujian. Melakukan refleksi atau membuat rangkuman dengan melibatkan siswa, memberi kesempatan untuk pelatihan lanjutan dan penerapan, guru memberikan soal sebagai latihan lanjutan agar pemahaman siswa lebih mantap.

Aspek siswa yang mengalami perbaikan adalah siswa sudah disiplin mengikuti kegiatan pembelajaran, memperhatikan tujuan dan motivasi, serta penjelasan guru. Aspek lain yang mulai mengalami perubahan adalah siswa sudah teliti dalam mengerjakan latihan dan LKS, mulai berani mengemukakan pendapat, dan tampil di depan kelas, mengerjakan sendiri tugas yang diberikan, memperhatikan latihan dan tugas-tugasnya yang telah diperiksa. Siswa yang tadinya malas mencatat dan mengumpulkan tugas sudah mulai rajin. Bila dihubungkan antara motivasi dengan hasil belajar siswa tampak bahwa semakin baik motivasi yang dimiliki siswa semakin bagus pula hasil belajar yang diperoleh. Hubungan ini dapat dilihat pada Tabel 1 dan Tabel 2.

Tabel 1

Kategori	Frekuensi	Persen	Frekuensi	Persen	Keterangan
Sangat mudah	9	29,03	0	0	Naik 29,03
Rendah	7	22,58	3	9,68	Naik 12,90
Sedang	1	3,23	3	9,68	Naik 6,45
Tinggi	14	45,16	19	61,29	Naik 16,13
Sangat tinggi	0	0	6	19,35	Naik 19,35
Jumlah	31	100	31	100	

Berdasarkan Tabel 1, diperoleh informasi bahwa, siswa yang memiliki hasil belajar sangat rendah pada siklus I sebanyak sembilan orang (29,03%) pada siklus II menjadi tidak ada, berarti terjadi kenaikan 29,03%. Siswa yang memiliki hasil belajar rendah pada siklus I sebanyak tujuh orang (22,58 %) pada siklus II menjadi tiga orang (9,68%), berarti terjadi kenaikan 12,90%. Siswa yang memiliki hasil belajar sedang pada siklus I sebanyak satu orang (3,23 %) (9,68%), berarti terjadi kenaikan 6,45%). Siswa yang memiliki hasil belajar tinggi pada siklus I sebanyak 14 orang (45,16 %) pada siklus II menjadi 19 orang (61,29%), berarti terjadi kenaikan 16,13%. Siswa yang memiliki hasil belajar sangat tinggi pada siklus I sebanyak tidak ada, pada siklus II menjadi enam orang (19,35%), berarti terjadi kenaikan 19,35%.

Tabel 2

Kategori	Frekuensi	Persen	Frekuensi	Persen	Frekuensi (%)
Sangat rendah	0	0	0	0	tetap
Rendah	7	22,58	3	9,67	Naik 12,91
Sedang	5	16,13	-	0	Naik 16,13
Tinggi	11	35,48	13	41,94	Naik 6,46
Sangat tinggi	8	25,81	15	48,39	Naik 22,58
Jumlah	31	100	31	100	

Berdasarkan Tabel 2 dapat diperoleh informasi bahwa pada siklus I dan II tidak ada siswa yang memiliki motivasi belajar kategori sangat rendah. Siswa yang memiliki motivasi rendah pada siklus I tujuh orang (22,58%) pada siklus II tiga orang (9,67%) berarti terjadi kenaikan 12,91%. Siswa yang memiliki motivasi sedang pada siklus I lima orang (16,13%) pada siklus II tidak ada, berarti terjadi kenaikan 16,13%. Siswa yang memiliki motivasi tinggi pada siklus I sebelas orang (35,48%) pada siklus II, 13 orang (41,94%) berarti terjadi kenaikan 6,46%. Siswa yang memiliki motivasi sangat tinggi pada siklus I, delapan orang (25,81%) pada siklus II 15 orang (48,39%) berarti terjadi kenaikan 22,58%. Hal ini menunjukkan terdapat peningkatan motivasi belajar dalam pembelajaran kimia dengan pemberian tugas terstruktur disertai umpan balik.⁸⁸

⁸⁸ Jurnal Chemica Vol. 13 Nomor 2 Desember 2012, 39 -46

BAB IV

TEKNIK TES DAN TEKNIK NONTES SEBAGAI ALAT EVALUASI HASIL BELAJAR

A. Teknis Tes

1. Pengertian Tes

Secara harfiah, kata “tes” berasal dari bahasa Perancis Kuno yakni *testum*, yang berarti “piring untuk menyisihkan logam-logam mulia” (maksudnya dengan menggunakan alat berupa piring itu akan dapat diperoleh jenis-jenis logam-logam mulia yang nilainya sangat tinggi). Dalam bahasa Inggris ditulis dengan *test* yang dalam bahasa Indonesia diterjemahkan dengan “tes”, ujian atau percobaan. Dalam bahasa Arab: *Imtihan*.

Dari segi istilah, terdapat beberapa definisi tentang istilah tes, di antaranya adalah Drs. Amir Daien Indrakusuma dalam bukunya yang berjudul *Evaluasi Pendidikan*, mengatakan bahwa tes adalah suatu alat atau prosedur yang sistematis dan objektif untuk memperoleh data-data atau keterangan yang diinginkan tentang seseorang, dengan cara yang boleh dikatakan tepat dan cepat.

Definisi lain tentang tes juga dikutip dari Webster’s Collegiate, bahwa “*test = any series of questions or exercises or other means of measuring the skill, knowledge, intelligence, capacities or aptitudes of an individual or group*”. Dari beberapa definisi tentang tes di atas, tampak jelas bahwa pada hakikatnya tidak ada perbedaan. Jadi

seorang tester dalam melakukan kegiatan penilaian membutuhkan suatu perangkat yang berupa pertanyaan, tugas, dan lain-lain. Perangkat tersebut biasa kita kenal dengan sebutan tes.

Ada beberapa istilah yang memerlukan penjelasan sehubungan dengan uraian di atas yaitu, tes adalah alat untuk prosedur yang dipergunakan dalam rangka pengukuran dan penilaian; testing berarti saat dilaksanakannya atau peristiwa berlangsungnya pengukuran dan penilaian. Jadi dapat disimpulkan bahwa tes adalah suatu alat pengumpul informasi yang bersifat lebih resmi bila dibandingkan alat-alat yang lain karena penuh dengan batasan-batasan. Adapun pelaksanaannya dapat dilaksanakan secara lisan maupun secara tes tulis.⁸⁹ Tes adalah alat yang direncanakan untuk mengukur kemampuan, keahlian, atau pengetahuan. Dari pengertian ini maka tes adalah:

1. Merupakan alat.
2. Harus direncanakan.
3. Berfungsi sebagai pengukur kemampuan, kecakapan, dan pengetahuan anak.

Adapun yang dimaksud dengan teknis tes adalah suatu teknik dalam evaluasi yang digunakan untuk mengetahui hasil belajar murid dengan mempergunakan alat tes.

Sehingga dari definisi-definisi di atas kiranya dapat dipahami bahwa dalam dunia evaluasi pendidikan, yang dimaksud dengan tes adalah cara (yang dapat dipergunakan) atau prosedur (yang perlu ditempuh) dalam rangka pengukuran dan penilaian di bidang pendidikan.

2. Fungsi Tes

Secara umum, ada lima macam fungsi yang dimiliki oleh teknik tes, yaitu:

1. Sebagai alat pengukur terhadap peserta didik. Dalam hubungan ini tes berfungsi mengukur tingkat perkembangan atau kemajuan yang telah dicapai oleh peserta didik setelah mereka menempuh proses belajar mengajar dalam jangka waktu.

⁸⁹ Sudijono, Anas. 2011. *Pengantar Evaluasi Pendidikan*. Jakarta: Persada

2. Sebagai alat pengukur keberhasilan program pengajaran, sebab melalui tes tersebut akan dapat diketahui sudah seberapa jauh program pengajaran yang telah ditentukan, telah dapat dicapai.
3. Tes berfungsi sebagai motivator dalam pembelajaran.
4. Tes dapat berfungsi dalam upaya perbaikan kualitas pembelajaran.
5. Tes dapat berfungsi untuk menentukan keberhasilan siswa untuk melanjutkan pendidikan pada jenjang yang lebih tinggi.⁹⁰

Sedangkan menurut Suharsimi Arikunto dalam bukunya *Dasar-Dasar Evaluasi Pendidikan*, fungsi tes dapat ditinjau dari tiga hal:

- a. Fungsi untuk kelas: mengadakan diagnosis terhadap kesulitan dalam belajar siswa.
- b. Fungsi untuk bimbingan: menentukan arah pembicaraan dengan orang tua tentang anak-anak mereka.
- c. Fungsi untuk administrasi: memberi petunjuk dalam mengelompokkan siswa.
- f. Penempatan siswa baru.
- g. Membantu siswa memiliki kelompok.
- h. Menilai kurikulum.
- i. Memperluas hubungan masyarakat.
- j. Menyediakan informasi untuk badan lain di luar sekolah.

3. Penggolongan Tes

Tes digolongkan menjadi 5 golongan di antaranya adalah sebagai berikut:

- a. Menurut sifatnya, tes dapat dikelompokkan menjadi:

- a. Tes verbal

Yang mana tes dengan cara ini menggunakan bahasa sebagai alat untuk melakukan tes. Tes verbal terdiri dari:

- Tes lisan (*oral test*)
- Tes tulisan (*written test*)

- b. Tes Nonverbal

Yaitu tes yang tidak menggunakan bahasa sebagai alat untuk melakukan tes, tetapi menggunakan gambar, memberikan tugas dan lainnya, atau dengan tes ini tester menghendaki adanya

⁹⁰ Sudijono, Anas. 2011. *Pengantar Evaluasi Pendidikan*. Jakarta: Persada

respons dan *testee* bukan berupa ungkapan kata-kata atau kalimat.⁹¹

- b. Menurut tujuannya, tes dapat dikelompokkan menjadi:
 - a. Tes bakat (*Aptitude Test*) yaitu tes yang digunakan untuk menyelidiki bakat seseorang.
 - b. Tes intelegensi (*Intelligenci Test*) yakni tes yang dilakukan dengan tujuan untuk mengungkap atau mengetahui tingkat kecerdasan seseorang.
 - c. Tes prestasi belajar (*Achievement Test*) yaitu tes yang dilakukan untuk mengetahui prestasi seseorang murid dari mata pelajaran yang telah diberikan.
- c. Menurut pembagiannya:
 1. Tes Standar
Tes standar atau tes yang dibakukan mengandung prosedur yang seragam untuk menentukan nilai dan administrasinya.
 2. Tes Buatan Guru
Tes buatan guru cenderung difokuskan pada tujuan instruksional untuk kelas tertentu.
- d. Menurut bentuk soalnya, tes dikelompokkan menjadi:
 1. Tes Uraian (*Eassy Test*)
Yaitu tes yang bentuk soalnya sedemikian rupa, sehingga memberi kesempatan kepada murid untuk menjawab secara bebas dengan uraian. Bentuk tes ini terdiri dari:
 - a. Uraian Bebas
 - b. Uraian Terbatas
 2. Tes Objektif
 - a. Variasi yang mana *testee* harus menyuplai jawabannya sehingga hampir tidak berbeda dengan *essay test*.
 - b. Variasi yang mana *testee* hanya memilih di antara jawaban yang telah disediakan bersama soalnya. Pada variasi ini ada beberapa bentuk tes, di mana tester harus tahu:
 1. Menyatakan apakah pernyataan itu benar atau salah.
 2. Memilih jawaban yang lain benar.

⁹¹ Thoha. M. Chabib. 2001. *Teknik Evaluasi Pendidikan*. Jakarta: PT. RajaGrafindo Persada.

3. Menjodohkan dua rentetan kata-kata yang tersedia sesuai dengan jawaban yang benar.
 4. Memilih di antara alternatif-alternatif jawaban yang disediakan untuk setiap soal.⁹²
- c. Ditinjau dari objek yang dites:
1. Tes Individual
Yaitu tes yang dalam pelaksanaannya memerlukan waktu yang cukup panjang.
 2. Tes Kelompok
Yaitu tes yang dilakukan terhadap beberapa murid dalam waktu yang sama.

B. Teknik Nontes

1. Pengamatan

a. Observasi

Observasi atau pengamatan merupakan teknik untuk merekam data atau keterangan atau informasi tentang diri seseorang yang dilakukan secara langsung atau tidak langsung terhadap kegiatan-kegiatan yang sedang berlangsung, sehingga diperoleh data tingkah laku seseorang yang menampak (*behavior observable*), apa yang dikatakan dan apa yang diperbuatnya.⁹³

A. Menurut cara tujuannya, observasi dapat dibedakan menjadi 3 macam:

- Observasi partisipatif, yaitu observasi yang dilakukan oleh *observer* (pengamat) dengan turut mengambil bagian dalam kegiatan yang dilakukan oleh objek yang diobservasi.
- Observasi sistematis, yaitu observasi yang direncanakan terlebih dahulu aspek-aspek yang akan diobservasi sesuai dengan tujuan, waktu, dan alat yang dipakai.

⁹² http://husnilmubaroq.blogspot.com/2014/12/teknik-tes-dan-teknik-non-tes-sebagai_31.html?m=1

⁹³ http://husnilmubaroq.blogspot.com/2014/12/teknik-tes-dan-teknik-non-tes-sebagai_31.html?m=1

- Observasi eksperimental, yaitu observasi yang dilakukan untuk mengetahui perubahan-perubahan atau gejala-gejala sebagai akibat dari situasi yang sengaja diadakan.

B. Berdasarkan hubungan *observer* dengan kelompok yang diamatinya, observasi dapat dilakukan dengan cara sebagai berikut:

- a) Partisipan penuh, yaitu *observer* menyamakan diri dengan orang yang diobservasi.
- b) *Observer* sebagai pengamat, yaitu masing-masing pihak menyadari peranannya. *Observer* sebagai pengamat membatasi diri dalam berpartisipasi sebagai pengamat, dan *observer* menyadari bahwa dirinya adalah objek pengamatan.
- c) *Observer* sebagai partisipan, yaitu *observer* hanya berpartisipasi sepanjang yang dibutuhkan dalam penelitiannya.
- d) Pengamat sempurna (*complete observer*), yaitu *observer* hanya menjadi pengamat tanpa partisipasi dengan yang diamati.

Agar data yang dikumpulkan melalui observasi ini dapat dicatat dengan sebaik-baiknya, maka diperlukan pedoman observasi.

C. Bentuk-bentuk pedoman observasi

1. Daftar cek (*checklist*)

Daftar cek adalah suatu daftar pernyataan yang memuat aspek-aspek yang mungkin terdapat dalam suatu situasi, tingkah laku, atau kegiatan individu yang sedang diamati. Semua aspek yang akan diobservasi dijabarkan dalam suatu daftar sehingga pada waktu observasi, *observer* (pengamat) tinggal membubuhkan tanda cek terhadap ada atau tidak adanya aspek-aspek yang menjadi pusat perhatian bagi diri individu atau kejadian yang diobservasi. Daftar cek ini dapat

digunakan untuk mengobservasi individu atau kelompok individu.⁹⁴

Gejala-gejala perilaku atau tingkah laku seseorang yang dapat diobservasi dengan teknik ini antara lain kebiasaan belajar, aktivitas belajar dan bekerja, kepemimpinan dan kerja sama, pergaulan, dan topik lain yang relevan dengan kegiatan akademik dan nonakademik dalam kehidupan sekolah.

2. Skala penilaian (*rating scale*)

Skala penilaian merupakan alat pengumpul data yang dipergunakan dalam observasi untuk menjelaskan, menggolongkan, dan menilai individu atau situasi. Dalam skala penilaian, aspek yang diobservasi dijabarkan dalam bentuk skala.

Skala penilaian pada umumnya terdiri dari suatu daftar yang berisi ciri-ciri tingkah laku atau sifat yang harus dicatat secara bertingkat sehingga *observer* hanya memberikan tanda cek pada tingkat mana gejala atau ciri-ciri tingkah laku itu muncul. Berdasarkan pada alternatif skala yang dipakai untuk menilai dan menggolongkan gejala perilaku individu atau situasi, maka skala penilaian dapat dibedakan menjadi tiga bentuk: kuantitatif, deskriptif, dan grafis. Skala penilaian deskriptif adalah suatu alat observasi yang digunakan untuk mengamati gejala atau ciri-ciri tingkah laku individu atau situasi dalam mana alternatif skalanya dijabarkan dalam bentuk kata-kata. Skala penilaian grafis adalah suatu alat observasi yang digunakan untuk mengamati gejala atau ciri-ciri tingkah laku individu atau situasi di mana alternatif skalanya dijabarkan dalam bentuk grafis (garis).

Observasi merupakan cara menghimpun keterangan data atau yang dilakukan dengan mengadakan pengamatan dan pencatatan secara sistematis terhadap fenomena yang sedang dijadikan sasaran pengamatan. Observasi dalam dunia

⁹⁴ <http://googleweblight.com/i?u=http://makalahpendidikanislamlengkap.blogspot.com/2015/06/tes-dan-nontes-evaluasi-pembelajaran.html?m%3D1&hl=id-ID>

pendidikan dan pengajaran dipergunakan untuk menilai ranah psikomotor siswa.

2. Wawancara

Wawancara merupakan suatu cara yang dilakukan secara lisan yang berisikan pertanyaan-pertanyaan yang sesuai dengan tujuan informasi yang hendak digali. Wawancara dibagi dalam 2 kategori, yaitu:

- a. Wawancara bebas, yaitu si penjawab (responden) diperkenankan untuk memberikan jawaban secara bebas sesuai dengan yang ia ketahui tanpa diberikan batasan oleh pewawancara.⁹⁵
- b. Wawancara terpimpin, di mana pewawancara telah menyusun pertanyaan-pertanyaan terlebih dahulu yang bertujuan untuk menggiring penjawab pada informasi-informasi yang diperlukan saja.

Wawancara adalah suatu teknik penilaian yang dilakukan dengan jalan percakapan (dialog) baik secara langsung (*face to face relation*) secara langsung apabila wawancara itu dilakukan kepada orang lain misalnya kepada orang tuannya atau kepada temannya. Keberhasilan wawancara sebagai alat penilaian sangat dipengaruhi oleh beberapa hal:

- a. Hubungan baik pewawancara dengan anak yang diwawancarai. Dalam hal ini hendaknya pewawancara dapat menyesuaikan diri dengan orang yang diwawancarai.
- b. Keterampilan pewawancara
Keterampilan pewawancara sangat besar pengaruhnya terhadap hasil wawancara yang dilakukan, karena guru perlu melatih diri agar memiliki keterampilan dalam melaksanakan wawancara.
- c. Pedoman wawancara

Keberhasilan wawancara juga sangat dipengaruhi oleh pedoman yang dibuat oleh guru sebelum guru melaksanakan wawancara harus membuat pedoman-pedoman secara terperinci, tentang pertanyaan yang akan diajukan.⁹⁶

⁹⁵ <http://googleweblight.com/i?u=http://makalahpendidikanislamlengkap.blogspot.com/2015/06/tes-dan-nontes-evaluasi-pembelajaran.html?m%3D1&hl=id-ID>

⁹⁶ Thoha. M. Chabib. 2001. *Teknik Evaluasi Pendidikan*. Jakarta: PT RajaGrafindo Persada.

A. Langkah-Langkah Penyusunan Wawancara:

- a) Perumusan tujuan
- b) Perumusan kegiatan atau aspek-aspek yang dinilai
- c) Penyusunan kisi-kisi
- d) Penyusunan pedoman wawancara
- e) Lembaran penilaian

B. Kelebihan dan Kelemahan Wawancara

- Kelebihan wawancara yaitu:
 - a. Wawancara dapat memberikan keterangan keadaan pribadi. Hal ini tergantung pada hubungan baik antara pewawancara dengan objek.
 - b. Wawancara dapat dilaksanakan untuk setiap umur dan mudah dalam pelaksanaannya.
 - c. Wawancara dapat dilaksanakan serempak dengan observasi data tentang keadaan individu lebih banyak diperoleh dan lebih tepat dibandingkan dengan observasi dan angket.
 - d. Wawancara dapat menimbulkan hubungan yang baik antara si pewawancara dengan objek.
- Sedangkan kelemahan wawancara:
 - 1) Keberhasilan wawancara dapat dipengaruhi oleh kesediaan, kemampuan individu yang diwawancarai.
 - 2) Kelancaran wawancara dapat dipengaruhi oleh keadaan sekitar pelaksanaan wawancara.
 - 3) Wawancara menuntut penguasaan bahasa yang baik dan sempurna dari pewawancara.
 - 4) Adanya pengaruh subjektif dari pewawancara dapat memengaruhi hasil wawancara.

Ada dua jenis wawancara yang dapat dipergunakan sebagai alat evaluasi, yaitu:

- Wawancara terpimpin (*Guided Interview*) yang juga sering dikenal dengan istilah wawancara berstruktur (*Structur Interview*) atau wawancara sistematis (*System Interview*).
- Wawancara tidak terpimpin (*Un-Guided Interview*) yang sering dikenal dengan istilah wawancara sederhana (*Simple Interview*)

atau wawancara tidak sistematis (*Non-Systematic Interview*), atau wawancara bebas.

C. Tahap Awal Pelaksanaan Wawancara

Perhatikan terhadap aspek yang pertama ini bertujuan untuk mengondisikan situasi wawancara. Pernyataan ini mengandung arti bahwa evaluator harus menciptakan situasi yang mengungkapkan suasana keakraban sehingga siswa tidak merasa takut, yang pada akhirnya dapat mengemukakan jawaban atau pendapatnya secara bebas.⁹⁷

D. Penggunaan Pertanyaan

Aspek kedua ini mengandung pengertian bahwa pertanyaan diajukan secara bertahap dan sistematis berdasarkan kisi-kisi yang telah dibuat sebelumnya. Apabila menggunakan wawancara terpimpin, maka pewawancara mengajukan pertanyaan beserta dengan alternatif jawabannya. Siswa kemudian mengemukakan pendapatnya, lalu pendapat tersebut diklasifikasikan ke dalam alternatif jawaban yang telah pada. Lain halnya ketika menggunakan wawancara tidak terpimpin, evaluator mengajukan pertanyaan dan siswa menjawab dengan bebas.

E. Pencatatan hasil wawancara

Pencatatan hasil wawancara sebaiknya dilakukan saat wawancara berlangsung. Terhadap wawancara terpimpin, pencatatannya cukup mudah sebab tinggal memberikan tanda pada alternatif jawaban. Sedangkan pada wawancara tidak terpimpin, pokok-pokok isi jawaban siswa perlu dicatat pada lembaran tersendiri.⁹⁸

Pedoman wawancara tersebut disusun dengan menempuh langkah-langkah sebagai berikut.

- a. Tentukan tujuan yang ingin dicapai dari kegiatan wawancara tersebut.
- b. Tentukan aspek-aspek apa saja yang ingin diungkap melalui kegiatan tersebut.

⁹⁷ <https://journal.unnes.ac.id/nju/index.php/edukasi/article/download/957/894>
⁹⁸ http://husnilmubaroq.blogspot.com/2014/12/teknik-tes-dan-teknik-non-tes-sebagai_31.html?m=1

- c. Tentukan jenis wawancara apa yang akan digunakan, terpimpin atau tidak terpimpin atau kombinasi dari kedua jenis tersebut.
- d. Buat pertanyaan sesuai dengan jenis wawancara yang akan digunakan
- e. Sebaiknya buat pedoman mengolah dan menafsirkan hasil wawancara.

Dengan memperhatikan dan mengikuti pedoman wawancara yang disebutkan di atas, diharapkan pelaksanaan wawancara dapat berjalan dengan baik sehingga fungsinya sebagai alat bantu dalam evaluasi pendidikan dan pengajaran dapat tercapai dengan baik pula.

3. Angket

- a. Angket atau kuesioner adalah serangkaian pertanyaan atau pernyataan tertulis yang diajukan kepada responden untuk memperoleh jawaban secara tertulis pula. Pertanyaan/pernyataan dalam angket tergantung pada maksud serta tujuan yang ingin dicapai. Maksud dan tujuan tersebut berpengaruh terhadap bentuk pertanyaan yang ada dalam angket itu.⁹⁹

Pada umumnya di dalam angket itu kita mendapat dua bagian pokok, yaitu bagian yang mengandung data identitas, dan bagian yang mengandung pertanyaan-pertanyaan yang ingin diperoleh jawabannya.

- b. Kelebihan Teknik Angket
 - a) Teknik angket lebih efisien karena dapat mengumpulkan data dalam jumlah responden yang besar dalam waktu yang singkat.
 - b) Dapat mengungkap data yang memerlukan perkembangan dan pemikiran, dan bukan jawaban spontan.
 - c) Dapat mengungkap keterangan yang mungkin bersifat pribadi dan tidak akan diberikan secara langsung.
 - d) Data yang dikumpulkan dapat lebih mudah dianalisis, karena pertanyaan yang diajukan sama

⁹⁹ <https://ejournal.staindirundeng.ac.id/index.php/tadib/article/download/138/96>

c. Kelemahan Teknik Angket

1. Tidak akan dapat menjangkau data yang sebenarnya jika petunjuk pengisian tidak jelas.
2. Tidak dapat diketahui dengan pasti bahwa responden sungguh-sungguh dalam mengisi angket. Sering terjadi angket juga diisi oleh orang lain (bukan responden yang sebenarnya), karena dilakukan tidak secara langsung berhadapan muka antara peneliti dan responden.
3. Tidak dapat ditambah keterangan yang dapat diperoleh lewat observasi dan angket diberikan terbatas kepada orang yang melek huruf.¹⁰⁰

d. Jenis-Jenis Angket

1. Dilihat dari sumber datanya, angket dibedakan sebagai berikut:
 - Angket langsung, yaitu angket tersebut diberikan kepada orang yang dimintai pendapat atau jawabannya atau responden yang ingin diselidiki.
 - Angket tidak langsung, yaitu angket disampaikan kepada orang lain yang dimintai pendapat tentang keadaan seseorang.
2. Dilihat dari strukturnya, angket dapat dibedakan sebagai berikut:
 - Angket tak berstruktur, yaitu angket yang berisi pertanyaan-pertanyaan yang menghendaki jawaban yang bebas dan uraian yang panjang lebar dari responden.
 - Angket berstruktur, yaitu angket yang berisi pertanyaan-pertanyaan beserta jawabannya yang jelas, singkat dan konkret.
3. Berdasarkan jenis pertanyaannya, angket dibedakan sebagai berikut:
 - Pertanyaan terbuka, yaitu angket yang memberikan kesempatan yang seluas-luasnya kepada responden untuk memberikan jawaban atau tanggapannya.
 - Pertanyaan tertutup, yaitu pertanyaan-pertanyaan yang membuat responden tinggal memilih jawaban yang telah disediakan di dalam angket itu.

¹⁰⁰ <https://osf.io>

- Kombinasi terbuka dan tertutup, yaitu jika jawabannya sudah ditentukan kemudian disusun pertanyaan terbuka.
4. Menurut bentuk jawabannya, angket dibedakan menjadi:
 - a. Jawaban tabuler, yaitu responden diminta menjawab dengan mengisi kolom-kolom pada tabel yang sudah tersedia.
 - b. Jawaban berskala, yaitu jawaban terhadap pertanyaan disusun berjenjang di mana responden diminta menyatakan membenaran atau penolakan terhadap setiap pertanyaan sikap, sehingga diperoleh gambaran tentang derajat kecakapan, keadaan sikap, dan keadaan diri responden.
 - c. Jawaban dengan cek, yaitu responden menjawab dengan cara memilih salah satu dari pilihan yang tersedia.
 - e. Langkah-Langkah Penyelenggaraan Angket
Langkah-langkah penyelenggaraan angket:

A. Tahap persiapan

Langkah pertama yang dilakukan dalam menyusun angket ialah memerinci atau menjabarkan variabel-variabel yang akan diukur.¹⁰¹ Langkah kedua menetapkan model jawaban yang ditentukan oleh bentuk jawaban yang dikehendaki dari variabel angket tertentu. Adapun langkah-langkah menyusun angket, yaitu:

a. Pengantar

Isi dari pengantar ini ialah mengadakan pendekatan terhadap responden agar bersedia memberi keterangan yang dibutuhkan.

b. Petunjuk pengisian

Petunjuk pengisian ini harus dirancang dengan baik dan jelas sebab akan mempermudah responden dalam mengisi setiap butir pertanyaan. Petunjuk pengisian angket hendaknya dirumuskan dengan bahasa yang sederhana, singkat dan mudah dimengerti, petunjuk memuat tentang cara mengisi angket.

¹⁰¹ Thoha. M. Chabib. 2001. *Teknik Evaluasi Pendidikan*. Jakarta: PT RajaGrafindo Persada.

c. Penyusunan butir pertanyaan

Yang harus diperhatikan dalam menyusun butir pertanyaan adalah susunan kalimat hendaknya sederhana dan jelas, gunakan kata-kata yang tidak mempunyai arti ganda, pertanyaan hendaknya disesuaikan dengan kemampuan responden, hindarkan kata-kata yang bersifat sugestif, pertanyaan jangan bersifat memaksa untuk dijawab, pertanyaan jangan menuntut siswa/responden untuk berpikir terlalu berat, gunakan kata-kata yang netral, hindarkan kata-kata yang tidak berguna atau tidak perlu.

d. Penutup: berisi ucapan terima kasih kepada responden atau siswa karena dedikasinya dalam bekerja sama untuk kepentingan bimbingan.

B. Tahap Pelaksanaan

Pada tahap ini kita mempersiapkan instrumen angket dan lembar jawaban yang diperlukan. Kemudian membagikan instrumen tersebut untuk diisi responden. Selanjutnya kita membacakan petunjuk pengisiannya dan mengecek jumlah responden yang sudah mengembalikan angket dan lembar jawabannya.

C. Tahap Analisis Hasil

Pada tahap ini dilakukan penskoran terhadap jawaban responden. Penskoran ini dibedakan atas penskoran terhadap pertanyaan-pertanyaan tertutup atau berstruktur dengan model jawaban yang sudah tersedia dan terbatas serta penskoran terhadap pertanyaan terbuka atau tidak berstruktur yang memerlukan jawaban uraian bebas. Selanjutnya, akan diperoleh gambaran menyeluruh tentang responden.¹⁰²

4. Pemeriksaan Dokumen

Pemeriksaan dokumen merupakan teknik nontes untuk mengevaluasi kemajuan, perkembangan atau keberhasilan belajar siswa

¹⁰² http://journal.uimalauddin.ac.id/index.php/lentera_pondidikan/article/download/3754/3428

tanpa menguji baik itu berupa informasi mengenai daftar riwayat hidup siswa maupun daftar riwayat hidup orang tua siswa.

Berbagai informasi baik mengenai siswa, orang tua siswa, dan lingkungannya pada saat-saat tertentu sangat diperlukan sebagai bahan pelengkap bagi guru dalam melakukan evaluasi hasil belajar siswa. Informasi-informasi tersebut berupa evaluasi mengenai kemajuan, perkembangan atau keberhasilan belajar peserta didik tanpa menguji (teknik nontes) juga dapat dilengkapi atau diperkaya dengan cara melakukan pemeriksaan terhadap dokumen-dokumen, misalnya dokumen yang memuat informasi mengenai riwayat hidup (autobiografi). Dari uraian tersebut di atas dapatlah dipahami, bahwa dalam rangka evaluasi itu tidak harus semata-mata dilakukan dengan menggunakan alat berupa tes-tes hasil belajar. Teknik nontes juga menempati kedudukan yang penting dalam rangka evaluasi hasil belajar, lebih-lebih evaluasi yang berhubungan dengan kondisi kejiwaan peserta didik, seperti persepsinya terhadap mata pelajaran tertentu, persepsinya terhadap guru, minatnya, bakatnya, tingkah laku atau sikapnya dan sebagainya, yang kesemuanya itu tidak mungkin dievaluasi dengan menggunakan tes sebagai alat pengukurnya.¹⁰³

Evaluasi mengenai kemajuan, perkembangan atau keberhasilan belajar peserta didik tanpa menguji (teknik nontes) juga dapat dilengkapi atau diperkaya dengan cara melakukan pemeriksaan terhadap dokumen-dokumen, misalnya dokumen yang memuat informasi mengenai riwayat hidup (autobiografi). Dari uraian tersebut dapat dipahami bahwa dalam rangka evaluasi itu tidak harus semata-mata dilakukan dengan menggunakan alat berupa tes-tes hasil belajar.

¹⁰³ Thoha. M. Chabib. 2001. *Teknik Evaluasi Pendidikan*. Jakarta: PT. RajaGrafindo Persada.

BAB V

TEKNIK PENGUJIAN VALIDITAS TES DAN VALIDITAS ITEM TES HASIL BELAJAR

A. Pengertian Validitas

Validitas merupakan suatu standar atau dasar ukuran yang menunjukkan ketetapan (*appropriateness*), kemanfaatan (*usefulness*) dan kesahihan yang mengarah pada ketepatan interpretasi suatu prosedur evaluasi sesuai dengan tujuan pengukurannya.

Pengertian Validitas Menurut para Ahli

Di bawah ini adalah definisi dari validitas menurut ahlinya.

1. Sudjana (2004: 12)

Pengertian validitas menurut Sudjana adalah ketepatan alat penilaian terhadap konsep yang dinilai sehingga betul-betul menilai apa yang harus dinilai.

2. Suryabrata (2000: 41)

Pengertian validitas menurut Suryabrata adalah derajat fungsi pengukuran suatu tes, atau derajat kecermatan ukur suatu tes. Validitas suatu tes mempermasalahkan apakah tes tersebut benar-benar mengukur atap yang hendak diukur.

3. Azwar (1987: 173)

Pengertian validitas menurut Azwar adalah sejauh mana ketepatan dan kecermatan suatu instrumen pengukur (tes) dalam melakukan fungsi ukurnya. Suatu tes dikatakan memiliki validitas yang tinggi apabila alat tersebut menjalankan fungsi ukur secara tepat atau memberikan hasil ukur yang sesuai dengan maksud dilakukannya pengukuran tersebut.

4. Arikunto (1999: 65)

Pengertian validitas menurut Arikunto adalah suatu ukuran yang menunjukkan tingkat kesahihan suatu tes. Suatu tes dikatakan valid apabila tes tersebut mengukur apa yang hendak diukur. Tes memiliki validitas yang tinggi jika hasilnya sesuai dengan kriteria, dalam arti memiliki kesejajaran antara tes dan kriteria.

5. Kusaeri (2012: 75)

Pengertian validitas menurut Kusaeri adalah ketepatan (*appropriateness*), kebermaknaan (*meaningfull*) dan kemanfaatan (*usefulness*) dari sebuah kesimpulan yang didapatkan dari interpretasi skor tes.

6. Nursalam (2003)

Pengertian validitas menurut Nursalam adalah suatu ukuran yang menunjukkan tingkat kevalidan atau kesahihan suatu instrumen.

7. Neuman (2007)

Pengertian validitas menurut Neuman adalah menunjukkan keadaan yang sebenarnya dan mengacu pada kesesuaian antara konstruk atau cara seorang peneliti mengonseptualisasikan ide dalam definisi konseptual dan suatu ukuran. Hal ini mengacu pada seberapa baik ide tentang realitas “sesuai” dengan realitas aktual. Dalam istilah sederhana, validitas membahas pertanyaan mengenai seberapa baik realitas sosial yang diukur melalui penelitian sesuai dengan konstruk yang peneliti gunakan untuk memahaminya.

8. Anastasia dan Urbina (1998)

Pengertian validitas menurut Anastasia dan Urbina adalah mengenai apa dan seberapa baik suatu alat tes dapat mengukur, sedangkan reliabilitas merujuk pada konsistensi skor yang dicapai oleh orang yang sama ketika diuji berulang kali dengan tes yang sama pada kesempatan yang berbeda, atau dengan seperangkat butir-butir ekuivalen (*equivalent items*) yang berbeda, atau di bawah kondisi pengujian yang berbeda.

9. Gronlund dan Linn (1990)

Pengertian validitas menurut Gronlund dan Linn adalah ketepatan interpretasi yang dibuat dari hasil pengukuran atau evaluasi.

10. Sukadji (2000)

Pengertian validitas menurut Sukadji adalah derajat yang menyatakan suatu tes mengukur apa yang seharusnya diukur.

11. KBBI (Kamus Besar Bahasa Indonesia)

Pengertian validitas menurut KBBI adalah sifat benar menurut bahan bukti yang ada, logika berpikir atau kekuatan hukum; sifat valid; kesahihan.¹⁰⁴

Sebuah alat pengukur dapat dikatakan valid apabila alat pengukur tersebut dapat mengukur apa yang hendak diukur secara tepat. Demikian pula dalam alat-alat evaluasi. Suatu tes dapat dikatakan memiliki validitas yang tinggi apabila tes tersebut betul-betul dapat mengukur hasil belajar. Jadi bukan sekadar mengukur daya ingatan atau kemampuan bahasa saja misalnya.

Validitas berasal dari kata *validity* yang mempunyai arti sejauh mana ketepatan dan kecermatan suatu alat ukur dalam melakukan fungsi ukurnya (Azwar, 1986). Suatu skala atau instrumen pengukur dapat dikatakan mempunyai validitas yang tinggi apabila instrumen tersebut menjalankan fungsi ukurnya, atau memberikan hasil ukur yang sesuai dengan maksud dilakukannya pengukuran tersebut. Sedangkan tes yang

¹⁰⁴ Diakses di <https://www.seputarpengetahuan.co.id/2020/03/validitas-adalah.html>

memiliki validitas rendah akan menghasilkan data yang tidak relevan dengan tujuan pengukuran.

Validitas tes biasa juga disebut sebagai kesahihan suatu tes adalah mengacu pada kemampuan suatu tes untuk mengukur karakteristik atau dimensi yang dimaksudkan untuk diukur. Valid berarti cocok atau sesuai. Suatu tes dikatakan valid, apabila tes tersebut benar-benar menyasar kepada apa yang dituju. Tes tersebut benar-benar dapat memberikan keterangan atau gambaran tentang apa yang diinginkan. Jika tes itu bahasa, maka tes tersebut harus diberikan gambaran tentang kemampuan dan kecakapan anak dalam hal bahasa, dan bukan menunjukkan gambaran kecakapan anak dalam hal ekonomi, ilmu bumi, dan sebagainya. Guna menjelaskan pengertian valid ini, dapat kita ambil contohnya jika kita ingin mengetahui berat dari suatu benda, maka kita pergunakan alat pengukuran timbangan. Jika ingin mengetahui panjang sesuatu, maka kita pergunakan alat pengukur meteran. Dan jika kita ingin mengetahui suhu sesuatu, kita pergunakan alat pengukur termometer.

Sifat valid diperlihatkan oleh tingginya validitas hasil ukur suatu tes. Suatu alat ukur yang tidak valid akan memberikan informasi yang keliru mengenai keadaan subjek atau individu yang dikenai tes itu. Apabila informasi yang keliru itu dengan sadar atau tidak dengan sadar digunakan sebagai dasar pertimbangan dalam pengambilan suatu keputusan, maka keputusan itu tentu bukan merupakan suatu keputusan yang tepat.

B. Teknik-Teknik Pengujian Validitas Tes Hasil Belajar

Penganalisan terhadap tes hasil belajar sebagai suatu totalitas dapat dilakukan dengan dua cara. Pertama, penganalisan yang dilakukan dengan jalan berpikir secara rasional atau penganalisan dengan menggunakan logika (*logical analysis*). Kedua, penganalisan yang dilakukan dengan mendasarkan diri kepada kenyataan empiris, di mana penganalisan dilaksanakan dengan menggunakan *empirical analysis*.

a) Pengujian Validitas Tes secara Rasional

Tes hasil belajar yang setelah dilakukan penganalisan secara rasional ternyata memiliki daya ketepatan mengukur, disebut tes hasil belajar yang telah memiliki validitas logika (*logical validity*). Validitas rasional adalah validitas yang diperoleh atas dasar hasil pemikiran,

validitas yang diperoleh dengan berpikir secara logis. Validitas rasional adalah validitas yang diperoleh atas dasar hasil pemikiran atau berpikir secara logis (Anas Sudijono, 1995). Dengan demikian, suatu tes hasil belajar dapat dikatakan telah memiliki validitas rasional apabila tes hasil belajar itu memang (secara rasional) dengan tepat telah dapat mengukur apa yang seharusnya diukur.

Untuk dapat menentukan apakah tes hasil belajar sudah memiliki validitas rasional ataukah belum, dapat dilakukan penelusuran dari dua segi, yaitu dari segi isinya (*content*) dan dari segi susunan atau konstruksinya (*construct*).

1. Validitas Isi (*Content Validity*)

Validitas isi adalah validitas yang ditilik dari segi isi tes itu sendiri sebagai alat pengukur hasil belajar peserta didik, isinya telah dapat mewakili secara representatif terhadap keseluruhan materi atau bahan pelajaran yang seharusnya diteskan (diujikan). Jadi, validitas isi sebenarnya identik dengan pembicaraan tentang populasi dan sampel. Kalau saja keseluruhan materi pelajaran yang telah diberikan kepada peserta didik atau sudah diperintahkan untuk dipelajari oleh peserta didik kita anggap sebagai populasi, dan isi tes hasil belajar dalam mata pelajaran yang sama kita anggap sebagai sampelnya, maka tes hasil belajar dalam mata pelajaran tersebut dapat dikatakan telah memiliki validitas isi, apabila isi tes tersebut (sebagai sampel), dapat menjadi wakil yang representatif (layak = memadai) bagi seluruh materi pelajaran yang telah diajarkan atau telah diperintahkan untuk dipelajari (sebagai populasi).

Upaya lain yang dapat ditempuh dalam rangka mengetahui validitas isi dari tes hasil belajar adalah dengan jalan menyelenggarakan diskusi panel. Dalam forum diskusi tersebut, para pakar yang dipandang memiliki keahlian yang ada hubungannya dengan mata pelajaran yang diujikan, diminta pendapat dan rekomendasinya terhadap isi atau materi yang terkandung dalam tes hasil belajar yang bersangkutan. Hasil-hasil diskusi itu selanjutnya dijadikan pedoman atau bahan acuan untuk memperbaiki dan menyempurnakan isi atau materi hasil belajar

tersebut. Jadi kegiatan menganalisis validitas isi dapat dilakukan baik sesudah maupun sebelum tes hasil belajar dilaksanakan.

2. Validitas Konstruksi (*Construct Validity*)

Tes hasil belajar dapat dinyatakan sebagai tes yang telah memiliki validitas konstruksi apabila tes hasil belajar tersebut—ditilik dari segi susunan, kerangka atau rekaannya—secara tepat dapat mencerminkan suatu konstruksi dalam teori psikologis. Tes hasil belajar baru dapat dikatakan telah memiliki validitas susunan apabila butir-butir soal atau *item* yang membangun tes tersebut benar-benar telah dapat dengan secara tepat mengukur aspek-aspek berpikir (seperti aspek kognitif, aspek afektif, aspek psikomotorik dan sebagainya) sebagaimana telah ditentukan dalam tujuan instruksional khusus.

Seperti halnya pada penganalisan validitas isi, maka penganalisan validitas konstruksi juga dapat dilakukan dengan jalan menyelenggarakan diskusi panel. Pengujian validitas konstruksi tes ini pun dapat dilakukan baik sesudah maupun sebelum tes hasil belajar tersebut dilaksanakan.

b) Pengujian Validitas Tes secara Empiris.

Validitas empiris adalah validitas yang bersumber atau diperoleh atas dasar pengamatan di lapangan (Anas Sudijono, 1998). Jadi, tes hasil belajar dapat dikatakan telah memiliki validitas empiris apabila berdasarkan hasil analisis yang dilakukan terhadap data hasil pengamatan di lapangan, terbukti bahwa tes hasil belajar itu dengan secara tepat telah dapat mengukur hasil belajar yang seharusnya diungkap atau diukur melalui tes hasil belajar tersebut.

Untuk dapat menentukan apakah tes hasil belajar sudah memiliki validitas empiris atukah belum, dapat dilakukan penelusuran dari dua segi, yaitu dari segi daya ketepatan meramalnya (*predictive validity*) dan daya ketepatan bandingannya (*concurrent validity*).¹⁰⁵

¹⁰⁵ diakses di <http://hamiddarmadi.blogspot.com/2011/05/validitas-tes-dan-item-tes-hasil.html?m=1>

c) Validitas Ramalan (*Predictive Validity*)

Istilah “ramalan” jika dikaitkan dengan validitas tes, maka yang dimaksud dengan validitas ramalan dari suatu tes adalah suatu kondisi yang menunjukkan seberapa jauh sebuah tes secara tepat menunjukkan kemampuannya untuk meramalkan apa yang bakal terjadi pada masa mendatang. Untuk mengetahui apakah suatu tes hasil belajar dapat dinyatakan sebagai tes yang telah memiliki validitas ramalan ataukah belum, dapat ditempuh dengan cara mencari korelasi antara tes hasil belajar yang sedang diuji validitas ramalannya dengan kriterium yang ada. Jika di antara kedua variabel tersebut terdapat korelasi positif yang signifikan, maka tes hasil belajar yang sedang diuji validitas ramalannya itu, dapat dinyatakan sebagai tes hasil belajar yang telah memiliki daya ramal yang tepat, artinya apa yang telah diramalkan, betul-betul telah terjadi secara nyata dalam praktik.

Dalam rangka mencari korelasi antara tes hasil belajar yang sedang diuji validitas ramalannya dengan kriteria yang telah ditentukan itu, cara sederhana yang paling sering digunakan adalah dengan menerapkan Teknik Analisis Korelasional Product Moment dari Karl Pearson. Hipotesis nihil (H_0) yang akan diuji, dirumuskan dalam susunan kalimat sebagai berikut: “Tidak terdapat korelasi positif yang signifikan, antara tes hasil belajar yang sedang diuji validitas ramalannya (= variabel X), dengan kriterium yang telah ditentukan (= variabel Y).”

Tes hasil belajar dapat dikatakan bersifat valid atau memiliki validasi, apabila tes tersebut dengan secara tepat, benar, dan sah atau absah dapat mengukur apa yang seharusnya diukur. Dengan kata lain tes hasil belajar dapat dinyatakan valid apabila tes hasil belajar tersebut (sebagai alat pengukur keberhasilan belajar peserta didik) dengan secara tepat, benar, sah atau absah telah dapat mengukur atau mengungkap hasil-hasil belajar yang telah dicapai oleh peserta didik, setelah mereka menempuh proses belajar mengajar dalam jangka waktu tertentu.

Untuk menetapkan apakah sebuah tes hasil belajar dapat dinyatakan sebagai tes yang telah memiliki validitas atau daya ketepatan mengukur, ataukah belum, dapat dilakukan penganalisan secara rasional atau secara logika (*logical analysis*) dan dapat pula dilakukan penganalisan secara empiris (*empirical analysis*).

C. Teknik Pengujian Validitas *Item* Tes Hasil Belajar

a. Pengertian validitas *item*

Validitas *item* dari sebuah tes adalah ketepatan mengukur yang dimiliki oleh sebuah *item* (yang merupakan bagian yang tak terpisahkan dari tes bagian suatu totalitas), dalam mengukur apa yang seharusnya diukur lewat butir *item* tersebut.

b. Teknik Pengujian Validitas *Item* Tes Hasil Belajar

Berdasarkan uraian di atas maka cukup jelas bahwa sebuah *item* dapat dikatakan telah memiliki validitas tinggi atau dapat dinyatakan valid, jika skor-skor pada butir *item* yang bersangkutan memiliki kesesuaian atau kesejajaran arah dengan skor totalnya, skor total di sini berkedudukan sebagai variabel terikat sedangkan variabel *item* merupakan variabel bebasnya.

Permasalahannya adalah bagaimana memilih dan menentukan jenis teknik dalam rangka menguji validitas *item* itu. Seperti yang diketahui pada tes objektif maka hanya ada dua kemungkinan yaitu betul atau salah.

Setiap butir soal yang dapat dijawab dengan benar diberikan skor 1 (satu) sedangkan untuk setiap jawaban yang salah diberikan skor 0 (nol) jenis data seperti ini biasanya merupakan tes benar – salah, ya – tidak dan sejenisnya dalam ilmu statistik dikenal dengan diskret murni atau data dikotomis. Sedangkan, skor total yang dimiliki oleh masing-masing *testee* adalah merupakan penjumlahan dari setiap skor itu merupakan data kontinu.

Berdasarkan teori yang ada apabila variabel I berupa data dikotomis sedangkan variabel II data kontinu maka, teknik korelasi yang tepat untuk digunakan dalam mencari korelasi dua variabel adalah teknik korelasi *point biserial*, di mana angka indeks korelasi diberi lambang r_{pbi} dapat diperoleh dengan rumus:

$$R_{pbi} = \frac{M_p - M_t}{SD_t} \text{ dikali akar } p/r$$

Di mana:

R_{pbi} = koefisien korelasi *point biserial* yang melambangkan kekuatan korelasi antara variabel I dan II yang dalam hal ini sebagai koefisien validitas *item*.

- M_p = skor rata-rata hitung yang dimiliki oleh *testee*
 M_t = skor rata-rata dari skor total
 SD_t = standar deviasi dari skor total
 p = proporsi *testee* yang menjawab dengan benar terhadap butir *item* yang diuji validitasnya.
 q = proporsi *testee* yang menjawab salah terhadap butir *item* yang diuji validitasnya.¹⁰⁶

D. Faktor-Faktor yang Memengaruhi Validitas Tes

Pengaruh latar belakang kebudayaan siswa seperti keadaan status ekonomi, perbedaan struktur sosial (status sosial tinggi/rendah) perbedaan jenis kelamin tersebut akan memengaruhi validitas tes. Sebagai contoh tes inteligensi Binet Simon, yang dipersiapkan untuk anak-anak Eropa, akan tidak valid jika dikenakan terhadap anak Indonesia yang mempunyai latar belakang kebudayaan/pengetahuan/kehidupan yang berbeda-beda. Demikian juga apabila tes tersebut sama-sama dikenakan sama anak desa, barangkali akan memberikan hasil rata-rata yang berbeda, apabila *group*/kelompok tersebut terdiri dari anak-anak yang mempunyai status sosial ekonomi rendah. Perbedaan tersebut disebabkan antara lain di samping perbedaan-perbedaan latar belakang sosial ekonomi akan berpengaruh terhadap pembentukan pola tingkah laku anak, juga akan berpengaruh terhadap persepsi anak dan pola responsnya dalam menanggapi suatu masalah.

Keadaan tipe tes; tiap-tiap tes akan berpengaruh terhadap validitas jawaban anak. Tes subjektif (*essay test*) akan mempunyai validitas yang berbeda dibandingkan dengan tes objektif. Tes subjektif akan mempunyai validitas yang lebih rendah, karena dengan tes subjektif kurang dapat mengungkapkan materi tes secara tuntas dibandingkan dengan tes objektif. Demikian juga di antara macam-macam tipe tes objektif, tes yang terdiri dari 2 alternatif (benar-salah, setuju-tidak setuju) akan mempunyai validitas yang lebih rendah dibandingkan dengan tes objektif tipe pilihan ganda. Hal itu disebabkan karena tes benar salah mempunyai tendensi untuk menerka-nerka yang lebih besar dibandingkan tes pilihan ganda.

¹⁰⁶ Diakses di https://www.academia.edu/31379497/TEKNIK_PENGUJIAN_VALIDITAS_TES_BAB_I.docx

Usaha menambah butir soal dengan maksud ingin mempertinggi reliabilitas tes; memang salah satu cara untuk mempertinggi reliabilitas tes adalah dengan menambah butir-butir soal tes secara otomatis akan dapat mempertinggi reliabilitas, salahnya justru akan dapat merendahkan validitas tes. Oleh karena itu, jika ingin meningkatkan reliabilitas tes dengan tidak akan merendahkan validitas tes, hendaknya penambahan butir-butir soal tes tersebut dengan menambahkan faktor-faktor baru dari materi tesnya, karena hal itu berarti memperbanyak/memperbesar sampel bahan tes.

Kurang jelasnya petunjuk mengerjakan tes; adanya petunjuk/pengarahan bagaimana cara mengerjakan tes dimaksudkan agar tidak terjadi kesalahan-kesalahan interpretasi tentang bagaimana cara mengerjakan tes tersebut. Jika petunjuknya kurang jelas, akan terjadi siswa yang memperoleh skor rendah tersebut barangkali disebabkan karena mengalami kesulitan dalam menginterpretasi petunjuknya, bukan karena tidak menguasai materinya.

BAB VI

CONTOH-CONTOH PENGOLAHAN EVALUASI TES DAN NONTES

A. Evaluasi

1. Pengertian Evaluasi

Secara etimologi evaluasi terambil dari kata Inggris yaitu *evaluation*, dan secara bahasa Arabnya التقدير, sedangkan dari bahasa Indonesia bermakna *penilaian*. Yang akar katanya adalah (*value*), sedangkan dari bahasa Arabnya القيمة, dan dari bahasa Indonesianya berarti *nilai*. Jika dari istilah evaluasi merupakan sebuah tindakan proses dalam menentukan nilai dari sesuatu tersebut. Dan belajar merupakan proses yang dilakukan tiap individu untuk memperoleh suatu perubahan yang positif. Yang berproses terdapat evaluasi adalah sebagai penentu dari dampak serta keefektifan penggunaannya dalam kegiatan pembelajaran.¹⁰⁷

Untuk memperoleh gambaran lengkap perlu dilakukan evaluasi baik terhadap proses maupun hasilnya. Aspek yang ingin diketahui dalam proses antara lain dampak media dan metode yang digunakan dalam proses pembelajaran.¹⁰⁸ Sedangkan dari hasilnya, yang ingin dinilai adalah

¹⁰⁷ Anas Sudijono, *Pengantar Evaluasi Pendidikan*, (Jakarta: PT RajaGrafindo Persada, 2007). hlm. 1.

¹⁰⁸ Suharsimi Arikunto, *Dasar-Dasar Evaluasi Pendidikan*, (Jakarta: Bumi Aksara, 2005). hlm. 17.

ketercapaian kompetensi atau tujuan yang telah ditetapkan untuk peserta didik. Selain itu, masalah pertama yang harus dilakukan dalam pelaksanaan adalah merumuskan tujuan evaluasi yang hendak dilaksanakan dalam suatu proses pelaksanaan program pembelajaran di kelas didasarkan atas tujuan yang hendak dicapai dalam program tersebut. Jadi, evaluasi belajar adalah suatu proses yang dilakukan untuk menentukan penilaian terhadap individu peserta didik guna mencapai perubahan yang positif.

2. Tujuan Evaluasi

a. Tujuan Umum Evaluasi

Pada umumnya evaluasi merupakan bagian dari sebuah rangkaian suatu kegiatan untuk memotivasi ke tahap kerja yang lebih baik tertuju kepada pelaksanaan program lembaga. *Pertama*, tujuan utama evaluasi fokus kepada proses pelaksanaan dalam belajar dan mengajar. *Kedua*, setelah melakukan evaluasi maka akan memperoleh hasil pencapaian dalam sebuah pembelajaran apakah sudah berhasil atau belum berhasil. *Ketiga*, kegunaan dalam pengevaluasian maka secara langsung akan mendapatkan berbagai informasi yang bersumber dari kelas yang telah diajar serta dari hasil evaluasi terhadap pendidik dalam mengayomi mutu peningkatan kualitas dan kuantitas pada pemberian materi ajar. *Keempat*, dengan adanya penggunaan evaluasi maka di dalam pengajaran yang dilakukan oleh tenaga pendidik dapat diketahui berhasil atau tidak proses pembelajaran yang telah terencana tersebut. *Kelima*, penggunaan evaluasi yang diterapkan pada pendidikan maka secara langsung akan mendapatkan sebuah hasil yang dilihat dari sistem program evaluasi itu sendiri.

b. Tujuan Khusus Evaluasi

Dilihat pada tujuan secara khusus pada evaluasi pendidikan, maka dapat diketahui, di antaranya, untuk memberikan klarifikasi tentang sifat hasil pembelajaran yang telah dilaksanakan, memberikan informasi tentang ketercapaian tujuan jangka pendek yang telah dilaksanakan, memberikan masukan untuk kemajuan pembelajaran, memberikan informasi tentang kesulitan dalam

pembelajaran dan untuk memilih pengalaman pembelajaran di masa yang akan datang.

Pada prinsipnya tujuan evaluasi pendidikan adalah untuk melihat dan mengetahui proses yang terjadi dalam proses pembelajaran. Dalam kapasitasnya proses pembelajaran memiliki tiga hal penting yaitu input, transformasi dan *output*, untuk dievaluasi. a. Input adalah peserta didik yang telah dinilai kemampuannya dan siap menjalani proses pembelajaran. b. Transformasi adalah segala unsur yang terkait dengan proses pembelajaran yaitu guru, media dan bahan belajar, metode pengajaran, sarana penunjang dan sistem administrasi. c. *Output* adalah capaian yang dihasilkan dari proses pembelajaran. Jika kita ingin melakukan kegiatan evaluasi, terlepas dari jenis evaluasi apa yang digunakan, terdapat tiga hal yang perlu diperhatikan, antara lain: a. Guru harus mengetahui dan memahami terlebih dahulu tentang tujuan dan fungsi evaluasi. Bila tidak, maka guru akan mengalami kesulitan merencanakan dan melaksanakan evaluasi. Hampir setiap orang yang membahas evaluasi pula tentang tujuan dan fungsi evaluasi. b. Tujuan evaluasi pembelajaran adalah untuk mengetahui keefektifan dan efisiensi sistem pembelajaran, baik yang menyangkut tentang tujuan materi, metode, media sumber belajar, lingkungan maupun sistem penilaian itu sendiri. c. Tujuan khusus evaluasi pembelajaran disesuaikan dengan jenis evaluasi pembelajaran itu sendiri, seperti evaluasi perencanaan dan pengembangan, evaluasi *monitoring*, evaluasi dampak, evaluasi efisiensi ekonomi, dan evaluasi program komprehensif.

Proses evaluasi dalam pendidikan apabila sekolah diumpamakan sebagai tempat untuk proses produksi, dan calon peserta didik diumpamakan sebagai bahan mentah, maka lulusan dari sekolah itu hampir sama dengan produk hasil olahan yang sudah siap digunakan disebut juga dengan ungkapan transformasi dimasukkan ke dalam transformasi. Dalam dunia sekolah maka yang dimaksud dengan bahan mentah adalah calon peserta didik yang baru akan memasuki sekolah. Sebelum memasuki sesuatu tingkat sekolah (*institusi*) calon peserta didik itu dinilai dahulu kemampuannya. Dengan penelitian

itu diketahui apakah kelak akan mampu mengikuti pelajaran dan melaksanakan tugas-tugas yang akan diberikan kepadanya.¹⁰⁹

Ouput adalah bahan jadi yang dihasilkan oleh transformasi. Yang dimaksud dalam pembicaraan ini adalah peserta didik lulusan sekolah yang bersangkutan untuk dapat menentukan apakah peserta didik berhak lulus atau tidak, perlu diadakan kegiatan penilaian. Transformasi adalah mesin yang bertugas mengubah bahan mentah menjadi bahan jadi. Dalam dunia sekolah, sekolah itulah yang dimaksud dengan transformasi. Sekolah itu sendiri terdiri dari beberapa mesin yang menyebabkan berhasil atau gagalnya sebagai transformasi.

Bahan jadi yang diharapkan dalam hal ini peserta didik lulusan sekolah ditentukan oleh beberapa faktor sebagai akibat pekerjaannya unsur-unsur yang ada. Unsur-unsur transformasi sekolah tersebut antara lain: a. Guru dan personal lainnya. b. Metode mengajar dan sistem evaluasi. c. Sarana penunjang. d. Sistem administrasi. Umpan balik (*feedback*) adalah segala informasi baik yang menyangkut *output* maupun transformasi. Umpan balik ini diperlukan sekali untuk memperbaiki input maupun transformasi. Lulusan yang kurang bermutu atau yang tidak siap pakai yang belum memenuhi harapan, akan menggugah semua pihak untuk mengambil tindakan yang berhubungan dengan penyebab kurang bermutunya lulusan. Penyebab-penyebab tersebut antara lain: a. Input yang kurang baik kualitasnya. b. Guru dan personal yang kurang tepat (kualitas). c. Materi yang tidak atau kurang cocok.

3. Ciri-Ciri Evaluasi

Suharsimi menjelaskan ciri-ciri yang terdapat pada sebuah evaluasi dapat dilakukan dengan tindakan secara langsung, yang berkaitan kepada kemampuan yang dimiliki seorang anak serta mengukur kelebihan pada

¹⁰⁹ Zainal Arifin. *Evaluasi instruksional*. (Bandung: PT. Remaja Rosdakarya.1990). hlm. 24.

kemampuannya di setiap penyelesaian pada soal-soal yang telah diberikan.¹¹⁰

Mengacu kepada, kemampuan anak mengerjakan setiap soal, anak mampu menempatkan penggunaan bahasa lisan dan tulisan dengan baik, anak mempunyai kemampuan di dalam menempatkan sebuah pengetahuan baru yang didapatkan melalui informasi yang didengarkannya, anak mempunyai kemampuan dalam mengingat materi ajar yang telah diberikan, anak mempunyai kemampuan di dalam membedakan antara keseriusan di dalam memahami pelajaran kemudian menyandingkannya pada sebuah penganalisisan terhadap materi ajar yang diberikan sehingga anak mampu menjabarkannya dalam bentuk cerita dan lain-lain. Untuk melihat sejauh mana kemampuan peserta didik, maka dapat dilakukan dengan cara mengukur kemampuan dalam berpikir yang terdapat pada inteligensi maka dapat dipaparkan sebagai berikut:

Pengukuran Kemampuan Intelegensi						
Luar Biasa	Dungu	Bodoh	Normal	Pandai	Sangat pandai	Jenius
1 %, IQ 30-70	5 %, IQ 70-80	14%, IQ 80-90	60%, IQ 90-110	14%, IQ 110-120	5%, IQ 120-130	1%, IQ lebih dari 130

Dari penjelasan tabel di atas tentang pengukuran kemampuan yang terkait kepada IQ dapat dipahami dengan adanya gambaran yang terdapat pada angka-angka, kepada ciri-ciri yang terdapat pada sebuah evaluasi pendidikan dalam penggunaan pengukuran evaluasi bersifat kuantitatif yang bersimbol kepada bilangan penghasilan pertama pengukuran. Selanjutnya diaplikasikan kepada pengukuran peserta didik seperti berikut, Tina mempunyai IQ 125, sedangkan IQ yang dimiliki Tika berukuran 105. Dalam hal ini maka dapat dikategorikan kedua peserta didik tersebut memiliki kemampuan yang lebih terlihat pada ukuran angka IQ nya masing-masing. Tina termasuk anak yang pandai, begitupun Tika termasuk juga anak yang normal. Maka dengan penjelasan ini terlihat bahwa

¹¹⁰ Suharsimi Arikunto, *Dasar-Dasar Evaluasi Pendidikan*, (Jakarta: Bumi Aksara, 2002). hlm. 11.

evaluasi di lembaga pendidikan telah menggunakan unit satuan yang menempatkan IQ angka 105 termasuk kategori anak normal yaitu bahwa penilaian pendidikan menggunakan unit-unit untuk satuan-satuan yang tetap karena IQ 105 termasuk anak normal.¹¹¹

Selanjutnya terlihat bahwa evaluasi pendidikan relatif tidak sama, atau tidak serupa dan bahkan tidak tetap dari waktu ke waktu. Sebagai contoh hasil ulangan yang diperoleh Sonia hari Selasa adalah 80, hasil hari Rabu 90, tetapi hasil ulangan hari Jumat hanya 50. Pada perubahan keadaan tidak tetap ini dilihat dari berbagai faktor. Karena evaluasi pada pendidikan ini sering terdapat kesalahan-kesalahan yang bersumber kepada letak pengukurannya. Alat ukur yang digunakan mestilah bagus sehingga tidak terdapat kesalahan seperti pengukuran pada panjang meja. Pada pengukuran menggunakan pita yang berukuran meja itu sendiri dan pita pun terbuat dari bahan yang elastis dan mudah ditarik-tarik. Dengan demikian, pita tersebut tergolong kepada alat ukur yang baik dikarenakan gambaran pada panjangnya didapatkan secara pasti, cara-cara penggunaan alat ukur di lembaga pendidikan. Kemudian terlihat kepada orang yang melakukan evaluasi, dalam hal ini, disebut kesalahan pada waktu melakukan penilaian, karena faktor subjektif penilai telah berpengaruh pada hasil pengukuran. Tulisan jelek dan tidak jelas, mau tidak mau sering memengaruhi subjektivitas penilai, jika pada waktu mengerjakan koreksi, penilai itu sendiri sedang risau. Itulah sebabnya pendidik harus sejauh mungkin dari hal itu. Kecenderungan dari penilai untuk memberikan nilai secara “murah” atau “mahal”.

Ada guru yang memberi nilai dua untuk peserta didik yang menjawab salah dengan alasan untuk upah menulis. Tetapi ada yang memberikan nol untuk jawaban yang serupa. Adanya (*hallo/effect*), yakni adanya kesan menilai terhadap peserta didik. Kesan-kesan itu dapat berasal dari guru yang lain maupun dari guru itu sendiri pada kesempatan memegang mata pelajaran itu, Adanya pengaruh hasil yang telah diperoleh terdahulu. Seorang peserta didik pada ulangan pertama mendapat angka sepuluh sebanyak dua belas kali. Untuk ulangan yang ketiga belas dan seterusnya, guru sudah terpengaruh ingin memberi angka lebih banyak dari

¹¹¹ Asrul, dkk. *Evaluasi Pembelajaran*. (Bandung: Cita pustaka media. 2014). hlm. 4

sebenarnya pada waktu ulangan tersebut, ia sedang mengalami nasib sial, yakni salah mengerjakan. Kesalahan yang disebabkan oleh kekeliruan menjumlah angka hasil penilaian.

Objek evaluasi pendidikan aspek-aspek yang diperlukan dalam evaluasi terhadap peserta didik meliputi, aspek-aspek tentang berpikir, termasuk di dalamnya: intelegensi, ingatan, cara menginterupsi data, prinsip-prinsip pengerjaan pemikiran logis. Perasaan sosial termasuk di dalamnya: cara bergaul, cara pemecahan nilai-nilai sosial, cara menghadapi dan cara berpartisipasi dalam kenyataan sosial. Keyakinan sosial dan kewarganegaraan menyangkut pandangan hidupnya terhadap masalah-masalah sosial, politik dan ekonomi, apresiasi seni dan budaya, minat, bakat dan hobi, perkembangan sosial dan personal.¹¹²

Ada dua yang digunakan sebagai alat untuk mengukur hasil belajar dan mengajar yang terdapat pada tes dan nontes. Pada hasil belajar tidak hanya dinilai oleh tes saja akan tetapi terlihat juga dari sisi objektifnya. Dengan bantuan kepada alat tes dan alat nontes. Alat yang dimaksud seperti alat wawancara, alat kuesioner, alat skala “skala penilaian dan skala sikap”. Selanjutnya, melalui alat observasi/pengamatan, alat studi kasus, dan alat sosiometri. Pada wawancara dan kuesioner pada umumnya digunakan untuk menilai aspek kognitif seperti pendapat atau pandangan seorang serta harapan dan aspirasinya di samping aspek afektif dan perilaku individu. Skala bisa digunakan untuk menilai aspek afektif seperti skala sikap dan skala minat serta aspek kognitif seperti skala penilaian.

Observasi pada umumnya digunakan untuk memperoleh data dalam kegiatan pendidikan mempunyai arti yang sangat utama, karena evaluasi merupakan alat ukur atau proses untuk mengetahui tingkat pencapaian keberhasilan yang telah dicapai peserta didik atas bahan ajar atau materi-materi yang telah disampaikan. Dengan adanya evaluasi maka tujuan dari pembelajaran akan terlihat secara akurat dan meyakinkan. Berdasarkan beberapa pengertian mengenai evaluasi kita dapat menarik kesimpulan bahwa evaluasi adalah suatu alat yang dijadikan sebagai patokan dalam menilai sesuatu agar kita mengetahui baik atau tidaknya atau apakah rencana yang kita miliki ini dapat berkembang atau tidak.

¹¹² Anas Sudijono. *Pengantar Evaluasi Pendidikan*. hlm. 71-72.

Dalam pembelajaran evaluasi dapat menentukan apakah pembelajaran yang kita laksanakan di kelas dapat mencapai tujuan yang kita inginkan atau tidak. Maka dari itu evaluasi ini sangat penting untuk kita mengetahui suatu nilai. Menurut Mardapi dalam berpendapat bahwa dalam kegiatan pembelajaran evaluasi memiliki peranan penting dalam terwujudnya tujuan pembelajaran. Oleh karena itu, maka evaluasi dalam pendidikan mencakup istilah-istilah yang sering digunakan dalam pelaksanaannya yaitu tes, pengukuran, dan penilaian.

Tes merupakan salah satu cara untuk menaksir besarnya kemampuan seseorang secara tidak langsung, yaitu melalui respons seseorang terhadap stimulus atau pertanyaan. Oleh karena itu, tes dijadikan salah satu alat untuk melakukan pengukuran, yaitu alat untuk mengumpulkan informasi karakteristik suatu objek. Secara garis besar ada dua kategori alat penilaian hasil belajar yaitu tes dan nontes. Hasil belajar dan proses belajar tidak hanya dinilai oleh tes, baik melalui bentuk tes uraian maupun tes objektif, tetapi juga dapat dinilai oleh alat-alat nontes atau bukan tes. Alat-alat bukan tes yang sering digunakan antara lain ialah wawancara, kuesioner, skala (skala penilaian, skala sikap), observasi atau pengamatan, studi kasus, dan sosiometri. Wawancara dan kuesioner pada umumnya digunakan untuk menilai aspek kognitif seperti pendapat atau pandangan seorang serta harapan dan aspirasinya di samping aspek afektif dan perilaku individu. Skala bisa digunakan untuk menilai aspek afektif seperti skala sikap dan skala minta serta aspek kognitif seperti skala penilaian.

B. Tes

1. Pengertian Tes

Secara sederhana tes diartikan sebagai sekumpulan soal-soal yang mesti dijawab, dan pernyataan-pernyataan yang mesti dipilih ataupun ditanggapi, serta tugas-tugas yang mesti dilakukan oleh peserta tes dengan tujuan untuk mengukur suatu aspek tertentu dari peserta tes. Yang berkaitan kepada proses pembelajaran yang tertuju kepada indikator tercapainya kompetensi. Sedangkan secara bahasa yang diambil dari bahasa Inggris kata tes itu *testum*, yang bermakna piring yang dipisahkan melalui logam mulia bahannya berasal dari material seperti pasir, batu, dan

tanah. Kemudian dimaknai pada ilmu psikologi dalam pendidikan yang menjelaskan sebuah instrumen yang akan digunakan sebagai pengukuran seberapa jauh kemampuan yang dimiliki peserta didik. Cronbach (dalam Azwar, 2005) mendefinisikan tes sebagai (*a systematic procedure for observing a person's behavior and describing it with the aid of a numerical scale or category system*). Menurut Ebster's Collegiate tes adalah serangkaian pertanyaan atau latihan atau alat lain yang digunakan untuk mengukur keterampilan, pengetahuan, intelegensia, kemampuan atau bakat yang dimiliki oleh individu atau kelompok.¹¹³

Dari dua definisi tersebut dapat diuraikan maknanya, *pertama*, tes adalah prosedur pengukuran yang sengaja dirancang secara sistematis, untuk mengukur atribut tertentu, dilakukan dengan prosedur administrasi dan pemberian angka yang jelas dan spesifik, sehingga hasilnya relatif baik bila dilakukan dalam kondisi yang relatif sama. *Kedua* tes pada umumnya berisi sampel perilaku, cakupan butir tes yang bisa dibuat dari suatu materi tidak terhingga jumlahnya, yang secara keseluruhan mungkin mustahil dapat tercakup dalam tes, sehingga tes harus dapat mewakili kawasan "*domain*" perilaku yang diukur, untuk itu perlu pembatasan yang jelas. *Ketiga* tes menghendaki subjek agar menunjukkan apa yang diketahui atau apa yang dipelajari dengan cara menjawab atau mengerjakan tugas dalam tes.

Respons subjek atas tes merupakan perilaku yang ingin diketahui dari penyelenggaraan tes, karena tes memang mengukur perilaku, sebagai manifestasi atribut psikologis yang mau diukur. Tes pada dasarnya adalah alat ukur atribut psikologis yang objektif atas sampel perilaku tertentu. Dalam psikologi, tes dapat diklasifikasikan menjadi empat, yaitu: tes yang mengukur intelegensia umum yang dirancang untuk mengukur kemampuan umum seseorang dalam suatu tugas, tes yang mengukur kemampuan khusus atau tes bakat yang dibuat untuk mengungkap kemampuan potensial dalam bidang tertentu, tes yang ditujukan untuk mengukur prestasi yang digunakan untuk mengungkapkan kemampuan aktual sebagai hasil belajar, tes yang mengungkap aspek kepribadian

¹¹³ Anas Sudijono. *Pengantar Evaluasi Pendidikan*. hlm. 82.

“*personality assessment*” yang bertujuan mengungkap karakteristik individual subjek dalam aspek yang diukur.¹¹⁴

Dengan melihat penggolongan di atas, tes dalam pembelajaran di kelas yang menjadi pembahasan ini adalah tes prestasi atau hasil belajar. Tes sebagai alat ukur dapat menyediakan informasi-informasi objektif yang dapat digunakan sebagai pertimbangan dalam penentuan keputusan yang harus diambil pendidik terhadap proses dan hasil belajar yang dilakukan siswa dapat dibagi menjadi tiga kelompok besar yaitu: a. keputusan yang diambil pada permulaan proses pembelajaran.

Penggunaan tes sebagai dasar pengambilan keputusan pada permulaan proses pembelajaran bermuara pada dua pertanyaan yang harus dijawab oleh pendidik sebelum memulai proses pembelajaran yaitu sejauh manakah pengetahuan, keterampilan, dan kemampuan yang harus dimiliki oleh siswa sebelum mengikuti Asesmen Pembelajaran di SD 4-5 proses pembelajaran yang berupa kemampuan awal yang diperlukan untuk mengikuti proses pembelajaran, sejauh manakah kemampuan dan keterampilan yang telah dicapai peserta didik terhadap pembelajaran yang direncanakan. Keduanya akan menentukan keputusan guru dalam merancang materi dan metode pembelajaran yang direncanakan.

2. Jenis-Jenis Tes

Ada lima jenis pembagian tes, yaitu pembagian jenis tes berdasarkan tujuan penyelenggaraan, jenis tes berdasarkan waktu penyelenggaraan, pembagian jenis tes berdasarkan cara mengerjakan, pembagian jenis tes berdasarkan cara penyusunan, dan pembagian jenis tes berdasarkan bentuk jawaban.

Uraian selengkapnya adalah jenis tes berdasarkan tujuan penyelenggaraan Saudara. Untuk mengawali pembahasan tentang jenis-jenis tes, Anda akan diminta untuk menjawab satu pertanyaan di bawah ini. (*Untuk apakah Anda menyelenggarakan tes?*). Saudara, cobalah berpikir sejenak untuk merenungkan jawaban atas pertanyaan tadi. Anda diminta untuk tidak tergesa-gesa melanjutkan membaca 4-6 Unit 4 penjelasan berikut ini sebelum Anda menentukan jawaban.

¹¹⁴ Mulyadi. *Evaluasi Pendidikan Pengembangan Model Evaluasi Pendidikan di Sekolah*. (Malang: UIN Maliki Press.2010). Hal 63.

Jika Anda sudah menemukan jawaban, barangkali jawaban Anda akan lebih dari satu jawaban, seperti yang tertera di bawah ini. Tes diselenggarakan dengan tujuan untuk keperluan seleksi, untuk menempatkan orang pada kelas-kelas tertentu, untuk mengetahui hasil belajar, untuk keperluan diagnostik, dan untuk keperluan uji coba. Penjelasan secara rinci adalah tes seleksi “*selection test*” Saudara, Anda bisa memahami hakikat dari tes seleksi ini dari arti kata “seleksi” itu sendiri, yaitu memilih.

Jadi, tes seleksi diselenggarakan untuk memilih peserta guna diikutsertakan dalam kegiatan yang menuntut kemampuan tertentu. Penentuan jenis kemampuan dan tingkat penguasaan pada tes seleksi, sepenuhnya tergantung pada kebutuhan akan kemampuan yang dibutuhkan untuk dapat mengikuti kegiatan. Dengan demikian, berdasarkan hasil tes seleksi, seseorang dapat dinyatakan diterima atau berhasil dan tidak diterima atau tidak lolos untuk mengikuti program kegiatan yang direncanakan. Sebagai contoh, jika kita menyelenggarakan tes seleksi untuk pemandu wisata, maka akan lebih baik menitikberatkan kemampuan berbicara daripada kemampuan menulis.

Seperti latihan *sekarang cobalah Anda tentukan, kemampuan manakah yang lebih Anda pentingkan dalam tes seleksi untuk seorang yang akan melakukan tugas redaksional dalam sebuah media cetak? Kemampuan menulis, kemampuan berbicara atau kemampuan menyimak? Mengapa kemampuan yang Saudara tentukan itu penting?* tes penempatan “*placement test*” Saudara, adalah suatu keniscayaan bahwa kemampuan seseorang tidaklah bisa sama. Sekelompok orang barangkali memiliki kemampuan lebih tinggi dari pada kelompok lainnya. Permasalahan yang muncul adalah, bagaimanakah jika kemampuan siswa dalam satu kelas relatif beragam? Hal ini akan bisa mempersulit jalannya proses pengajaran yang Anda lakukan. Untuk itu perlu dilakukan tes penempatan. Tes penempatan umumnya diselenggarakan menjelang dimulainya suatu program pengajaran, dengan maksud untuk menempatkan seseorang pada kelompok yang sesuai dengan tingkat kemampuan yang dimilikinya.

Tes hasil belajar “*achievement test*” tentu tidak lagi asing bagi Saudara. Brown memberikan pengertian tes hasil belajar merupakan (*a test to see how far students achieve materials addressed in a curriculum within*

a particular time frame). Hasil belajar yang diungkap lewat tes hasil belajar dapat mengacu pada hasil pengajaran secara keseluruhan pada akhir penyelenggaraan atau pada kurun waktu tertentu. Sebagai tes yang memfokuskan pada hasil yang telah dapat dicapai oleh suatu bentuk pengajaran, tes hasil belajar memiliki kaitan yang erat dengan apa yang telah diajarkan “kurikulum”. Kaitan itu terutama dalam hal isi tes. Isi tes harus secara jelas mencerminkan isi pengajaran yang secara nyata telah diselenggarakan.

Tes diagnostik “*diagnostic test*” secara bahasa, diagnostik diambil dari bahasa Inggris (*diagnostic*). Bentuk kata kerjanya adalah (*to diagnose*), yang artinya (*to determine the nature of disease from observation of symptoms*). Mendiagnosis berarti melakukan observasi terhadap penyakit tertentu, sebagai dasar menentukan macam atau jenis penyakitnya. Jadi, tes diagnostik sengaja dirancang sebagai alat untuk menemukan kesulitan belajar yang sedang dihadapi siswa. Hasil tes diagnostik dapat digunakan sebagai dasar penyelenggaraan pengajaran yang lebih sesuai dengan kemampuan siswa sebenarnya, termasuk kesulitan-kesulitan belajarnya.

Tes ini dilakukan apabila diperoleh informasi bahwa sebagian besar peserta didik gagal dalam mengikuti proses pembelajaran pada mata pelajaran tertentu. Hasil tes diagnostik memberikan 4-8 unit 4 informasi tentang konsep-konsep yang belum dipahami dan yang telah dipahami. Oleh karenanya, tes ini berisi materi yang dirasa sulit oleh siswa, namun tingkat kesulitan tes ini cenderung rendah. tes uji coba apabila sebagai seorang guru pasti pernah menggunakan tes. Tes yang digunakan belum tentu memenuhi kualifikasi sebagai tes yang (*baik*) dalam arti luas. Untuk mengetahui apakah tes yang dikembangkan bagus, perlu serangkaian uji coba, untuk memperoleh informasi, tidak hanya tentang ciri-ciri tes yang penting, seperti validitas, reliabilitas, tingkat kesulitan, dan tingkat pembeda, melainkan juga segi-segi lain, seperti kecukupan waktu, kejelasan tulisan maupun perintah tes, dan lain sebagainya.

Jenis tes berdasarkan tahapan atau waktu penyelenggaraan selanjutnya diajak untuk memperhatikan jenis tes berdasar waktu penyelenggaraan tes, yang terbagi menjadi empat yaitu, pertama tes masuk “*entrance test*”. Tes masuk diselenggarakan sebelum dan menjelang suatu

program pengajaran dimulai. Sama dengan tes seleksi, tes masuk diselenggarakan untuk menentukan apakah seorang calon dapat diterima sebagai peserta program pengajaran karena ia memiliki jenis dan kemampuan yang dipersyaratkan. Tes masuk dirancang secara khusus dan disesuaikan dengan tujuan program pengajaran. Semakin sesuai isi tes masuk itu dengan tujuan pokok program pengajaran, maka akan semakin tinggi tingkat relevansi serta efektivitas dari tes masuk tersebut.

Tes formatif "*formative test*" dilakukan pada saat program pengajaran sedang berlangsung (*progress*), tujuannya untuk memperoleh informasi tentang jalannya pengajaran sampai tahap tertentu. Informasi tersebut penting untuk mengetahui apakah program pengajaran berjalan sesuai dengan format yang ditentukan sehingga dipertahankan atau program pembelajaran memerlukan perubahan atau penyesuaian, hasilnya berguna untuk memperbaiki strategi mengajar.

Tes ini dilakukan secara periode sepanjang rentang proses pembelajaran, materi tes dipilih berdasarkan tujuan pembelajaran tiap pokok bahasan atau subpokok materi. Jadi tes untuk Asesmen Pembelajaran menentukan keberhasilan belajar dan untuk mengetahui keberhasilan proses pembelajaran. Selanjutnya, tes sumatif "*summative test*", kata dari sumatif adalah sum yang berarti *total obtained by adding together items, numbers or amounts*. Dengan demikian, tes sumatif diselenggarakan untuk mengetahui hasil pengajaran secara keseluruhan "*total*".

Risiko yang mesti didapatkan ketika melakukan tes yang menekankan hasil pengajaran secara keseluruhan, maka *item* tes sumatif atau bahan cakupannya meliputi seluruh materi yang telah disampaikan. Tes sumatif diberikan di akhir suatu pelajaran, atau akhir semester. Hasilnya untuk menentukan keberhasilan belajar peserta didik. Tingkat keberhasilan dinyatakan dengan skor atau nilai, pemberian sertifikat, dan sejenisnya. *Pre-test* dan *post test* untuk mengetahui kemampuan yang dimiliki seorang siswa di awal program pengajaran, kadang-kadang diselenggarakan *pre-test*. Sehingga hasil *pre-test* digunakan untuk mengetahui tingkat kemampuan siswa pada awal program pengajaran.

Tingkat kemampuan awal ini penting untuk menentukan sejauh mana kemajuan seorang siswa. Kemajuan yang dicapai bisa dilihat dari

perbandingan hasil par-tes dengan hasil tes yang diselenggarakan di akhir program pengajaran "*post test*". Jenis tes dilihat secara umumnya, dapat dilakukan dengan cara tertulis ataupun lisan. Kemudian dapat dicermati serta dipahami pada tes tertulis evaluasinya dilakukan secara tertulis baik dalam hal soal maupun jawabannya, namun tes yang disampaikan secara lisan dan dikerjakan secara tertulis masih digolongkan ke dalam jenis tes tertulis. Sebaliknya, tes yang soalnya diberikan dalam bentuk tulisan sedangkan jawabannya berbentuk lisan tidak dapat dikategorikan ke dalam bentuk tes tertulis. 4-10 Unit 4.

Tes lisan yang terdapat pada pertanyaan dan jawaban "*response*" semuanya dalam bentuk lisan. Dikarenakan pada tes lisan lebih kepada tidak memiliki rambu-rambu penyelenggaraan tes yang baku, karena itu, hasil dari tes lisan biasanya tidak menjadi informasi pokok tetapi pelengkap dari instrumen asesmen yang lain. Tes unjuk kerja pada tes ini peserta didik diminta untuk melakukan sesuatu sebagai indikator pencapaian kompetensi yang berupa kemampuan psikomotor. Jenis tes berdasarkan cara penyusunan berdasarkan kriteria ini, tes dapat dibedakan menjadi dua, yaitu tes buatan guru terstandar, tes buatan guru "*teacher made test*".

Untuk melakukan tugas evaluasi itu, seorang guru harus mengembangkan alat ukur, salah satunya tes. Tes yang dikembangkan sendiri oleh guru disebut tes buatan guru "*teacher made test*". Jadi tes buatan guru adalah tes yang dirancang dan dipersiapkan oleh guru, tetap dengan mengacu pada karakteristik tes yang baik dan dilakukan secara cermat, untuk tetap menjamin validitas maupun reliabilitasnya. Tes terstandar "*standardized test*" dapat dikembangkan dengan mengikuti prosedur serta prinsip pengembangan tes secara ketat. Semua prosedur pengembangan tes diikuti sehingga ciri-ciri tes sebagai alat ukur yang baik senantiasa dapat dipenuhi. Dengan demikian, tingkat validitas, reliabilitas, kepraktisan, maupun daya beda sudah bukan menjadi masalah lagi.

Bagaimana cara mengembangkan tes sebagai alat ukur yang baik? Jenis tes berdasarkan bentuk jawaban, jika kita melihat bentuk jawaban yang diberikan oleh peserta tes, kita dapat membedakan tiga jenis tes, yaitu tes *essay*, tes jawaban pendek, dan tes objektif.

Untuk lebih jelasnya, dapat dipahami pada penjelasan di bawah ini:

- a) Tes *essay* atau tes bentuk uraian merupakan sebuah tuntutan kepada peserta didik dalam mengorganisasikan gagasan-gagasan tentang apa yang telah dipelajarinya dengan cara mengemukakannya dalam bentuk tulisan. Keunggulan tes uraian, guru dapat mengukur kemampuan siswa dalam hal mengorganisasikan pikirannya, mengemukakan pendapatnya, dan mengekspresikan gagasan dengan menggunakan kata-kata atau kalimat sendiri. Sedang keterbatasannya adalah cakupan materi pelajaran yang terbatas, waktu pemeriksaan jawaban yang lama, penskorannya cenderung subjektif dan umumnya kurang andal dalam pengukuran.
- b) Tes jawaban pendek, tes dapat digolongkan menjadi tes jawaban pendek jika peserta tes diminta menuangkan jawabannya bukan dalam bentuk esai, tetapi memberikan jawaban-jawaban pendek, dalam bentuk rangkaian kata-kata pendek, kata-kata lepas, maupun angka-angka. Termasuk ke dalam tes jenis ini adalah tes yang mewajibkan siswa untuk mengisi bagian yang kosong dari sebuah kalimat atau teks. Sehingga diharapkan dapat memberikan jawabannya sesingkat mungkin.
- c) Tes objektif adalah tes yang keseluruhan informasi yang diperlukan untuk menjawab tes telah tersedia. Oleh karenanya sering pula disebut dengan istilah tes pilihan jawaban (*selected response test*). Butir soal telah mengandung kemungkinan 4-12 Unit 4 jawaban yang harus dipilih atau dikerjakan oleh peserta tes. Menurut Subino (1987) perbedaan yang khas bentuk soal objektif dibanding dengan soal esai adalah tugas peserta tes (*testee*) dalam merespons tes. Pada tes objektif, tugas *testee* adalah memanipulasikan data yang telah ada dalam butir soal.

Oleh karenanya, tes objektif adalah tes yang dalam pemeriksaannya dapat dilakukan secara objektif. Karena sifatnya yang objektif maka penskorannya dapat dilakukan dengan bantuan mesin. Soal ini tidak memberi peluang untuk memberikan penilaian yang bergradasi karena dia hanya mengenal benar dan salah. Soal tes objektif sangat bermanfaat untuk mengukur hasil belajar kognitif tingkat rendah. Hasil-hasil belajar kompleks seperti menciptakan dan mengorganisasikan gagasan kurang

cocok diukur menggunakan soal bentuk ini. Soal objektif sangat bervariasi bentuknya. Variasi yang bisa dibuat dari soal objektif adalah benar-salah, pilihan ganda, menjodohkan, melengkapi dan jawaban singkat. Rangkuman tes secara sederhana dapat diartikan sebagai himpunan pertanyaan yang harus dijawab, pernyataan-pernyataan yang harus dipilih/ditanggapi, atau tugas-tugas yang harus dilakukan oleh peserta tes dengan tujuan untuk mengukur suatu aspek tertentu dari peserta tes.

Jenis-jenis tes dapat dikelompokkan menjadi beberapa model klasifikasi yaitu: pembagian jenis tes berdasarkan tujuan penyelenggaraan, jenis tes berdasarkan waktu penyelenggaraan, pembagian jenis tes berdasarkan cara mengerjakan, pembagian jenis tes berdasarkan cara penyusunan, dan pembagian jenis tes berdasarkan bentuk jawaban. Dicantumkan tes formatif yang bertujuan untuk mengukur pemahaman Anda mengenai uraian, contoh, dan rangkuman yang tercantum dalam subunit soal-soal:

- a) Jawablah pertanyaan berikut sesuai dengan permintaan!
- b) Jelaskan pembagian jenis tes dilihat dari tujuan penyelenggaraannya!
- c) Jelaskan pembagian jenis tes berdasarkan bentuk jawabannya!
- d) Jelaskan dan berikan contoh tes yang melandasi pengambilan keputusan guru di awal proses pembelajaran!
- e) Jelaskan dan berikan contoh jenis tes yang melandasi pengambilan keputusan guru selama proses pembelajaran berlangsung!
- f) Jelaskan dengan contoh perbedaan antara tes objektif dengan tes *essay* serta umpan balik dan tindak lanjut!

Jawablah pertanyaan tes formatif di atas, setelah selesai baru cocokkan dengan kunci jawabannya. Diskusikan dengan teman bila jawaban belum sesuai atau Anda belum merasa masih ada hal-hal yang meragukan. Hal ini sangat diperlukan karena kesepahaman tentang pengertian ini akan mendasari dan memengaruhi langkah dan kegiatan dalam menyelesaikan mata kuliah ini.

3. Langkah-Langkah Menyusun Tes

Pokok mengembangkan tes Anda sebagai pendidik seringkali kurang menyadari bahwa, mengembangkan tes sebagai instrumen asesmen

proses dan hasil belajar adalah menyusun alat ukur suatu gejala yang bersifat abstrak yaitu pemahaman dan penguasaan anak terhadap materi yang berupa seperangkat kompetensi dipersyaratkan, dan dalam kenyataan di lapangan sebagian besar tenaga pengajar memang menggunakan teknik tes sebagai upaya untuk mengukur hasil belajar tersebut. Karena demikian seringnya pengajar menyusun tes hasil belajar, justru sering menimbulkan kecerobohan, karena menganggap hal ini sebagai hal yang sudah biasa atau umum dilakukan, dan kurang perlu 4-16 Unit 4 mempersiapkannya secara cermat.

Padahal penyusunan tes sangat besar pengaruhnya terhadap siswa yang akan mengikuti tes, untuk mengurangi kesalahan dalam pengukuran maka tes harus direncanakan secara cermat. Secara umum ada lima langkah pokok yang harus dilewati yaitu perencanaan tes dalam langkah perencanaan tes ada beberapa kegiatan yang harus dilakukan guru sebagai pendidik dalam menentukan cakupan materi yang akan diukur yang menyangkut penetapan cakupan materi dan aspek *ranah* kemampuan yang akan diukur.

Penetapan ini penting mengingat bahwa kemampuan belajar merupakan proses yang kompleks dan menyangkut pemahaman yang bersifat abstrak, sehingga harus jelas pada bagian mana cakupan materi yang akan diukur dan dikembangkan dalam soal tes, langkah ini biasanya dilakukan dengan menyusun kisi-kisi soal yaitu daftar spesifikasi. Ada tiga langkah dalam mengembangkan kisi-kisi tes dalam sistem penilaian berbasis kompetensi dasar, yaitu menulis kompetensi dasar, menulis materi pokok, menentukan indikator, dan menentukan jumlah soal. Bentuk tes berupa pemilihan bentuk tes akan dapat dilakukan dengan tepat bila didasarkan pada tujuan tes, jumlah peserta tes, waktu yang tersedia untuk memeriksa lembar jawaban tes, cakupan materi tes, dan karakteristik mata pelajaran yang diujikan. Misalnya, bentuk tes objektif pilihan ganda dan bentuk tes benar salah cocok digunakan bila jumlah peserta tes banyak, waktu koreksi singkat, dan cakupan materi yang diujikan banyak. Bentuk tes objektif lebih cocok digunakan pada mata pelajaran yang batasnya jelas, misalnya mata pelajaran Matematika, Biologi, dan sebagainya.

Ketika dalam memilih cara tes mana yang akan digunakan pendidik yang mesti diperhatikan ciri indikatornya, contohnya apabila tuntutan

indikator melakukan sesuatu, maka cara di dalam penilaiannya adalah tes unjuk kerja “*performance*”, sedangkan jika tuntutan indikator tersebut berkaitan dengan pemahaman konsep, maka teknik penilaiannya adalah tes tertulis. Tingkat berpikir yang digunakan dalam mengerjakan tes harus mencakup mulai yang rendah sampai yang tinggi, dengan proporsi yang sebanding sesuai dengan jenjang lembaga pendidikan, menetapkan panjang tes dan langkah menetapkan panjang tes, meliputi berapa waktu disediakan untuk melakukan tes, hal ini berhubungan erat dengan penetapan jumlah *item-item* tes yang akan dikembangkan dan digunakan.¹¹⁵

Apabila pendidik telah membuat materi yang akan diajarkan kemudian dinilai yang soal-soal dibuat tersebut dilihat dari tingkat kesulitan ataupun tingkat kemudahan, pendidik dapat memberikan pembobotan yang berbeda dari setiap soal yang disusun. Ada tiga hal yang mesti diperhatikan menentukan jumlah pembuatan soal, yaitu bobot masing-masing bagian yang telah ditentukan dalam kisi-kisi, keandalan yang diinginkan, dan waktu yang tersedia. Menulis butir pertanyaan setelah selesai mencermati dan menjabarkan setiap indikator menjadi deskriptor-deskriptor, dan telah ditetapkan ukurannya, maka pendidik mulai dapat mengembangkan atau menulis butir pertanyaan sesuai dengan kisi-kisi yang telah ditetapkan. Ada tiga kegiatan pokok dalam menulis butir soal yaitu, menulis *draft* soal menulis soal bagi Anda pasti sudah menjadi pekerjaan rutin sebelum ulangan, tetapi seharusnya Anda perlu mencermatinya karena langkah ini juga memerlukan kecermatan dalam memilih kalimat-kalimat yang mudah dimengerti dan tidak menimbulkan interpretasi ganda. Ada dua hal yang perlu mendapat perhatian dalam penulisan butir pertanyaan yaitu format pertanyaan dan alternatif jawaban.

Dalam hal ini perlu diperhatikan beberapa hal yaitu, *apakah pertanyaan mudah dimengerti?*, *apakah sudah sesuai dengan indikator?*, *apakah tata letak keseluruhan baik?*, *apakah perlu pembobotan?*, *apakah kunci jawaban sudah benar?* Memantapkan validitas isi “*content validity*” pada dasarnya merupakan koefisien yang menunjukkan kesesuaian antara

¹¹⁵ Bisri dan M. Ichsan. *Penilaian Otentik dengan Teknik Nontes di Sekolah Dasar Authentic Assessment with Nontest Technique in Primary School*. Jurnal Sosial Humaniora ISSN 2087-4928 Volume 6 Nomor 2, Oktober 2015. hlm. 34-35.

draft tes yang telah disusun dengan isi dari konsep dan kisi-kisi yang telah disusun, apakah semua materi telah terjabar dalam *item*, dan apakah soal yang disusun telah pula sesuai ranah atau kawasan yang akan diukur.

Langkah ini dapat dilakukan dengan beberapa cara misalnya diskusi dengan sesama pendidik ataupun dengan cara mencermati kembali substansi dari konsep yang akan diukur, melakukan uji coba “*try out*” Mungkin Anda mengira bahwa *try out* hanya digunakan untuk tes standar dan tidak perlu dilakukan untuk tes buatan guru. Anggapan itu kurang benar karena uji coba tetap diperlukan dalam penyusunan tes buatan guru. *Try out* tidak harus dilakukan secara formal dan dalam skala besar, yang perlu Anda perhatikan adalah bahwa *try out* dapat dilakukan untuk berbagai kepentingan di antaranya adalah untuk, analisis *item*, bagaimana 4-18 Unit 4 pada perencanaan dan pelaksanaan, memperkirakan penggunaan waktu pengerjaan, kejelasan format tes, kejelasan petunjuk pengisian, dan pemahaman bahasa yang digunakan.

Revisi soal hasil dari uji coba kemudian dilakukan analisis untuk mencari tingkat kesulitan soal dan penggunaan bahasa yang kurang komunikatif, untuk kemudian dilakukan revisi sesuai dengan kebutuhan. Misalnya revisi dilakukan untuk eliminasi butir-butir yang jelek, menambah butir-butir baru, memperjelas petunjuk, dan memodifikasi format dan urutan, dsb. Melakukan pengukuran dengan tes, sekarang tes sebagai instrumen sudah selesai disusun, apa saja yang harus Anda lakukan selanjutnya? Ada beberapa langkah yang harus diperhatikan pada saat menyelenggarakan tes untuk siswa yaitu, menjaga objektivitas pelaksanaan tes dengan kegiatan pengukuran yang berupa penyelenggaraan tes juga sudah menjadi kegiatan Anda sehari-hari. Meskipun demikian, pendidik tetap harus menjaga objektivitas dalam pengawasan, menjaga kerahasiaan soal, dan berbagai kode etik penyelenggaraan tes yang lain.

Setelah ujian dilaksanakan maka langkah berikut adalah koreksi dan interpretasi dari hasil ujian tersebut. Kemudian berdasarkan data hasil analisis tersebut akan diambil keputusan dalam berbagai kepentingan. Memberikan skor pada hasil tes yaitu memeriksa hasil jawaban dari para siswa, untuk memberikan skor atau angka sebagai penghargaan terhadap setiap poin soal yang dapat dikerjakan, hasilnya berupa angka yang disebut

skor mentah, angka yang menunjukkan berapa soal yang bisa dijawab benar oleh siswa. Penentuan jumlah soal yang bisa dijawab benar ini tidak menjadi masalah untuk tes objektif. Namun untuk bentuk soal tes uraian masalah ini akan menjadi persoalan, karena setiap siswa akan mengemukakan argumentasi yang berbeda-beda untuk menjawab soal dan permasalahan tes. Sehingga dalam melakukan langkah ini harus pula dijaga objektivitas dengan selalu menggunakan kunci jawaban dan indikator keberhasilan.

4. Fungsi Tes

Fungsi tes secara umum, ada dua yaitu, pertama, sebagai alat pengukur terhadap peserta didik. Dalam hubungan ini tes berfungsi mengukur tingkat perkembangan atau kemajuan yang telah dicapai oleh peserta didik setelah mereka menempuh proses belajar mengajar dalam jangka waktu tertentu. Kedua, sebagai alat pengukur keberhasilan program pengajaran, sebab melalui tes tersebut akan dapat diketahui sudah seberapa jauh program pengajaran yang telah ditentukan, telah dapat dicapai.

5. Contoh dalam Pengolahan Tes

Tes tertulis bagian dari sebuah aktivitas yang berlandaskan kepada ide, gagasan, pikiran serta perasaan pada simbol sebuah kebahasaan. Yang terdiri dari penempatan pada penggunaan tanda baca serta ejaannya pada sebuah kosakata, struktur kalimat, dan pengembangan pada paragraf, pengolahan gagasan serta pengembangan model karangan. Pada pendeskripsian tes tertulis berkaitan kepada temuan hasil penelitian berkaitan secara langsung dengan ide-ide untuk dipraktikkan.

Selain itu, aktivitas tes tertulis merupakan bentuk adanya kemampuan dalam berbahasa yang mesti dikuasai peserta didik, karena untuk menghasilkan bahasa maka tahap yang dilalui berdasarkan pendengaran, pembicaraan, serta bahan bacaan. Jika dihubungkan dengan 3 kemampuan keterampilan berbahasa dari segi menulis lebih sulit dikuasai meskipun yang bersangkutan penutur asli dari bahasa tersebut karena dipengaruhi oleh kemampuan tes tertulis yang menekankan pada pemahaman di berbagai unsur kebahasaan dari dalam ataupun unsur dari luar kebahasaan itu sendiri. Menghasilkan isi karangan atau tulisan, baik

unsur bahasa maupun unsur isi, sehingga terjalin hubungan sedemikian rupa atas hasil karangan yang tertata rapi.

Beberapa para ahli menjelaskan Stolz dan Bruck, di dalam pembagian tes bahasa yang terdiri atas dua macam, baik secara langsung ataupun tidak langsung, di antaranya adalah:

a) Tes Tulisan Langsung

Untuk menghasilkan penglihatan mengenai apakah tes tulisan langsung tersebut valid ketika menggunakannya, maka dapat dilihat secara langsung dari sampel tes yang diambil dari evaluasi sikap atau perilaku. Tes dianggap tes langsung serta bisa digunakan jika tes itu menginginkan peserta didik melakukan tes tersebut dengan tepat sesuai hasil pengukurannya. Tes tersebut diartikan sebagai tes langsung karena berkaitan dengan isi yang butuh akan pemisahan dari sisi bahasanya, situasinya, kebudayaannya, serta faktor kendala afektif yang menimbulkan proses interaksi pada proses komunikasi.

Dari penjelasan pemaparan di atas, bertujuan sebagai panduan untuk memahami tes tulisan yang dilakukan langsung oleh peserta didik yang tertuju kepada pekerjaan menulis tugas yang tidak hanya melibatkan komponen bahasa seperti *grammar* saja, akan tetapi juga memperhatikan ejaan dan tanda baca.

b) Tulisan Tidak Langsung

Hughes memberikan penjelasan tentang tes tidak langsung yang lebih luas yang berbeda dari tugas mempergunakan bahasa yang biasa atau umum. Yang tertuju kepada pemberian tes tulisan tidak langsung dengan tes poin yang berbeda berhubungan dengan bentuk seperti *grammar*, ejaan, kosakata, ejaan serta tanda baca tersebut dapat secara terpisah dengan menggunakan tes objektif. Dapat disimpulkan bahwa tes *cloze* prosedur pada pilihan ganda yang dikelompokkan sebagai penilaian tes tidak langsung pada tulisan. Penggunaan pendekatan tes tulisan masing-masing mempunyai kelebihan dan kelemahan. Tes tulisan tidak langsung dapat dinilai secara objektif dikarenakan hasil yang didapatkan melalui tes tulisan tersebut dipercaya dan merupakan tes yang lebih praktis, karena dapat diberi skor dengan cepat dan lebih mudah. Tes tulisan tidak langsung mempunyai kelemahan pada penentuan validitasnya,

karena dalam penyusunan tes menghadapi masalah untuk memilih misalnya, kosakata yang hendak diuji. Oleh karena itu, tidak dapat dipastikan apakah jumlah unsur-unsur yang memisahkan tes itu sebenarnya, menunjukkan kemampuan berpikir peserta didik dalam berbicara. Siswa yang pandai misalnya, dalam hal kosakata dan tata bahasa mungkin tidak pandai menulis. Sebaliknya, tes langsung lebih disukai karena menjamin validitasnya dan tes langsung kurang praktis karena biaya dan pelaksanaannya besar untuk dilakukan oleh sebagian besar siswa.

C. Nontes

1. Pengertian Nontes

Nontes adalah cara penilaian hasil belajar peserta didik yang dilakukan tanpa menguji peserta didik tetapi dengan melakukan pengamatan secara sistematis. Teknik evaluasi nontes berarti melaksanakan penilaian dengan tidak menggunakan tes. Teknik penilaian ini umumnya untuk menilai kepribadian anak secara menyeluruh meliputi sikap, tingkah laku, sifat, sikap sosial dan lain-lain. Yang berhubungan dengan kegiatan belajar dalam pendidikan, baik secara individu maupun secara kelompok. Dengan teknik nontes maka penilaian atau evaluasi hasil belajar peserta didik dapat dilakukan dengan pengamatan secara sistematis “*observasi*”, melakukan wawancara “*interview*” dan menyebar angket “*questionnaire*”.¹¹⁶

Bentuk-bentuk teknik nontes berupa observasi “pengamatan” merupakan salah satu bentuk teknik nontes yang biasa dipergunakan untuk menilai sesuatu melalui pengamatan terhadap objeknya secara langsung, saksama, dan sistematis. Pengamatan memungkinkan untuk melihat dan mengamati sendiri kemudian mencatat perilaku dan kejadian yang terjadi pada keadaan sebenarnya. *Interview* “wawancara” adalah cara menghimpun bahan-bahan keterangan yang dilaksanakan dengan cara melakukan tanya jawab lisan secara sepihak, berhadapan muka, dengan arah serta tujuan yang telah ditentukan. Angket “*questionnaire*” dapat digunakan sebagai alat bantu dalam rangka penilaian hasil belajar. Angket

¹¹⁶ Nana Sudjana. *Penilaian Hasil Proses Belajar Mengajar*. (Bandung: PT. Remaja Rosdakarya.1995). hlm. 60.

adalah teknik pengumpulan data yang dilakukan dengan cara memberi seperangkat pertanyaan tertulis kepada responden untuk dijawabnya, sehingga angket berbeda dengan wawancara. Prinsip penulisan angket isi dan tujuan pertanyaan jelas, bahasa yang digunakan mudah dipahami, tipe, dan bentuk pertanyaan “*terbuka atau tertutup*”, pertanyaan tidak mendua, tidak menanyakan yang sudah lupa, panjang pertanyaan maksimal 30 pertanyaan, urutan pertanyaan “*dari mudah ke sulit*”, prinsip pengukuran, dan penampilan fisik angket.

Pengukuran penilaian hasil belajar menggunakan instrumen nontes untuk mengevaluasi hasil belajar aspek afektif dan keterampilan motorik. Bentuk penilaian yang menggunakan alat ukur atau instrumen nontes yaitu penilaian unjuk kerja *performance*, penilaian proyek atau produk, penilaian portofolio, dan penilaian sikap. Alat penilaian yang tergolong teknik nontes di antaranya kuesioner atau angket, wawancara, daftar cocok “*checklist*”, pengamatan atau observasi, penugasan, portofolio, jurnal, inventori, penilaian diri “*self assessment*”, penilaian oleh teman sejawat “*peer assessment*”, adapun penjelasannya akan dipaparkan di bawah ini:

a) Penilaian unjuk kerja “*performance assessment*”

Pada tes bentuk perbuatan “*unjuk kerja*”, umumnya dilakukan dengan cara menyuruh peserta tes untuk melakukan sesuatu pekerjaan yang bersifat fisik “*praktik*”. Tes bentuk perbuatan ini sangat cocok untuk melakukan penilaian dalam pelajaran praktik/keterampilan atau praktikum di laboratorium. Alat yang digunakan untuk melakukan penilaian pada umumnya berupa lembar pengamatan “*lembar observasi*”. Tes bentuk perbuatan ini pada umumnya dapat digunakan untuk menilai proses maupun hasil “*produk*” dari suatu kegiatan praktik. Mengukur dimaksudkan memberi bentuk kuantitatif dari suatu kegiatan atau kemampuan yang dimiliki, yaitu dalam bentuk angka.¹¹⁷

Pada pengukuran unjuk kerja yang digunakan adalah lembar pengamatan. Pengukuran unjuk kerja dipergunakan untuk mencocokkan kesesuaian antara pengetahuan mengenai teori dan keterampilan di dalam praktik sehingga hasil evaluasinya menjadi

¹¹⁷ Nasution, *Metode Research*, (Jakarta: Bumi Aksara, 2000). hlm. 45.

lebih jelas. Penilaian penguasaan kompetensi aspek keterampilan atau psikomotor yang dimiliki oleh seseorang atau peserta didik, hanya ada satu bentuk tes yang tepat yaitu tes perbuatan “*performance assessment*”. Artinya orang yang akan dinilai kemampuan *skill*-nya harus menampilkan atau melakukan *skill* yang dimilikinya di bawah persyaratan-persyaratan kerja yang berlaku.

Hal yang penting dalam pembelajaran keterampilan adalah diperolehnya penguasaan keterampilan praktis, serta pengetahuan dan perilaku yang berhubungan langsung dengan keterampilan tersebut. Sehubungan dengan itu, maka para ahli telah mengembangkan berbagai metode pembelajaran keterampilan yang berbeda-beda, tergantung pada sasaran atau maksud yang hendak dicapai di dalam pembelajaran tersebut. Model yang sederhana untuk pembelajaran keterampilan kerja adalah metode empat tahap “*training within industry*”. Tahap-tahap tersebut meliputi: *pertama*, persiapan dalam hal ini, pendidik atau instruktur mengutarakan sasaran-sasaran latihan kerja, menjelaskan arti pentingnya latihan, membangkitkan minat para peserta pelatihan untuk menerapkan pengetahuan yang sudah dimilikinya dalam situasi yang riil.¹¹⁸ *Kedua*, peragaan pada tahap ini, instruktur memperagakan keterampilan yang dipelajari oleh peserta didik, menjelaskan cara kerja dan proses kerja yang benar. Dalam hal ini, instruktur harus mengambil posisi sedemikian rupa sehingga para peserta pelatihan akan dapat mengikuti demonstrasi mengenai proses kerja dengan baik. *Ketiga*, peniruan pada tahap ini, peserta pelatihan menirukan aktivitas kerja yang telah diperagakan oleh instruktur. Dalam hal ini, instruktur mengamati peniruan yang dilakukan oleh peserta pelatihan, menyuruh melakukannya secara berulang-ulang dan membantu serta mendorong para peserta pelatihan agar dapat melakukan pekerjaannya dengan benar, keempat praktik.

Untuk mengamati setiap sikap atau perilaku yang dikerjakan peserta didik, maka dapat digunakan alat instrumen seperti, daftar

¹¹⁸ Sitti Mania. *Teknik Non Tes: Telaah atas Fungsi Wawancara dan Kuesioner dalam Evaluasi Pendidikan*, Lentera Pendidikan Vol. 11 No. 1 JUNI 2008. hlm. 45-54 46.

cek evaluasi unjuk kerja yang dapat dikerjakan dengan menggunakan daftar cek peserta didik mendapat nilai apabila kriteria penguasaan kemampuan tertentu dapat diamati oleh penilai. Jika tidak dapat diamati, peserta didik tidak memperoleh nilai. Kelemahan cara ini hanya dari segi benar atau salah, dapat diamati atau tidak dapat diamati, baik atau tidak baik. Maka dengan hal yang demikian secara langsung tidak terdapat nilai tengah. Tapi daftar cek yang dilakukan lebih mudah apabila dipakai dengan melihat subjek ataupun objeknya dalam jumlah skala yang lebih besar.¹¹⁹ Seperti contoh:

1. Lembar Evaluasi yang Terdapat pada Pembelajaran PAI

No	Aspek yang dinilai	Nilai				Keterangan
		1	2	3	4	
A	Perencanaan					
1	Pelaksanaan					
2	Evaluasi					
Jumlah 10%						
B	Proses					
1	Disiplin					
2	Pemanfaatan waktu					
3	Pemberian materi ajar					
Jumlah 30%						
C	Hasil					
1	Kesesuaian dalam perencanaan					
2	Kesesuaian dalam pelaksanaan					
3	Kesesuaian dalam pengevaluasian					
	Jumlah	60%				
Jumlah 100%						

¹¹⁹ Sukirman. *Tes Kemampuan Keterampilan Menulis dalam Pembelajaran Bahasa Indonesia di Sekolah*. (Institut Agama Islam Negeri Palopo: Jurnal Konsepsi, Vol. 9, No. 2, Agustus 2020). hlm. 56.

2. Contoh Lembar Observasi untuk Mengamati Kemampuan Peserta Didik dalam Memahami Materi Ajar yang Diberikan

No	Aspek yang dinilai	Hasil Observasi	
		Benar	Salah
1	Kedisiplinan dalam kelas		
2	Keaktifan dalam kelas		
3	Kebersihan dalam kelas		

3. Lembar Telaah Butir untuk Soal Psikomotor

No	Kriteria Telaah
Materi	Pada pembuatan soal mesti harus sesuai dengan kemampuan peserta didik, serta tepat dan jelas.
Konstruksi	Petunjuk cara dalam mengerjakan soal-soal harus jelas dan tepat serta mudah dipahami.
Bahasa	Reduksi bahasa mesti mudah dipahami oleh peserta didik dari bahasa yang sederhana ke bahasa yang sulit, pada pembuatan butir-butir soal mesti menggunakan bahasa Indonesia yang baik dan benar.

4. Contoh-Contoh Halaman Sampul pada Buku Catatan Harian

<p>BUKU CATATAN HARIAN TENTANG PESERTA DIDIK (Nama Sekolah)</p>	
Mata Pelajaran	:.....
Kelas	:.....
Program	:.....
Tahun Pelajaran	:.....
Nama Guru	:.....
Padang, 2021	

5. Contoh Isi Buku Catatan Harian

No	Hari/Tanggal	Peserta didik	Kejadian Positif/Negatif

6. Contoh Format Penilaian Sikap dalam Pembelajaran PAI

No	Peserta didik	Perilaku				Nilai	Keterangan
		Disiplin	Rajin	Sopan	aktif		
1.	Tina						
2.	Zakiah						
3.	Hasanah						

Catatan:

- a. Pada kolom penilaian sikap diisi dengan angka sesuai dengan kriteria berikut:
 - a) Sangat kurang
 - b) Kurang
 - c) Sedang
 - d) Baik
 - e) Amat Baik
- b. Jumlah nilai diambil dari skor tiap-tiap indikator
- c. Keterangan diisi sesuai kriteria yang ditetapkan
 - Nilai 18-20 berarti amat baik
 - Nilai 14-17 berarti baik
 - Nilai 10-13 berarti sedang
 - Nilai 6-9 berarti kurang
 - Nilai 0-5 berarti sangat kurang

7. Contoh Lembar Penilaian untuk Menilai Minat Peserta Didik pada Mata Pelajaran PAI

No	Pertanyaan	Jawaban				
		1	2	3	4	5
1	Kerajinan peserta didik dalam mengikuti proses pembelajaran					
2	Kedisiplinan peserta didik ketika mengikuti proses pembelajaran					
3	Keaktifan peserta didik dalam mengikuti proses pembelajaran					

8. Contoh Lembar Penilaian Portofolio

Kompetensi Dasar	Nama : Tanggal :				
Kelancaran Pelafalan Makhorijul huruf					
Indikator	Penilaian				
	Jelek sekali	Jelek	Sedang	Baik	Baik sekali
Merumuskan indikator berdasarkan SK dan KD					
Memilih materi yang sesuai dengan SK dan KD					
Dicapai melalui	Komentar guru:				
1. Bimbingan guru					
2. Semua kelas					
3. Per kelompok					
4. Per individu					
Komentar orang tua					

2. Langkah Pengembangan Instrumen Nontes

Seperti halnya pengembangan instrumen tes, pengembangan instrumen nontes juga memiliki langkah-langkah yang harus diikuti, menentukan spesifikasi instrumen, menulis instrumen, menentukan skala instrumen, menentukan sistem penskoran, menelaah instrumen, merakit instrumen, melakukan uji coba, menganalisis hasil uji coba, memperbaiki instrumen, melaksanakan pengukuran, dan menafsirkan hasil pengukuran.

a. Spesifikasi Instrumen terdiri atas tujuan dan kisi-kisi instrumen.

Tujuan pengembangan instrumen nontes sangat tergantung pada data yang akan dihimpun. Instrumen nontes mencakup afektif dan psikomotorik. Ditinjau dari tujuannya, instrumen ranah afektif dibedakan menjadi lima, yaitu instrumen sikap, minat, konsep diri, nilai, dan moral. Ada empat hal yang perlu diperhatikan ketika menyusun spesifikasi instrumen, tujuan pengukuran, kisi-kisi instrumen, bentuk dan format instrumen, dan panjang instrumen. Instrumen minat bertujuan untuk memperoleh informasi tentang minat siswa terhadap mata pelajaran. Selanjutnya hasil pengukuran terhadap minat digunakan untuk meningkatkan minat siswa terhadap mata pelajaran. Instrumen sikap bertujuan untuk mengetahui sikap siswa terhadap suatu objek. Misalnya, sikap siswa terhadap kegiatan sekolah, guru, dll. Sikap terhadap mata pelajaran bisa positif bisa negatif. Hasil pengukuran sikap berguna untuk menentukan strategi pembelajaran yang tepat bagi siswa. Instrumen konsep diri bertujuan untuk mengetahui kekuatan dan kelemahan diri sendiri.

Siswa melakukan evaluasi secara objektif terhadap potensi yang ada dalam dirinya. Karakteristik potensi siswa sangat penting untuk menentukan jenjang kariernya. Informasi kekuatan dan kelemahan siswa digunakan untuk menentukan program yang sebaiknya ditempuh oleh siswa. Instrumen nilai bertujuan untuk mengungkap nilai dan keyakinan individu. Informasi yang diperoleh bisa positif bisa negatif. Hal-hal yang positif diperkuat, sedangkan yang negatif diperlemah dan akhirnya dihilangkan. Instrumen moral bertujuan untuk mengungkap moral. Informasi moral seseorang diperoleh melalui pengamatan atas perbuatan yang ditampilkan dan laporan diri, yaitu dengan mengisi kuesioner. Informasi hasil pengamatan

bersamaan dengan hasil kuesioner menjadi informasi penting tentang moral seseorang. Setelah tujuan pengukuran afektif ditetapkan, kegiatan berikutnya adalah menyusun kisi-kisi instrumen.

Kisi-kisi juga disebut *blue print*. Kisi-kisi pada tabel matriks yang berisi spesifikasi instrumen yang akan ditulis. *Langkah pertama* menentukan kisi-kisi adalah menentukan definisi konseptual yang berasal dari teori-teori yang diambil dari referensi. *Langkah kedua*, mengembangkan definisi operasional berdasarkan definisi konseptual. *Langkah ketiga* ditentukan aspek dan dimensi yang mengonstruksikan instrumen yang dikembangkan. Aspek atau dimensi ini kemudian dijabarkan menjadi sejumlah indikator, yang digunakan sebagai pedoman dalam menulis instrumen. Tiap indikator dapat terdiri atas dua atau lebih butir instrumen. Menulis instrumen disusun berdasarkan kisi-kisi yang telah dibuat. Instrumen dapat berbentuk pernyataan atau pertanyaan. Hal-hal yang penting dalam pelaksanaan ketika menulis butir-butir soal, adalah sebagai berikut:¹²⁰

- a. Hindari kalimat yang mengandung banyak interpretasi
- b. Rumusan pernyataan/pertanyaan singkat
- c. Satu pernyataan hanya mengandung satu pikiran yang lengkap
- d. Pernyataan dirumuskan dengan kalimat sederhana
- e. Hindari penggunaan kata-kata selalu, semua, tidak pernah, dan sejenisnya.
- f. Hindari pernyataan tentang fakta, atau yang dapat diinterpretasikan sebagai fakta, menentukan skala instrumen. Ada beberapa skala yang biasa digunakan dalam mengukur ranah afektif, di antaranya adalah skala Likert, Thrustone, dan Beda Semantik. Langkah-langkah pengembangan skala dalam menentukan objek sikap atau tingkah laku yang akan dikembangkan skalanya, menyusun kisi-kisi instrumen (*skala sikap*), menulis butir pernyataan serta melengkapi butir

¹²⁰ Djaali & Mulyono. *Pengukuran Dalam Bidang Pendidikan*. Grasindo, Jakarta, 2008). hlm. 86.

pernyataan dengan skala sikap *bisa genap, empat, enam bahkan bisa menggunakan angka ganjil lima, tujuh dan sembilan.*¹²¹

3. Sistem Penskoran

Sistem penskoran yang digunakan tergantung pada skala yang digunakan. Misalnya, apabila digunakan skala Thrustone, maka skor tertinggi tiap butir adalah angka tujuh dan terendah angka satu. Selanjutnya dilakukan analisis untuk tingkat siswa dan tingkat kelas, yaitu dengan mencari nilai rata-rata pada simpangan baku skor. Hasil analisis pada data tersebut digunakan untuk menafsirkan ranah afektif dari setiap siswa dan kelas terhadap suatu objek. Hasil tafsiran perlu ditindaklanjuti oleh guru dengan melakukan perbaikan-perbaikan, seperti perbaikan pada media, strategi dan metode pembelajaran, pada penggunaan alat peraga, dll.

4. Telaah Instrumen

Kegiatan pada telaah instrumen adalah meneliti tentang, kesesuaian antara butir pertanyaan ataupun pernyataan dengan indikator, kekomunikatifan bahasa yang digunakan, kebenaran dari tata bahasa yang digunakan, ada tidaknya bias pada pertanyaan/pernyataan, kemenarikan format instrumen, kecukupan butir instrumen, sehingga tidak membosankan. Telaah dilakukan oleh pakar dalam bidang yang diukur dan akan lebih baik bila ada pakar penilaian. Telaah bisa juga dilakukan oleh teman sejawat. Panjang instrumen berhubungan dengan masalah kebosanan. Lama pengisian instrumen sebaiknya tidak lebih dari 30 menit. Pertanyaan/pernyataan yang diajukan jangan sampai bias, yaitu mengarahkan jawaban responden pada arah tertentu, positif atau negatif.

Contoh pernyataan biasa: Sebagian besar responden setuju bahwa masyarakat berhak menerima layanan kesehatan. Apakah Anda setuju bila semua masyarakat menerima layanan kesehatan? Hasil telaah selanjutnya digunakan untuk memperbaiki instrumen. Perbaikan dilakukan terhadap konstruksi instrumen, yaitu kalimat yang digunakan, waktu yang

¹²¹ Ina Maghdalena. Dkk. *Penggunaan Evaluasi Non tes dan hambatannya dalam pembelajaran di SDS Sari Putra Jakarta Barat*, Pensa: Jurnal Pendidikan dan Ilmu sosial, 2021. Hal 3-6.

diperlukan untuk mengisi instrumen, cara pengisian, dll. Merakit instrumen setelah instrumen diperbaiki, selanjutnya dirakit dengan memperhatikan format, tata letak, urutan pernyataan dan pertanyaan. Format harus menarik. Urutan pernyataan sesuai dengan aspek yang akan diukur. Uji coba instrumen setelah dirakit, instrumen diujicobakan. Sampel uji coba dipilih yang karakteristiknya mewakili populasi yang ingin dinilai. Ukuran sampel minimal 30 orang, bisa berasal dari satu sekolah atau lebih. Pada saat uji coba, yang perlu dicatat adalah saran-saran dari responden atas kejelasan pedoman pengisian instrumen, kejelasan kalimat, waktu yang digunakan, dll.¹²²

5. Analisis Hasil Uji Coba

Analisis hasil uji coba ruang lingkupnya mengayomi di setiap macam jawaban pada tiap-tiap butir pertanyaan ataupun pernyataan. Ketika skala instrumen satu sampai lima, maka bila jawaban bervariasi dari satu sampai lima berarti instrumen tersebut baik. Namun apabila jawaban semua responden sama, misalnya tiga semua, maka instrumen tergolong tidak baik. Indikator yang digunakan adalah besarnya daya beda atau korelasi antara skor butir dengan skor totalnya. Jika daya beda butir lebih dari angka (0,3) maka instrumen tersebut tergolong baik. Indikator lain yang diperhatikan adalah indeks keandalan atau reliabilitas. Besarnya indeks reliabilitas sebaiknya minimal angka 0,7.¹²³

6. Perbaikan Instrumen

Pelaksanaan pengukuran pelaksanaan pengukuran sebaiknya dilakukan pada saat responden tidak lelah. Ruang untuk pelaksanaan pengukuran harus representatif, baik kondisi ruang, tempat duduk, ataupun yang lain. Diusahakan responden tidak saling bertanya ketika pengukuran dilaksanakan. Pengisian instrumen dimulai dengan penjelasan tujuan pengisian, manfaat bagi responden, dan pedoman pengisian instrumen. Penafsiran hasil pengukuran berupa skor atau angka. Menafsirkan hasil pengukuran disebut dengan penilaian. Untuk menafsirkan hasil

¹²² Sumarna, *Pedoman Pengembangan Penilaian Portofolio*. (Jakarta: Puspendik Balitbang Depdiknas, 2006). hlm. 37.

¹²³ Popham. W. James, *Classroom Assessment*. (Boston: Allyn & Bacon, 1996). hlm. 67.

pengukuran diperlukan suatu kriteria. Kriteria yang digunakan tergantung pada skala dan jumlah butir yang digunakan.

7. Jenis Instrumen Nontes

Telah disampaikan pada pendahuluan, bahwa teknik penilaian nontes mengacu pada penilaian kelas yang meliputi unjuk kerja, sikap, proyek, produk, portofolio, dan penilaian diri pada unjuk kerja penilaiannya dikerjakan melalui hasil dari pengamatan proses kegiatan belajar peserta didik dalam melakukan sesuatu. Dan penilaian ini cocok dipakai sebagai penilaian hasil pada pencapaian kompetensi yang menuntut peserta didik dalam mengerjakan tugas tertentu seperti praktik di laboratorium, praktik salat, praktik OR, presentasi, diskusi, bermain peran, memainkan alat musik, bernyanyi, membaca puisi (*deklamasi*), dll. Cara penilaian ini bisa dikatakan lebih autentik dari pada tes tertulis karena hasil yang didapat tersebut atas apa yang dilihat berdasarkan kepada *skill* atau kemampuan yang dimiliki peserta didik itu sendiri.

Pengamatan unjuk kerja perlu dilakukan dalam berbagai konteks untuk menetapkan tingkat pencapaian kemampuan tertentu. Untuk menilai kemampuan melakukan kinerja ilmiah peserta didik, dilakukan pengamatan atau observasi yang beragam, seperti mempersiapkan alat, merangkai percobaan, dan mengamati hasil percobaan. Untuk mengamati unjuk kerja peserta didik dapat menggunakan alat atau instrumen berupa daftar cek (*checklist*), skala bertingkat (*rating scale*), catatan pengamatan. Pertama, daftar cek (*checklist*) penilaian unjuk kerja dapat dilakukan dengan menggunakan daftar cek (*baik-tidak baik*). Dengan menggunakan daftar cek, peserta didik mendapat nilai bila kriteria penguasaan kompetensi tertentu dapat diamati oleh penilai. Jika tidak dapat diamati, peserta didik tidak memperoleh nilai. Kelemahan cara ini adalah penilai hanya mempunyai dua pilihan mutlak, misalnya benar-salah, dapat diamati-tidak dapat diamati, baik-tidak baik. Dengan demikian tidak terdapat nilai tengah, namun daftar cek lebih praktis digunakan mengamati subjek dalam jumlah besar. Penilaian unjuk kerja yang menggunakan skala penilaian memungkinkan penilai memberi nilai tengah terhadap penguasaan kompetensi tertentu, karena pemberian nilai secara kontinu di

mana pilihan kategori nilai lebih dari dua. Skala penilaian terentang dari tidak sempurna sampai sangat sempurna.

Untuk memperkecil faktor subjektivitas, perlu dilakukan penilaian oleh lebih dari satu orang, agar hasil penilaian lebih akurat. Agar para penilai memiliki *frame* yang sama maka perlu dibuat rubrik penilaiannya. Rubrik penilaian dapat berupa generik dan spesifik. Skala Penilaian “*rating scale*” merupakan bagian dari pengukuran psikologi. Karena menyangkut sikap manusia, maka hasil pengukuran tidak pernah mencapai hasil yang sempurna. Pengukuran sikap sangat sukar bahkan mungkin tidak pernah dapat dilakukan dengan validitas, reliabilitas, dan objektivitas yang tinggi. Hal ini, antara lain dikarenakan, atribut psikologi bersifat tidak tampak, indikator-indikator perilaku jumlahnya terbatas, karena tanggapan dapat terpengaruh oleh variabel-variabel tidak relevan seperti suasana hati, kondisi dan situasi sekitar, dan banyak sumber kesalahan, baik dari penilai, yang dinilai, alat yang digunakan, cara analisis.

Penilaian proyek merupakan kegiatan penilaian terhadap suatu tugas yang harus diselesaikan dalam periode atau kurun waktu tertentu. Tugas tersebut berupa suatu investigasi sejak dari perencanaan, pengumpulan data, pengorganisasian, pengolahan, dan penyajian data. Penilaian proyek dapat digunakan untuk mengetahui pemahaman, kemampuan mengaplikasikan, kemampuan penyelidikan dan kemampuan menginformasikan peserta didik pada mata pelajaran tertentu secara jelas. Dalam penilaian proyek setidaknya ada tiga hal yang perlu dipertimbangkan yaitu: pertama, kemampuan pengelolaan yaitu kemampuan peserta didik dalam memilih topik, mencari informasi dan mengelola waktu pengumpulan data serta penulisan laporan. Kedua, relevansi kesesuaian dengan mata pelajaran, dengan mempertimbangkan tahap pengetahuan, pemahaman, dan keterampilan dalam pembelajaran. Ketiga, keaslian proyek yang dilakukan peserta didik harus merupakan hasil karyanya, dengan mempertimbangkan kontribusi guru berupa petunjuk dan dukungan terhadap proyek peserta didik.

Penilaian produk dilakukan mulai dari perencanaan, proses pengerjaan, sampai produk tersebut selesai dikerjakan. Maka, guru perlu menetapkan hal-hal atau tahapan yang perlu dinilai, seperti penyusunan desain, pengumpulan data, analisis data, dan penyiapan laporan tertulis.

Laporan tugas atau hasil penelitian juga dapat disajikan dalam bentuk poster. Pelaksanaan penilaian dapat menggunakan alat ataupun instrumen penilaian berupa daftar cek ataupun skala penilaian. Beberapa contoh kegiatan peserta didik dalam penilaian produk hasil penelitian secara sederhana tentang penggunaan daya listrik di rumah.

Penelitian yang telah dibuat secara sederhana tentang perkembangan harga sembako, dll. Maka penilaian evaluasi produk terhadap penilaian proses dari segi pembuatan dan kualitas suatu produk. Karena penilaian produk meliputi penilaian kemampuan peserta didik dalam membuat produk teknologi seni, di antaranya makanan, pakaian, hasil karya seni (*patung, lukisan, gambar*), barang-barang terbuat dari kayu, keramik, plastik, dan logam. Pengembangan produk meliputi tiga tahap dan setiap tahap perlu diadakan penilaian yaitu, tahap persiapan, meliputi penilaian kemampuan peserta didik merencanakan, menggali, dan mengembangkan gagasan, dan mendesain produk. Kedua, tahap pembuatan produk "*proses*", yang meliputi pada penilaian *skill* kemampuan peserta didik dalam pemilihan penggunaan bahan, alat, serta tekniknya. Karena pada tahap penilaian produk (*appraisal*), terdapat sebuah data penilaian produk yang dihasilkan peserta didik sesuai kriteria yang ditetapkan. Penilaian produk yang digunakan secara holistik dan analitik.

- a. Cara analitik, yaitu berdasarkan aspek-aspek produk, biasanya dilakukan terhadap semua kriteria yang terdapat pada semua tahap proses pengembangan (tahap: persiapan, pembuatan produk, penilaian produk).
- b. Cara holistik, yaitu berdasarkan kesan keseluruhan dari produk, biasanya dilakukan hanya pada tahap penilaian produk "*appraisal*". Penilaian portofolio merupakan penilaian berkelanjutan yang berdasarkan pada sekumpulan informasi yang tertuju kepada pelaksanaan yang telah dilakukan berdasarkan pada *skill* atau kemampuan yang ada pada peserta didik dalam satu periode tertentu. Informasi tersebut bersumber pada karya yang telah dibuat peserta didik dari proses pembelajaran yang telah dikatakan baik oleh peserta didik, maka lampiran lembar jawaban tes yang menunjukkan soal yang mampu dan tidak mampu untuk dijawab

“*bukan nilai*” maka informasi yang didapatkan antara lain yang berhubungan dengan kompetensi tertentu pada satu mata pelajaran.

Penilaian portofolio pada dasarnya menilai karya-karya siswa secara individu pada satu periode untuk suatu mata pelajaran. Akhir suatu periode hasil karya tersebut dikumpulkan dan dinilai oleh guru dan peserta didik sendiri. Berdasarkan informasi perkembangan tersebut, guru dan peserta didik sendiri dapat menilai perkembangan kemampuan peserta didik dan terus melakukan perbaikan. Dengan demikian, portofolio dapat memperlihatkan perkembangan kemajuan belajar peserta didik melalui karyanya, antara lain karangan, puisi, surat, komposisi musik, gambar, foto, lukisan, resensi buku atau literatur, laporan penelitian, sinopsis, dsb. Hal-hal yang perlu diperhatikan dan dijadikan pedoman dalam penggunaan penilaian portofolio di sekolah, antara lain karya siswa adalah benar-benar karya peserta didik itu sendiri.

Guru melakukan penelitian atas hasil karya peserta didik yang dijadikan bahan penilaian portofolio agar karya tersebut merupakan hasil karya yang dibuat oleh peserta didik itu sendiri. Saling percaya antara guru dan peserta didik. Dalam proses penilaian guru dan peserta didik harus memiliki rasa saling percaya, saling memerlukan dan saling membantu sehingga terjadi proses pendidikan berlangsung dengan baik, kerahasiaan bersama antara guru dan peserta didik. Kerahasiaan hasil pengumpulan informasi perkembangan peserta didik perlu dijaga dengan baik dan tidak disampaikan kepada pihak-pihak yang tidak berkepentingan sehingga memberi dampak negatif proses pendidikan, milik bersama “*joint ownership*” antara peserta didik dan guru. Guru dan peserta didik perlu mempunyai rasa memiliki berkas portofolio sehingga peserta didik akan merasa memiliki karya yang dikumpulkan dan akhirnya akan berupaya terus meningkatkan kemampuannya. Kepuasan hasil kerja portofolio sebaiknya berisi keterangan dan atau bukti yang memberikan dorongan peserta didik untuk lebih meningkatkan diri. Kesesuaian hasil kerja yang dikumpulkan adalah hasil kerja yang sesuai dengan kompetensi yang tercantum dalam kurikulum.

Penilaian proses dan hasil penilaian portofolio menerapkan prinsip proses dan hasil. Proses belajar yang dinilai misalnya didapatkan dari catatan dokumen guru tentang kinerja dan karya peserta didik. Penilaian

dan pembelajaran penilaian portofolio merupakan hal yang tak terpisahkan dari proses pembelajaran. Manfaat utama penilaian ini sebagai diagnostik yang sangat berarti bagi guru untuk melihat kelebihan dan kekurangan peserta didik. Teknik penilaian portofolio di dalam kelas memerlukan langkah-langkah sebagai berikut:

- a) Jelaskan kepada peserta didik bahwa penggunaan portofolio, tidak hanya merupakan kumpulan hasil kerja peserta didik yang digunakan oleh guru untuk penilaian, tetapi digunakan juga oleh peserta didik sendiri. Dengan melihat portofolionya peserta didik dapat mengetahui kemampuan, keterampilan, dan minatnya. Proses ini tidak akan terjadi secara spontan, tetapi membutuhkan waktu bagi peserta didik untuk belajar meyakini hasil penilaian mereka sendiri.
- b) Tentukan bersama peserta didik sampel-sampel portofolio apa saja yang akan dibuat. Portofolio antara peserta didik yang satu dan yang lain bisa sama bisa berbeda. Misalnya, untuk kemampuan menulis peserta didik mengumpulkan karangan-karangannya. Sedangkan untuk kemampuan menggambar, peserta didik mengumpulkan gambar-gambar buatannya.
- c) Kumpulkan dan simpanlah karya-karya tiap peserta didik dalam satu map atau folder di rumah masing-masing atau loker masing-masing di sekolah.
- d) Berilah tanggal pembuatan pada setiap bahan informasi perkembangan peserta didik sehingga dapat terlihat perbedaan kualitas dari waktu ke waktu.
- e) Sebaiknya tentukan kriteria penilaian sampel portofolio dan bobotnya dengan para peserta didik sebelum mereka membuat karyanya. Diskusikan cara penilaian kualitas karya para peserta didik. Contoh kriteria penilaian kemampuan menulis karangan yaitu penggunaan tata bahasa, pemilihan kosakata, kelengkapan gagasan, dan sistematika penulisan. Dengan demikian, peserta didik mengetahui harapan “standar” guru dan berusaha mencapai standar tersebut. Minta peserta didik menilai karyanya secara berkesinambungan. Guru dapat membimbing peserta didik, bagaimana cara menilai dengan memberi keterangan tentang

kelebihan dan kekurangan karya tersebut, serta bagaimana cara memperbaikinya. Hal ini dapat dilakukan pada saat membahas portofolio.

- f) Setelah suatu karya dinilai dan nilainya belum memuaskan, maka peserta didik diberi kesempatan untuk memperbaiki. Namun, antara peserta didik dan guru perlu dibuat (kontrak) atau perjanjian mengenai jangka waktu perbaikan, misalnya dua minggu karya yang telah diperbaiki harus diserahkan kepada guru. Bila perlu, jadwalkan pertemuan untuk membahas portofolio. Jika perlu, undang orang tua peserta didik dan diberi penjelasan tentang maksud serta tujuan portofolio, sehingga orang tua dapat membantu dan memotivasi anaknya. Penilaian diri "*self assessment*" adalah suatu teknik penilaian di mana peserta didik diminta untuk menilai dirinya sendiri berkaitan dengan status, proses, dan tingkat pencapaian kompetensi yang dipelajarinya dalam mata pelajaran tertentu didasarkan atas kriteria atau acuan yang telah disiapkan. Tujuan utama dari penilaian diri adalah untuk mendukung atau memperbaiki proses dan hasil belajar. Meskipun demikian, hasil penilaian diri dapat digunakan guru sebagai bahan pertimbangan untuk memberikan nilai. Peran penilaian diri menjadi penting bersamaan dengan bergesernya pusat pembelajaran dari guru ke siswa yang didasarkan pada konsep belajar mandiri "*autonomous learning*".

Ada beberapa jenis penilaian diri, di antaranya, penilaian langsung dan spesifik, yaitu penilaian secara langsung, pada saat atau setelah selesai melakukan tugas, untuk menilai aspek-aspek kompetensi tertentu dari suatu mata pelajaran, penilaian tidak langsung dan holistik, yaitu penilaian yang dilakukan dalam kurun waktu yang panjang, untuk memberikan penilaian secara keseluruhan. Penilaian sosio-afektif, yaitu penilaian terhadap unsur-unsur afektif atau emosional. Misalnya, peserta didik dapat diminta untuk membuat tulisan yang memuat curahan perasaannya terhadap suatu objek tertentu. Penggunaan teknik ini dapat memberi dampak positif terhadap perkembangan kepribadian seseorang. Kelebihan pada penggunaan evaluasi diri di dalam kelas maka akan menimbulkan rasa percaya diri peserta didik, karena mereka dipercaya untuk dapat

menilai dirinya sendiri, dan peserta didik menyadari kekuatan dan kelemahan dirinya, karena ketika mereka melakukan penilaian, mesti melakukan introspeksi terhadap kelebihan dan kekurangan yang dimilikinya, yang akan membiasakan diri peserta didik untuk berlaku jujur. Sehingga bisa mendorong, membiasakan, serta melatih peserta didik untuk berlaku jujur, karena mereka diminta untuk jujur dan objektif di setiap melakukan penilaian pada diri mereka masing-masing. Pada penentuan kualitas instrumen nontes yang telah digunakan maka pada instrumen nontes yang dilakukan mesti melengkapi data yang dilihat dari kriteria kualitas dan kuantitas instrumen seperti reliabilitas, validitas, serta daya pembedanya. Karena instrumen nontes tidak ada nilai benar dan salah jika tidak ada jawaban responden. Maka tidak perlu mengukur tingkat kesukaran soal seperti pada instrumen tes.

1. Reliabilitas teknik penentuan reliabilitas dapat menggunakan tes *retest*, tes paralel, tes belah dua. Karena rumus yang bisa menjadi pedoman antara lain *korelasi product moment*, *Sperman Brown*, *Alpha*, dll. Dan reliabilitas instrumen ini didapat dari besarnya koefisien korelasi, koefisien reliabilitas memperlihatkan hubungan skor skala yang diperoleh x dengan skor sesungguhnya “*skor murni*”. Koefisien reliabilitas yang menghasilkan nilai angka 0,9 maka dapat diartikan skala yang bisa memperlihatkan besaran angka 90% dari variansi skor murni
2. Validitas atau kevalidan instrumen nontes yang digunakan dan dikembangkan maka dapat dilihat dari validitas isi dan konstruk teori yang mendasarinya. Maka ketepatan dalam menentukan definisi konseptual, definisi operasional, dan penetapan aspek serta indikator yang direncanakan. Faktor yang dapat melemahkan validitas, serta identifikasi kawasan ukur tidak cukup jelas Operasionalisasi secara konsepnya “*perumusan indikator*” tidak tepat penulisan jika butir nontes tidak mengikuti kaidah, administrasi skala yang tidak hati-hati “*kondisi subjek, kondisi testing*”, pemberian skor tidak cermat, dan interpretasi yang keliru.
3. Daya beda pertanyaan atau pernyataan pada instrumen nontes harus dapat membedakan sikap positif dan negatif. Indeks daya beda juga merupakan indikator keselarasan atau konsistensi antara fungsi butir

dengan fungsi skala secara keseluruhan. Rumus yang dapat digunakan antara lain skor interval, korelasi (*product moment Pearson*).

DAFTAR PUSTAKA

- Al Imam Albukhari Terjemahan Bukhari
- Al Qur'an dan terjemahan serta kitab-kitab hadis
- Analisis Tingkat Kesukaran Daya Pembeda dan Fungsi Soal Ujian. Jurnal 2015
- Analisis Butir Soal Jurnal Tajdid 2012
- Arifin, Zainal. *Evaluasi Instruksional*. Bandung: PT. Remaja Rosdakarya.1990.
- Arikunto, Suharsimi. *Dasar-Dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara. 2005.
- Asrul, dkk. *Evaluasi Pembelajaran*. Bandung: Cita pustaka media. 2014.
- Bukhari Umar, *Hadis Tarbawi* 2012
- _____. *Hadis Tarbawi*. Jakarta: Grafika Offset. 2012
- Djaali & Mulyono. *Pengukuran dalam Bidang Pendidikan*. Grasindo. Jakarta. 2008.
- Gronlund, N.E. *Measurement and Evaluationin Testing* 5 th Ed. New York: Macmillan Publising Co.Inc.
- Jurnal Chemical Volume 13 Tahun 2012
- Kadar M. Yusuf. *Tafsir Tarbawi Pesan-Pesan Al Qur'an Tentang Pendidikan*.
- Mania, Sitti. *Teknik Non Tes: Telaah atas Fungsi Wawancara dan Kuesioner dalam Evaluasi Pendidikan*, Lentera Pendidikan Vol. 11 No. 1 JUNI 2008.
- Menyusun dan Menganalisa Tes Hasil Belajar Jurnal 2015
- Mulyadi. *Evaluasi Pendidikan Pengembangan Model Evaluasi Pendidikan di Sekolah*. Malang: UIN Maliki Press. 2010.

- Nana, Sudjana. *Penilaian Hasil Proses Belajar Mengajar*. Bandung: PT. Remaja Rosdakarya. 1995.
- Nasution. *Metode Research*, Jakarta: Bumi Aksara. 2000.
- Ngalim, Purwanto. *Prinsip-Prinsip dan Teknik Evaluasi Pengajaran*. Bandung: PT. Remaja Rosdakarya. 1987.
- Nurwadjah Ahmad. *Tafsir Ayat-Ayat Pendidikan 2007*
- Popham, W. James, *Classroom Assessment*. Boston: Allyn & Bacon. 1996.
- Sudijono, Anas. *Pengantar Evaluasi Pendidikan*. Jakarta: PT RajaGrafindo Persada. 2007.
- Sukirman. *Tes Kemampuan Keterampilan Menulis dalam Pembelajaran Bahasa Indonesia di Sekolah*. (Institut Agama Islam Negeri Palopo: Jurnal Konsepsi, Vol. 9, No. 2, Agustus 2020).
- Sumadi. *Pengembangan Tes Hasil Belajar*. Jakarta: Rajawali. 1987.
- Sumarna. *Pedoman Pengembangan Penilaian Portofolio*. Jakarta: Puspendik Balitbang Depdiknas. 2006.
- Sunah Abu Daud Jus 3
- T. Raka, Joni. *Pengukuran dan Penilaian Pendidikan*. (Surabaya: Karya Anda. 1984.
- Uni Salamah, Penjaminan Mutu Penilaian. *Jurnal Evaluasi Volume 2 Tahun 2018*
- Willeiam, A. Mohrens. *Measurement and Evaluation in Education and Psychology*. New York: Rinchart and Wionston. 1984.
- Zainul, Asmawi dan Nasution N. *Penilaian Hasil Belajar*. Jakarta: PAU Direktorat Jenderal Pendidikan Tinggi, Departemen P dan K. 1993.

TENTANG PENULIS



Nama : Dr. Rahmi, MA
Tempat Mengajar : 1. Pasca UM Sumbar
2. Dosen Luar Biasa di UNAND
3. IAIN Bukittinggi
NIDN : 1020018804
Tempat & Tanggal Lahir : Kampung Kering, 20 Januari 1988
Jenis Kelamin : Perempuan
Suku : Melayu
Agama : Islam
Golongan/Pangkat : 3C
Jabatan Akademik : Lektor
Perguruan Tinggi : UMSB
Alamat : Pasir Kandang Padang
Telp./HP/Fax : 081365010405
Alamat Email : drrahmi15@gmail.com

RIWAYAT PENDIDIKAN PERGURUAN TINGGI			
Tahun Lulus	Program Pendidikan (Diploma, Sarjana, Magister, Spesialis dan Doktor)	Perguruan Tinggi	Jurusan/Program Studi
2018	Doktor	UIN Imam Bonjol	Pendidikan Islam
2014	Magister	UMSB	Pendidikan Islam

RIWAYAT PENDIDIKAN PERGURUAN TINGGI			
Tahun Lulus	Program Pendidikan (Diploma, Sarjana, Magister, Spesialis dan Doktor)	Perguruan Tinggi	Jurusan/Program Studi
2012	Sarjana	IAIN Imam Bonjol	Pendidikan Bahasa Arab
	dst		



Prof. Dr. Martin Kustati is an English lecturer at the Faculty of Tarbiyah and Teacher Training, UIN Imam Bonjol Padang. She received her Doctoral degree from the Faculty of Education, Universiti Kebangsaan Malaysia. Her ID Scopus, SINTA ID, google scholars ID are 57204519113, 257234, and ZjqIIfMAAAAJ. She was a recipient of a national outstanding learner achievement award from the Department of Education of the Republic of Indonesia and also a recipient of a Southeast Asia fellowship at the Asia Research Institute NUS Singapore. She currently becomes a member of ASWGI (Association of the Indonesian Women and Gender Studies), ICMI, TEFLIN, ADRI, and ELITERS. She is also the editor of Kafa'ah journal: Journal of Gender Studies and Editor-in-Chief of Al-ta'lim journal in UIN Imam Bonjol Padang, West Sumatra. Her works on TEFL, TESL, Semantics-Pragmatics, Language Learning Strategies, Research Methodology, and Gender Studies. She is a rector of UIN Imam Bonjol Padang, West Sumatra.

E-mail: martinkustati@uinib.ac.id

Orcid number: 0000-0002-2221-0127

Phone: (+62)8126713546



Saya, Hadeli (Drs. M.A., M.Pd.), lahir di Bukit Tinggi, Sumatera Barat pada tanggal 1 Februari 1966. Ayah bernama Bermawi Dt. Majo Basa (almarhum) dan Ibu bernama Zulfidar Nurdin (almarhumah). Setelah menamatkan Sekolah Dasar di kampung halaman pada tahun 1977, pada tahun 1978 sd. 1984 saya melanjutkan pendidikan pada Madrasah Tsanawiyah Negeri dan Madrasah Aliyah Negeri Koto Baru, Padang Panjang. Pada tahun 1985 saya melanjutkan pendidikan pada Jurusan Tadris Pendidikan Bahasa Inggris Fakultas Tarbiyah dan Keguruan IAIN Imam Bonjol Padang, tamat tahun 1989.

Pada tahun 1991 saya mengikuti IAIN *Lecturer Pre-Departure Training Programme for Overseas Studies* di UIN Syahid, Ciputat, Tangerang, Banten. Setelah memperoleh sertifikat *TOEFL*, sertifikat *IELTS*, *Letter of Acceptance*, dan beasiswa dari *FCO British Council*, pada pertengahan tahun 1992 saya berangkat melanjutkan pendidikan pada *Centre for Applied Research in Education (CARE)*, *School of Education*, *University of East Anglia (UEA)*, *Norwich, Norfolk, England, United Kingdom* dan mendapatkan gelar *Master of Art (MA)* dalam bidang *Applied Research in Education* pada tahun 1994. Pada tahun yang sama (1994) saya menikah dengan Hj. Afifah, S.E., M.M. yang saat ini bekerja pada Direktorat Jendral Penyelenggaraan Haji dan Umrah, Kementerian Agama RI. Magister ke dua (M.Pd.) saya peroleh dari Program Pascasarjana (PPs), Universitas Negeri Padang (UNP), Program Studi Pendidikan Bahasa, Konsentrasi Pendidikan Bahasa Inggris pada tahun 2008.

Terhitung sejak tanggal 1 Maret 1992, saya diangkat sebagai Calon Pegawai Negeri Sipil (CPNS) dan selanjutnya diangkat sebagai Dosen Tetap pada Fakultas Tarbiyah dan Keguruan IAIN Imam Bonjol Padang. Beberapa jabatan yang pernah dipercayakan kepada saya adalah: 1) Staf Ahli Rektor IAIN Imam Bonjol Padang 1995-1998, 2) Kepala Bagian Tata Usaha (Kabag TU) PPs IAIN Imam Bonjol Padang 1998-2000, dan 3) Wakil Dekan Bidang Akademik Fakultas Tarbiyah dan Keguruan IAIN Imam Bonjol Padang 2012-2014.

Evaluasi merupakan bagian yang sangat penting dalam komponen pendidikan, karena untuk melihat mutu pendidikan suatu bangsa dapat dilihat dari hasil evaluasinya. Evaluasi dan komponen pendidikan lainnya sangat berpengaruh dalam kondisi politik suatu negara dan juga berpengaruh terhadap tujuan suatu negara.

Jika pembelajaran merupakan proses transfer ilmu dari pendidik kepada peserta didik yang bertujuan adanya perubahan dan pendewasaan dari diri peserta didik, maka evaluasi merupakan komponen yang akan menjadi penambah dan pengubah perilaku peserta didik.

Buku ini terdiri atas bab: Konsep Evaluasi Pendidikan Perspektif Islam, Teknik Penyusunan dan Pelaksanaan Hasil Belajar, Teknik Penganalisisan *Item* Tes Hasil Belajar, Teknik Tes dan Teknik Nontes sebagai Alat Evaluasi Hasil Belajar, Teknik Pengujian Validitas Tes dan Validitas *Item* Hasil Belajar, dan Contoh-Contoh Pengolahan Evaluasi Tes dan Nontes.

EVALUASI PENDIDIKAN PERSPEKTIF ISLAM

Penerbit Deepublish (CV BUDI UTAMA)
Jl. Kaliurang Km 9,3 Yogyakarta 55581
Telp/Fax : (0274) 4533427
Anggota IKAPI (076/DIY/2012)

✉ cs@deepublish.co.id
📘 Penerbit Deepublish
@ @penerbitbuku_deepublish
🌐 www.penerbitdeepublish.com



Kategori : Islam dan Pendidikan

ISBN 978-623-02-4381-3



9 786230 243813